LOWER PREVISIONS INDUCED BY FILTER MAPS

GERT DE COOMAN AND ENRIQUE MIRANDA

ABSTRACT. We investigate under which conditions a transformation of an imprecise probability model of a certain type (coherent lower previsions, n-monotone capacities, minitive measures) produces a model of the same type. We give a number of necessary and sufficient conditions, and study in detail a particular class of such transformations, called filter maps. These maps include as particular models multi-valued mappings as well as other models of interest within imprecise probability theory, and can be linked to filters of sets and $\{0,1\}$-valued lower probabilities.

Keywords: Coherent lower previsions, n-monotone capacities, minitive measures, filters, Choquet integral.

1. INTRODUCTION

In many practical problems, it is not uncommon to encounter situations with vague or imprecise knowledge about the probabilistic information of the variables involved; this could be due to deficiencies in the observational process or conflicts between the opinions of several experts, amongst other things. In such cases, it may advisable to consider imprecise probability models as a more robust alternative to the classical models based on probability measures. These models include as particular cases sets of probability measures [14], coherent lower previsions [19], n-monotone capacities [1], and necessity measures [11].

All these models are mathematically related to each other: sets of probability measures and coherent lower previsions are equivalent, and they include n-monotone capacities as special cases. Necessity measures are in particular n-monotone for any natural number n.

Interestingly, a transformation of such models—under a permutation of the possibility space, or more generally, under another map that connects two possibility spaces—need not preserve their character. To give an example, it is not uncommon that a transformation of a necessity measure goes beyond the framework of necessity measures, and produces only a coherent lower prevision (probability). In this paper, we investigate under which conditions a transformation of the possibility space, or of the associated space of real-valued functions defined on it, preserves different consistency notions: avoiding sure loss, coherence, n-monotonicity or being minimum preserving.

We show that for the first two we must consider so-called *coherence preserving* transformations, which are closely related to (but as the same time more general than) coherent lower previsions. We investigate their properties in Sections 3 and 4. We show in Section 7 that as particular cases of coherence preserving mappings, we have the notion of probability induced by a random variable, but also the Markov operators considered in [17]. For the last two conditions, we show in Sections 5 and 6 that a conditional model preserves n-monotonicity in addition to coherence if and only if it it is a $\wedge$-homomorphism, i.e. minimum preserving. We characterise minimum preserving coherent lower previsions, and show that they are in a one-to-one correspondence with filters of subsets of the possibility space.

This leads us to the second part of the paper, where we study in detail a particular type of mappings, called *filter maps*, which we introduce in Section 8. They are maps that assign to any element of an initial space a filter of subsets of a final space. But they can also be interpreted as minimum preserving transformations. We show that they include as particular cases the lower previsions induced by multi-valued mappings as well as lower oscillation models, and we study the properties of the lower prevision that results from combining a filter map with a lower prevision on the initial space.

2. PRELIMINARIES

2.1. **Coherent lower previsions.** We begin with a brief discussion of coherent lower previsions. We refer to [19] for more details and background, and for the proofs of all results mentioned in this section.

Consider a possibility space $\mathscr{X}$. A *gamble* f on $\mathscr{X}$ is any bounded real-valued map on $\mathscr{X}$. This is for instance the case for the indicator I_A of a subset A of $\mathscr{X}$, where $I_A(x)$ takes the value 1 when $x \in A$ and 0 when $x \notin A$.

The set of all gambles on $\mathscr{X}$ is denoted by $\mathscr{G}(\mathscr{X})$. This is a linear space: it is closed under point-wise addition and multiplication by real numbers. It is moreover a *lattice*, that is, closed under point-wise maxima $\vee$ and minima $\wedge$. We will be interested in some particular transformations between sets of gambles.

Definition 1. Given two spaces $\mathscr{X}$ and $\mathscr{Y}$, a map $r\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ is called a $\wedge$-*homomorphism* when $r(f_1 \wedge f_2) = r(f_1) \wedge r(f_2)$ for all $f_1, f_2 \in \mathscr{G}(\mathscr{X})$.

A real-valued map defined on a set of gambles is called a lower prevision. It can be given a behavioural interpretation: the lower prevision of a gamble f is the supremum price μ such that the transaction $f - \mu$ is desirable for a given subject. This interpretation lies at the basis of the following definitions:

Definition 2. A *lower prevision* on a set of gambles $\mathscr{K} \subseteq \mathscr{G}(\mathscr{X})$ is a functional $\underline{P}\colon \mathscr{K} \to \mathbb{R}$. It is said to *avoid sure loss* when for every $n \in \mathbb{N}_0$ and $f_1, \dots, f_n \in \mathscr{K}$:

$$\sum_{i=1}^{n} \underline{P}(f_i) \leq \sup \left[\sum_{i=1}^{n} f_i \right],$$

where $\mathbb{N}_0$ denotes the set of natural numbers (with zero). It is called *coherent* if and only if for all numbers $n, m \in \mathbb{N}_0$ and $f_0, f_1, \dots, f_n \in \mathscr{K}$:

$$\sum_{i=1}^{n} \underline{P}(f_i) - m\underline{P}(f_0) \leq \sup \left[\sum_{i=1}^{n} f_i - m f_0 \right].$$

One example of coherent lower previsions are the so-called *vacuous* ones. For any non-empty subset A of $\mathscr{X}$, the vacuous lower prevision relative to A is defined on $\mathscr{G}(\mathscr{X})$ as

$$\underline{P}_A(f) := \inf_{x \in A} f(x) \text{ for all } f \in \mathscr{G}(\mathscr{X}).$$

A lower prevision $\underline{P}$ on the set $\mathscr{G}(\mathscr{X})$ of all gambles turns out to be coherent if and only if:

C1. $\underline{P}(f) \geq \inf f$ for all gambles f on $\mathscr{X}$;
C2. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ for all gambles f on $\mathscr{X}$ and all real $\lambda \geq 0$;
C3. $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$ for all gambles f and g on $\mathscr{X}$.

On the other hand, a coherent lower prevision defined on indicators of events only is called a *coherent lower probability*. To simplify the notation, we will sometimes use the same symbol A to denote a set A and its indicator I_A, so we will write $\underline{P}(A)$ instead of $\underline{P}(I_A)$. Of particular interest are the $\{0,1\}$-valued coherent lower probabilities, which are related to filters of events:

Definition 3. Let $\mathscr{P}(\mathscr{X})$ denote the power set of a space $\mathscr{X}$. A subset $\mathscr{F}$ of $\mathscr{P}(\mathscr{X})$ is called a (proper) *filter* when it satisfies the following properties:

F1. $\emptyset \notin \mathscr{F}$;
F2. if $A, B \in \mathscr{F}$ then $A \cap B \in \mathscr{F}$;
F3. if $A \in \mathscr{F}$ and $A \subseteq B$ then $B \in \mathscr{F}$.

A filter $\mathscr{F}$ is called *fixed*[1] if there is a non-empty subset A of $\mathscr{X}$ such that $B \in \mathscr{F}$ if and only if $A \subseteq B$, and it is called *free* otherwise.

If $\underline{P}$ is a coherent lower probability on $\mathscr{P}(\mathscr{X})$ that takes only the values 0 and 1, then the class $\{A \subseteq \mathscr{X} : \underline{P}(A) = 1\}$ forms a filter; and, conversely, for any filter $\mathscr{F}$ of $\mathscr{P}(\mathscr{X})$, the lower probability $\underline{P}$ given by $\underline{P}(A) = 1$ if A belongs to $\mathscr{F}$, and 0 otherwise, is coherent. We denote by $\mathbb{F}(\mathscr{X})$ the set of all filters over the space $\mathscr{X}$.

If a lower prevision $\underline{P}$ avoids sure loss, then we can 'correct' it into a coherent lower prevision. This is done by means of the procedure of *natural extension*, whose main properties are summarised in the following theorem:

Theorem 1. [19] *Let $\underline{P}$ be a lower prevision with domain $\mathscr{K}$ that avoids sure loss. Let $\underline{E}_{\underline{P}}$ be the lower prevision on the set $\mathscr{G}(\mathscr{X})$ of all gambles given by*

$$\underline{E}_{\underline{P}}(f) := \sup\left\{\mu \in \mathbb{R} \colon f - \mu \geq \sum_{j=1}^{n} \lambda_j (f_j - \underline{P}(f_j)) \text{ for some } n \in \mathbb{N}, \lambda_j > 0, f_j \in \mathscr{K}\right\}, \tag{1}$$

where $\mathbb{N}$ is the set of natural numbers (without zero).

(i) *$\underline{E}_{\underline{P}}$ is a coherent lower prevision.*
(ii) *$\underline{E}_{\underline{P}}$ is the smallest coherent lower prevision that satisfies $\underline{E}_{\underline{P}}(f) \geq \underline{P}(f)$ for all $f \in \mathscr{K}$.*
(iii) *$\underline{P}$ is coherent if and only if $\underline{E}_{\underline{P}}(f) = \underline{P}(f)$ for all $f \in \mathscr{K}$.*

Another particular case of coherent lower previsions with domain $\mathscr{G}(\mathscr{X})$ are the *linear* previsions:

Definition 4. A lower prevision $\underline{P}$ defined on the set $\mathscr{G}(\mathscr{X})$ of all gambles is called a *linear prevision* when it satisfies $\underline{P}(f+g) = \underline{P}(f) + \underline{P}(g)$ for every pair of gambles f, g, and moreover $\underline{P}(f) \geq \inf f$ for every gamble f.

A linear prevision coincides with the expectation functional associated with its restriction to (indicators of) events, which is a finitely additive probability. Moreover, a lower prevision $\underline{P}$ defined on the set $\mathscr{G}(\mathscr{X})$ of all gambles is coherent if and only if

$$\underline{P}(f) = \min\{P(f) \colon P \in \mathbb{M}(\underline{P})\}, \tag{2}$$

where

$$\mathbb{M}(\underline{P}) := \{P \geq \underline{P} \colon P \text{ linear prevision}\}$$

is the *credal set* associated with $\underline{P}$. In other words, a coherent lower prevision $\underline{P}$ is always the lower envelope of the set $\mathbb{M}(\underline{P})$ of all linear previsions that dominate it, and as such it can be seen as a lower expectation functional. Moreover, a lower prevision $\underline{P}$ with domain $\mathscr{K}$ avoids sure loss if and only if $\mathbb{M}(\underline{P}) \neq \emptyset$, and in that case its natural extension $\underline{E}_{\underline{P}}$ can be obtained by taking the lower envelope of $\mathbb{M}(\underline{P})$:

$$\underline{E}_{\underline{P}}(f) = \min\{P(f) \colon P \in \mathbb{M}(\underline{P})\}, \tag{3}$$

and $\mathbb{M}(\underline{P}) = \mathbb{M}(\underline{E}_{\underline{P}})$.

In particular, the vacuous lower prevision relative to a set A is the lower envelope of the set of linear previsions P that assign probability 1 to A:

$$\underline{P}_A(f) = \min\{P(f) \colon P(A) = 1 \text{ and } P \text{ linear prevision}\}, \tag{4}$$

and $\mathbb{M}(\underline{P}_A) = \{P \text{ linear prevision} \colon P(A) = 1\}$.

[1] Sometimes fixed filters are called *principal* or *degenerate*.

2.2. n**-Monotone lower previsions.** A particular case of coherent lower previsions are those called n-monotone:

Definition 5. Let $n \in \mathbb{N}$. A lower prevision defined on the set $\mathscr{G}(\mathscr{X})$ of all gambles is *n-monotone* if for all $p \in \mathbb{N}$, $p \leq n$, and all $f, f_1, \ldots, f_p$ in $\mathscr{G}(\mathscr{X})$:

$$\sum_{I \subseteq \{1,\ldots,p\}} (-1)^{|I|} \underline{P}\left(f \wedge \bigwedge_{i \in I} f_i\right) \geq 0,$$

where, as before, $\wedge$ is used to denote the point-wise minimum.

n-monotone lower previsions have been studied in detail in [8, 18], to which we refer for more details and background, and for the proofs of all results mentioned in this section.

When $n \geq 2$, an n-monotone lower prevision $\underline{P}$ is in particular coherent provided that $\underline{P}(0) = 0$ and $\underline{P}(1) = 1$, and it corresponds to the Choquet integral with respect to its restriction to events, which is called an *n-monotone lower probability*. A lower prevision that is n-monotone for all natural n is called *∞-monotone* or *completely monotone*. This is for instance the case for the vacuous lower prevision $\underline{P}_A$ considered above, and more generally for all minimum or infimum preserving lower previsions:

Definition 6. A lower prevision defined on the set $\mathscr{G}(\mathscr{X})$ of all gambles is *minitive* when $\underline{P}(f \wedge g) = \underline{P}(f) \wedge \underline{P}(g)$ for every pair of gambles f, g on $\mathscr{X}$, and *infimum preserving* when $\underline{P}(\wedge_{i \in I} f_i) = \wedge_{i \in I} \underline{P}(f_i)$ for any non-empty family $\{f_i \colon i \in I\}$ of gambles on $\mathscr{X}$. Similarly, a lower probability on $\mathscr{P}(\mathscr{X})$ is called *minitive* when $\underline{P}(A \cap B) = \min\{\underline{P}(A), \underline{P}(B)\}$ for all $A, B \subseteq \mathscr{X}$, and a *necessity measure* when $\underline{P}(\cap_{i \in I} A_i) = \inf_{i \in I} \underline{P}(A_i)$ for any non-empty family $\{A_i \colon i \in I\}$ of subsets of $\mathscr{X}$.

Any linear prevision P on $\mathscr{G}(\mathscr{X})$ is always completely monotone, but it need not be minitive.

2.3. **Conditional lower previsions.** We may also consider lower previsions in a conditional context. Within the framework of this paper, we consider two spaces $\mathscr{X}$ and $\mathscr{Y}$, and call *conditional lower prevision* any functional $\underline{P}(\cdot|Y)$ defined on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ that to any gamble f on $\mathscr{X} \times \mathscr{Y}$ assigns a gamble $\underline{P}(f|Y)$ on $\mathscr{Y}$, where the value that $\underline{P}(f|Y)$ assumes in $y \in \mathscr{Y}$ is denoted by $\underline{P}(f|y)$ and called the conditional lower prevision of f given that $Y = y$.

Definition 7. A conditional lower prevision $\underline{P}(\cdot|Y)$ on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ is called *separately coherent* when it satisfies the following conditions:

SC1. $\underline{P}(f|y) \geq \inf_{x \in \mathscr{X}} f(x, y)$,
SC2. $\underline{P}(\lambda f|y) = \lambda \underline{P}(f|y)$,
SC3. $\underline{P}(f + g|y) \geq \underline{P}(f|y) + \underline{P}(g|y)$,

for any gambles $f, g \in \mathscr{G}(\mathscr{X} \times \mathscr{Y})$, all real $\lambda \geq 0$ and all $y \in \mathscr{Y}$.

If $\underline{P}(\cdot|Y)$ is a separately coherent conditional lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ and $\underline{P}$ is a coherent lower prevision on $\mathscr{G}(\mathscr{Y})$, their *marginal extension* is the coherent lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ given by $\underline{P}(\underline{P}(\cdot|Y))$. It corresponds to the smallest lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ that is coherent with both $\underline{P}$ and $\underline{P}(\cdot|Y)$, in the sense considered in [19]. Marginal extension generalises the Law of Total Probability to lower previsions.

3. Coherence preserving transformations

In this section, we investigate how to transform a lower prevision on some set of gambles over a space $\mathscr{Y}$ into a lower prevision on gambles defined on another space $\mathscr{X}$ while still keeping most of its properties.

Definition 8. Let $\underline{T} \colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ be a transformation between two sets of gambles. It is called *coherence preserving* when it satisfies the following properties:

T1. $\inf(\underline{T}f) \geq \inf f$ for all $f \in \mathscr{G}(\mathscr{X})$;
T2. $\underline{T}(\lambda f) = \lambda \underline{T}f$ for all $f \in \mathscr{G}(\mathscr{X})$ and all real $\lambda \geq 0$;
T3. $\underline{T}(f+g) \geq \underline{T}f + \underline{T}g$ for all $f, g \in \mathscr{G}(\mathscr{X})$.

A coherence preserving transformation $\underline{T}$ automatically satisfies the following properties:

T4. if $f \leq g$ then $\underline{T}f \leq \underline{T}g$ for all $f, g \in \mathscr{G}(\mathscr{X})$;
T5. $\underline{T}(f+\mu) = \underline{T}f + \mu$ for all $f \in \mathscr{G}(\mathscr{X})$ and all real μ.

The interest of this type of transformations lies in the following result:

Proposition 2. *Let $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ be a transformation between two sets of gambles. Then the following statements are equivalent:*

(i) *$\underline{T}$ is coherence preserving;*
(ii) *for every coherent lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$, the composition $\underline{P} \circ \underline{T}$ is a coherent lower prevision on $\mathscr{G}(\mathscr{X})$.*

Moreover, if $\underline{T}$ is coherence preserving, and the lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$ avoids sure loss, then the lower prevision $\underline{P} \circ \underline{T}$ on $\mathscr{G}(\mathscr{X})$ avoids sure loss as well.

Proof. Let us begin the proof of the equivalence by showing that the first statement implies the second. It suffices to prove that $\underline{P} \circ \underline{T}$ satisfies conditions C1–C3.

C1. For any gamble f in $\mathscr{G}(\mathscr{X})$, $(\underline{P} \circ \underline{T})(f) = \underline{P}(\underline{T}f) \geq \underline{P}(\inf \underline{T}f) \geq \underline{P}(\inf f) = \inf f$, where the first inequality follows from the coherence of $\underline{P}$ and the second from T1.

C2. Given f in $\mathscr{G}(\mathscr{X})$ and $\lambda > 0$, $\underline{P}(\underline{T}(\lambda f)) = \underline{P}(\lambda \underline{T}f) = \lambda \underline{P}(\underline{T}f)$, where the first equality follows from T2 and the second from the coherence of $\underline{P}$.

C3. Given $f, g \in \mathscr{G}(\mathscr{X})$, $\underline{P}(\underline{T}(f+g)) \geq \underline{P}(\underline{T}f + \underline{T}g) \geq \underline{P}(\underline{T}f) + \underline{P}(\underline{T}g)$, where the first inequality follows from T3 and the second from the coherence of $\underline{P}$.

Conversely, assume that the second statement holds, consider any $y \in \mathscr{Y}$, and take the degenerate linear prevision $\underline{P}_{\{y\}}$, given by $\underline{P}_{\{y\}}(g) \coloneqq g(y)$ for all gambles g on $\mathscr{Y}$. Then $\underline{P}_{\{y\}} \circ \underline{T}$ is a coherent lower prevision by assumption, and $(\underline{P}_{\{y\}} \circ \underline{T})(f) = (\underline{T}f)(y)$. Let us use this to show that $\underline{T}$ satisfies T1–T3.

T1. Since $\underline{P}_{\{y\}} \circ \underline{T}$ is a coherent lower prevision for all $y \in \mathscr{Y}$, we infer from C1 that $\inf f \leq (\underline{P}_{\{y\}} \circ \underline{T})(f) = (\underline{T}f)(y)$, whence indeed $\inf(\underline{T}f) \geq \inf f$, for any gamble f on $\mathscr{X}$.

T2. Since $\underline{P}_{\{y\}} \circ \underline{T}$ is a coherent lower prevision for all $y \in \mathscr{Y}$, we infer from C2 that $(\underline{T}(\lambda f))(y) = (\underline{P}_{\{y\}} \circ \underline{T})(\lambda f) = \lambda (\underline{P}_{\{y\}} \circ \underline{T})(f) = \lambda (\underline{T}f)(y)$, whence indeed $\underline{T}(\lambda f) = \lambda \underline{T}f$, for any gamble f on $\mathscr{X}$ and any real $\lambda \geq 0$.

T3. Since $\underline{P}_{\{y\}} \circ \underline{T}$ is a coherent lower prevision for all $y \in \mathscr{Y}$, we infer from C3 that $(\underline{T}(f+g))(y) = (\underline{P}_{\{y\}} \circ \underline{T})(f+g) \geq (\underline{P}_{\{y\}} \circ \underline{T})(f) + (\underline{P}_{\{y\}} \circ \underline{T})(g) = (\underline{T}f)(y) + (\underline{T}g))(y)$, whence indeed $\underline{T}(f+g) \geq \underline{T}f + \underline{T}g$, for all gambles f and g on $\mathscr{X}$.

To complete the proof, assume that $\underline{T}$ is coherence preserving and consider a lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$ that avoids sure loss. Then there is a linear prevision P on $\mathscr{G}(\mathscr{Y})$ such that $P \geq \underline{P}$. As a consequence, $P \circ \underline{T} \geq \underline{P} \circ \underline{T}$, and we have just proved that $P \circ \underline{T}$ is a coherent lower prevision, which is in therefore in particular dominated by some linear prevision. This implies that $\underline{P} \circ \underline{T}$ is dominated by some linear prevision, and as a consequence it avoids sure loss. □

A transformation $\underline{T}$ takes any 'vector' f in the linear space $\mathscr{G}(\mathscr{X})$ to a 'vector' $\underline{T}f$ in the linear space $\mathscr{G}(\mathscr{Y})$, whose 'components' are the real numbers $(\underline{T}f)(y)$, $y \in \mathscr{Y}$. This also means that we can interpret $\underline{T}$ as a 'vector map' with as 'component maps' the real functionals $\underline{T}_y\colon \mathscr{G}(\mathscr{X}) \to \mathbb{R}$, $y \in \mathscr{Y}$, defined by

$$\underline{T}_y(f) \coloneqq (\underline{T}f)(y) = (\underline{P}_{\{y\}} \circ \underline{T})(y) \text{ for all } f \in \mathscr{G}(\mathscr{X}). \tag{5}$$

We infer at once from our argumentation in the proof of Proposition 2 that:

Proposition 3. *A transformation $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ is coherence preserving if and only if each of its component maps $\underline{T}_y = \underline{P}_{\{y\}} \circ \underline{T}$, $y \in \mathscr{Y}$ is a coherent lower prevision on $\mathscr{G}(\mathscr{X})$.*

In other words, a coherence preserving transformation $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ can be seen as a family $\underline{T}_y$, $y \in \mathscr{Y}$ of coherent lower previsions on $\mathscr{G}(\mathscr{X})$.

We can use this simple idea to prove a number of basic properties for coherence preserving transformations. The first one is an envelope result akin to the one mentioned for coherent lower previsions in Eq. (2):

Proposition 4. *A transformation $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ is coherence preserving if and only if it is the component-wise lower envelope of a set of coherence preserving linear transformations.*

Proof. 'only if'. Let $\underline{T}$ be a coherence preserving transformation. Then for every $y \in \mathscr{Y}$ the lower prevision $\underline{T}_y$ defined by Eq. (5) is coherent. As a consequence, it is the lower envelope of the set of linear previsions $\mathbb{M}(\underline{T}_y)$. Define now the set $\mathscr{H}$ of transformations

$$\mathscr{H} := \left\{T\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})\colon (\forall y \in \mathscr{Y}) T_y \in \mathbb{M}(\underline{T}_y)\right\},$$

then obviously $\underline{T}$ is the component-wise lower envelope of $\mathscr{H}$. From Proposition 3, any $T \in \mathscr{H}$ is a coherence preserving transformation, because its associated component maps are linear previsions, and therefore in particular coherent lower previsions. To see that T is a linear transformation, note that for any $f, g \in \mathscr{G}(\mathscr{X})$ and all $y \in \mathscr{Y}$:

$$T(f+g)(y) = T_y(f+g) = T_y f + T_y g = (Tf + Tg)(y),$$

and therefore $T(f+g) = Tf + Tg$.

'if'. That $\underline{T}$ is a component-wise lower envelope of coherence preserving linear transformations means that for every $y \in \mathscr{Y}$ the lower prevision $\underline{T}_y$ defined by Eq. (5) is a lower envelope of linear previsions, and is therefore a coherent lower prevision. Now use Proposition 3. □

Next, we show that a number of combinations of coherence-preserving transformations produce another coherence-preserving transformation.

Proposition 5. (i) *Let $\{\underline{T}_i\colon i \in I\}$ be a family of coherence preserving transformations, and let $\underline{T} := \inf_{i\in I} \underline{T}_i$ be their component-wise lower envelope. Then $\underline{T}$ is coherence preserving.*

(ii) *Let $\underline{T}_n$ be a sequence of coherence preserving transformations that converges pointwise towards some $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$. Then $\underline{T}$ is coherence preserving.*

(iii) *Let $\underline{T}_1$ and $\underline{T}_2$ be coherence preserving transformations and let $\alpha \in (0,1)$. Then $\underline{T} := \alpha\underline{T}_1 + (1-\alpha)\underline{T}_2$ is coherence preserving.*

Proof. Taking into account Proposition 3, it suffices to establish the corresponding properties for the coherent component maps. These have been proved in [19, Theorems 2.6.3–2.6.5]. □

In a similar vein, and taking into account the characterisation of coherence-preserving functionals in Proposition 3, we can extend most of the properties established for coherent lower previsions in [19, Section 2] to coherence-preserving functionals.

4. Coherence preserving transformations and conditional lower previsions

Consider a coherence preserving transformation $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$. We have seen in Proposition 3 that its component maps $\underline{T}_y$ are coherent lower previsions, and it follows from the discussion in Section 2.3 that if we let

$$\underline{P}(h|y) = \underline{P}(h(\cdot,y)|y) := \underline{T}_y(h(\cdot,y)) = (\underline{T}h(\cdot,y))(y) \text{ for all } y \in \mathscr{Y} \text{ and } h \in \mathscr{G}(\mathscr{X}\times\mathscr{Y}), \tag{6}$$

then $\underline{P}(\cdot|Y)$ is a conditional lower prevision on $\mathscr{G}(\mathscr{X}\times\mathscr{Y})$ that is separately coherent. Conversely, with a separately coherent conditional lower prevision $\underline{P}(\cdot|Y)$ on $\mathscr{G}(\mathscr{X}\times\mathscr{Y})$, we can associate a transformation $\underline{T}$ defined by

$$(\underline{T}f)(y) := \underline{P}(f|y) \text{ for all } y\in\mathscr{Y} \text{ and } f\in\mathscr{G}(\mathscr{X}) \tag{7}$$

that is coherence preserving. In other words, Proposition 3 can be reformulated to state that there is a one-to-one correspondence between coherence preserving transformations and separately coherent conditional lower previsions, expressed by Eqs. (6) and (7).

Interestingly, if we have a coherent lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$, and a coherent preserving transformation $\underline{T}\colon\mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{Y})$ with associated conditional lower prevision $\underline{P}(\cdot|Y)$ via Eq. (6), then the transformed coherent lower prevision $\underline{P}\circ\underline{T}=\underline{P}(\underline{P}(\cdot|Y))$ is the marginal extension of $\underline{P}$ and $\underline{P}(\cdot|Y)$, as discussed in Section 2.3.

This correspondence points towards an interesting way of extending a transformation $\underline{T}\colon\mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{Y})$ to a transformation $\tilde{\underline{T}}\colon\mathscr{G}(\mathscr{X}\times\mathscr{Y})\to\mathscr{G}(\mathscr{Y})$, defined by

$$(\tilde{\underline{T}}h)(y) := (\underline{T}h(\cdot,y))(y) \text{ for all } y\in\mathscr{Y} \text{ and } h\in\mathscr{G}(\mathscr{X}\times\mathscr{Y}). \tag{8}$$

This simple idea will turn out to be of crucial importance for the discussion further on.

5. $\wedge$-HOMOMORPHISMS

Before we turn to the study of n-monotonicity, let us take a look at the transformations that preserve minitivity.

Proposition 6. *Let $\underline{T}\colon\mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{Y})$ be a transformation between two sets of gambles. Then the following statements are equivalent:*

(i) *$\underline{T}$ is a $\wedge$-homomorphism;*
(ii) *for every minitive lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$, the composition $\underline{P}\circ\underline{T}$ is a minitive lower prevision on $\mathscr{G}(\mathscr{X})$.*

Proof. Let us begin the proof of the equivalence by showing that the first statement implies the second. Given $f,g\in\mathscr{G}(\mathscr{X})$ and a minitive $\underline{P}$, we have indeed that $\underline{P}(\underline{T}(f\wedge g)) = \underline{P}(\underline{T}f\wedge\underline{T}g)=\underline{P}(\underline{T}f)\wedge\underline{P}(\underline{T}g)$, so $\underline{P}\circ\underline{T}$ is minitive too.

Conversely, assume that the second statement holds, consider any $y\in\mathscr{Y}$, and consider the degenerate linear prevision $\underline{P}_{\{y\}}$, given by $\underline{P}_{\{y\}}(g) := g(y)$ for all gambles g on $\mathscr{Y}$. It is clearly minitive, and therefore its component map $\underline{T}_y=\underline{P}_{\{y\}}\circ\underline{T}$ is minitive by assumption. Hence $\underline{T}(f\wedge g)(y)=\underline{T}_y(f\wedge g)=\underline{T}_yf\wedge\underline{T}_yg$, and therefore $\underline{T}(f\wedge g)=\underline{T}f\wedge\underline{T}g$ for all gambles f and g on $\mathscr{G}(\mathscr{X})$, so $\underline{T}$ is a $\wedge$-homomorphism. □

The argumentation in this proof leads at once to the following proposition:

Proposition 7. *A transformation $\underline{T}\colon\mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{Y})$ is a $\wedge$-homomorphism if and only if each of its component maps $\underline{T}_y=\underline{P}_{\{y\}}\circ\underline{T}$, $y\in\mathscr{Y}$ is a minitive lower prevision on $\mathscr{G}(\mathscr{X})$.*

Because minitive lower previsions are completely monotone, they are also coherent if and only if they satisfy $\underline{P}(0)=0$ and $\underline{P}(1)=1$. The following result provides a further characterisation.

Theorem 8. *Let $\underline{P}$ be a coherent lower prevision on $\mathscr{G}(\mathscr{X})$. Then the following statements are equivalent:*

(i) *$\underline{P}$ is minitive;*
(ii) *the restriction of $\underline{P}$ to events is a $\{0,1\}$-valued lower probability;*
(iii) *there is a filter $\mathscr{F}:=\{A\subseteq\mathscr{X}:\underline{P}(A)=1\}$ such that $\underline{P}(f)=\sup_{F\in\mathscr{F}}\inf_{x\in F}f(x)$ for all gambles f on $\mathscr{X}$.*

Proof. We begin by showing that the first statement implies the second. Assume that $\underline{P}$ is a minitive coherent lower prevision. Then it follows that its restriction to events is a minitive lower probability. Assume *ex absurdo* that there is some event A such that $\underline{P}(A)\in(0,1)$.

Consider the gambles $f := 1$ and $g := \frac{2}{\underline{P}(A)} I_A$. Then $\underline{P}(f) = 1$ and $\underline{P}(g) = \frac{2}{\underline{P}(A)}\underline{P}(A) = 2$, so $\underline{P}(f) \wedge \underline{P}(g) = 1$. On the other hand, $f \wedge g = I_A$, so $\underline{P}(f \wedge g) = \underline{P}(A) < 1$. This contradicts that $\underline{P}$ is minitive.

To show that the second statement implies the third, assume that the restriction $\underline{Q}$ of $\underline{P}$ to events is a $\{0,1\}$-valued lower probability on $\mathscr{P}(\mathscr{X})$, which is coherent as a restriction of the coherent lower prevision $\underline{P}$. It then follows from the discussion in [19, Sections 2.9.8 and 3.2.6] that the set $\mathscr{F} := \{A \subseteq \mathscr{X} : \underline{P}(A) = 1\}$ is a (proper) filter, and that $\underline{Q}$ has a unique coherent extension to $\mathscr{G}(\mathscr{X})$—which must therefore coincide with $\underline{P}$—given by:

$$\underline{P}(f) = \sup_{F \in \mathscr{F}} \inf_{x \in F} f(x) \text{ for all gambles } f \text{ on } \mathscr{X}. \tag{9}$$

Finally, to show that the third statement implies the first, we must prove that the lower prevision $\underline{P}$ defined by Eq. (9) is minitive. Consider any two gambles f and g on $\mathscr{X}$. Then for any $\varepsilon > 0$ there are two events F_1 and F_2 in $\mathscr{F}$ such that $\underline{P}(f) \leq \inf_{x \in F_1} f(x) + \varepsilon$ and $\underline{P}(g) \leq \inf_{x \in F_2} g(x) + \varepsilon$, and therefore also $\underline{P}(f) \leq \inf_{x \in F_1 \cap F_2} f(x) + \varepsilon$ and $\underline{P}(g) \leq \inf_{x \in F_1 \cap F_2} g(x) + \varepsilon$. Since $\mathscr{F}$ is a filter, we deduce that $F_1 \cap F_2 \in \mathscr{F}$ and therefore

$$\begin{aligned}\underline{P}(f \wedge g) &\geq \inf_{x \in F_1 \cap F_2} (f \wedge g)(x) = \inf_{x \in F_1 \cap F_2} f(x) \wedge \inf_{x \in F_1 \cap F_2} g(x) \\ &\geq (\underline{P}(f) - \varepsilon) \wedge (\underline{P}(g) - \varepsilon) \geq (\underline{P}(f) \wedge \underline{P}(g)) - \varepsilon.\end{aligned}$$

We deduce that $\underline{P}(f \wedge g) \geq \underline{P}(f) \wedge \underline{P}(g)$. The converse inequality follows from the monotonicity (due to coherence) of $\underline{P}$. □

We deduce from Theorem 8 that there is a one-to-one correspondence between filters of subsets of $\mathscr{X}$ and minitive coherent lower previsions. If in addition we require that the lower prevision should be *infimum preserving*, we have the following result:

Theorem 9 ([2, Section 8]). *Let $\underline{P}$ be a coherent lower prevision on $\mathscr{G}(\mathscr{X})$. Then $\underline{P}$ is infimum preserving if and only if its restriction to events is a $\{0,1\}$-valued necessity measure.*

Taking into account the discussion above, we deduce that an infimum preserving coherent lower prevision $\underline{P}$ is also associated with a filter $\mathscr{F}$ of subsets of $\mathscr{A}$; but since $\underline{P}$ is infimum and not just minimum preserving, it follows that

$$\underline{P}\left(\bigcap \mathscr{F}\right) = \inf_{A \in \mathscr{F}} \underline{P}(A) = 1,$$

and as a consequence there is a smallest subset $\bigcap \mathscr{F}$ of $\mathscr{X}$ with lower probability 1, so $\bigcap \mathscr{F} \in \mathscr{F}$, meaning that $\mathscr{F}$ is fixed. This shows that infimum preserving coherent lower previsions are those associated with *fixed* filters, while more generally minitive coherent lower previsions can also be associated with a free filter. It also shows that the only infimum preserving coherent lower previsions are the vacuous lower previsions we have introduced in Section 2. Finally, when $\mathscr{X}$ is finite, the properties of being infimum preserving and minimum preserving are equivalent, because in that case all filters are fixed; hence, in that case the vacuous lower previsions are the only minimum preserving ones.

6. n-MONOTONICITY PRESERVING TRANSFORMATIONS

From transformations that preserve coherence, and transformations that preserve minitivity, we now turn to the study of transformations that preserve n-monotonicity. We begin by showing that a coherence preserving transformation does not preserve n-monotonicity in general:

Example 1. Let $\underline{P}$ be a coherent lower prevision on $\mathscr{X}$ that is not n-monotone. Consider the coherence preserving transformation $\underline{T}$ defined by $(\underline{T} f)(y) := \underline{P}(f)$ for all $f \in \mathscr{G}(\mathscr{X})$ and all $y \in \mathscr{Y}$. Then for any n-monotone lower prevision $\underline{Q}$ on $\mathscr{G}(\mathscr{Y})$, $\underline{Q} \circ \underline{T}$ coincides with $\underline{P}$, which is not n-monotone.

The second type of transformations we have studied so far are the $\wedge$-homomorphisms we have characterised in Section 5. It is not difficult to show that this type of transformations do preserve n-monotonicity:

Proposition 10 ([8, Lemma 6])**.** *Let $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ be a $\wedge$-homomorphism and let $\underline{P}$ be a lower prevision on $\mathscr{G}(\mathscr{Y})$. If $\underline{P}$ is n-monotone, then $\underline{P} \circ \underline{T}$ is n-monotone too.*

Our next example shows that being a $\wedge$-homomorphism, although sufficient, is not necessary for a coherence preserving transformation to preserve n-monotonicity:

Example 2. Let $\underline{P}$ be an n-monotone lower prevision on $\mathscr{X}$ that is not minitive. Consider the coherence preserving transformation $\underline{T}$ defined by $(\underline{T}f)(y) := \underline{P}(f)$ for all $f \in \mathscr{G}(\mathscr{X})$ and all $y \in \mathscr{Y}$. Then for any n-monotone lower prevision $\underline{Q}$ on $\mathscr{G}(\mathscr{Y})$, $\underline{Q} \circ \underline{T} = \underline{P}$ is n-monotone, so $\underline{T}$ preserves n-monotonicity, but it is not a $\wedge$-homomorphism, because its component maps $\underline{T}_y = \underline{P}$ are not minitive; see Proposition 6.

Nevertheless, it is possible, with some modifications, to prove a characterisation of transformations that preserve n-monotonicity as $\wedge$-homomorphisms. For this we need to look at the extended transformations $\tilde{\underline{T}}\colon \mathscr{G}(\mathscr{X} \times \mathscr{Y}) \to \mathscr{G}(\mathscr{Y})$ that can be used to turn a lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$ into a lower prevision $\underline{P} \circ \tilde{\underline{T}}$ on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$, and that are given by Eq. (8). We begin with some basic observations:

Proposition 11. *Let $\underline{T}$ be a coherence preserving transformation between $\mathscr{G}(\mathscr{X})$ and $\mathscr{G}(\mathscr{Y})$.*

- (i) *If $\underline{P}$ avoids sure loss, then so does $\underline{P} \circ \tilde{\underline{T}}$.*
- (ii) *If $\underline{P}$ is coherent, then so is $\underline{P} \circ \tilde{\underline{T}}$.*
- (iii) *If $\underline{P}$ is n-monotone and $\underline{T}$ is a $\wedge$-homomorphism, then $\underline{P} \circ \tilde{\underline{T}}$ is n-monotone as well.*
- (iv) *If $\underline{P}$ is minitive and $\underline{T}$ is a $\wedge$-homomorphism, then $\underline{P} \circ \tilde{\underline{T}}$ is minitive as well.*

Proof. If we consider that $\tilde{\underline{T}}$ is a transformation between $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ and $\mathscr{G}(\mathscr{Y})$, we can consider its component maps $\tilde{\underline{T}}_y$, which are the lower previsions on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ given by:

$$\tilde{\underline{T}}_y(h) = (\tilde{\underline{T}}h)(y) = (\underline{T}h(\cdot,y))(y) = \underline{T}_y h(\cdot,y) \text{ for all gambles } h \text{ on } \mathscr{X} \times \mathscr{Y} \text{ and all } y \in \mathscr{Y}.$$

This shows, taking into account Propositions 3 and 7, that $\tilde{\underline{T}}$ is coherence preserving if $\underline{T}$ is, and that $\tilde{\underline{T}}$ is a $\wedge$-homomorphism if $\underline{T}$ is. Now invoke Proposition 2 for the first two statements, Proposition 10 for the third, and Proposition 6 for the fourth. □

We can show that a transformation $\tilde{\underline{T}}$ preserves n-monotonicity if and only if $\underline{T}$ is a $\wedge$-homomorphism:

Proposition 12. *Let $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ be a transformation between two sets of gambles. Then the following statements are equivalent:*

- (i) *$\underline{T}$ is a $\wedge$-homomorphism;*
- (ii) *for every n-monotone lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$, the composition $\underline{P} \circ \tilde{\underline{T}}$ is an n-monotone lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$.*

Proof. The direct implication follows from Proposition 11.

To prove the converse implication, assume that for every n-monotone lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$, the composition $\underline{P} \circ \tilde{\underline{T}}$ is an n-monotone lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$.

We first prove that then $\underline{T}_y$ must be monotone for every $y \in \mathscr{Y}$. Assume *ex absurdo* that there are $y_o \in \mathscr{Y}$ and gambles $f \leq g$ on $\mathscr{X}$ such that $\underline{T}_{y_o} f = \underline{T}f(y_o) > \underline{T}g(y_o) = \underline{T}_{y_o} g$. This means that with $\underline{P} := \underline{P}_{\{y_o\}}$—the completely monotone vacuous lower prevision on $\mathscr{G}(\mathscr{Y})$ relative to $\{y_o\}$—the composition $\underline{P} \circ \tilde{\underline{T}}$ is not completely monotone, because it is not even monotone: if we define the gambles $\tilde{f} \leq \tilde{g}$ on $\mathscr{X} \times \mathscr{Y}$ by letting $\tilde{f}(x,y) := f(x)$ and $\tilde{g}(x,y) := g(x)$ for all $x \in \mathscr{X}$ and $y \in \mathscr{Y}$, then $\tilde{\underline{T}}_y \tilde{f} = (\underline{T}\tilde{f}(\cdot,y))(y) = \underline{T}f(y) = \underline{T}_y f$ and similarly $\tilde{\underline{T}}_y \tilde{g} = \underline{T}_y g$, which leads to the contradiction:

$$(\underline{P} \circ \tilde{\underline{T}})(\tilde{f}) = \underline{T}_{y_o}(f) > \underline{T}_{y_o}(g) = (\underline{P} \circ \tilde{\underline{T}})(\tilde{g}).$$

Next, assume *ex absurdo* that $\underline{T}$ is not a $\wedge$-homomorphism, then there is, by Proposition 7, some $y \in \mathscr{Y}$ such that $\underline{T}_y$ is not minitive. Then there are gambles f_1 and f_2 on $\mathscr{X}$ such that $\underline{T}_y(f_1 \wedge f_2) < \underline{T}_y f_1 \wedge \underline{T}_y f_2$, since the monotonicity of $\underline{T}_y$ already implies that $\underline{T}_y(f_1 \wedge f_2) \leq \underline{T}_y f_1 \wedge \underline{T}_y f_2$.

Consider any $y' \neq y$, and the vacuous lower prevision $\underline{P}_{\{y,y'\}}$ relative to $\{y,y'\}$, which is a completely monotone coherent lower prevision on $\mathscr{G}(\mathscr{Y})$. We only need to show that $\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}}$ is not 2-monotone. Consider any real number a in the non-empty interval $(\underline{T}_y(f_1 \wedge f_2), \underline{T}_y f_1)$ and any real number b in the non-empty interval $(\underline{T}_y(f_1 \wedge f_2), \underline{T}_y f_2)$, and define the gambles f, g on $\mathscr{X} \times \mathscr{Y}$ by:

$$f := aI_{\{y'\}} + f_1 I_{\{y\}} \text{ and } g := bI_{\{y'\}} + f_2 I_{\{y\}}.$$

Then

$$\underline{\tilde{T}} f(z) = \underline{T} f(\cdot, z) = \begin{cases} \underline{T}_y f_1 & z = y \\ a & z = y' \\ 0 & \text{elsewhere} \end{cases} \text{ and } \underline{\tilde{T}} g(z) = \underline{T} g(\cdot, z) = \begin{cases} \underline{T}_y f_2 & z = y \\ b & z = y' \\ 0 & \text{elsewhere,} \end{cases}$$

and similarly

$$\underline{\tilde{T}}(f \wedge g)(z) = \begin{cases} \underline{T}_y(f_1 \wedge f_2) & z = y \\ a \wedge b & z = y' \\ 0 & \text{elsewhere} \end{cases} \text{ and } \underline{\tilde{T}}(f \vee g)(z) = \begin{cases} \underline{T}_y(f_1 \vee f_2) & z = y \\ a \vee b & z = y' \\ 0 & \text{elsewhere,} \end{cases}$$

so

$$\begin{aligned} (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(f) &= \min\{a, \underline{T}_y f_1\} = a \\ (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(g) &= \min\{b, \underline{T}_y f_2\} = b \\ (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(f \wedge g) &= \min\{a \wedge b, \underline{T}_y(f_1 \wedge f_2)\} = \underline{T}_y(f_1 \wedge f_2) < a \wedge b \\ (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(f \vee g) &= \min\{a \vee b, \underline{T}_y(f_1 \vee f_2)\} = a \vee b, \end{aligned}$$

taking into account that

$$\begin{aligned} &a > \underline{T}_y(f_1 \wedge f_2),\ b > \underline{T}_y(f_1 \wedge f_2) \Rightarrow a \wedge b > \underline{T}_y(f_1 \wedge f_2) \text{ and} \\ &a < \underline{T}_y f_1,\ b < \underline{T}_y f_2 \Rightarrow a \vee b < \underline{T}_y f_1 \vee \underline{T}_y f_2 \leq \underline{T}_y(f_1 \vee f_2). \end{aligned}$$

As a consequence,

$$\begin{aligned} (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(f \vee g) + (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(f \wedge g) &< (a \vee b) + (a \wedge b) = a + b \\ &= (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(f) + (\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}})(g), \end{aligned}$$

and this implies that $\underline{P}_{\{y,y'\}} \circ \underline{\tilde{T}}$ is not 2-monotone. □

We want to stress that, although the only $\underline{\tilde{T}}$ that always preserve n-monotonicity are those for which $\underline{T}$ is a $\wedge$-homomorphism, it does not hold that $\underline{T}$ preserves n-monotonicity if and only if it is a $\wedge$-homomorphism, as we have already seen in Example 2. This is because when $\underline{T}$ is a $\wedge$-homomorphism between $\mathscr{G}(\mathscr{X})$ and $\mathscr{G}(\mathscr{Y})$ then $\underline{\tilde{T}}$ is also a $\wedge$-homomorphism between $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ and $\mathscr{G}(\mathscr{Y})$, but the converse is not necessarily true.

7. Examples

In this section, we discuss a number of particular cases of coherence preserving transformations.

7.1. **Liftings.** Consider a map $t\colon \mathscr{Y}\to\mathscr{X}$ and its associated lifting $\underline{T}\colon \mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{Y})$ given by $\underline{T}(f):=f\circ t$ for gambles f on $\mathscr{X}$. Then $\underline{T}$ is a coherence preserving transformation, which moreover satisfies T3 with equality, and given $y\in\mathscr{Y}$, the component map $\underline{T}_y$ of $\underline{T}$, given by Eq. (5), is the degenerate linear prevision $\underline{P}_{\{t(y)\}}$.

In the special case that $\mathscr{Y}=\mathscr{X}$, the relationship between a coherent lower prevision $\underline{P}$ and its transformation $\underline{P}\circ\underline{T}$ lies at the basis of the notions of weak and strong invariance discussed in [5]. Two particular cases are of interest: the shift transformations, leading to a treatment of time-invariance, where $\mathscr{X}=\mathbb{N}$ and $t(n):=n+1$ ([5, Section 8] and [19, Section 2.9.5]); and the permutations of a finite space $\mathscr{X}$ that give rise to a discussion of *exchangeable* lower previsions ([7] and [19, Section 9.5]).

The lower prevision induced by a lifting transformation can be seen as a generalisation to an imprecise-probabilistic context to the notion of probability induced by a map. In this sense, these notions also lead to the convergence in distribution results for coherent lower previsions in [7, Theorem 6].

7.2. **Transition operators.** In a series of papers on the dynamics of Markov chains in an imprecise probability context [3, 4, 12], De Cooman et al. introduce and study lower transition operators $\underline{T}\colon\mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{X})$, which are, essentially, coherence preserving transformations between the set $\mathscr{G}(\mathscr{X})$ of all gambles on a finite state space $\mathscr{X}$ of a Markov chain and itself. These operators are the imprecise-probabilistic counterparts of transition matrices, and they allow for a complete characterisation of the dynamical behaviour of a Markov chain with imprecise transition probabilities. They are defined as follows: $\underline{T}f(x):=\underline{P}(f|x)$ is the conditional lower prevision of a gamble $f(X_{n+1})$ on the state X_{n+1} at time $n+1$, when the Markov chain is on state $X_n=x$ at time n.

We deduce from Section 4 that these transition operators can be seen as coherence preserving transformations, and from Section 6 that they preserve n-monotonicity when they are minitive.

7.3. **Markov operators.** In [17], Škulj defines the following operators, which generalise conditional (linear) expectations:

Definition 9 ([17, Definition 1])**.** Let $\mathscr{H},\mathscr{K}$ be two linear subspaces of $\mathscr{G}(\mathscr{X})$ that include all constant gambles. A linear operator $T\colon\mathscr{H}\to\mathscr{K}$ that is monotone $[f\leq g\Rightarrow T(f)\leq T(g)]$ and such that $T(1)=1$ is called a *Markov operator*.

Škulj discusses how to model risk and uncertainty aversion in a context of imprecision by means of a notion of invariance with respect to a set of Markov operators. When $\mathscr{H}=\mathscr{K}=\mathscr{G}(\mathscr{X})$, a Markov operator is a particular case of a coherence preserving transformation: T2 and T3 hold because $\underline{T}$ is a linear operator. With respect to T1, note that for any gamble $f\in\mathscr{G}(\mathscr{X})$, $T(f)\geq T(\inf f)=\inf f$, where the inequality follows from the monotonicity of T and the equality from its linearity and the fact that $T(1)=1$.

7.4. **Multi-valued maps.** Let $\Gamma\colon\mathscr{Y}\to\mathscr{P}(\mathscr{X})$ be a multi-valued map, and use it to define the transformation $\underline{T}\colon\mathscr{G}(\mathscr{X})\to\mathscr{G}(\mathscr{Y})$ by $\underline{T}f(y):=\inf_{x\in\Gamma(y)}f(x)$. This transformation is coherence preserving, because its associated component maps $\underline{T}_y$ are the vacuous lower previsions relative to the sets $\Gamma(y)$, which are in particular coherent and infimum preserving: they are associated with the fixed filters determined by the sets $\Gamma(y)$.

These types of coherence preserving transformations were investigated in detail in [16], and will be generalised by the filter maps we will introduce in Section 8. We can identify $\underline{T}$ with a conditional lower prevision $\underline{P}(\cdot|Y)$, in such a way that each $\underline{P}(\cdot|y)$ is the vacuous lower prevision relative to $\Gamma(y)$. Taking into account Proposition 11, we see that the composition of this coherence preserving transformation with a coherent lower prevision is again a coherent lower prevision. In fact, since $\underline{T}$ is a $\wedge$-homomorphism, the same result shows that this coherence preserving transformation preserves n-monotonicity too.

This approach allows us to generalise some of the main concepts from random set theory to a context of imprecise-probabilistic information about the initial probability space $\mathscr{Y}$. In Section 8, we will see that this generalisation can be taken one step further by means of filter maps.

8. FILTER MAPS

When we combine the results of Propositions 7 and 12 with Theorem 8, we see that the transformations $\underline{T}$ that are guaranteed to preserve n-monotonicity are the ones whose component maps $\underline{T}_y$ are minitive coherent lower previsions, meaning that there is some filter $\Phi(y)$ of subsets of $\mathscr{X}$ such that

$$\underline{T}_y h = \sup_{F\in\Phi(y)} \inf_{x\in F} h(x,y) \text{ for all gambles } h \text{ on } \mathscr{X}\times\mathscr{Y}.$$

In this section, we study these transformations in some detail. But we begin by providing yet another way of motivating their study.

8.1. **Motivation.**

Consider two variables X and Y assuming values in the non-empty (but not necessarily finite) sets $\mathscr{X}$ and $\mathscr{Y}$.

Let us first look at a single-valued map γ between the spaces $\mathscr{Y}$ and $\mathscr{X}$. Given a linear prevision P on $\mathscr{G}(\mathscr{Y})$, such a map induces a precise probability P_γ on $\mathscr{G}(\mathscr{X})$ by $P_\gamma(A) := P(\gamma^{-1}(A)) = P(\{y\in\mathscr{Y}\colon \gamma(y)\in A\})$ for all $A\subseteq\mathscr{X}$, or equivalently

$$P_\gamma(f) := P(f\circ\gamma) \text{ for all gambles } f \text{ on } \mathscr{X}, \tag{10}$$

which is a well-known 'change of variables result' for previsions (or expectations): If we have a variable Y, and the variable X is given by $X := \gamma(Y)$, then any gamble $f(X)$ on the value of X can be translated back to a gamble $f(\gamma(Y)) = (f\circ\gamma)(Y)$ on the value of Y, which explains where Eq. (10) comes from: *if the uncertainty about Y is represented by the model P, then the uncertainty about $X=\gamma(Y)$ is represented by P_γ.* This can also be seen as an application of the lifting procedure from Section 7.1.

There is another way of motivating the same formula, which lends itself more readily to generalisation. We can interpret the map γ as conditional information: *if we know that $Y=y$, then we know that $X=\gamma(y)$.* This conditional information can be represented by a so-called conditional linear prevision $P(\cdot|Y)$ on $\mathscr{G}(\mathscr{X})$, defined by

$$P(f|y) := f(\gamma(y)) = (f\circ\gamma)(y) \text{ for all gambles } f \text{ on } \mathscr{X}. \tag{11}$$

It states that conditional on $Y=y$, all probability mass for X is located in the single point $\gamma(y)$. If $P(f|Y)$ is the gamble on $\mathscr{Y}$ that assumes the value $P(f|y)$ in y, then clearly $P(f|Y) = (f\circ\gamma)(Y)$, which allows us to rewrite Eq. (10) as:

$$P_\gamma(f) = P(P(f|Y)) \text{ for all gambles } f \text{ on } \mathscr{X}, \tag{12}$$

which shows that Eq. (10) is actually a special case of the Law of Iterated Expectations, the expectation form of the Law of Total Probability, in classical probability (see, for instance, [9, Theorem 4.7.1]).

Assume now that, more generally, the relation between X and Y is determined as follows. There is a so-called *multi-valued map* $\Gamma\colon\mathscr{Y}\to\mathscr{P}(\mathscr{X})$ that associates with any $y\in\mathscr{Y}$ a non-empty subset $\Gamma(y)$ of $\mathscr{X}$, and *if we know that $Y=y$, then all we know about X is that it can assume any value in $\Gamma(y)$.* There is no immediately obvious way of representing this conditional information using a precise probability model. If we want to remain within the framework of precise probability theory, we must abandon the simple and powerful device of interpreting the multi-valued map Γ as conditional information. But if we work with the theory of imprecise probabilities, as we are doing here, it is still perfectly possible to interpret Γ as conditional information that can be represented by a special conditional *lower* prevision $\underline{P}(\cdot|Y)$ on $\mathscr{G}(\mathscr{X})$, where

$$\underline{P}(f|y) = \underline{P}_{\Gamma(y)}(f) = \inf_{x\in\Gamma(y)} f(x) \text{ for all gambles } f \text{ on } \mathscr{X} \tag{13}$$

is the vacuous lower prevision relative to the event $\Gamma(y)$; see Section 7.4 for more details. Given information about Y in the form of a coherent lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$, it follows from Walley's Marginal Extension Theorem (see [19, Section 6.7]) that the corresponding information about X is the lower prevision $\underline{P}_\Gamma$ on $\mathscr{G}(\mathscr{X})$ defined by

$$\underline{P}_\Gamma(f) = \underline{P}(\underline{P}(f|Y)) \text{ for all gambles } f \text{ on } \mathscr{X}, \tag{14}$$

which is an immediate generalisation of Eq. (12). This formula provides a well-justified method for using the conditional information embodied in the multi-valued map Γ to turn the uncertainty model $\underline{P}$ about Y into an uncertainty model $\underline{P}_\Gamma$ about X. This approach has been introduced and explored in great detail by Miranda et al. [16].

What we intend to do here, is take this idea of conditional information one useful step further. To motivate going even further that multi-valued maps, assume that the information about the relation between X and Y is the following: *If we know that $Y = y$, then all we know about X is that it lies arbitrarily close to $\gamma(y)$,* in the sense that X lies inside any neighbourhood of $\gamma(x)$. We are assuming that, in order to capture what 'arbitrarily close' means, we have provided $\mathscr{X}$ with a topology $\mathscr{T}$ of open sets. We can model this type of conditional information using the conditional *lower* prevision $\underline{P}(\cdot|Y)$ on $\mathscr{G}(\mathscr{X})$, where

$$\underline{P}(f|y) = \underline{P}_{\mathscr{N}_{\gamma(y)}}(f) = \sup_{N\in\mathscr{N}_{\gamma(y)}} \inf_{x\in N} f(x) \text{ for all gambles } f \text{ on } \mathscr{X} \tag{15}$$

is the lower prevision associated with the neighbourhood filter $\mathscr{N}_{\gamma(y)}$ of $\gamma(y)$. Information about Y in the form of a coherent lower prevision $\underline{P}$ on $\mathscr{G}(\mathscr{Y})$ can now be turned into information $\underline{P}(\underline{P}(\cdot|Y))$ about X, via this conditional model, using Eq. (14). This is the idea behind the notion of filter maps we define next.

8.2. **Definition and main properties.** So, given this motivation, let us try and capture these ideas in an abstract model.

Definition 10. A *filter map* Φ from $\mathscr{Y}$ to $\mathscr{X}$, is a map $\Phi\colon \mathscr{Y} \to \mathbb{F}(\mathscr{X})$ that associates a proper[2] filter $\Phi(y)$ with each element y of $\mathscr{Y}$.

We can use filter maps to model some type of conditional information, which can be represented by a (specific) conditional lower prevision. In order to do this, given a filter map Φ, we associate with any gamble f on $\mathscr{X} \times \mathscr{Y}$ a *lower inverse* $f_\circ$ (under Φ), which is the gamble on $\mathscr{Y}$ defined by

$$f_\circ(y) = \underline{P}_{\Phi(y)}(f(\cdot,y)) = \sup_{F\in\Phi(y)} \inf_{x\in F} f(x,y) \text{ for all } y \text{ in } \mathscr{Y}, \tag{16}$$

where, of course, $\underline{P}_{\Phi(y)}$ is the lower prevision on $\mathscr{G}(\mathscr{X})$ associated with the filter $\Phi(y)$ of $\mathscr{X}$. Let us check that this lower prevision is coherent:

Proposition 13. *The lower prevision $\underline{P}_{\Phi(y)}$ on $\mathscr{G}(\mathscr{X})$ associated with the filter $\Phi(y)$ of $\mathscr{X}$ is coherent.*

Proof. This is an immediate consequence of Theorem 8, taking into account the one-to-one correspondence between filters and minitive coherent lower previsions. □

[2]Although in this paper we are only dealing with *proper* filters—i.e. we assume that $\emptyset \notin \Phi(y)$—, it is also possible to manage without this and similar assumptions, following the ideas out in [16, Technical Remarks 1 and 2].

Similarly, we define for any gamble g on $\mathscr{X}$ its *lower inverse* $g_\bullet$ (under Φ) as the gamble on $\mathscr{Y}$ defined by

$$g_\bullet(y) = \underline{P}_{\Phi(y)}(g) = \sup_{F\in\Phi(y)} \inf_{x\in F} g(x) \text{ for all } y \text{ in } \mathscr{Y}. \tag{17}$$

Eqs. (16) and (17) are obviously very closely related to, and inspired by, the expressions (11), (13) and (15). In particular, we find for any $A \subseteq \mathscr{X}\times\mathscr{Y}$ that $(I_A)_\circ = I_{A_\circ}$, where we let

$$A_\circ = \{y \in \mathscr{Y} : (\exists F \in \Phi(y))F \times \{y\} \subseteq A\}$$

denote the so-called *lower inverse* of A (under Φ). And if $B \subseteq \mathscr{X}$, then

$$(B\times\mathscr{Y})_\circ = B_\bullet = \{y \in \mathscr{Y} : (\exists F \in \Phi(y))F \subseteq B\},$$

is the set of all y for which B occurs eventually with respect to the filter $\Phi(y)$.

If we consider the transformation $\underline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{Y})$ with component maps $\underline{T}_y := \underline{P}_{\Phi(y)}$, then $f_\circ = \tilde{\underline{T}} f$ for any gamble f on $\mathscr{X}\times\mathscr{Y}$, and $g_\bullet = \underline{T} g$ for any gamble g on $\mathscr{X}$. We know from the discussion in the previous sections that $\underline{T}$ and $\tilde{\underline{T}}$ preserve coherence and n-monotonicity.

Now consider any lower prevision $\underline{P}$ on Y that avoids sure loss. Then we can consider its natural extension $\underline{E}_{\underline{P}}$, and use it together with the filter map Φ to construct an *induced lower prevision* $\underline{P}_\circ := \underline{E}_{\underline{P}} \circ \tilde{\underline{T}}$ on $\mathscr{G}(\mathscr{X}\times\mathscr{Y})$:

$$\underline{P}_\circ(f) = \underline{E}_{\underline{P}}(f_\circ) \text{ for all gambles } f \text{ on } \mathscr{X}\times\mathscr{Y}. \tag{18}$$

The so-called $\mathscr{X}$-marginal $\underline{P}_\bullet := \underline{E}_{\underline{P}} \circ \underline{T}$ of this lower prevision is the lower prevision on $\mathscr{G}(\mathscr{X})$ given by

$$\underline{P}_\bullet(g) = \underline{E}_{\underline{P}}(g_\bullet) \text{ for all gambles } g \text{ on } \mathscr{X}. \tag{19}$$

Eqs. (18) and (19) are very closely related to, and inspired by, the expressions (10), (12), and (14). Induced lower previsions are what results if we use the conditional information embodied in the filter map to turn an uncertainty model about Y into an uncertainty model about X. This is because filter maps are in a one-to-one correspondence with minitive conditional lower previsions, taking into account the results in Sections 5 and 6.

We immediately deduce the following result:

Proposition 14. *Let $n \in \mathbb{N}$ and let $\underline{P}$ be a lower prevision that avoids sure loss, and is defined on a subset of $\mathscr{G}(\mathscr{Y})$. Let Φ be a filter map from $\mathscr{Y}$ to $\mathscr{X}$, and let $\underline{P}_\circ$ be the lower prevision given by Eq.* (18). *Then the following statements hold:*

(i) *$\underline{P}_\circ$ is a coherent lower prevision.*
(ii) *If $\underline{E}_{\underline{P}}$ is n-monotone, then so is $\underline{P}_\circ$.*

Proof. Since $\underline{E}_{\underline{P}}$ is coherent, we infer from Proposition 11 that $\underline{P}_\circ = \underline{E}_{\underline{P}} \circ \tilde{\underline{T}}$ is coherent as well. The same result guarantees that if $\underline{E}_{\underline{P}}$ is n-monotone, then so is $\underline{P}_\circ$, because $\underline{T}$ is a $\wedge$-homomorphism by Proposition 7 and Theorem 8. □

As a particular case of the above results we have filter maps that associate with any $y \in \mathscr{Y}$ a *fixed* filter $\Phi(y)$ of subsets of $\mathscr{X}$. If we denote by $\Gamma(y)$ the smallest subset in this filter, then we can define a multi-valued map $\Gamma\colon \mathscr{Y} \to \mathscr{P}(\mathscr{X})$. The conditional lower prevision $\underline{P}(\cdot|Y)$ associated with this filter map is precisely the one considered in [16], that we have already discussed in Section 7.4; many of the results we establish in this section constitute generalisations of results in [16] to arbitrary filter maps.

Remark 1. Another interesting particular case is that when $\mathscr{Y} = \mathscr{X}$ and our filter map associates, with any $y \in \mathscr{Y}$, the filter $\Phi(y)$ of all neighbourhoods of y under a certain topology in $\mathscr{Y}$. Then the lower inverse defined in Eq. (17) produces

$$g_\bullet(y) = \underline{P}_{\Phi(y)}(g) = \sup_{F\in\Phi(y)} \inf_{z\in F} g(x) \text{ for all } y \text{ in } \mathscr{Y}.$$

This is called a *lower oscillation model* in [6], and it can be seen to model the assessment that all the probability mass is concentrated around y. This was the idea behind the motivation of filter maps we gave in Eq. (15).

Now, given a linear prevision π on the class of continuous gambles over a compact metric space, it can be proved that its natural extension to all gambles is uniquely determined by its value on lower semi-continuous gambles, and that these are precisely the lower oscillations defined in the above equation. This natural extension is moreover the lower envelope of the linear extensions of the linear prevision π to the class of gambles, and we can use the lower oscillations to characterise those gambles for which there is a unique extension. See [6] for more details. ⧫

8.3. **Equivalent representations.** Next, let us define a *prevision kernel from $\mathscr{Y}$ to $\mathscr{X}$* as any map K from $\mathscr{Y} \times \mathscr{G}(\mathscr{X})$ to $\mathbb{R}$ such that $K(y,\cdot)$ is a linear prevision on $\mathscr{G}(\mathscr{X})$ for all y in $\mathscr{Y}$. Prevision kernels are clear generalisations of probability or Markov kernels [13, p. 20], but without the measurability conditions. They are in a one-to-one correspondence with separately coherent conditional *linear* previsions. We can extend $K(y,\cdot)$ to a linear prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ by letting $K(y, f) = K(y, f(\cdot,y))$ for all gambles f on $\mathscr{X} \times \mathscr{Y}$. For any lower prevision $\underline{P}$ on Y that avoids sure loss, we denote by $\underline{P}K$ the lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$ defined by

$$\underline{P}K(f) = \underline{E}_{\underline{P}}(K(\cdot, f)) \text{ for all gambles } f \text{ on } \mathscr{X} \times \mathscr{Y}.$$

Obviously, if $\underline{P}$ is a linear prevision on $\mathscr{G}(\mathscr{Y})$, then $\underline{P}K$ is a linear prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$. As an immediate consequence, $\underline{P}K$ is always a coherent lower prevision on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$, as a lower envelope of linear previsions (Eq. (2)). We also use the following notation:

$$\begin{aligned}\mathbb{K}(\Phi) &= \Big\{K\colon\ (\forall y \in \mathscr{Y})K(y,\cdot) \in \mathbb{M}(\underline{P}_{\Phi(y)})\Big\} \\ &= \{K\colon\ (\forall y \in \mathscr{Y})(\forall A \in \Phi(y))K(y,A) = 1\},\end{aligned} \tag{20}$$

where the last equality follows from Eq. (4). The set $\mathbb{K}(\Phi)$ can be seen as the set of conditional linear previsions that dominate the conditional lower prevision associated with the filter map Φ. The next proposition can be seen as a special case of Walley's lower envelope theorem for marginal extension [19, Theorem 6.7.4]. Our proof closely follows Walley's original proof.

Proposition 15. *Let $\underline{P}$ be a lower prevision that avoids sure loss, and is defined on a subset of $\mathscr{G}(\mathscr{Y})$. Let Φ be a filter map from $\mathscr{Y}$ to $\mathscr{X}$, and let $\underline{P}_\circ$ be the lower prevision given by Eq.* (18)*. Then for all $K \in \mathbb{K}(\Phi)$ and $P \in \mathbb{M}(\underline{P})$, $PK \in \mathbb{M}(\underline{P}_\circ)$; and for all gambles f on $\mathscr{X} \times \mathscr{Y}$ there are $K \in \mathbb{K}(\Phi)$ and $P \in \mathbb{M}(\underline{P})$ such that $\underline{P}_\circ(f) = PK(f)$.*

Proof. First, fix any P in $\mathbb{M}(\underline{P})$ and any $K \in \mathbb{K}(\Phi)$. Consider any gamble f on $\mathscr{X} \times \mathscr{Y}$. We infer from Eqs. (16) and (20) that $K(y, f) \geq f_\circ(y)$ for all $y \in \mathscr{Y}$, and therefore $PK(f) = P(K(\cdot, f)) \geq P(f_\circ) \geq \underline{E}_{\underline{P}}(f_\circ) = \underline{P}_\circ(f)$, where the first inequality follows from the coherence [monotonicity] of the linear prevision P, and the second from Eq. (3). This shows that $PK \in \mathbb{M}(\underline{P}_\circ)$.

Next, fix any gamble f on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$. We infer from Eqs. (2) and (20) and the coherence of the lower previsions $\underline{P}_{\Phi(y)}$, $y \in \mathscr{Y}$ [Proposition 13] that there is some $K \in \mathbb{K}(\Phi)$ such that $f_\circ = K(\cdot, f)$. Similarly, there is some $P \in \mathbb{M}(\underline{P})$ such that $\underline{E}_{\underline{P}}(f_\circ) = P(f_\circ)$, and therefore $\underline{P}_\circ(f) = P(K(\cdot, f)) = PK(f)$. □

The above result can be summarised by means of the following figure:

$$\begin{array}{ccc} \underline{P} & \xrightarrow{\text{credal set}} & \mathbb{M}(\underline{P}) \\ {\scriptstyle\text{lower inverse}}\big\downarrow & & \big\downarrow{\scriptstyle\text{composition with } \mathbb{K}(\Phi)} \\ \underline{P}_\circ & \xrightarrow{\text{credal set}} & \mathbb{M}(\underline{P}_\circ) \end{array}$$

The above result allows us to discuss an interesting particular case of filter maps: those associated with ultrafilters.

Definition 11. Let $\mathscr{P}(\mathscr{X})$ denote the power set of $\mathscr{X}$. A subset $\mathscr{F}$ of $\mathscr{P}(\mathscr{X})$ is an *ultrafilter* when it is a filter and moreover for any subset A of $\mathscr{X}$, either A or A^c belongs to $\mathscr{F}$.

Similarly to filters, ultrafilters are in a one-to-one correspondence to $\{0,1\}$-valued *finitely additive* probabilities. If the corresponding filter is fixed, then there is some $x \in \mathscr{X}$ such that $A \in \mathscr{F}$ if and only if $x \in A$, and the associated linear prevision corresponds to the degenerate probability measure on x; on the other hand, if the filter is free (i.e., it has no minimal element) then the associated probability is finitely additive but not countably additive.

Any filter of subsets of $\mathscr{X}$ can be obtained as the intersections of the ultrafilters that include it [19, Theorem 3.6.6]; taking this into account, if we define

$$\mathbb{K}_1(\Phi) := \{K \in \mathbb{K}(\Phi) \colon K(y,A) \in \{0,1\}\ \forall y \in \mathscr{Y}, A \subseteq \mathscr{X}\}$$

it is easy to establish the following:

Corollary 16. *Let $\underline{P}$ be a lower prevision that avoids sure loss, and is defined on a subset of $\mathscr{G}(\mathscr{Y})$. Let Φ be a filter map from $\mathscr{Y}$ to $\mathscr{X}$, and let $\underline{P}_\circ$ be the lower prevision given by Eq.* (18)*. Then $\underline{P}_\circ$ is the lower envelope of the set*

$$\{PK(f) \colon P \in \mathbb{M}(\underline{P}), K \in \mathbb{K}_1(\Phi)\}.$$

Proof. The result follows from Proposition 15 and [19, Theorem 3.6.6], also taking into account the one-to-one correspondence between ultrafilters that include Φ and prevision kernels in $\mathbb{K}_1(\Phi)$. □

If we restrict our attention to the $\mathscr{X}$-marginal $\underline{P}_\bullet$ of the lower prevision $\underline{P}_\circ$ on $\mathscr{G}(\mathscr{X} \times \mathscr{Y})$, we can go somewhat further: the following simple proposition is a considerable generalisation of a result mentioned by Wasserman [20, Section 2.4], see also [21, Section 2].[3] Here, we use the notation $C\int f\,\mathrm{d}\mu$ for the Choquet integral of the gamble f with respect to a functional μ ([1, 10]).

Proposition 17. *Let $n \in \mathbb{N}$ and let $\underline{P}$ be a lower prevision that avoids sure loss, and is defined on a subset of $\mathscr{G}(\mathscr{Y})$. If $\underline{E}_{\underline{P}}$ is n-monotone, then so is $\underline{P}_\bullet$, and moreover*

$$\underline{P}_\bullet(g) = C\int g\,\mathrm{d}\underline{P}_\bullet = C\int g_\bullet\,\mathrm{d}\underline{E}_{\underline{P}} = \underline{E}_{\underline{P}}(g_\bullet) \text{ for all } g \in \mathscr{G}(\mathscr{X}). \tag{21}$$

Proof. That $\underline{P}_\bullet$ is n-monotone follows at once from Proposition 14(ii), taking into account that it is a restriction of $\underline{P}_\circ$. To prove Eq. (21), it suffices to prove that the first and last equalities hold, because of Eq. (19). To this end, use that both $\underline{E}_{\underline{P}}$ and $\underline{P}_\bullet$ are n-monotone, and apply [18, p. 56]. □

This means that the procedures of natural extension and taking the lower inverse commute, as we represent in the following figure:

$$\begin{array}{ccc}
\underline{P} & \xrightarrow{\text{lower inverse of P}} & \underline{P}_\bullet \\
\text{natural extension}\downarrow & & \downarrow\text{natural extension} \\
\underline{E}_{\underline{P}} & \xrightarrow{\text{lower inverse of g}} & \underline{E}_{\underline{P}}(g_\bullet) = \underline{P}_\bullet(g)
\end{array}$$

Now, if the lower prevision $\underline{P}$ we start with is defined on $\mathscr{K} \subseteq \mathscr{G}(\mathscr{Y})$, then it is also useful to consider the lower prevision $\underline{P}_\circ^{\mathrm{r}}$, defined on the set of gambles

$${}_\circ\mathscr{K} := \{f \in \mathscr{G}(\mathscr{X} \times \mathscr{Y}) \colon f_\circ \in \mathscr{K}\}$$

[3]Wasserman considers the special case that $\underline{P}$ is a probability measure and that g satisfies appropriate measurability conditions, so the right-most Choquet integral coincides with the usual expectation of $g_\bullet$. See also [15, Theorem 14].

as follows:

$$\underline{P}^{\mathrm{r}}_{\circ}(f) \coloneqq \underline{P}(f_{\circ}) \text{ for all } f \text{ such that } f_{\circ} \in \mathscr{K}. \tag{22}$$

If $\underline{P}$ is coherent, then of course $\underline{P}^{\mathrm{r}}_{\circ}$ is the restriction of $\underline{P}_{\circ}$ to ${}_{\circ}\mathscr{K}$, because then $\underline{E}_{\underline{P}}$ and $\underline{P}$ coincide on $\mathscr{K}$ [see Theorem 1(iii)]. We now show that, interestingly and perhaps surprisingly, all the 'information' present in $\underline{P}_{\circ}$ is then already contained in the restricted model $\underline{P}^{\mathrm{r}}_{\circ}$.

Proposition 18. *Let $\underline{P}$ be a coherent lower prevision, defined on a set of gambles $\mathscr{K} \subseteq \mathscr{G}(\mathscr{Y})$. Then the following statements hold:*

(i) *$\underline{P}^{\mathrm{r}}_{\circ}$ is the restriction of $\underline{P}_{\circ}$ to the set of gambles ${}_{\circ}\mathscr{K}$, and therefore a coherent lower prevision on ${}_{\circ}\mathscr{K}$.*

(ii) *The natural extension $\underline{E}_{\underline{P}^{\mathrm{r}}_{\circ}}$ of $\underline{P}^{\mathrm{r}}_{\circ}$ coincides with the induced lower prevision: $\underline{E}_{\underline{P}^{\mathrm{r}}_{\circ}} = \underline{P}_{\circ}$.*

Proof. (i). It follows from Theorem 1(iii) that $\underline{E}_{\underline{P}}$ and $\underline{P}$ coincide on $\mathscr{K}$. For any $f \in {}_{\circ}\mathscr{K}$, we have $f_{\circ} \in \mathscr{K}$ and therefore $\underline{P}_{\circ}(f) = \underline{E}_{\underline{P}}(f_{\circ}) = \underline{P}(f_{\circ}) = \underline{P}^{\mathrm{r}}_{\circ}(f)$, also using Eqs. (18) and (22).

(ii). Since $\underline{P}_{\circ}$ is coherent [Proposition 14(i)] and coincides with $\underline{P}^{\mathrm{r}}_{\circ}$ on ${}_{\circ}\mathscr{K}$, we infer from Theorem 1(ii) that $\underline{E}_{\underline{P}^{\mathrm{r}}_{\circ}} \leq \underline{P}_{\circ}$. Conversely, consider any gamble f on $\mathscr{X} \times \mathscr{Y}$, then we must show that $\underline{E}_{\underline{P}^{\mathrm{r}}_{\circ}}(f) \geq \underline{P}_{\circ}(f)$.

Fix $\varepsilon > 0$. If we use the definition of natural extension [for $\underline{P}$] in Eq. (1), we see that there are n in $\mathbb{N}$, non-negative $\lambda_1, \dots, \lambda_n$ in $\mathbb{R}$, and gambles $g_1, \dots, g_n$ in $\mathscr{K}$ such that

$$f_{\circ}(y) - \underline{P}_{\circ}(f) + \frac{\varepsilon}{2} \geq \sum_{k=1}^{n} \lambda_k [g_k(y) - \underline{P}(g_k)] \text{ for all } y \in \mathscr{Y}. \tag{23}$$

It also follows from the definition of the lower inverse $f_{\circ}$ in Eq. (16) that for each $y \in \mathscr{Y}$, there is some set $F(y) \in \Phi(y)$ such that

$$\inf_{x \in F(y)} f(x,y) \geq f_{\circ}(y) - \frac{\varepsilon}{2}. \tag{24}$$

Now define the corresponding gambles h_k on $\mathscr{X} \times \mathscr{Y}$, $k = 1, \dots, n$ by

$$h_k(x,y) = \begin{cases} g_k(y) & \text{if } y \in \mathscr{Y} \text{ and } x \in F(y) \\ L & \text{if } y \in \mathscr{Y} \text{ and } x \notin F(y), \end{cases}$$

where L is some real number strictly smaller than $\min_{k=1}^{n} \inf g_k$, to be determined shortly. Then for any $y \in \mathscr{Y}$:

$$\begin{aligned} (h_k)_{\circ}(y) &= \sup_{F \in \Phi(y)} \inf_{x \in F} h_k(x,y) = \sup_{F \in \Phi(y)} \inf_{x \in F} \begin{cases} g_k(y) & \text{if } x \in F(y) \\ L & \text{otherwise} \end{cases} \\ &= \sup_{F \in \Phi(y)} \begin{cases} g_k(y) & \text{if } F \subseteq F(y) \\ L & \text{otherwise} \end{cases} = g_k(y), \end{aligned}$$

because $L \leq g_k(y)$ and we can select $F = F(y)$. Hence, $(h_k)_{\circ} = g_k \in \mathscr{K}$ and therefore $h_k \in {}_{\circ}\mathscr{K}$ and $\underline{P}^{\mathrm{r}}_{\circ}(h_k) = \underline{P}(g_k)$. This, together with Eqs. (23) and (24), allows us to infer that

$$f(x,y) - \underline{P}_{\circ}(f) + \varepsilon \geq \sum_{k=1}^{n} \lambda_k [h_k(x,y) - \underline{P}^{\mathrm{r}}_{\circ}(h_k)] \text{ for all } y \in \mathscr{Y} \text{ and } x \in F(y).$$

Moreover, by an appropriate choice of L [small enough], we can always make sure that the inequality above holds for *all* $(x,y) \in \mathscr{X} \times \mathscr{Y}$ (note that once L is strictly smaller than $\min_{k=1}^{n} \inf g_k$ decreasing it does not affect $\underline{P}^{\mathrm{r}}_{\circ}(h_k) = \underline{P}(g_k)$). Then the definition of natural extension [for $\underline{P}^{\mathrm{r}}_{\circ}$] in Eq. (1) guarantees that $\underline{E}_{\underline{P}^{\mathrm{r}}_{\circ}}(f) \geq \underline{P}_{\circ}(f) - \varepsilon$. Since this inequality holds for any $\varepsilon > 0$, the proof is complete. □

9. Conclusions

The above results show that a number of transformations considered in the literature can all be seen as coherence preserving transformations of a possibility space, and as a consequence they can be combined with an imprecise probability model of a certain type in order to produce a model of the same type.

If we want to preserve n-monotonicity, which is necessary in order for a lower prevision to have a representation by means of a Choquet integral and for it to be uniquely determined by the restriction to events, we have seen that in a definite sense, it becomes necessary to work with $\wedge$-homomorphisms, which in this context imply working with minitive lower previsions. We have proved that these minimum preserving lower previsions are in a one-to-one correspondence with filters of events.

Our results given rise to the consideration of filter maps: maps that assign to any element of the initial space a filter of subsets of the final space, and that are in a one-to-one correspondence with conditional lower previsions that preserve n-monotonicity. We have studied them in detail, and have shown that they possess a number of interesting properties: for instance, they commute with other operations such as taking lower inverses and considering the associated credal sets. In doing this, we have extended a number of results from the literature.

As future lines of research, we would like to point out the application of coherence preserving mappings and filter maps in decision making, as well as the investigation of their connection with other models, such as p-boxes or (fuzzy) set-valued random variables.

Acknowledgements

The research in this paper has been supported by the SBO project 060043 of the IWT-Vlaanderen and by project MTM2010-17844.

References

[1] Gustave Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953–1954.
[2] Gert de Cooman and Dirk Aeyels. Supremum preserving upper probabilities. *Information Sciences*, 118:173–212, 1999.
[3] Gert de Cooman and Filip Hermans. Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence*, 172(11):1400–1427, 2008.
[4] Gert de Cooman, Filip Hermans, and Erik Quaeghebeur. Imprecise Markov chains and their limit behaviour. *Probability in the Engineering and Informational Sciences*, 23(4):597–635, 2009.
[5] Gert de Cooman and Enrique Miranda. Symmetry of models versus models of symmetry. In W. L. Harper and G. R. Wheeler, editors, *Probability and Inference: Essays in Honor of Henry E. Kyburg, Jr.*, pages 67–149. King's College Publications, 2007.
[6] Gert de Cooman and Enrique Miranda. The F. Riesz representation theorem and finite additivity. In Didier Dubois, María Asunción Lubiano, Henri Prade, María Ángeles Gil, Przemyslaw Grzegorzewski, and Olgierd Hryniewicz, editors, *Soft Methods for Handling Variability and Imprecision*, Advances in Soft Computing, pages 243–252. Springer, 2008.
[7] Gert de Cooman, Erik Quaeghebeur, and Enrique Miranda. Exchangeable lower previsions. *Bernoulli*, 15(3):721–735, 2009.
[8] Gert de Cooman, Matthias C. M. Troffaes, and Enrique Miranda. n-Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347:143–156, 2008.
[9] Morris H. DeGroot and Mark J. Schervisch. *Probability and Statistics*. Pearson, fourth edition, 2011.
[10] Dieter Denneberg. *Non-additive measure and integral*. Kluwer, Dordretch, 1994.
[11] Didier Dubois and Henri Prade. *Possibility Theory – An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.
[12] Filip Hermans and Gert de Cooman. Characterisation of ergodic upper transition operators. *International Journal of Approximate Reasoning*, 53(4):573–583, 2012.
[13] Olav Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, second edition, 2002.
[14] Isaac Levi. *The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability, and Chance*. MIT Press, Cambridge, 1983.

[15] Enrique Miranda, Inés Couso, and Pedro Gil. Approximation of upper and lower probabilities by measurable selections. *Information Sciences*, 180(8):1407–1417, 2010.
[16] Enrique Miranda, Gert de Cooman, and Inés Couso. Lower previsions induced by multi-valued mappings. *Journal of Statistical Planning and Inference*, 133(1):177–197, 2005.
[17] Damjan Škulj. The use of Markov operators to constructing generalised probabilities. *International Journal of Approximate Reasoning*, 52(9):1392–1408, 2011.
[18] Peter Walley. Coherent lower (and upper) probabilities. Statistics Research Report 22, University of Warwick, Coventry, 1981.
[19] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
[20] Larry A. Wasserman. Belief functions and statistical inference. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 18(3):183–196, 1990.
[21] Larry A Wasserman. Prior envelopes based on belief functions. *Annals Of Statistics*, 18(1):454–464, 1990.

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

UNIVERSITY OF OVIEDO, DEP. OF STATISTICS AND O.R.–C-CALVO SOTELO, S/N 33007 OVIEDO, SPAIN.
E-mail address: mirandaenrique@uniovi.es