A Particular Upper Expectation as Global Belief Model for Discrete-Time Finite-State Uncertain Processes

Natan T'Joens, Jasper De Bock, Gert de Cooman

FLip, Ghent University, Technologiepark 125, 9052 Zwijnaarde, Belgium

Abstract

To model discrete-time finite-state uncertain processes, we argue for the use of a global belief model in the form of an upper expectation that is the most conservative one under a set of basic axioms. Our motivation for these axioms, which describe how local and global belief models should be related, is based on two possible interpretations for an upper expectation: a behavioural one similar to Walley's, and an interpretation in terms of upper envelopes of linear expectations. We show that the most conservative upper expectation satisfying our axioms, that is, our model of choice, coincides with a particular version of the game-theoretic upper expectation introduced by Shafer and Vovk. This has two important implications: it guarantees that there is a unique most conservative global belief model satisfying our axioms; and it shows that Shafer and Vovk's model can be given an axiomatic characterisation and thereby provides an alternative motivation for adopting this model, even outside their game-theoretic framework. Finally, we relate our model to the upper expectation resulting from a traditional measure-theoretic approach. We show that this measure-theoretic upper expectation also satisfies the proposed axioms, which implies that it is dominated by our model or, equivalently, the game-theoretic model. Moreover, if all local models are precise, all three models coincide.

Keywords: upper expectations, uncertain processes, coherence, game-theoretic probability, measure-theoretic probability, global uncertainty model

1. Introduction

There are various ways to describe discrete-time uncertain processes, such as Markov processes, mathematically. For many, measure theory is and has been the preferred framework to describe the uncertain dynamics of such processes [1–3]. Others have used martingales or a game-theoretic approach [4–6]. However, the common starting point for all these approaches are the local belief models. They describe the dynamics of the process from one time instant to the next and they are typically learned from measurements or elicited from experts. For instance, in some specific cases, one typically may have information about 'the probability of throwing heads on the next coin toss', 'the expected number of goods that are sold by a certain shop on a single day', 'the probability of rain tomorrow', ... In a measure-theoretic context, such local belief models are presented in the form of probability charges or probability measures on the local state space; in a game-theoretic context, sets of allowable bets are used. The latter belong to a family of so-called imprecise probability models [7–9], e.g. upper and lower previsions (or expectations), sets of desirable gambles, credal sets... Such imprecise probability models generalise traditional precise models, in the sense that they allow for indecision and for a representation that only expresses partial knowledge about the probabilities. In order to allow for imprecision in a measure-theoretic context, we will often consider sets of probability measures (or charges) on the local state space. Then, if local state-spaces are assumed to be finite, as in our case, the game-theoretic and measure-theoretic local descriptions are mathematically equivalent. This will be clarified in Sections 2 and 3.

In practice however, we are typically interested in more general inferences such as 'the (lower and upper) expected number of tosses until the first tails is thrown', 'the (lower and upper) probability of being out of stock on a given day', ... It is less straightforward how we should directly learn about such more complicated inferences and, even if we could in principle do so, it is often not possible or feasible to gather sufficient information because of

time and budget limitations. Hence, the question arises: 'How and to which extent, can we extend the information captured in the local models towards global information about the entire process?'.

Measure theory relies on countable additivity to do this elegantly, and this results in a mathematically powerful, but rather abstract framework. The mathematical arguments are often encumbered by measurability constraints, and it is not always clear how to extend classical measure-theoretic notions to an imprecise probabilities setting. On the other hand, Shafer and Vovk's game-theoretic framework defines global upper expectations constructively using the concept of a 'supermartingale'. Though it allows us to directly model imprecision and does not require measurability constraints, the involved upper expectations thus far lack a characterisation in terms of mathematical properties. This impacts the generality of their model in the negative, since it can therefore only be motivated from a game-theoretic point of view.

We propose a different, third approach. Our aim is to establish a global belief model, in the form of a conditional upper expectation, that extends the information contained in the local models by using a number of mathematical properties. Notably, this model will not be bound to a single interpretation. Indeed, its characterising properties can be justified starting from a number of different interpretations. We consider and discuss two of the most significant ones in Section 4; a behavioural interpretation that is more in line with a game-theoretic approach to uncertainty, and an interpretation in terms of upper envelopes of linear expectation operators. We then define the desired global model as the most conservative—the unique model that does not contain any additional information—under this particular set of properties. In other words, it is the so-called natural extension [7, 9, 10] under this set of properties.

Care should be taken here though. While the concept of natural—that is, most conservative—extension is well-established in the field of imprecise probabilities, it usually refers to the natural extension *under coherence*. The advantage of this notion of coherence, and the related natural extension, is that it can be applied to general uncertainty models. However, the scope and properties of the resulting extension can sometimes be rather weak. A first issue with natural extension under coherence is that it traditionally only allows us to work with *bounded* variables. The problem of how to extend this method to a broader class of variables was addressed in [11, Chapter 13], yet, only for the class of *unbounded real-valued* variables, whereas we wish to consider extended real-valued variables as well. There is a second, more profound issue though, which is that the natural extension under coherence lacks some basic continuity properties that we feel are desirable. For example, in our context of uncertain processes, natural extension under coherence would lead to a global upper expectation that does not necessarily satisfy upward monotone convergence, even if we were to restrict attention to bounded variables.[1] Furthermore, and related to this, the corresponding inferences are often very conservative, or even vacuous. If we would only be interested in *bounded* variables and uncertain processes with a *finite* time horizon, all of these issues would not come to the fore, and the natural extension under coherence would then be a valuable model; see for example [12]. However, in our—vastly—more general setting, where we will be looking at *extended real-valued* variables and uncertain processes with *infinite* time horizons, we find this model unsatisfactory for the reasons described above. We will therefore not be considering the natural extension under coherence (alone); instead, we will consider the natural extension under a different set of axioms. Unlike the coherence axioms, our axioms are specifically designed with the purpose of extending local models in the context of uncertain processes (with infinite time horizons). As we will see, the resulting natural—most conservative—extension yields more satisfying results when it comes to the mathematical properties of the resulting global model.

In Section 5, we give an alternative characterisation of our most conservative model by showing that it is the unique operator that satisfies a specific set of properties. This alternative characterisation is particularly interesting from a technical point of view, since it allows us to establish our main result in Section 6: that our

---

[1]The experienced (or interested) reader can verify this using the following example (the notation and terminology is introduced further on in the paper). Interestingly, it shows that upward monotone convergence fails even for increasing sequences of n-measurable gambles that converge to a limit that is itself a gamble. Fix any $x \in \mathcal{X}$ and let $\overline{Q}_s(f) := f(x)$ for all $s \in \mathcal{X}^*$; this means that in every situation, the next state will be x with (lower) probability 1. Consider now the increasing sequence $\{f_n\}_{n \in \mathbb{N}}$ of n-measurable gambles defined, for all $n \in \mathbb{N}$, by $f_n(\omega) := 0$ if $\omega_i = x$ for all $i \in \{1, \cdots, n\}$ and $f_n(\omega) := 1$ for all other $\omega \in \Omega$. Then it can be checked that the value of the natural extension under coherence is 0 for all gambles f_n, but that it is 1—and therefore vacuous—for the limit $f := \lim_{n \to +\infty} f_n$. This means that the (upper) probability of observing a state that is not x in the first n steps, for any n, is zero, but that the upper probability of ever observing such a state is 1. We are indebted to Matthias C. M. Troffaes for this example—or rather, a very similar one that inspired it. He presented it in a talk at WPMSIIP 2013 in Lugano, and recently also discussed it with us in private communication; it seems to be unpublished though.

model coincides with a version of the game-theoretic upper expectation introduced by Shafer and Vovk [4, 5]. On the one hand, this serves as an additional motivation for our model. On the other hand, it establishes a concrete axiomatisation for the game-theoretic upper expectation. Section 8 aims at strengthening the relevance of the most conservative model that results from our axioms by listing its most salient properties.

In the final section, we couch our findings in a more measure-theoretic language. As a first step, we put forward a sensible way of applying measure theory in our imprecise probabilities setting. To do this, we draw inspiration from earlier work done in our group [13]. Subsequently, we show that the measure-theoretic upper expectation this results in, satisfies the axioms presented in Section 4 and is therefore always dominated by the most conservative one satisying these axioms—which is our model. Moreover, the fact that this model, entirely based on measure-theoretic principles, also satisfies our proposed axioms, is to be seen as an additional motivation to adopt them. Furthermore, in a context where all local models are precise, meaning that every single one of them can be represented by a single probability mass function on the local state space, both upper expectations—and hence, also the game-theoretic one—turn out to be one and the same model.

As a final remark, we want to mention that our current work extends upon an earlier conference contribution [14]. The added material is mainly gathered in the penultimate section; it concerns the study of how our model is related to a possible measure-theoretic model. Apart from that, another important difference with the conference contribution lies in how local models are extended in order to become consistent with the game-theoretic approach; see the first part of Section 6 and compare this with [14, Section 2].

2. Upper Expectations

We denote the set of all natural numbers, without 0, by $\mathbb{N}$, and let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. The set of extended real numbers is denoted by $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$ and is endowed with the usual order and the usual order topology (corresponding to the two-point compactification of $\mathbb{R}$). The set of positive real numbers is denoted by $\mathbb{R}_{>0}$ and the set of non-negative real numbers by $\mathbb{R}_{\geq 0}$. We also adopt the usual conventions for the addition between the reals and $+\infty$ and $-\infty$, and the conventions that $+\infty - \infty = -\infty + \infty = +\infty$ and $0 \cdot (+\infty) = 0 \cdot (-\infty) = 0$.

Informally, we consider a subject who is uncertain about the value that some variable Y assumes in a non-empty set $\mathscr{Y}$. More formally, we call any map on $\mathscr{Y}$ a *variable*; our informal Y is a special case: it corresponds to the identity map on $\mathscr{Y}$. A subject's uncertainty about the unknown value of Y can then be represented by an *upper expectation* $\overline{\mathrm{E}}$: an extended real-valued map defined on some subset $\mathscr{D}$ of the set $\overline{\mathscr{L}}(\mathscr{Y})$ of all extended real-valued variables on $\mathscr{Y}$. An element f of $\overline{\mathscr{L}}(\mathscr{Y})$ is simply called an *extended real variable*. We say that f is bounded below if there is some real c such that $f(y) \geq c$ for all $y \in \mathscr{Y}$, and we say that f is bounded above if $-f$ is bounded below. We call a sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of extended real variables *uniformly bounded below* if there is a real c such that $f_n(y) \geq c$ for all $n \in \mathbb{N}_0$ and $y \in \mathscr{Y}$. An important role will be reserved for elements f of $\overline{\mathscr{L}}(\mathscr{Y})$ that are bounded, meaning that they are bounded above *and* below. Bounded real-valued variables on $\mathscr{Y}$ are called *gambles* on $\mathscr{Y}$, and we use $\mathscr{L}(\mathscr{Y})$ to denote the set of all of them. The set of all bounded below elements of $\overline{\mathscr{L}}(\mathscr{Y})$ is denoted by $\overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{Y})$.

Consider now the special case that $\overline{\mathrm{E}}$ is defined on at least the set of all bounded real-valued variables; so $\mathscr{L}(\mathscr{Y}) \subseteq \mathscr{D}$. Then we call $\overline{\mathrm{E}}$ *boundedly coherent* [7–9] if it satisfies the following three coherence axioms:

C1. $\overline{\mathrm{E}}(f) \leq \sup f$ for all $f \in \mathscr{L}(\mathscr{Y})$;   [upper bound]

C2. $\overline{\mathrm{E}}(f + g) \leq \overline{\mathrm{E}}(f) + \overline{\mathrm{E}}(g)$ for all $f, g \in \mathscr{L}(\mathscr{Y})$;   [sub-additivity]

C3. $\overline{\mathrm{E}}(\lambda f) = \lambda \overline{\mathrm{E}}(f)$ for all $\lambda \in \mathbb{R}_{\geq 0}$ and $f \in \mathscr{L}(\mathscr{Y})$.   [non-negative homogeneity]

In the particular case where $\mathscr{D} = \mathscr{L}(\mathscr{Y})$, we will simply say that $\overline{\mathrm{E}}$ is *coherent*. For any (boundedly) coherent $\overline{\mathrm{E}}$, if we let $\underline{\mathrm{E}}$ be the conjugate lower expectation defined by $\underline{\mathrm{E}}(f) := -\overline{\mathrm{E}}(-f)$ for all $f \in -\mathscr{D}$, then the following additional properties follow from C1–C3:

C4. $f \leq g \Rightarrow \overline{\mathrm{E}}(f) \leq \overline{\mathrm{E}}(g)$ for all $f, g \in \mathscr{L}(\mathscr{Y})$;   [monotonicity]

C5. $\inf f \leq \underline{\mathrm{E}}(f) \leq \overline{\mathrm{E}}(f) \leq \sup f$ for all $f \in \mathscr{L}(\mathscr{Y})$;   [bounds]

C6. $\overline{\mathrm{E}}(f + \mu) = \overline{\mathrm{E}}(f) + \mu$ for all $\mu \in \mathbb{R}$ and all $f \in \mathscr{L}(\mathscr{Y})$; [constant additivity]

C7. for any $f \in \mathscr{L}(\mathscr{Y})$ and any sequence $\{f_n\}_{n \in \mathbb{N}_0}$ in $\mathscr{L}(\mathscr{Y})$: [uniform continuity]

$$\lim_{n \to +\infty} \sup |f - f_n| = 0 \Rightarrow \lim_{n \to +\infty} \overline{\mathrm{E}}(f_n) = \overline{\mathrm{E}}(f).$$

Proof. We only prove C5, which clearly implies that $\overline{\mathrm{E}}$ is real-valued on $\mathscr{L}(\mathscr{Y})$, and the remaining properties then follow from the standard argumentation in [7, Section 2.6.1.].

First, note that $\overline{\mathrm{E}}(0) = 0$ because of C3 and our convention that $0 \cdot (+\infty) = 0 \cdot (-\infty) = 0$. Therefore, for all $f \in \mathscr{L}(\mathscr{Y})$, it follows from C2 that $0 \leq \overline{\mathrm{E}}(f) + \overline{\mathrm{E}}(-f)$, or equivalently, that $-\overline{\mathrm{E}}(-f) \leq \overline{\mathrm{E}}(f)$. Applying C1 to both sides, we find that $\inf f = -\sup(-f) \leq -\overline{\mathrm{E}}(-f) \leq \overline{\mathrm{E}}(f) \leq \sup f$. The result now follows readily from the definition of $\underline{\mathrm{E}}$. $\qquad\square$

A gamble f is typically interpreted as an uncertain reward or gain that depends on the value that Y takes in $\mathscr{Y}$; if Y takes the value y, the (possibly negative) gain is $f(y)$. Then, according to Walley's behavioural interpretation [7], the upper expectation $\overline{\mathrm{E}}(f)$ of a gamble f is a subject's infimum selling price for the gamble f, implying that, for any $\alpha > \overline{\mathrm{E}}(f)$, the subject is willing to accept the gamble $\alpha - f$.[2] Axioms C1–C3 are then called rationality axioms, since they ensure that these selling prices are chosen rationally. Alternatively, any coherent upper expectation on $\mathscr{L}(\mathscr{Y})$ can equivalently be represented by a set of *linear expectations* on $\mathscr{L}(\mathscr{Y})$: coherent upper expectations on $\mathscr{L}(\mathscr{Y})$ that are self-conjugate, meaning that $\overline{\mathrm{E}}(f) = -\overline{\mathrm{E}}(-f)$ for all $f \in \mathscr{L}(\mathscr{Y})$. More precisely, it follows from [7, Section 3.3.3] that any coherent upper expectation $\overline{\mathrm{E}}$ on $\mathscr{L}(\mathscr{Y})$ is the upper envelope of the set of all linear expectations on $\mathscr{L}(\mathscr{Y})$ that are dominated by $\overline{\mathrm{E}}$; so $\overline{\mathrm{E}}(f) = \sup\{\mathrm{E}(f) : (\forall g \in \mathscr{L}(\mathscr{Y}))\mathrm{E}(g) \leq \overline{\mathrm{E}}(g)\}$, where E ranges over the linear expectations on $\mathscr{L}(\mathscr{Y})$. Moreover, according to [9, Theorem 8.15], linear expectations on $\mathscr{L}(\mathscr{Y})$ are in a one-to-one relation with *probability charges* on the power set $\mathscr{P}(\mathscr{Y})$ of $\mathscr{Y}$, being maps $p : \mathscr{P}(\mathscr{Y}) \to \mathbb{R}_{\geq 0}$ that are finitely additive and where $p(\mathscr{Y}) = 1$ and $p(\emptyset) = 0$. Such probability charges are more general than the more conventional *probability measures*, which additionally require σ-additivity and typically also that the domain should be restricted to a σ-algebra on $\mathscr{Y}$. However, when $\mathscr{Y}$ is finite, this distinction disappears, and we can simply limit ourselves to working with probability mass functions on $\mathscr{Y}$; these can be seen as probability charges or measures restricted to the domain of all singletons. In that case, we have that $\mathrm{E}(f) = \sum_{y \in \mathscr{Y}} f(y)p(y)$ for any linear expectation E on $\mathscr{L}(\mathscr{Y})$ and any $f \in \mathscr{L}(\mathscr{Y})$, where p is the unique probability mass function on $\mathscr{Y}$ defined by $p(y) := \mathrm{E}(\mathbf{1}_y)$ for all $y \in \mathscr{Y}$. Here, we used $\mathbf{1}_A$ to denote the *indicator* of $A \subseteq \mathscr{Y}$: the gamble on $\mathscr{Y}$ assuming the value 1 on A and 0 elsewhere, and where we will not distinguish in notation between y and $\{y\}$. Hence, if $\mathscr{Y}$ is finite, any coherent upper expectation $\overline{\mathrm{E}}$ on $\mathscr{L}(\mathscr{Y})$ can alternatively be represented by the corresponding set of probability mass functions

$$\mathbb{P}_{\overline{\mathrm{E}}} := \Big\{ p \in \mathbb{P} : \big(\forall f \in \mathscr{L}(\mathscr{Y})\big) \sum_{y \in \mathscr{Y}} f(y)p(y) \leq \overline{\mathrm{E}}(f) \Big\}, \tag{1}$$

where $\mathbb{P}$ denotes the set of all probability mass functions on $\mathscr{Y}$. It then follows from [15, Theorem 2][3] that

$$\overline{\mathrm{E}}(f) = \max\Big\{ \sum_{y \in \mathscr{Y}} f(y)p(y) : p \in \mathbb{P}_{\overline{\mathrm{E}}} \Big\} \text{ for all } f \in \mathscr{L}(\mathscr{Y}). \tag{2}$$

This alternative representation of a coherent upper expectation will be our starting point in Section 9, where we aim to establish a measure-theoretic global belief model in the form of an upper expectation.

It follows from the discussion above that, in general, (boundedly coherent) upper expectations can be interpreted in at least two possible ways: in a direct behavioural manner in terms of selling prices for gambles, or as a supremum over—an upper envelope of—a set of linear expectations. As will be discussed in Section 6, the

---

[2]Strictly speaking, the interpretation as infimum selling price only guarantees that there are $\alpha > \overline{\mathrm{E}}(f)$ *arbitrarily close to* $\overline{\mathrm{E}}(f)$ for which our subject accepts $\alpha - f$ (by definition of the infimum). Our claim that this is true for *all* $\alpha > \overline{\mathrm{E}}(f)$ is based on an additional implicit assumption that if a subject is willing to sell a gamble for some price, then she is also willing to sell that gamble for any higher price.

[3]More specifically, Equation (2) does not follow from [15, Theorem 2] itself, but rather from the last paragraph of its proof.

behavioural interpretation allows us to motivate the game-theoretic approach proposed by Shafer and Vovk. On the other hand, an interpretation in terms of upper envelopes of linear expectations seems more natural from a measure-theoretic point of view, where probability measures or mass functions are regarded as the primary objects for describing uncertainty. Linear expectations are then typically obtained by integration and limiting the domain to measurable functions. In this paper, we will not allow ourselves to be bound to any of these two interpretations. Instead, we will motivate the defining properties of our proposed global belief model in terms of either of these interpretations.

3. Upper Expectations in Discrete-Time Uncertain Processes

We consider a *discrete-time uncertain process*: a sequence $X_1, X_2, ..., X_n, ...$ of uncertain states, where the state X_k at each discrete time $k \in \mathbb{N}$ takes values in a fixed non-empty set $\mathscr{X}$, called the *state space*. We will assume that this state space $\mathscr{X}$ is *finite*. Let a *situation* $x_{1:n}$ be any string $(x_1, ..., x_n) \in \mathscr{X}_{1:n} := \mathscr{X}^n$ of possible state values with finite length $n \in \mathbb{N}_0$. In particular, the unique empty string $x_{1:0}$, denoted by $\square$, is called the *initial situation*: $\mathscr{X}_{1:0} := \{\square\}$. We denote the set of all situations by $\mathscr{X}^* := \cup_{n \in \mathbb{N}_0} \mathscr{X}_{1:n}$. We will also use the generic notations s and t to denote situations. Furthermore, when we write $x_{1:n} \in \mathscr{X}^*$, we implicitly assume that $n \in \mathbb{N}_0$.

To model our uncertainty about the dynamics of an uncertain process, we associate, with every situation $x_{1:n} \in \mathscr{X}^*$, a coherent upper expectation $\overline{Q}_{x_{1:n}}$ on $\mathscr{L}(\mathscr{X})$. This upper expectation expresses a subject's beliefs about the uncertain value of the next state X_{n+1}, when she has observed that $X_1 = x_1, X_2 = x_2, \cdots, X_{n-1} = x_{n-1}$ and $X_n = x_n$. Hence, it gives information about how the process changes from one time instant to the next. We will therefore also refer to $\overline{Q}_{x_{1:n}}$ as the *local* model or the *local* upper expectation associated with $x_{1:n}$. We gather all local models in an *imprecise probability tree*: a function $\overline{Q}_\bullet$ that maps any situation $s \in \mathscr{X}^*$ to its corresponding coherent upper expectation $\overline{Q}_s$. Hence, such an imprecise probability tree $\overline{Q}_\bullet$, which we will also simply denote by $\overline{Q}$, represents the dynamics of the uncertain process as a whole.

As discussed in the previous section, we consider two possible ways of interpreting the coherent upper expectations $\overline{Q}_s$: in terms of upper envelopes of linear expectations and in terms of betting behaviour. The latter is, by nature, more oriented towards a game-theoretic approach, whereas the interpretation in terms of upper envelopes of linear expectations will seem more natural for a reader with a measure-theoretic background. In fact, since $\mathscr{X}$ is assumed to be finite, the coherent upper expectation $\overline{Q}_s$, for any $s \in \mathscr{X}^*$, can be represented equivalently by the set $\mathbb{P}_s := \mathbb{P}_{\overline{Q}_s}$ of probability mass functions on $\mathscr{X}$, related to it by Equations (1) and (2). In measure-theoretic contexts, this representation in terms of $\mathbb{P}_s$ is then typically preferred over the representation in terms of $\overline{Q}_s$.

The mathematical equivalence between the different approaches—the game-theoretic, the measure-theoretic and ours—on a local level is crucial, because only then, we can meaningfully compare the differences of their associated uncertainty models on a global level. Indeed, though it is locally only a matter of interpretational differences, it is not immediately clear how these different approaches relate on a global level, where we look at the entire dynamics of the process and no longer only at the transition from one time instant to the next. Each approach will extend the local models using different concepts, assumptions and methods, resulting in possibly different global models. Our upper expectation approach will use a limited set of simple properties, that, as we will argue, are desirable for a global model to have, under both of the interpretations we consider. Before we proceed to do so, we finish this section by introducing some further notation about uncertain processes.

An infinite sequence $x_1 x_2 x_3 \cdots$ of state values is called a *path*, which we denote by $\omega := x_1 x_2 x_3 \cdots$. We gather all paths in the *sample space* $\Omega := \mathscr{X}^{\mathbb{N}}$. For any path $\omega := x_1 x_2 \cdots \in \Omega$, the situation $x_{1:n} := x_1 x_2 \cdots x_n$ that consists of its first n state values is denoted by $\omega^n \in \mathscr{X}_{1:n}$. The state value x_n at time n is denoted by $\omega_n \in \mathscr{X}$. An *event* $A \subseteq \Omega$ is a collection of paths, and in particular, the *cylinder event* $\Gamma(x_{1:n}) := \{\omega \in \Omega : \omega^n = x_{1:n}\}$ of some situation $x_{1:n} \in \mathscr{X}^*$, is the set of all paths $\omega \in \Omega$ that 'go through' the situation $x_{1:n}$.

A variable on Ω is called a *global* variable and we gather all extended real(-valued) global variables in the set $\overline{\mathbb{V}} := \overline{\mathscr{L}}(\Omega)$. Similarly, we let $\overline{\mathbb{V}}_{\mathrm{b}} := \overline{\mathscr{L}}_{\mathrm{b}}(\Omega)$ and $\mathbb{V} := \mathscr{L}(\Omega)$. For any natural $k \leq \ell$, we denote by $X_{k:\ell}$ the global variable defined by $X_{k:\ell}(\omega) := (\omega_k, ..., \omega_\ell)$ for all $\omega \in \Omega$. As such, the state $X_k := X_{k:k}$ at time k can also be regarded as a global variable. Moreover, for any natural $k \leq \ell$ and any map $f : \mathscr{X}^{\ell-k+1} \to \overline{\mathbb{R}}$, we will write $f(X_{k:\ell})$ to denote the global extended real variable defined by $f(X_{k:\ell}) := f \circ X_{k:\ell}$. We call a global extended real variable f *n-measurable*

for some $n \in \mathbb{N}_0$, if it only depends on the initial n state values; so $f(\omega_1) = f(\omega_2)$ for any two paths ω_1 and ω_2 such that $\omega_1^n = \omega_2^n$. We will then use the notation $f(x_{1:n})$ for its constant value $f(\omega)$ on all paths $\omega \in \Gamma(x_{1:n})$. Clearly, the indicator $\mathbb{I}_{x_{1:n}} \coloneqq \mathbf{1}_{\Gamma(x_{1:n})}$ of the cylinder event $\Gamma(x_{1:n})$ for $x_{1:n} \in \mathscr{X}^*$ is an n-measurable variable. Finally, we call any $f \in \overline{\mathbb{V}}$ *finitary* if it is n-measurable for some $n \in \mathbb{N}_0$. We gather all finitary *gambles* in $\mathbb{V}_{\mathrm{fin}}$.

4. In Search of a Global Belief Model

Any extended real-valued map $\overline{\mathrm{E}} \colon \overline{\mathbb{V}} \times \mathscr{X}^* \to \overline{\mathbb{R}} \colon (f, s) \mapsto \overline{\mathrm{E}}(f|s)$ will be called a *global upper expectation*. The corresponding *global lower expectation* $\underline{\mathrm{E}} \colon \overline{\mathbb{V}} \times \mathscr{X}^* \to \overline{\mathbb{R}} \colon (f, s) \mapsto \underline{\mathrm{E}}(f|s)$ is then defined by conjugacy: $\underline{\mathrm{E}}(f|s) \coloneqq -\overline{\mathrm{E}}(-f|s)$ for all $f \in \overline{\mathbb{V}}$ and all $s \in \mathscr{X}^*$. These lower and upper expectations also give rise to upper and lower probabilities. In particular, for any event $A \subseteq \Omega$ and any $s \in \mathscr{X}^*$, $\overline{\mathrm{P}}(A|s) \coloneqq \overline{\mathrm{E}}(\mathbf{1}_A|s)$ is the *upper probability* of A conditional on s and, similarly, $\underline{\mathrm{P}}(A|s) \coloneqq \underline{\mathrm{E}}(\mathbf{1}_A|s)$ is the *lower probability* of A conditional on s. Throughout this work, however, we will focus almost entirely on (global) upper expectations, and will regard (global) lower expectations and upper and lower probabilities as derived notions.

Given an imprecise probability tree $\overline{Q}$ that associates a local upper expectation $\overline{Q}_s$ with every situation $s \in \mathscr{X}^*$, we aim to define a global upper expectation $\overline{\mathrm{E}}$ that extends the information included in these local models in a 'rational' manner. To do so, we impose the following properties, where we adopt the notation that $\overline{\mathrm{E}}(f|X_{1:n}) \coloneqq \overline{\mathrm{E}}(f|\cdot) \circ X_{1:n}$ for all $f \in \overline{\mathbb{V}}$ and all $n \in \mathbb{N}_0$, and where limits of variables are intended to be taken pointwise.

P1. $\overline{\mathrm{E}}(f(X_{n+1})|x_{1:n}) = \overline{Q}_{x_{1:n}}(f)$ for all $f \in \mathscr{L}(\mathscr{X})$ and all $x_{1:n} \in \mathscr{X}^*$.

P2. $\overline{\mathrm{E}}(f|s) = \overline{\mathrm{E}}(f\,\mathbb{I}_s|s)$ for all $f \in \mathbb{V}_{\mathrm{fin}}$ and all $s \in \mathscr{X}^*$.

P3. $\overline{\mathrm{E}}(f|X_{1:n}) \le \overline{\mathrm{E}}(\overline{\mathrm{E}}(f|X_{1:n+1})|X_{1:n})$ for all $f \in \mathbb{V}_{\mathrm{fin}}$ and all $n \in \mathbb{N}_0$.

P4. $f \le g \Rightarrow \overline{\mathrm{E}}(f|s) \le \overline{\mathrm{E}}(g|s)$ for all $f, g \in \overline{\mathbb{V}}$ and all $s \in \mathscr{X}^*$.

P5. For any sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of finitary gambles that is uniformly bounded below and any $s \in \mathscr{X}^*$:

$$\lim_{n \to +\infty} f_n = f \Rightarrow \limsup_{n \to +\infty} \overline{\mathrm{E}}(f_n|s) \ge \overline{\mathrm{E}}(f|s).$$

To motivate P1–P5, we need to attach some interpretation to $\overline{\mathrm{E}}$. We will consider two particular ones, similar to what we have done for the (boundedly) coherent upper expectations in Section 2.

We start from the interpretation of a global *gamble* $f \in \mathbb{V}$ as an uncertain reward depending on the uncertain path ω that the uncertain process takes in Ω. However, it is not clear what this means operationally if the gamble f depends on the entire length of the path. Indeed, the gamble f depends on an infinite number of subsequent state values, so there is no point in time when we can determine the actual reward linked to the gamble f. The same interpretational problem arises when considering unbounded or extended real variables (on a general set $\mathscr{Y}$ such as Ω). The simple interpretation of an uncertain reward does not suffice there because the reward itself can be unbounded or infinite, which is unrealistic—or even meaningless—in operational practice. For this reason, we prefer to only attach a direct operational meaning to the value $\overline{\mathrm{E}}(f|s)$ of a global upper expectation $\overline{\mathrm{E}}$ for a *finitary gamble* $f \in \mathbb{V}_{\mathrm{fin}}$ conditional on a situation $s \in \mathscr{X}^*$. Such finitary gambles can be given an operationally meaningful behavioural interpretation as uncertain rewards because they take real values[4] and only depend on the state at a finite number of time instances.

We distinguish the following two ways for interpreting the global upper expectation $\overline{\mathrm{E}}(f|s)$ of a finitary gamble $f \in \mathbb{V}_{\mathrm{fin}}$ conditional on a situation $s \in \mathscr{X}^*$:

---

[4]One could also question the meaning of a direct behavioural interpretation for some specific (real-valued) gambles. For instance, how do we exchange money if the gamble's value is equal to π? In this case, we can only give an indirect interpretation in terms of simpler gambles that are 'sufficiently' close.

- **Behavioural interpretation.** It is a subject's infimum selling price for f *contingent* on the event $\Gamma(s)$, implying that,[5] for any $\alpha > \overline{\mathrm{E}}(f|s)$, she is willing to accept the uncertain reward associated with the gamble $\mathbb{I}_s(\alpha - f)$.

- **Interpretation as an upper envelope.** It is the supremum value of $\mathrm{E}(f|s)$, where E belongs to some given set $\mathbb{E}$ of conditional linear expectation operators: $\overline{\mathrm{E}}(f|s) = \sup\{\mathrm{E}(f|s)\colon \mathrm{E} \in \mathbb{E}\}$.

Since P1–P3 only apply to finitary gambles, a direct justification for these axioms can be given quite easily in each of the above interpretations. Property P1 requires that the global model $\overline{\mathrm{E}}$ should be compatible with the local models $\overline{\mathrm{Q}}_s$. The desirability of this property is self-evident, no matter which interpretation is used. Property P2 requires that the upper expectation of a finitary gamble f conditional on s should only depend on the value of f on the paths $\omega \in \Gamma(s)$. This property is clearly desirable when using the behavioural interpretation, because $\mathbb{I}_s(\alpha - f)$ only depends on the restriction of f to $\Gamma(s)$. It is also quite evident that this property is desirable for an upper envelope of conditional linear expectations, because P2 should in particular also hold for such conditional linear expectations themselves. Similarly, that P3 should hold under the upper envelope interpretation, is motivated by the fact that conditional linear expectations satisfy P3 with equality—then better known as the law of iterated expectations; an upper envelope of conditional expectations is therefore guaranteed to satisfy P3, under the assumption that P4 holds for conditional linear expectations. In order to see that property P3 is also desirable according to the behavioural interpretation, one requires a conditional version of the notion of coherence that we discussed in Section 2 [7, 9, 10].[6] Explaining in more detail why this is the case, and what this conditional notion of coherence exactly entails, would however lead us too far.

Having attached an interpretation to finitary gambles and their conditional upper expectations, we now proceed to do the same for more general variables. We have already argued that no direct operational meaning can be given to such variables. However, this should not be taken to imply that an uncertainty model should not be able to deal with them. In fact, they can serve as useful abstract idealisations of (sequences of) variables that can be given a direct operational meaning. In particular, we will regard any extended real variable f that is bounded below and that can be written as the pointwise limit $\lim_{n \to +\infty} f_n$ of some sequence of finitary gambles $\{f_n\}_{n \in \mathbb{N}_0}$, as an abstract idealisation of f_n for *large n*. We gather these limits in the set

$$\overline{\mathbb{V}}_{\mathrm{b},\lim} := \left\{ f \in \overline{\mathbb{V}}_{\mathrm{b}} \colon f = \lim_{n \to +\infty} f_n \text{ for some sequence } \{f_n\}_{n \in \mathbb{N}_0} \text{ in } \mathbb{V}_{\mathrm{fin}} \right\}.$$

Since P5 applies to precisely these kinds of variables, this axiom can be justified by extending the above idealisation from the variables f to their upper expectations $\overline{\mathrm{E}}(f|s)$. Basically, since f is an abstract idealisation of f_n for large n, $\overline{\mathrm{E}}(f|s)$ should be an abstract idealisation of $\overline{\mathrm{E}}(f_n|s)$ for large n. The practical benefit of this is that we can then use $\overline{\mathrm{E}}(f|s)$ to reason about $\overline{\mathrm{E}}(f_n|s)$ for a generic large value of n, without having to specify the specific value of n. The problem, however, is that the sequence $\{\overline{\mathrm{E}}(f_n|s)\}_{n \in \mathbb{N}_0}$ may not converge. What we then do know for sure, however, is that as n approaches infinity, $\overline{\mathrm{E}}(f_n|s)$ will oscillate between the limit superior and inferior of the sequence $\{\overline{\mathrm{E}}(f_n|s)\}_{n \in \mathbb{N}_0}$. Since we want $\overline{\mathrm{E}}(f|s)$ to serve as an idealisation of $\overline{\mathrm{E}}(f_n|s)$ for generic large values of n, $\overline{\mathrm{E}}(f|s)$ should therefore definitely not exceed the limit superior, as this would result in an unwarranted loss of 'information', in the sense that $\overline{\mathrm{E}}(f|s)$ would be too high—this will be clarified shortly. We therefore impose P5.

The final property that we impose is P4, which states that $\overline{\mathrm{E}}$ should be monotone. For finitary gambles f and g, this follows easily from either of our two different interpretations for $\overline{\mathrm{E}}(f|s)$ and $\overline{\mathrm{E}}(g|s)$. Under a behavioural interpretation, since the reward associated with g is guaranteed to be at least as high as that of f, the same should be true for a subject's infimum selling prices for these two gambles. Under an interpretation in terms of upper envelopes of expectations, monotonicity of the envelope is implied by the monotonicity of each of the individual expectations. If f and g are more general variables—so not necessarily finitary gambles—the motivation for P4 is that, still, higher rewards—even if abstract and idealized—should correspond to higher upper expectations. It is also worth noting that the combination of P4 and P5 implies that $\overline{\mathrm{E}}$ is continuous with respect to non-decreasing sequences of finitary gambles. In a measure-theoretic context, this kind of continuity is usually obtained as a consequence of the assumption of σ-additivity [1, 16].

---

[5]The comment in Footnote 2, applies here as well, suitable adapted to the conditional setting.

[6]There are several versions of conditional coherence [7, 9, 10], however, in the case where variables take values in a finite set, all these different versions are mathematically equivalent.

We will show in Section 6 that P1–P5 are consistent, in the sense that if the local models $\overline{Q}_s$ are coherent, there always is at least one global upper expectation $\overline{E}$ satisfying P1–P5. However, there may be more than one global upper expectation $\overline{E}$ satisfying P1–P5. In that case, the best thing to do, we believe, is to choose the most conservative model among those that satisfy P1–P5, as choosing any other would mean adding 'information' that is not simply implied by our axioms. We will denote this most conservative global upper expectation by $\overline{E}_A$ and let $\underline{E}_A$ be its conjugate lower expectation. As we will see in Section 6, $\overline{E}_A$ is guaranteed to exist, and furthermore coincides with a particular version of the game-theoretic upper expectation defined by Shafer and Vovk [4, 17].

Of course, in order for our definition of $\overline{E}_A$ to make sense, we need to know what it means for an upper expectation $\overline{E}'$ to be more conservative than some other upper expectation $\overline{E}$. We here take this to mean that $\overline{E}'$ is higher than $\overline{E}$. So higher upper expectations are more conservative, or less informative. This is why, in our motivation for P5, we said that a too high $\overline{E}(f|s)$ corresponds to an unwarranted loss of 'information'. That it is indeed reasonable to regard higher expectations as more conservative, can again be motivated using either of the two interpretations that we considered before. Under the behavioural interpretation, higher upper expectations mean higher selling prices, which is clearly more conservative. Using an interpretation in terms of upper envelopes of expectations, higher upper expectations correspond to larger sets of expectations, so weaker constraints on expectations, which is again less informative and hence more conservative.

Having said all this, we would like to stress that—despite our extensive use of them to motivate our axioms—none of the results that we are about to develop hinge on any particular interpretation for upper expectations. Other interpretations could also be adopted, or perhaps even no interpretation at all. All that is needed is that we agree on P1–P5 and on the fact that higher upper expectations are more conservative.

5. An Axiomatisation of $\overline{E}_A$

For a given imprecise probability tree $\overline{Q}$, let $\overline{\mathbb{E}}_{1-2}(\overline{Q})$ denote the set of all global upper expectations satisfying P1–P2, and similarly for $\overline{\mathbb{E}}_{1-4}(\overline{Q})$ and $\overline{\mathbb{E}}_{1-5}(\overline{Q})$. In this section, we introduce sufficient conditions for a global upper expectation $\overline{E}$ to be the most conservative among all upper expectations in $\overline{\mathbb{E}}_{1-5}(\overline{Q})$. We start by considering the domain of finitary gambles and then, step by step, extend the domain and introduce additional conditions on $\overline{E}$, in such a way that it becomes the most conservative on this extended domain.

For any situation $x_{1:n} \in \mathscr{X}^*$ and any $(n+1)$-measurable (global) gamble f, we use $f(x_{1:n}\cdot)$ to denote the gamble on $\mathscr{X}$ that assumes the value $f(x_{1:n+1})$ in $x_{n+1} \in \mathscr{X}$, and then use $f(x_{1:n}X_{n+1})$ as a shorthand for the global gamble $f(x_{1:n}\cdot) \circ X_{n+1}$. The following lemma establishes compatibility with the local models in a stronger way than P1 does.

Lemma 1. *Consider any* $\overline{E} \in \overline{\mathbb{E}}_{1-2}(\overline{Q})$. *Then, for any situation* $x_{1:n} \in \mathscr{X}^*$ *and any* $(n+1)$*-measurable gamble* f,

$$\overline{E}(f|x_{1:n}) = \overline{Q}_{x_{1:n}}(f(x_{1:n}\cdot)).$$

Proof. Fix any $x_{1:n} \in \mathscr{X}^*$ and any $(n+1)$-measurable gamble f. Note that $f(x_{1:n}X_{n+1})\mathbb{I}_{x_{1:n}} = f\mathbb{I}_{x_{1:n}}$ and hence, because of P2, $\overline{E}(f|x_{1:n}) = \overline{E}(f\mathbb{I}_{x_{1:n}}|x_{1:n}) = \overline{E}(f(x_{1:n}X_{n+1})\mathbb{I}_{x_{1:n}}|x_{1:n}) = \overline{E}(f(x_{1:n}X_{n+1})|x_{1:n})$. P1 therefore implies that, indeed, $\overline{E}(f|x_{1:n}) = \overline{Q}_{x_{1:n}}(f(x_{1:n}\cdot))$. $\square$

To make sure that some particular upper expectation $\overline{E} \in \overline{\mathbb{E}}_{1-4}(\overline{Q})$ is a most conservative upper expectation on the domain of all finitary gambles, we impose on it the following property, known as *the law of iterated upper expectations* [4, 18]:

P3$^=$. $\overline{E}(f|X_{1:n}) = \overline{E}(\overline{E}(f|X_{1:n+1})|X_{1:n})$ for all $f \in \mathbb{V}_{\text{fin}}$ and all $n \in \mathbb{N}_0$.

Proposition 2. *Consider any* $\overline{E} \in \overline{\mathbb{E}}_{1-4}(\overline{Q})$ *that satisfies* P3$^=$. *Then, for any* $\overline{E}' \in \overline{\mathbb{E}}_{1-4}(\overline{Q})$, *we have that*

$$\overline{E}(f|s) \geq \overline{E}'(f|s) \text{ for all } f \in \mathbb{V}_{\text{fin}} \text{ and all } s \in \mathscr{X}^*.$$

Proof. For any $p \in \mathbb{N}_0$ and any $(p+1)$-measurable gamble g, we let $\overline{Q}_{X_{1:p}}(g)$ be the p-measurable gamble defined by $\overline{Q}_{X_{1:p}}(g)(\omega) := \overline{Q}_{\omega^p}(g(\omega^p \cdot))$ for all $\omega \in \Omega$. Note that $\overline{Q}_{X_{1:p}}(g)$ is indeed a gamble because of (bounded) coherence [C5] and the fact that g is a gamble. Then, because both the operators $\overline{E}$ and $\overline{E}'$ satisfy P1 and P2, it follows from Lemma 1 that

$$\overline{E}'(g|X_{1:p}) = \overline{Q}_{X_{1:p}}(g) = \overline{E}(g|X_{1:p}). \tag{3}$$

Now fix any $f \in \mathbb{V}_{\text{fin}}$ and any $x_{1:m} \in \mathscr{X}^*$. Since f is finitary, it is n-measurable for some $n \in \mathbb{N}_0$. We can assume that $m + 2 < n$ without loss of generality, because f is obviously also p-measurable for every $p \geq n$. Now, it follows from Equation (3) and the fact that $\overline{Q}_{X_{1:n-1}}(f)$ is an $(n-1)$-measurable gamble, that $\overline{E}'(f|X_{1:n-1}) = \overline{E}(f|X_{1:n-1})$ is an $(n-1)$-measurable gamble. Subsequently, we can apply Equation (3) once more, where $\overline{E}'(f|X_{1:n-1}) = \overline{E}(f|X_{1:n-1})$ now takes the role of g, in order to find that $\overline{E}'(\overline{E}'(f|X_{1:n-1})|X_{1:n-2}) = \overline{E}(\overline{E}(f|X_{1:n-1})|X_{1:n-2})$ is an $(n-2)$-measurable gamble. Then we can continue in the same way, to finally obtain that

$$\overline{E}'(\overline{E}'(\cdots \overline{E}'(f|X_{1:n-1}) \cdots |X_{1:m+1})|x_{1:m}) = \overline{E}(\overline{E}(\cdots \overline{E}(f|X_{1:n-1}) \cdots |X_{1:m+1})|x_{1:m}).$$

By repeatedly applying P3 and P4, we infer that $\overline{E}'(\overline{E}'(\cdots \overline{E}'(f|X_{1:n-1}) \cdots |X_{1:m+1})|x_{1:m}) \geq \cdots \geq \overline{E}'(\overline{E}'(f|X_{1:m+1})|x_{1:m}) \geq \overline{E}'(f|x_{1:m})$. Hence, plugging this back into the previous expression, we have that

$$\overline{E}'(f|x_{1:m}) \leq \overline{E}(\overline{E}(\cdots \overline{E}(f|X_{1:n-1}) \cdots |X_{1:m+1})|x_{1:m}).$$

On the other hand, since $\overline{E}$ satisfies P3$^=$, we infer that $\overline{E}(\overline{E}(\cdots \overline{E}(f|X_{1:n-1}) \cdots |X_{1:m+1})|x_{1:m}) = \cdots = \overline{E}(\overline{E}(f|X_{1:m+1})|x_{1:m}) = \overline{E}(f|x_{1:m})$. So indeed, we find that $\overline{E}'(f|x_{1:m}) \leq \overline{E}(f|x_{1:m})$. $\qquad \square$

Next, we consider the domain $\overline{\mathbb{V}}_{\text{b,lim}} \subset \overline{\mathbb{V}}$ of all extended real variables that are bounded below and that can be written as the pointwise limit of a sequence of finitary gambles. The following condition, together with P3$^=$, turns out to be sufficient for an upper expectation $\overline{E}$ to be a most conservative one on the domain of $\overline{\mathbb{V}}_{\text{b,lim}}$ amongst all upper expectations in $\overline{\mathbb{E}}_{1-5}(\overline{Q})$.

P6. For any $f \in \overline{\mathbb{V}}_{\text{b,lim}}$ and any $s \in \mathscr{X}^*$, there is some sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of n-measurable gambles that is uniformly bounded below and that converges pointwise to f, such that moreover $\limsup_{n \to +\infty} \overline{E}(f_n|s) \leq \overline{E}(f|s)$.

Note that, for an upper expectation $\overline{E} \in \overline{\mathbb{E}}_{1-5}(\overline{Q})$, the inequality $\limsup_{n \to +\infty} \overline{E}(f_n|s) \leq \overline{E}(f|s)$ in P6 is in fact an equality because of P5.

Proposition 3. *Consider any $\overline{E} \in \overline{\mathbb{E}}_{1-5}(\overline{Q})$ that satisfies P3$^=$ and P6. Then, for all $\overline{E}' \in \overline{\mathbb{E}}_{1-5}(\overline{Q})$, we have that*

$$\overline{E}(f|s) \geq \overline{E}'(f|s) \text{ for all } f \in \overline{\mathbb{V}}_{\text{b,lim}} \text{ and all } s \in \mathscr{X}^*.$$

Proof. Fix any $f \in \overline{\mathbb{V}}_{\text{b,lim}}$, any $s \in \mathscr{X}^*$ and any $\overline{E}' \in \overline{\mathbb{E}}_{1-5}(\overline{Q})$. According to P6, there is some sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of n-measurable gambles that is uniformly bounded below and that converges pointwise to f in such a way that $\limsup_{n \to +\infty} \overline{E}(f_n|s) \leq \overline{E}(f|s)$ and therefore, due to P5, also that

$$\limsup_{n \to +\infty} \overline{E}(f_n|s) = \overline{E}(f|s). \tag{4}$$

Because all f_n are finitary gambles and both $\overline{E}$ and $\overline{E}'$ are upper expectations in $\overline{\mathbb{E}}_{1-4}(\overline{Q})$, with $\overline{E}$ additionally satisfying P3$^=$, we can apply Proposition 2 to find that

$$\limsup_{n \to +\infty} \overline{E}'(f_n|s) \leq \limsup_{n \to +\infty} \overline{E}(f_n|s). \tag{5}$$

Furthermore, since $\{f_n\}_{n \in \mathbb{N}_0}$ is a sequence of finitary gambles that is uniformly bounded below and that converges pointwise to f, P5 implies that $\overline{E}'(f|s) \leq \limsup_{n \to +\infty} \overline{E}'(f_n|s)$. Combining this with Equations (4) and (5), we find that, indeed, $\overline{E}'(f|s) \leq \overline{E}(f|s)$. $\qquad \square$

Finally, we consider the entire domain $\overline{\mathbb{V}}$. Now, in order for an upper expectation on $\overline{\mathbb{V}}$ to be the most conservative one, it suffices to additionally impose the following property, as we will show presently.

P7. For any $f \in \overline{\mathbb{V}}$ and any $s \in \mathscr{X}^*$,

$$\overline{\mathrm{E}}(f|s) = \inf\left\{\overline{\mathrm{E}}(g|s) : g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\right\}.$$

Theorem 4. *Consider any $\overline{\mathrm{E}} \in \overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$ that satisfies* P3$^=$, P6 *and* P7. *Then, for all $\overline{\mathrm{E}}' \in \overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$, we have that*

$$\overline{\mathrm{E}}(f|s) \geq \overline{\mathrm{E}}'(f|s) \text{ for all } f \in \overline{\mathbb{V}} \text{ and all } s \in \mathscr{X}^*.$$

Proof. Fix any $f \in \overline{\mathbb{V}}$, any $s \in \mathscr{X}^*$ and any $\overline{\mathrm{E}}' \in \overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$. According to P7, we have that

$$\overline{\mathrm{E}}(f|s) = \inf\left\{\overline{\mathrm{E}}(g|s) : g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\right\}.$$

Then, using Proposition 3, we get

$$\inf\left\{\overline{\mathrm{E}}'(g|s) : g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\right\} \leq \inf\left\{\overline{\mathrm{E}}(g|s) : g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\right\} = \overline{\mathrm{E}}(f|s).$$

Since, for any $g \in \overline{\mathbb{V}}_{\mathrm{b,lim}}$ such that $f \leq g$, we have that $\overline{\mathrm{E}}'(f|s) \leq \overline{\mathrm{E}}'(g|s)$ because $\overline{\mathrm{E}}'$ satisfies P4, it follows that

$$\overline{\mathrm{E}}'(f|s) \leq \inf\left\{\overline{\mathrm{E}}'(g|s) : g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\right\} \leq \overline{\mathrm{E}}(f|s),$$

which completes the proof. $\qquad\square$

We conclude that, according to Theorem 4, if there is some upper expectation $\overline{\mathrm{E}} \in \overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$ that satisfies P3$^=$, P6 and P7, it must be the unique most conservative upper expectation in $\overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$, and it must therefore be the global belief model $\overline{\mathrm{E}}_A$ that we are after. So it now only remains to show that there is at least one—then necessarily unique—global upper expectation in $\overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$ that additionally satisfies P3$^=$, P6 and P7.

6. Game-Theoretic Upper Expectations

In this section, we show that a particular version of the game-theoretic upper expectation developed by Shafer and Vovk [4, 5] belongs to $\overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$ and furthermore satisfies P3$^=$, P6 and P7, which will then—because of Theorem 4—imply the existence of the unique $\overline{\mathrm{E}}_A$ we are in pursuit of. This game-theoretic upper expectation relies on the concept of a *supermartingale*,[7] which is a capital process—the evolution of a subject's capital—that is obtained by betting against a 'forecasting system'. The forecasting system—called 'Forecaster' in Shafer and Vovk's framework—determines for each situation a collection of allowable bets that a subject—called 'Skeptic' in Shafer and Vovk's framework—can choose from. These allowable bets define Forecaster's uncertainty model and are also captured by our notion of an imprecise probability tree. We should point out though that Shafer and Vovk also allow for more general settings where local models need not be boundedly coherent, where state spaces can be infinite and where Forecaster his commitments—that is, the allowable bets for Skeptic—in each situation need not be announced beforehand, but may also depend on previous moves by Skeptic. In that respect, our discussion here will be more particular and limited.

As before, we start from an imprecise probability tree that specifies a local coherent upper expectation $\overline{\mathrm{Q}}_s$ on $\mathscr{L}(\mathscr{X})$ for every situation $s \in \mathscr{X}^*$. However, in their most recent work [5, Part II], Shafer and Vovk start from local models $\overline{\mathrm{E}}_s$ that are defined on the extended domain $\overline{\mathscr{L}}(\mathscr{X})$ of all extended real local variables, and they require these local models $\overline{\mathrm{E}}_s$ to satisfy a modified version of the bounded coherence axioms C1–C3, generalised to extended real variables. Concretely, for any $s \in \mathscr{X}^*$, they impose the following properties on the upper expectation $\overline{\mathrm{E}}_s$ on $\overline{\mathscr{L}}(\mathscr{X})$ that models the corresponding local beliefs:

---

[7] Shafer and Vovk drew inspiration from Jean Ville, whose work [19] constitutes the first major steps towards a theory of probability where (super)martingales, instead of probability measures, function as the fundamental, primary objects.

E1. $\overline{\mathrm{E}}_s(c) = c$ for all $c \in \mathbb{R}$;

E2. $\overline{\mathrm{E}}_s(f + g) \leq \overline{\mathrm{E}}_s(f) + \overline{\mathrm{E}}_s(g)$ for all $f, g \in \overline{\mathscr{L}}(\mathscr{X})$;

E3. $\overline{\mathrm{E}}_s(\lambda f) = \lambda \overline{\mathrm{E}}_s(f)$ for all $\lambda \in \mathbb{R}_{>0}$ and all $f \in \overline{\mathscr{L}}(\mathscr{X})$;

E4. $f \leq g \Rightarrow \overline{\mathrm{E}}_s(f) \leq \overline{\mathrm{E}}_s(g)$ for all $f, g \in \overline{\mathscr{L}}(\mathscr{X})$.

E5. $\lim_{n \to +\infty} \overline{\mathrm{E}}_s(f_n) = \overline{\mathrm{E}}_s\left(\lim_{n \to +\infty} f_n\right)$ for any non-decreasing and non-negative sequence $\{f_n\}_{n \in \mathbb{N}_0}$ in $\overline{\mathscr{L}}(\mathscr{X})$.

Hence, in order to place ourselves squarely in their framework, we need to extend the domain of our local models $\overline{\mathrm{Q}}_s$ from $\mathscr{L}(\mathscr{X})$ to $\overline{\mathscr{L}}(\mathscr{X})$ in such a way that they satisfy E1–E5. For any $f \in \overline{\mathscr{L}}(\mathscr{X})$ and any $c \in \mathbb{R}$, let $f^{\wedge c}$ and $f^{\vee c}$ be the variables in $\overline{\mathscr{L}}(\mathscr{X})$ respectively defined by $f^{\wedge c}(x) := \min\{f(x), c\}$ and $\underline{f}^{\vee c}(x) := \max\{f(x), c\}$ for all $x \in \mathscr{X}$. We propose the following two continuity properties for an upper expectation $\overline{\mathrm{Q}}_s^{\uparrow}$ on $\overline{\mathscr{L}}(\mathscr{X})$ that extends a coherent upper expectation $\overline{\mathrm{Q}}_s$ on $\mathscr{L}(\mathscr{X})$:

C8. $\overline{\mathrm{Q}}_s^{\uparrow}(f) = \lim_{c \to +\infty} \overline{\mathrm{Q}}_s^{\uparrow}(f^{\wedge c})$ for all $f \in \overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$;

C9. $\overline{\mathrm{Q}}_s^{\uparrow}(f) = \lim_{c \to -\infty} \overline{\mathrm{Q}}_s^{\uparrow}(f^{\vee c})$ for all $f \in \overline{\mathscr{L}}(\mathscr{X})$;

Properties C8 and C9 are respectively called *continuity with respect to upper* and *lower cuts* (or, alternatively, bounded above and below support). According to [20, Proposition 10 (ii)–(iii)] such an extension $\overline{\mathrm{Q}}_s^{\uparrow}$ always exists and is moreover unique. The upper expectation $\overline{\mathrm{Q}}_s^{\uparrow}$ is also an appropriate extension, in the sense that it satisfies E1–E5; this follows from [20, Proposition 10] together with [20, Proposition 1].[8] Hence, in order to adhere to Shafer and Vovk's approach, we can—and will—adopt this particular extension $\overline{\mathrm{Q}}_s^{\uparrow}$ of $\overline{\mathrm{Q}}_s$ for each situation $s \in \mathscr{X}^*$ in the following game-theoretic reasoning.

A map $\mathscr{M} : \mathscr{X}^* \to \overline{\mathbb{R}}$ is called a *supermartingale* if it satisfies $\overline{\mathrm{Q}}_s^{\uparrow}(\mathscr{M}(s\cdot)) \leq \mathscr{M}(s)$ for all $s \in \mathscr{X}^*$, where $\mathscr{M}(s\cdot)$ denotes the variable in $\overline{\mathscr{L}}(\mathscr{X})$ that assumes the value $\mathscr{M}(sx)$ for all $x \in \mathscr{X}$. As mentioned earlier, a supermartingale $\mathscr{M}$ can be interpreted as a capital process that results from betting against a forecasting system. In order for this to make sense, such an interpretation will only be attached to supermartingales that are *bounded below*, meaning that there is some real c such that $\mathscr{M}(s) \geq c$ for all $s \in \mathscr{X}^*$. This represents the constraint that we can never borrow an infinite or even unbounded amount of money. To see that a bounded below supermartingale $\mathscr{M}$ can indeed be interpreted as described above, we first need to know how a forecasting system allows us to play. In our case, the allowable bets in any given situation $s \in \mathscr{X}^*$ are those variables $f \in \overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$ for which $\overline{\mathrm{Q}}_s^{\uparrow}(f) \leq 0$—we will come back to this shortly. Now, if $\mathscr{M}(s)$ is real, then the condition that $\overline{\mathrm{Q}}_s^{\uparrow}(\mathscr{M}(s\cdot)) \leq \mathscr{M}(s)$ is equivalent to the condition that $\overline{\mathrm{Q}}_s^{\uparrow}(\mathscr{M}(s\cdot) - \mathscr{M}(s)) \leq 0$ because $\overline{\mathrm{Q}}_s^{\uparrow}$ satisfies constant additivity on the domain $\overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$; see [20, Proposition 3(E8)]. Hence, the incremental change $\mathscr{M}(s\cdot) - \mathscr{M}(s)$ of the capital process $\mathscr{M}$ in the situation s indeed represents an allowable bet. If on the other hand $\mathscr{M}(s) = +\infty$ (the case where $\mathscr{M}(s) = -\infty$ is impossible because we assumed $\mathscr{M}$ to be bounded below), the condition that $\overline{\mathrm{Q}}_s^{\uparrow}(\mathscr{M}(s\cdot)) \leq \mathscr{M}(s)$ does not impose any constraints on $\mathscr{M}(s\cdot)$. The interpretation that the increment $\mathscr{M}(s\cdot) - \mathscr{M}(s)$ represents an allowable bet is, in this case, somewhat questionable. However, one could argue that, since the process $\mathscr{M}$ has reached the maximum possible value—that is, $+\infty$—in the situation s, it cannot increase any further, hence there is no need for any constraints on the process' future value.

That the allowable bets f in some situation $s \in \mathscr{X}^*$ are characterised by the condition that $\overline{\mathrm{Q}}_s^{\uparrow}(f) \leq 0$, agrees with the behavioural interpretation of (boundedly) coherent upper expectations. To see this, we limit ourselves to the domain $\mathscr{L}(\mathscr{X})$ and recall that $\overline{\mathrm{Q}}_s^{\uparrow}(f) = \overline{\mathrm{Q}}_s(f)$, for any $f \in \mathscr{L}(\mathscr{X})$, can then be interpreted as a subject's infimum selling price for the gamble f. So the subject, being the forecasting system in this case, is willing to offer the gambles that have negative[9] upper expectation. We can then take the subject up on his commitments by accepting its offer and selecting one such gamble. Moreover, since the upper expectations $\overline{\mathrm{Q}}_s$ are assumed to be (boundedly) coherent, the associated sets of available gambles will also satisfy certain rationality axioms. We

---

[8]To see this, note that the definition of an upper expectation in Reference [20] is different from the one we adopt here.

[9]In the interest of brevity, we ignore in our discussion of the interpretation the special case where we have zero upper expectation. See [12] for more details about this special case.

refer to an earlier paper by one of us [12] for a more elaborate discussion of how Walley's behavioural approach can be related (in a slightly different context) to the game-theoretic one proposed by Shafer and Vovk.

Our definition of the global game-theoretic upper expectation—and the one considered by Shafer and Vovk—will be based on bounded below supermartingales, not only because it allows for a sensible interpretation as we have explained above, but also because of more technical reasons, which we discuss in [20, Section 8]. So let $\overline{\mathbb{M}}_b$ be the set of all such bounded below supermartingales. For any $\mathcal{M} \in \overline{\mathbb{M}}_b$, we then use $\liminf \mathcal{M}$ to denote the extended real global variable that assumes the value $\liminf_{n \to +\infty} \mathcal{M}(\omega^n)$ in each $\omega \in \Omega$. We also write, for any situation $s \in \mathcal{X}^*$ and any two f and g in $\overline{\mathbb{V}}$, that $f \geq_s g$ if $f(\omega) \geq g(\omega)$ for all paths $\omega \in \Gamma(s)$. The global game-theoretic upper expectation $\overline{\mathrm{E}}_V \colon \overline{\mathbb{V}} \times \mathcal{X}^* \to \overline{\mathbb{R}}$ can now be defined as

$$\overline{\mathrm{E}}_V(f|s) := \inf\left\{ \mathcal{M}(s) \colon \mathcal{M} \in \overline{\mathbb{M}}_b \text{ and } \liminf \mathcal{M} \geq_s f \right\} \text{ for all } f \in \overline{\mathbb{V}} \text{ and all } s \in \mathcal{X}^*.$$

The corresponding game-theoretic lower expectation $\underline{\mathrm{E}}_V$ is again defined by conjugacy; so $\underline{\mathrm{E}}_V(f|s) := -\overline{\mathrm{E}}_V(-f|s)$ for all $f \in \overline{\mathbb{V}}$ and all $s \in \mathcal{X}^*$. For a concrete and detailed motivation and interpretation for the operator $\overline{\mathrm{E}}_V$, we refer to [4, 5, 20]. Essentially, the upper expectation $\overline{\mathrm{E}}_V(f|x_{1:n})$ can once more be seen as an infimum selling price for a variable $f \in \overline{\mathbb{V}}$ conditional on the fact that we have observed some history $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$ of the process. It is based on the assumption that we should agree to sell f for any starting capital $\mathcal{M}(x_{1:n})$ such that, by appropriately betting against the forecasting system from $x_{1:n}$ onwards, we manage to end up with a higher capital than what we would receive from f, regardless of the path $\omega \in \Omega$ taken by the process. Indeed, it is reasonable to state that such a starting capital $\mathcal{M}(x_{1:n})$ is worth more than the uncertain (possibly negative) payoff corresponding to f. The global game-theoretic upper expectation $\overline{\mathrm{E}}_V(f|x_{1:n})$ is the infimum over all such starting capitals.

As we will briefly discuss in Section 7, the global game-theoretic upper expectation $\overline{\mathrm{E}}_V$ satisfies various, remarkably powerful properties. In particular, it satisfies P1–P5, P3$^=$, P6 and P7, which allows us to state our main result: that $\overline{\mathrm{E}}_V$ is the unique most conservative upper expectation $\overline{\mathrm{E}}_A$ in $\overline{\mathbb{E}}_{1-5}(\overline{Q})$. Our proofs for these results are heavily based on our work in [20]. The setting there is entirely the same as the one described in this section, with the exception that we generally do not impose C9 on the local models in [20]. It will soon be clarified why we have made this distinction; for the moment, it suffices to see that the setting in [20] is more general and therefore, that the results in [20] obviously remain valid here.

Proposition 5. $\overline{\mathrm{E}}_V$ *is an element of* $\overline{\mathbb{E}}_{1-5}(\overline{Q})$ *and furthermore satisfies* P3$^=$, P6 *and* P7.

Proof. We prove that $\overline{\mathrm{E}}_V$ satisfies P1–P5, P3$^=$, P6 and P7. Axiom P1 follows immediately from [20, Proposition 14] which, in particular, says that $\overline{\mathrm{E}}(f|x_{1:n}) = \overline{Q}_{x_{1:n}}(f(x_{1:n}\cdot))$ for any situation $x_{1:n} \in \mathcal{X}^*$ and any $(n+1)$-measurable gamble f. Indeed, for any $h \in \mathcal{L}(\mathcal{X})$, $h(X_{n+1})$ is an $(n+1)$-measurable gamble, so it follows from [20, Proposition 14] that $\overline{\mathrm{E}}_V(h(X_{n+1})|x_{1:n}) = \overline{Q}_{x_{1:n}}(h)$. To prove P2, observe that $\liminf \mathcal{M} \geq_s f$ if and only if $\liminf \mathcal{M} \geq_s f \mathbb{I}_s$ for all $f \in \overline{\mathbb{V}}$ and all $\mathcal{M} \in \overline{\mathbb{M}}_b$. The desired equality therefore follows directly from the definition of $\overline{\mathrm{E}}_V$. P3$^=$, and hence also P3, follows immediately from [20, Theorem 15]. Property P4 follows from [20, Proposition 13(V4)], which says that $\overline{\mathrm{E}}_V$ satisfies the stronger property that $\overline{\mathrm{E}}(f|s) \leq \overline{\mathrm{E}}(g|s)$ if $f \leq_s g$ for all $f, g \in \overline{\mathbb{V}}$ and all $s \in \mathcal{X}^*$.

Furthermore, P6 and P7 follow from, respectively, [20, Theorem 32] and [20, Proposition 28]. Note that, in these results, the definition of $\overline{\mathbb{V}}_{b,\lim}$ seems to be slightly more general, since it there also includes pointwise limits of possibly *extended* real finitary variables. However, according to [20, Proposition 30], any such limit of extended real finitary variables is in fact also a pointwise limit of n-measurable, and therefore finitary, gambles. So both our definitions of $\overline{\mathbb{V}}_{b,\lim}$ are equivalent, therefore indeed allowing us to apply the results in [20] here.

Finally, Property P5 follows from [20, Corollary 25], which says that, for any sequence $\{f_n\}_{n \in \mathbb{N}_0}$ in $\overline{\mathbb{V}}_b$ that is uniformly bounded below, $\overline{\mathrm{E}}_V(\liminf_{n \to +\infty} f_n|s) \leq \liminf_{n \to +\infty} \overline{\mathrm{E}}_V(f_n|s)$. Indeed, in the special case that $\{f_n\}_{n \in \mathbb{N}_0}$ is a sequence of finitary gambles that is uniformly bounded below and converges pointwise to some variable $f \in \overline{\mathbb{V}}_b$, this implies that $\overline{\mathrm{E}}_V(f|s) \leq \liminf_{n \to +\infty} \overline{\mathrm{E}}_V(f_n|s) \leq \limsup_{n \to +\infty} \overline{\mathrm{E}}_V(f_n|s)$. $\qquad\square$

Theorem 6. *There is a unique most conservative upper expectation* $\overline{\mathrm{E}}_A$ *in* $\overline{\mathbb{E}}_{1-5}(\overline{Q})$ *and it is the global game-theoretic upper expectation* $\overline{\mathrm{E}}_V$. *Furthermore, if an upper expectation* $\overline{\mathrm{E}}$ *in* $\overline{\mathbb{E}}_{1-5}(\overline{Q})$ *satisfies* P3$^=$, P6 *and* P7, *then it is equal to this most conservative upper expectation* $\overline{\mathrm{E}}_A$, *and therefore also equal to* $\overline{\mathrm{E}}_V$.

Proof. Immediately from Theorem 4 and Proposition 5. $\qquad\square$

7. On the Implications of Theorem 6

Shafer and Vovk define their game-theoretic upper expectations using supermartingales. Their definition has a constructive flavour and can be given a clear interpretation in terms of capital processes and betting behaviour. However, it requires that one allows unbounded and even infinite-valued bets, which we find more questionable from a behavioural point of view. Furthermore, a complete axiomatisation of $\overline{E}_V$ is, to the best of our knowledge, absent from the literature. Theorem 6 addresses both of these issues. First, it provides an abstract axiomatisation for $\overline{E}_V$ using P3=, P6 and P7 in addition to P1–P5. Most importantly, however, Theorem 6 provides an alternative definition— and interpretation—for $\overline{E}_V$ as the most conservative upper expectation $\overline{E}_A$ under a limited set of intuitive properties: P1–P5. This strengthens the argument in favour of using $\overline{E}_V$ as a global upper expectation, because it can now be motivated from both a game-theoretic point of view and from a purely axiomatic point of view. Readers need not even be familiar with the concepts of game-theoretic probability in order to use $\overline{E}_V$ as global upper expectation. Put simply, they only have to agree on the axioms P1–P5. And even if they wish to impose additional axioms, then $\overline{E}_V$ would still serve as a conservative upper bound for their desired (upper) expectation.

Still, recalling our assumptions about the extended local models $\overline{Q}_s^\uparrow$, one could regret the fact that we are only considering a special case of what is generally being done by Shafer and Vovk; an upper expectation $\overline{Q}_s^\uparrow$ on $\mathscr{L}(\mathscr{X})$ satisfying E1–E5 does not necessarily have to be an extension of a coherent upper expectation $\overline{Q}_s$ on $\mathscr{L}(\mathscr{X})$ satisfying C8 and C9, see [20, Example 1]. However, this restriction in generality actually turns out to be rather desirable. Indeed, as we show in [20], compatibility of local and global game-theoretic upper expectations—a property that we find essential—can only be obtained if we limit ourselves to extended local models $\overline{Q}_s^\uparrow$ of the form that we have described earlier, that is, boundedly coherent and satisfying C8 and C9.

That said, as mentioned before, we do not a priori impose C9 on the local models in [20]. In that paper, we mainly focus on the mathematical properties of global game-theoretic upper expectations and most of these do not require C9. Furthermore, by dropping this axiom, we can allow for more general local models than Shafer and Vovk, who impose E1–E5. Nevertheless, in this less mathematical and more philosophical contribution, we do impose C9 on the local models because, practically speaking, we always want local and global belief models to be compatible; see for example Section 4, where we considered it self-evident that P1 should be desirable. Note however that Theorem 6 actually does not depend on whether the extended local models $\overline{Q}_s^\uparrow$ satisfy C9. On the one hand, because $\overline{E}_A$ only depends on the (boundedly) coherent local models $\overline{Q}_s$, and on the other hand, because $\overline{E}_V$ only depends on what the values of the extended local models $\overline{Q}_s^\uparrow$ are on the restricted domain $\overline{\mathscr{L}}_b(\mathscr{X})$. The latter can easily be seen from the fact that $\overline{E}_V$ is defined using supermartingales that are bounded below; we refer to [20] for more details.

8. Additional Properties

Of course, there is more to an uncertainty model than an axiomatisation or an interpretation, however compelling they may be. In order to be practically useful, its mathematical properties should also be sufficiently powerful such that inferences can be computed or approximated efficiently. For instance, the popularity of the Lebesgue integral as a tool for defining expected values, is in part due to its strong continuity properties (e.g. the Dominated Convergence Theorem [1, 16]). What may perhaps be somewhat surprising, is that despite the simplicity of our axioms P1–P5, our most conservative model $\overline{E}_A$ also scores well on this count. For example, it satisfies the following generalisation of bounded coherence. Its proof, as well as the proofs of the other results in this section, follows immediately from the fact that $\overline{E}_A$ coincides with $\overline{E}_V$ and the fact that these properties hold for $\overline{E}_V$, as is shown in [20].

Proposition 7. [20, Proposition 13 and Theorem 15] *For all $f, g \in \overline{\mathbb{V}}$, all $\lambda \in \mathbb{R}_{\geq 0}$, all $\mu \in \mathbb{R}$, all $n \in \mathbb{N}_0$ and all $s \in \mathscr{X}^*$, $\overline{E}_A$ satisfies*

V1. $\inf_{\omega \in \Gamma(s)} f(\omega) \leq \underline{E}_A(f|s) \leq \overline{E}_A(f|s) \leq \sup_{\omega \in \Gamma(s)} f(\omega)$;

V2. $\overline{E}_A(f + g|s) \leq \overline{E}_A(f|s) + \overline{E}_A(g|s)$;

13

V3. $\overline{\mathrm{E}}_{\mathrm{A}}(\lambda f | s) = \lambda \overline{\mathrm{E}}_{\mathrm{A}}(f | s)$;

V4. $\overline{\mathrm{E}}_{\mathrm{A}}(f + \mu | s) = \overline{\mathrm{E}}_{\mathrm{A}}(f | s) + \mu$.

V5. $\overline{\mathrm{E}}_{\mathrm{A}}(f | X_{1:n}) = \overline{\mathrm{E}}_{\mathrm{A}}\big(\overline{\mathrm{E}}_{\mathrm{A}}(f | X_{1:n+1}) \big| X_{1:n}\big)$.

Another important result, better known as Fatou's Lemma, shows that the upper bound imposed by property P5 for a global upper expectation $\overline{\mathrm{E}}(f | s)$ of a variable $f \in \overline{\mathbb{V}}_{\mathrm{b,lim}}$ need not be attained by the most conservative upper expectation $\overline{\mathrm{E}}_{\mathrm{A}}$ in $\overline{\mathbb{E}}_{1-5}(\overline{\mathrm{Q}})$. Interestingly enough, it can be replaced by a generally tighter one, which is established in the following lemma.

Lemma 8. [20, Corollary 25] *For any $s \in \mathscr{X}^*$ and any sequence $\{f_n\}_{n \in \mathbb{N}_0}$ in $\overline{\mathbb{V}}_{\mathrm{b}}$ that is uniformly bounded below, $\overline{\mathrm{E}}_{\mathrm{A}}(\liminf_{n \to +\infty} f_n | s) \leq \liminf_{n \to +\infty} \overline{\mathrm{E}}_{\mathrm{A}}(f_n | s)$.*

The following theorem states that $\overline{\mathrm{E}}_{\mathrm{A}}$ satisfies continuity with respect to non-decreasing sequences of bounded below variables and with respect to non-increasing sequences of finitary, bounded above variables.

Theorem 9. [20, Theorem 23 and Theorem 31]

  i. *For any $s \in \mathscr{X}^*$ and any non-decreasing sequence $\{f_n\}_{n \in \mathbb{N}_0}$ in $\overline{\mathbb{V}}_{\mathrm{b}}$ that converges point-wise to a variable $f \in \overline{\mathbb{V}}_{\mathrm{b}}$, we have that $\overline{\mathrm{E}}_{\mathrm{A}}(f | s) = \lim_{n \to +\infty} \overline{\mathrm{E}}_{\mathrm{A}}(f_n | s)$.*

 ii. *For any $s \in \mathscr{X}^*$ and any non-increasing sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of finitary, bounded above variables that converges point-wise to a variable $f \in \overline{\mathbb{V}}$, we have that $\overline{\mathrm{E}}_{\mathrm{A}}(f | s) = \lim_{n \to +\infty} \overline{\mathrm{E}}_{\mathrm{A}}(f_n | s)$.*

Apart from the brief overview above, $\overline{\mathrm{E}}_{\mathrm{V}} = \overline{\mathrm{E}}_{\mathrm{A}}$ also satisfies a weak and a strong law of large numbers, a law of the iterated logarithm, Lévy's zero-one law, and many more surprisingly strong properties; we refer the interested reader to Shafer and Vovk's work [5, 17] and to ours [20, 21].

9. Measure-Theoretic Belief Models

In contradistinction with the game-theoretic approach, where the central concept is that of a supermartingale, the measure-theoretic approach relies on σ-additive probability measures in order to describe the global dynamics of an uncertain process. In a precise probabilities context, this global measure is usually obtained by extending the information that is included in the (precise) local models, now represented by probability mass functions on $\mathscr{X}$, using Ionescu Tulcea's extension theorem [1, Theorem 2.9.2]. This consists of two consecutive steps. First, a probability measure on the algebra generated by all cylinder events is constructed in a straightforward manner, by combining local (conditional) probabilities using the well-known product rule. Subsequently, this probability measure is extended to a probability measure on the σ-algebra $\mathscr{F}$ generated by all cylinder events, which, according to Carathéodory's extension theorem [6, Theorem 1.7], can be done uniquely. Linear expectations are then typically obtained by Lebesgue integration and by restricting the domain to (random) variables that are measurable with respect to the σ-algebra $\mathscr{F}$.

This traditional measure-theoretic approach has a few shortcomings however. First of all, it only considers the special case where each local model is represented by a single probability mass function on $\mathscr{X}$. Hence, we cannot directly apply it here in our more general, imprecise probabilities setting. Secondly, the domain of the expectation operators is restricted to measurable (random) variables and therefore, no explicit information is given about non-measurable variables. Thirdly, by condensing all local probability mass functions into a single global probability measure, we disregard part of the information that is incorporated in the local probability mass functions. More specifically, the part that describes the process' dynamics conditional on the fact that its path will lie in a set of probability zero (with respect to the global probability measure). Indeed, conditional probabilities and expectations are typically derived from the global probability measure using Bayes' rule, yet, for those cases where the conditioning event has probability zero, the conditional probabilities and expectations become ill-defined or are chosen arbitrarily; see Appendix A and [1, 3].

In this section, we aim to adapt the traditional measure-theoretic approach to our setting and deal with these problems in a satisfactory manner. As a first step, we will limit ourselves to the precise case and try to resolve

14

the last two issues mentioned above. The global probability measure on $\mathscr{F}$, which is the starting point for the traditional approach, will be replaced by a 'conditional probability measure' in the sense of [13, Definition 6]. Essentially, such a conditional measure specifies a probability measure on $\mathscr{F}$ for each situation $s \in \mathscr{X}^*$, which describes the global dynamics of the uncertain process if we are sure that the path ω taken by the process will pass through this situation s. This will enable us to also meaningfully define linear expectations conditional on situations that have probability zero. Subsequently, we will extend the domain of these conditional linear expectation operators to all global variables $\overline{\mathbb{V}}$ by taking a particular upper integral leading to an upper expectation. As a final step, we allow for imprecision in the local models by considering a set $\mathbb{P}_s$ of probability mass functions on $\mathscr{X}$ in each situation s. We can then apply the foregoing course of reasoning for each possible selection $p \colon s \in \mathscr{X}^* \mapsto p(\cdot|s)$ of local probability mass functions, where, in each situation s, the selected probability mass function $p(\cdot|s)$ belongs to $\mathbb{P}_s$. Hence, instead of a single upper expectation operator, we now obtain a set of them. The global measure-theoretic upper expectation $\overline{\mathrm{E}}_{\mathrm{meas}}$ will then be defined as the upper envelope of all such 'compatible' upper expectation operators, similar to what is done on a local level.

The resulting operator $\overline{\mathrm{E}}_{\mathrm{meas}}$ will satisfy P1–P5, which implies by Theorem 6 that the upper expectation $\overline{\mathrm{E}}_{\mathrm{A}}$—and therefore also $\overline{\mathrm{E}}_{\mathrm{V}}$—is an upper bound for $\overline{\mathrm{E}}_{\mathrm{meas}}$. Moreover, if all local models are precise, all upper expectation operators $\overline{\mathrm{E}}_{\mathrm{A}}$, $\overline{\mathrm{E}}_{\mathrm{V}}$ and $\overline{\mathrm{E}}_{\mathrm{meas}}$ will be shown to coincide.

Precise Probability Trees

In this subsection, we consider a *precise probability tree p*, that is, a map $p \colon s \in \mathscr{X}^* \mapsto p(\cdot|s)$ where each $p(\cdot|s)$ is a probability mass function on $\mathscr{X}$. Recalling our considerations in Section 2, each mass function $p(\cdot|s)$ is in a one-to-one correspondence with a linear expectation $\overline{\mathrm{Q}}_s$ on $\mathscr{L}(\mathscr{X})$, so we could just as well have started from an imprecise probability tree $\overline{\mathrm{Q}}$ where all local models $\overline{\mathrm{Q}}_s$ are linear on $\mathscr{L}(\mathscr{X})$. However, we prefer to start with p because it is in accordance with the traditional measure-theoretic view towards uncertainty, where probability measures or charges—and therefore also probability mass functions—are considered to be primary objects.

As mentioned above, we aim to construct a conditional probability measure that allows us to condition on any situation $s \in \mathscr{X}^*$ meaningfully. To do so, we follow the same reasoning as proposed in [13, Chapter 3]. First, we define, for any $x_{1:n} \in \mathscr{X}^*$, a probability charge $\mathrm{P}(\cdot|x_{1:n})$ on the algebra generated by all cylinder events. This can be done in a simple and intuitive manner [13, Lemma 14]: for any $m \in \mathbb{N}_0$ and any $C \subseteq \mathscr{X}^m$, we let

$$\mathrm{P}(C|x_{1:n}) := \sum_{z_{1:m} \in C} \mathrm{P}(z_{1:m}|x_{1:n}), \text{ where } \mathrm{P}(z_{1:m}|x_{1:n}) := \begin{cases} \prod_{i=n}^{m-1} p(z_{i+1}|z_{1:i}) & \text{if } n < m \text{ and } z_{1:n} = x_{1:n} \\ 1 & \text{if } n \geq m \text{ and } z_{1:m} = x_{1:m} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where we use the shorthand notation $\mathrm{P}(C|x_{1:n})$ for any $C \subseteq \mathscr{X}^m$ to mean $\mathrm{P}(\cup_{z_{1:m} \in C} \Gamma(z_{1:m})|x_{1:n})$, and where we did not make any particular distinction between the notations $z_{1:m}$ and $\{z_{1:m}\}$. Since each $\mathrm{P}(\cdot|x_{1:n})$ defined in this way is a (finitely additive) probability charge on the algebra generated by all cylinder events [13, Chapter 3], and therefore automatically a σ-addive probability measure on this algebra [3, Theorem 2.3], we can use Carathéodory's extension theorem [6, Theorem 1.7] to extend each $\mathrm{P}(\cdot|x_{1:n})$ to a unique σ-additive measure on the σ-algebra $\mathscr{F}$ generated by all cylinder events. Then, according to [13, Theorem 15], we can gather all such probability measures $\mathrm{P}(\cdot|x_{1:n})$ to obtain a single *conditional probability measure* $\mathrm{P} \colon \mathscr{F} \times \mathscr{X}^* \to \mathbb{R}$ [13, Definition 6], which is *coherent* in the sense of [13, Definition 5]. This notion of coherence is related to the bounded coherence condition in Section 2; both notions represent rational behaviour, but, the version used in [13, Definition 5] is adapted to probability measures and involves conditioning on events.

Before defining our global measure-theoretic (upper) expectation, we first recall the following concepts about measurability and integration; we refer to Appendix A for a more elaborate introduction. We say that a global variable $f \in \overline{\mathbb{V}}$ is $\mathscr{F}$-measurable if, for all $c \in \mathbb{R}$, the set $\{\omega \in \Omega \colon f(\omega) \leq c\}$ is $\mathscr{F}$-measurable, meaning that it is an element of the σ-algebra $\mathscr{F}$. We will commonly use the properties that $\mathscr{F}$-measurability of variables is preserved under taking maxima or minima (this follows immediately from the fact that the class of all $\mathscr{F}$-measurable sets is closed under countable unions and intersections [22, Section II.4]) and that $\mathscr{F}$-measurability is preserved under pointwise convergence of variables [22, Theorem II.4.2]. Moreover, recall that the Lebesgue integral $\int_\Omega f \, \mathrm{d}\mathrm{P}_{\mathrm{u}}$ [1, 3, 16] with respect to any (unconditional) probability measure P_{u} on $\mathscr{F}$ always exists if f

is $\mathscr{F}$-measurable and either bounded or non-negative. However, for a general $\mathscr{F}$-measurable variable $f \in \overline{\mathbb{V}}$, the integral only exists if $\min\{\int_\Omega f^+ dP_u, \int_\Omega f^- dP_u\} < +\infty$, where $f^+ := f^{\vee 0}$ and $f^- := -f^{\wedge 0}$. In that case, we let $\int_\Omega f dP_u := \int_\Omega f^+ dP_u - \int_\Omega f^- dP_u$. Then, because the Lebesgue integral is trivially real-valued for a bounded $\mathscr{F}$-measurable variable (this can for instance be deduced from M1 and M2 in Appendix A), it can easily be seen that $\int_\Omega f dP_u$ also always exists for an $\mathscr{F}$-measurable variable $f \in \overline{\mathbb{V}}_b$.

We now adopt an approach that is similar to what is done in a traditional context, defining expectations as Lebesgue integrals. An important difference, however, is that we integrate with respect to a probability measure that depends on the conditioning event. Concretely, we define the *measure-theoretic expectation* $E_{\mathrm{meas},p}(f|s)$ of an $\mathscr{F}$-measurable variable $f \in \overline{\mathbb{V}}$ conditional on the situation $s \in \mathscr{X}^*$ by the Lebesgue integral $\int_\Omega f dP_{|s}$ with respect to the (unconditional) probability measure $P_{|s} := P(\cdot|s)$, provided that $\int_\Omega f dP_{|s}$ exists. The expectation $E_{\mathrm{meas},p}(\cdot|\square)$ then corresponds to the unconditional expectation as introduced in traditional measure theory. The subscript p in $E_{\mathrm{meas},p}$ is used to remind us that $E_{\mathrm{meas},p}$ depends on the precise probability tree p. Since $E_{\mathrm{meas},p}$ is defined in terms of the Lebesgue integral, it inherits several properties of this integral; again, we refer to Appendix A for more information. The following is a subset of these properties—simplified and applied to $E_{\mathrm{meas},p}$—that we will need in the main text. We take $s \in \mathscr{X}^*$ to be any situation and implicitly assume that the considered expectations exist.

M1. $E_{\mathrm{meas},p}(af + b|s) = aE_{\mathrm{meas},p}(f|s) + b$ for all $\mathscr{F}$-measurable variables $f \in \overline{\mathbb{V}}$ and all $a, b \in \mathbb{R}$.

M2. $f \le g \Rightarrow E_{\mathrm{meas},p}(f|s) \le E_{\mathrm{meas},p}(g|s)$ for all $\mathscr{F}$-measurable variables $f, g \in \overline{\mathbb{V}}$.

M8. Consider any non-decreasing sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of $\mathscr{F}$-measurable variables in $\overline{\mathbb{V}}$. If there is an $\mathscr{F}$-measurable function f^* such that $E_{\mathrm{meas},p}(f^*|s) > -\infty$ and $f_n \ge f^*$ for all $n \in \mathbb{N}_0$, then

$$\lim_{n \to +\infty} E_{\mathrm{meas},p}(f_n|s) = E_{\mathrm{meas},p}(f|s) \text{ where } \lim_{n \to +\infty} f_n = f.$$

M9. Consider any non-increasing sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of $\mathscr{F}$-measurable variables in $\overline{\mathbb{V}}$. If there is an $\mathscr{F}$-measurable function f^* such that $E_{\mathrm{meas},p}(f^*|s) < +\infty$ and $f_n \le f^*$ for all $n \in \mathbb{N}_0$, then

$$\lim_{n \to +\infty} E_{\mathrm{meas},p}(f_n|s) = E_{\mathrm{meas},p}(f|s) \text{ where } \lim_{n \to +\infty} f_n = f.$$

Now that our measure-theoretic definitions are in place, we aim at relating the expectation $E_{\mathrm{meas},p}$ to our upper expectation operator $\overline{E}_A$ discussed in the earlier sections of this paper. For the sake of clarity, we will use $\overline{E}_{A,p}$ to denote our upper expectation $\overline{E}_A$ corresponding to the imprecise probability tree $\overline{Q}$ associated with p; so each $\overline{Q}_s$ is defined by $\overline{Q}_s(f) := \sum_{x \in \mathscr{X}} f(x)p(x|s)$ for all $f \in \mathscr{L}(\mathscr{X})$. We first focus on the domain of all $\mathscr{F}$-measurable variables in $\overline{\mathbb{V}}_b$. This seems appropriate because, though we cannot guarantee existence of the integral $\int_\Omega f dP_{|s}$—and therefore also $E_{\mathrm{meas},p}(f|s)$—for a general $\mathscr{F}$-measurable variable $f \in \overline{\mathbb{V}}$, it can always be guaranteed for an $\mathscr{F}$-measurable variable $f \in \overline{\mathbb{V}}_b$, as we have explained earlier. It will turn out that both operators $E_{\mathrm{meas},p}$ and $\overline{E}_{A,p}$ coincide on the domain of all $\mathscr{F}$-measurable variables in $\overline{\mathbb{V}}_b$.

To prove this, we base ourselves on [5, Theorem 9.3]; a result by Shafer and Vovk—yet the essential idea originates from Jean Ville—that shows that game-theoretic and measure-theoretic (upper) expectations coincide on $\mathscr{F}$-measurable gambles, in case we are considering precise probability trees. Proposition 10 below establishes the equality of $\overline{E}_{A,p}$ and $E_{\mathrm{meas},p}$ on this domain and therefore, due to Theorem 6, also of $\overline{E}_V$ and $E_{\mathrm{meas},p}$. Though the proof of Proposition 10 is based on the same principles and ideas as those that underly the proof of [5, Theorem 9.3], they are stated in a somewhat different setting and, for an unexperienced reader, it may therefore not be entirely clear how both are related to each other. Moreover, our result here also applies to conditional expectations, whereas [5, Theorem 9.3] only considers the unconditional case. We have therefore decided to include a self-contained proof of Proposition 10 in Appendix B, starting from Ville's martingale theorem.

Proposition 10. *For any $\mathscr{F}$-measurable $f \in \mathbb{V}$ and any $s \in \mathscr{X}^*$, we have that $E_{\mathrm{meas},p}(f|s) = \overline{E}_{A,p}(f|s)$.*

The next result extends the domain of the equality to all $\mathscr{F}$-measurable variables in $\overline{\mathbb{V}}_b$.

Theorem 11. *For any $\mathscr{F}$-measurable $f \in \overline{\mathbb{V}}_b$ and any $s \in \mathscr{X}^*$, we have that $E_{\mathrm{meas},p}(f|s) = \overline{E}_{A,p}(f|s)$.*

Proof. Consider any $f \in \overline{\mathbb{V}}_b$ that is $\mathscr{F}$-measurable and any $s \in \mathscr{X}^*$. Because f is bounded below and $\mathscr{F}$-measurable, $\mathrm{E}_{\mathrm{meas},p}(f|s)$ exists. We can moreover assume that f is non-negative without loss of generality because it is bounded below and both $\mathrm{E}_{\mathrm{meas},p}$ and $\overline{\mathrm{E}}_{\mathrm{A},p}$ are constant additive with respect to real constants; see M1 and V4. Consider now the non-decreasing sequence $\{f^{\wedge n}\}_{n \in \mathbb{N}_0}$ of upper cuts and note that each $f^{\wedge n}$ is bounded and $\mathscr{F}$-measurable; this follows from the fact that $\mathscr{F}$-measurability is preserved under taking minima. Using M8—which we are allowed to use because $\mathrm{E}_{\mathrm{meas},p}(f^{\wedge 0}|s)$ is real (since $f^{\wedge 0}$ is bounded) and because $\{f^{\wedge n}\}_{n \in \mathbb{N}_0}$ is non-decreasing—we have that $\mathrm{E}_{\mathrm{meas},p}(f|s) = \lim_{n \to +\infty} \mathrm{E}_{\mathrm{meas},p}(f^{\wedge n}|s)$. As a consequence,

$$\mathrm{E}_{\mathrm{meas},p}(f|s) = \lim_{n \to +\infty} \mathrm{E}_{\mathrm{meas},p}(f^{\wedge n}|s) = \lim_{n \to +\infty} \overline{\mathrm{E}}_{\mathrm{A},p}(f^{\wedge n}|s) = \overline{\mathrm{E}}_{\mathrm{A},p}(f|s),$$

where the second equality follows from Proposition 10 and the third one from Theorem 9(i). $\qquad\square$

Next, we want to drop the condition of $\mathscr{F}$-measurability and generalise towards the domain $\overline{\mathbb{V}}$ of all global extended real variables. So we are looking for an upper expectation operator that appropriately extends $\mathrm{E}_{\mathrm{meas},p}$ to the entire domain $\overline{\mathbb{V}} \times \mathscr{X}^*$. Inspired by Theorem 11, we simply suggest the following upper integral:

$$\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) := \inf\left\{\mathrm{E}_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}_b, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\right\} \text{ for all } f \in \overline{\mathbb{V}} \text{ and } s \in \mathscr{X}^*.$$

That it is valid to write $\mathrm{E}_{\mathrm{meas},p}(g|s)$ for any $g \in \overline{\mathbb{V}}_b$ that is $\mathscr{F}$-measurable, follows from the fact that $\mathrm{E}_{\mathrm{meas},p}(g|s) = \int_\Omega g \, \mathrm{dP}_{|s}$ is guaranteed to exist for such a variable. Let us now show that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is indeed an extension of $\mathrm{E}_{\mathrm{meas},p}$.

Proposition 12. *For any $f \in \overline{\mathbb{V}}$ and any $s \in \mathscr{X}^*$, we have that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \mathrm{E}_{\mathrm{meas},p}(f|s)$ if $\mathrm{E}_{\mathrm{meas},p}(f|s)$ exists.*

Proof. Suppose that $\mathrm{E}_{\mathrm{meas},p}(f|s)$ exists, which means that the variable f is $\mathscr{F}$-measurable and that $\mathrm{E}_{\mathrm{meas},p}(f^+|s)$ or $\mathrm{E}_{\mathrm{meas},p}(f^-|s)$ is real-valued. By the monotonicity [M2] of $\mathrm{E}_{\mathrm{meas},p}$, we immediately have that

$$\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \inf\left\{\mathrm{E}_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}_b, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\right\} \geq \mathrm{E}_{\mathrm{meas},p}(f|s). \tag{7}$$

If $\mathrm{E}_{\mathrm{meas},p}(f^+|s) = +\infty$ and therefore $\mathrm{E}_{\mathrm{meas},p}(f^-|s) \in \mathbb{R}$ and $\mathrm{E}_{\mathrm{meas},p}(f|s) = +\infty$, the desired equality follows trivially from Equation (7). We proceed to show that this is also the case if $\mathrm{E}_{\mathrm{meas},p}(f^+|s) < +\infty$.

Consider the non-increasing sequence $\{f^{\vee -n}\}_{n \in \mathbb{N}_0}$ of lower cuts of f. Then, for any $n \in \mathbb{N}_0$, the global variable $f^{\vee -n}$ is bounded below and $\mathscr{F}$-measurable. Hence,

$$\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \inf\left\{\mathrm{E}_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}_b, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\right\}$$
$$\leq \inf_{n \in \mathbb{N}_0} \mathrm{E}_{\mathrm{meas},p}(f^{\vee -n}|s) = \lim_{n \to +\infty} \mathrm{E}_{\mathrm{meas},p}(f^{\vee -n}|s) = \mathrm{E}_{\mathrm{meas},p}(f|s),$$

where the second equality follows from the non-increasing character of the sequence $\{f^{\vee -n}\}_{n \in \mathbb{N}_0}$ and the monotonicity [M2] of $\mathrm{E}_{\mathrm{meas},p}$, and the final equality follows from M9, which we can use because $f^{\vee -n} \leq f^{\vee 0} = f^+$ for all $n \in \mathbb{N}_0$ and $\mathrm{E}_{\mathrm{meas},p}(f^+|s) < +\infty$ by assumption. Combining this with Equation (7), the desired equality again follows. $\qquad\square$

Our next result shows that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is in fact the most conservative extension of $\mathrm{E}_{\mathrm{meas},p}$ that satisfies P4.

Proposition 13. $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ *is the most conservative extension of* $\mathrm{E}_{\mathrm{meas},p}$ *that satisfies* P4.

Proof. Fix any $f \in \overline{\mathbb{V}}$, any $s \in \mathscr{X}^*$ and any $\overline{\mathrm{E}}'_p : \overline{\mathbb{V}} \times \mathscr{X}^* \to \overline{\mathbb{R}}$ that satisfies P4 and that coincides with $\mathrm{E}_{\mathrm{meas},p}$ on its domain. Then,

$$\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \inf\left\{\mathrm{E}_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}_b, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\right\}$$
$$= \inf\left\{\overline{\mathrm{E}}'_p(g|s) : g \in \overline{\mathbb{V}}_b, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\right\} \geq \overline{\mathrm{E}}'_p(f|s),$$

where the second equality follows from the fact that $\mathrm{E}_{\mathrm{meas},p}(g|s)$ exists for any $\mathscr{F}$-measurable variable $g \in \overline{\mathbb{V}}_b$, and the inequality follows from the fact that $\overline{\mathrm{E}}'_p$ satisfies monotonicity [P4]. Hence, the proposition now follows immediately from the fact that, due to Proposition 12, $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ itself is an extension of $\mathrm{E}_{\mathrm{meas},p}$ and the fact that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ satisfies P4 because $\mathrm{E}_{\mathrm{meas},p}$ is monotone [M2]. $\qquad\square$

In Section 4, we argued that P1–P5 are desirable for a global upper expectation and, under those axioms, we put forward the most conservative upper expectation as our model of choice. Now, Proposition 13 guarantees that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is the most conservative extension of $\mathrm{E}_{\mathrm{meas},p}$ only under P4. So it may very well be that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ does not necessarily satisfy the axioms P1–P5. Since we consider these axioms to be desirable, that would mean that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is a too conservative extension of $\mathrm{E}_{\mathrm{meas},p}$, in the sense that it results in an unwarranted loss of information. Our following result assures that this is not the case; the extension $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ coincides with $\overline{\mathrm{E}}_{\mathrm{A},p}$ and hence, $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is guaranteed to satisfy P1–P5.

Theorem 14. *For any $f \in \overline{\mathbb{V}}$ and any $s \in \mathscr{X}^*$, we have that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \overline{\mathrm{E}}_{\mathrm{A},p}(f|s)$.*

Proof. For any $f \in \overline{\mathbb{V}}$ and any $s \in \mathscr{X}^*$, we immediately find that

$$\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \inf\Big\{\mathrm{E}_{\mathrm{meas},p}(g|s) \colon g \in \overline{\mathbb{V}}_{\mathrm{b}},\, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\Big\}$$
$$= \inf\Big\{\overline{\mathrm{E}}_{\mathrm{A},p}(g|s) \colon g \in \overline{\mathbb{V}}_{\mathrm{b}},\, g \text{ is } \mathscr{F}\text{-measurable and } g \geq f\Big\} \geq \overline{\mathrm{E}}_{\mathrm{A},p}(f|s),$$

where the second equality follows from Theorem 11 and the inequality from the monotonicity [P4] of $\overline{\mathrm{E}}_{\mathrm{A},p}$. To show that the converse inequality holds, we will use the fact that $\overline{\mathrm{E}}_{\mathrm{A},p}$ satisfies P7 (this follows from Proposition 5 and Theorem 6).

Consider any $g \in \overline{\mathbb{V}}_{\mathrm{b,lim}}$. Since g is the pointwise limit of a sequence $\{g_n\}_{n \in \mathbb{N}_0}$ of finitary gambles $g_n \in \mathbb{V}_{\mathrm{fin}}$ and since any finitary gamble is clearly $\mathscr{F}$-measurable, g is the pointwise limit of a sequence of $\mathscr{F}$-measurable global gambles. Then it follows that g itself is also $\mathscr{F}$-measurable ($\mathscr{F}$-measurability is preserved under pointwise convergence). Furthermore, by the definition of $\overline{\mathbb{V}}_{\mathrm{b,lim}}$, g is also bounded below. Hence, $\mathrm{E}_{\mathrm{meas},p}(g|s)$ exists and due to Theorem 11, we have that $\overline{\mathrm{E}}_{\mathrm{A},p}(g|s) = \mathrm{E}_{\mathrm{meas},p}(g|s)$. Since $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is an extension of $\mathrm{E}_{\mathrm{meas},p}$ by Proposition 12, this implies that $\overline{\mathrm{E}}_{\mathrm{A},p}(g|s) = \overline{\mathrm{E}}^1_{\mathrm{meas},p}(g|s)$. We conclude that

$$\overline{\mathrm{E}}_{\mathrm{A},p}(f|s) = \inf\Big\{\overline{\mathrm{E}}_{\mathrm{A},p}(g|s) \colon g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\Big\} = \inf\Big\{\overline{\mathrm{E}}^1_{\mathrm{meas},p}(g|s) \colon g \in \overline{\mathbb{V}}_{\mathrm{b,lim}} \text{ and } g \geq f\Big\} \geq \overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s),$$

where the first step follows from P7 and the last step follows from the monotonicity [P4] of $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ (as established by Proposition 13). $\qquad\square$

Theorem 14 does not only show us that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ satisfies P1–P5, but also, and more importantly, that it is the most conservative global upper expectation satisfying P1–P5. Furthermore, if we combine Theorem 14 with Proposition 13, it can also easily be seen that it is the most conservative extension of $\mathrm{E}_{\mathrm{meas},p}$ satisfying P1–P5.

Corollary 15. $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ *is the most conservative extension of* $\mathrm{E}_{\mathrm{meas},p}$ *that satisfies* P1–P5.

Proof. We know from Proposition 13 that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is the most conservative extension of $\mathrm{E}_{\mathrm{meas},p}$ that satisfies P4. This implies in particular that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is more (or at least as) conservative than any extension of $\mathrm{E}_{\mathrm{meas},p}$ that satisfies P1–P5. Since $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ itself satisfies P1–P5 due to Theorem 14, it follows that $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ is indeed the most conservative extension of $\mathrm{E}_{\mathrm{meas},p}$ satisfying P1–P5. $\qquad\square$

Another interesting consequence of Theorem 14 and Proposition 13 is that $\overline{\mathrm{E}}_{\mathrm{A},p}$ turns out to coincide with $\mathrm{E}_{\mathrm{meas},p}$ on the entire domain where it is defined, or, in other words, our model $\overline{\mathrm{E}}_{\mathrm{A},p}$ is an extension of the traditional measure-theoretic expectation $\mathrm{E}_{\mathrm{meas},p}$.

Corollary 16. *For any $f \in \overline{\mathbb{V}}$ and any $s \in \mathscr{X}^*$, we have that $\overline{\mathrm{E}}_{\mathrm{A},p}(f|s) = \mathrm{E}_{\mathrm{meas},p}(f|s)$ if $\mathrm{E}_{\mathrm{meas},p}(f|s)$ exists.*

Proof. Immediate consequence of Theorem 14 and Proposition 13. $\qquad\square$

Now, before we continue to consider the more general case of imprecise probability trees, we briefly want to come back to our choice of the extension $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$. Recall that we defined it as an upper approximation—called an upper integral—with respect to all $\mathscr{F}$-measurable variables in $\overline{\mathbb{V}}_{\mathrm{b}}$. However, one could rightfully wonder why we approximate with respect to this particular domain. It turned out to be a suitable choice, as Corollary 15

shows, but there was a priori no concrete reason for doing so. We could just as well have adopted the following alternative upper integral:[10]

$$\overline{\mathrm{E}}^2_{\mathrm{meas},p}(f|s) := \inf\left\{\mathrm{E}_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}, \mathrm{E}_{\mathrm{meas},p}(g|s) \text{ exists and } g \geq f\right\} \text{ for all } f \in \overline{\mathbb{V}} \text{ and } s \in \mathscr{X}^*.$$

The upper expectation $\overline{\mathrm{E}}^2_{\mathrm{meas},p}$ is an extension of $\mathrm{E}_{\mathrm{meas},p}$ because $\mathrm{E}_{\mathrm{meas},p}$ is monotone [M2]. Moreover, it is also clear that $\overline{\mathrm{E}}^2_{\mathrm{meas},p}(f|s)$ is dominated by $\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s)$ for all $f \in \overline{\mathbb{V}}$ and all $s \in \mathscr{X}^*$ because the infimum in $\overline{\mathrm{E}}^2_{\mathrm{meas},p}$ is taken over an equal or larger set than the set over which the infimum is taken in $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$. Hence, one could be tempted to adopt the upper integral $\overline{\mathrm{E}}^2_{\mathrm{meas},p}$ instead of $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ with the aim of obtaining a more informative extension—recall that we regard higher upper expectations as less informative. However, it can straightforwardly be shown that this is only vain hope.

Proposition 17. *We have that* $\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) = \overline{\mathrm{E}}^2_{\mathrm{meas},p}(f|s)$ *for all* $f \in \overline{\mathbb{V}}$ *and all* $s \in \mathscr{X}^*$.

Proof. Fix any $f \in \overline{\mathbb{V}}$ and any $s \in \mathscr{X}^*$. That $\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s) \geq \overline{\mathrm{E}}^2_{\mathrm{meas},p}(f|s)$ follows immediately from the fact that $\mathrm{E}_{\mathrm{meas},p}(g|s)$ exists for each $\mathscr{F}$-measurable variable $g \in \overline{\mathbb{V}}_{\mathrm{b}}$ and therefore, that the infimum in $\overline{\mathrm{E}}^2_{\mathrm{meas},p}(f|s)$ is taken over a set that is at least as large as the set over which the infimum is taken in $\overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s)$. On the other hand, we have that

$$\overline{\mathrm{E}}^2_{\mathrm{meas},p}(f|s) = \inf\left\{\mathrm{E}_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}, \mathrm{E}_{\mathrm{meas},p}(g|s) \text{ exists and } g \geq f\right\}$$

$$= \inf\left\{\overline{\mathrm{E}}^1_{\mathrm{meas},p}(g|s) : g \in \overline{\mathbb{V}}, \mathrm{E}_{\mathrm{meas},p}(g|s) \text{ exists and } g \geq f\right\} \geq \overline{\mathrm{E}}^1_{\mathrm{meas},p}(f|s),$$

where the second equality follows from Proposition 12 and the inequality follows from the monotonicity of $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ (see Proposition 13). □

Hence, both measure-theoretic extensions $\overline{\mathrm{E}}^1_{\mathrm{meas},p}$ and $\overline{\mathrm{E}}^2_{\mathrm{meas},p}$ are equal and therefore, by Theorem 14, both of them coincide with our model $\overline{\mathrm{E}}_{\mathrm{A},p}$ which, on its turn, coincides with the game-theoretic upper expectation $\overline{\mathrm{E}}_{\mathrm{V}}$ according to Theorem 6. In summary, we conclude that, if the local models are precise, all of the three possible approaches—the game-theoretic, the measure-theoretic and ours—are equivalent.

Imprecise Probability Trees

Let us now turn to the general case where we consider an imprecise probability tree $\overline{\mathrm{Q}}$ for which the local models $\overline{\mathrm{Q}}_s$ are (boundedly) coherent upper expectations that are not necessarily self-conjugate. As we have explained earlier, these local models can equivalently be represented by sets $\mathbb{P}_s$ of probability mass functions on $\mathscr{X}$, where $\mathbb{P}_s := \mathbb{P}_{\overline{\mathrm{Q}}_s}$ is related to $\overline{\mathrm{Q}}_s$ according to Equations (1) and (2) for all $s \in \mathscr{X}^*$. We call a precise probability tree $p : s \in \mathscr{X}^* \mapsto p(\cdot|s)$ *compatible* with the imprecise probability tree $\overline{\mathrm{Q}}$ and write that $p \sim \overline{\mathrm{Q}}$ if $p(\cdot|s) \in \mathbb{P}_s$ for all $s \in \mathscr{X}^*$. In other words, a compatible precise probability tree $p \sim \overline{\mathrm{Q}}$ corresponds to a possible selection, where a local probability mass function $p(\cdot|s)$ is chosen from the set $\mathbb{P}_s$ for each situation $s \in \mathscr{X}^*$. Now, for any such compatible tree $p \sim \overline{\mathrm{Q}}$, we can consider the global upper expectation $\overline{\mathrm{E}}_{\mathrm{meas},p} := \overline{\mathrm{E}}^1_{\mathrm{meas},p} = \overline{\mathrm{E}}^2_{\mathrm{meas},p}$ as defined in the previous section. Hence, instead of a single upper expectation, we now obtain a set of compatible global upper expectations. We define the *global measure-theoretic upper expectation* $\overline{\mathrm{E}}_{\mathrm{meas}}$ as the upper envelope over all such compatible upper expectations:

$$\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) := \sup\left\{\overline{\mathrm{E}}_{\mathrm{meas},p}(f|s) : p \sim \overline{\mathrm{Q}}\right\} \text{ for all } f \in \overline{\mathbb{V}} \text{ and } s \in \mathscr{X}^*.$$

This definition is in correspondence with the usual measure-theoretic view towards imprecision where it is regarded as partial information about a single, unknown probability measure. Upper (and lower) expectations

---

[10]This upper integral seems to be the one used by Shafer and Vovk in [5, Chapter 9].

are then considered as upper (and lower) bounds on the set of all 'possible' linear expectations. A similar, yet unconditional version of this measure-theoretic upper expectation is also adopted by Shafer and Vovk in [5, Chapter 9].

As we will now show, the measure-theoretic upper expectation $\overline{E}_{\text{meas}}$ satisfies our axioms P1–P5. This implies, by Theorem 6, that $\overline{E}_{\text{meas}}$ is always dominated by the global upper expectation $\overline{E}_A$, and therefore also by $\overline{E}_V$. Moreover, that the operator $\overline{E}_{\text{meas}}$ is entirely based on a measure-theoretic view towards uncertainty, yet turns out to satisfy our axioms P1–P5, can only be seen as an additional motivation for the use of these axioms.

Theorem 18. *The global measure-theoretic upper expectation $\overline{E}_{\text{meas}}$ is an element of $\overline{\mathbb{E}}_{1-5}(\overline{Q})$.*

Proof. To show that P1 holds, consider any $f \in \mathscr{L}(\mathscr{X})$ and any $x_{1:n} \in \mathscr{X}^*$. Since $f(X_{n+1})$ is bounded and clearly $\mathscr{F}$-measurable, we have, for any $p \sim \overline{Q}$, that $E_{\text{meas},p}(f(X_{n+1})|x_{1:n})$ exists and therefore that $\overline{E}_{\text{meas},p}(f(X_{n+1})|x_{1:n}) = E_{\text{meas},p}(f(X_{n+1})|x_{1:n})$, due to Proposition 12. Moreover, using $P_{|x_{1:n}}(x_{n+1})$ to denote $P_{|x_{1:n}}(\cup_{z_{1:n} \in \mathscr{X}^n} \Gamma(z_{1:n}x_{n+1}))$ for any $x_{n+1} \in \mathscr{X}$, it is clear that $E_{\text{meas},p}(f(X_{n+1})|x_{1:n}) = \sum_{x_{n+1} \in \mathscr{X}} f(x_{n+1})P_{|x_{1:n}}(x_{n+1})$. It also follows immediately from the construction (6) of $P_{|x_{1:n}} = P(\cdot|x_{1:n})$ that $P_{|x_{1:n}}(x_{n+1}) = P_{|x_{1:n}}(x_{1:n+1}) = p(x_{n+1}|x_{1:n})$ for all $x_{n+1} \in \mathscr{X}$ and $p \sim \overline{Q}$. Hence,

$$
\begin{aligned}
\overline{E}_{\text{meas}}(f(X_{n+1})|x_{1:n}) &= \sup\left\{\overline{E}_{\text{meas},p}(f(X_{n+1})|x_{1:n}): p \sim \overline{Q}\right\} \\
&= \sup\left\{E_{\text{meas},p}(f(X_{n+1})|x_{1:n}): p \sim \overline{Q}\right\} \\
&= \sup\left\{\sum_{x_{n+1} \in \mathscr{X}} f(x_{n+1})p(x_{n+1}|x_{1:n}): p \sim \overline{Q}\right\} \\
&= \sup\left\{\sum_{x_{n+1} \in \mathscr{X}} f(x_{n+1})p(x_{n+1}|x_{1:n}): p(\cdot|x_{1:n}) \in \mathbb{P}_{x_{1:n}}\right\} = \overline{Q}_{x_{1:n}}(f),
\end{aligned}
$$

where the last step follows from the relation between $\mathbb{P}_{x_{1:n}}$ and $\overline{Q}_{x_{1:n}}$ which is given by Equation (2).

The remaining properties P2–P5 can be easily derived from the definition of $\overline{E}_{\text{meas}}$ and the fact that, due to Corollary 15, these properties are satisfied by each compatible $\overline{E}_{\text{meas},p}$. For Properties P2 and P4 this is immediate. To see that P3 holds, consider any $f \in \mathbb{V}_{\text{fin}}$ and any $n \in \mathbb{N}_0$. Note that, for any compatible tree $p \sim \overline{Q}$ and any $x_{1:n} \in \mathscr{X}^n$, we have that $\overline{E}_{\text{meas},p}(f|x_{1:n}) \leq \overline{E}_{\text{meas},p}(\overline{E}_{\text{meas},p}(f|X_{1:n+1})|x_{1:n})$ because $\overline{E}_{\text{meas},p}$ satisfies P3. Hence,

$$
\begin{aligned}
\overline{E}_{\text{meas}}(f|x_{1:n}) = \sup\left\{\overline{E}_{\text{meas},p}(f|x_{1:n}): p \sim \overline{Q}\right\} &\leq \sup\left\{\overline{E}_{\text{meas},p}(\overline{E}_{\text{meas},p}(f|X_{1:n+1})|x_{1:n}): p \sim \overline{Q}\right\} \\
&\leq \sup\left\{\overline{E}_{\text{meas},p}(\overline{E}_{\text{meas}}(f|X_{1:n+1})|x_{1:n}): p \sim \overline{Q}\right\} \\
&= \overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:n+1})|x_{1:n}),
\end{aligned}
$$

where the second inequality follows from the monotonicity [P4] of $\overline{E}_{\text{meas},p}$ and the fact that $\overline{E}_{\text{meas},p}(f|z_{1:n+1}) \leq \overline{E}_{\text{meas}}(f|z_{1:n+1})$ for all $p \sim \overline{Q}$ and all $z_{1:n+1} \in \mathscr{X}^{n+1}$. The inequality above holds for any $x_{1:n} \in \mathscr{X}^n$, so we indeed find that $\overline{E}_{\text{meas}}(f|X_{1:n}) \leq \overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:n+1})|X_{1:n})$. To prove P5, we consider any $s \in \mathscr{X}^*$ and any sequence $\{f_n\}_{n \in \mathbb{N}_0}$ of finitary gambles that is uniformly bounded below and that converges pointwise to some variable $f \in \overline{\mathbb{V}}$. For any compatible tree $p \sim \overline{Q}$, $\overline{E}_{\text{meas},p}$ satisfies P5 and therefore, $\overline{E}_{\text{meas},p}(f|s) \leq \limsup_{n \to +\infty} \overline{E}_{\text{meas},p}(f_n|s)$. Then we also have that

$$
\begin{aligned}
\overline{E}_{\text{meas}}(f|s) = \sup\left\{\overline{E}_{\text{meas},p}(f|s): p \sim \overline{Q}\right\} &\leq \sup\left\{\limsup_{n \to +\infty} \overline{E}_{\text{meas},p}(f_n|s): p \sim \overline{Q}\right\} \\
&\leq \sup\left\{\limsup_{n \to +\infty}\left[\sup\left\{\overline{E}_{\text{meas},p'}(f_n|s): p' \sim \overline{Q}\right\}\right]: p \sim \overline{Q}\right\} \\
&= \limsup_{n \to +\infty}\left[\sup\left\{\overline{E}_{\text{meas},p'}(f_n|s): p' \sim \overline{Q}\right\}\right] = \limsup_{n \to +\infty} \overline{E}_{\text{meas}}(f_n|s). \qquad \square
\end{aligned}
$$

As mentioned before, the result above implies, in combination with Theorem 6, that $\overline{E}_{\text{meas}}$ is dominated by our model $\overline{E}_A$.

Corollary 19. *We have that* $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) \leq \overline{\mathrm{E}}_{\mathrm{A}}(f|s) = \overline{\mathrm{E}}_{\mathrm{V}}(f|s)$ *for all* $f \in \overline{\mathbb{V}}$ *and all* $s \in \mathscr{X}^*$.

Proof. Immediate consequence of Theorem 6 and Theorem 18. □

Corollary 19 extends upon an earlier result by Lopatatzidis et al. [13, Theorem 33], which says that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) \leq \overline{\mathrm{E}}_{\mathrm{V}}(f|s)$ for all $s \in \mathscr{X}^*$ and some specific variables f in $\overline{\mathbb{V}}_{\mathrm{b,lim}}$. The definition of $\overline{\mathrm{E}}_{\mathrm{V}}$ in [13] slightly differs from ours though. The version in [13] is always larger or equal than—so at least as conservative as—ours because they work with supermartingales that can only take real values; see [20, section 8] for more details. Hence, Corollary 19 certainly implies [13, Theorem 33]. Another existing result that is strongly related to Corollary 19 is due to Shafer and Vovk [5, Proposition 9.8], which seems to say that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|\square) \leq \overline{\mathrm{E}}_{\mathrm{V}}(f|\square)$ for all bounded below variables $f \in \overline{\mathbb{V}}_{\mathrm{b}}$. Since our result applies to all $f \in \overline{\mathbb{V}}$ and all $s \in \mathscr{X}^*$, our result would in that sense be stronger than theirs. However, it is not immediately clear whether we can actually draw such conclusions. Unlike in our case, Shafer and Vovk [5, Section 9.2] do not require that, for each possible situation, a local model should be specified beforehand. To put it in their terms: Forecaster need not declare to Skeptic beforehand—that is, before the game starts—which local upper expectations he is going to choose. Instead, Shafer and Vovk only start off with a set $\mathbb{Q}$ of local upper expectations from which, at each situation, Forecaster can choose. Since this set $\mathbb{Q}$ is moreover required to be constant throughout the whole game, this crucially implies that we, as subjects outside the game, are unable to specify Forecaster's moves—and hence, the choice of the local upper expectation at each situation. The global game-theoretic upper expectation will therefore not depend on Forecaster's moves, but rather only on the set $\mathbb{Q}$. Moreover, both of their global models—their global game-theoretic upper expectation and their global measure-theoretic upper expectation—are defined on variables with a domain that differs from Ω. Besides the fact that we find it far from trivial how this whole framework relates to ours here, they also impose some relatively strong toplogical conditions on the set $\mathbb{Q}$. Hence, all together, in order to draw a fair comparison, we feel like a closer look is required. We hope to clarify this in our future work.

Moving on, the inevitable question that comes to mind when looking at Corollary 19 is whether the inequality can be turned into an equality. We are aware of two partial answers. On the one hand, Shafer and Vovk seem to establish in [5, Theorem 9.7] that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|\square) = \overline{\mathrm{E}}_{\mathrm{V}}(f|\square)$ for all Suslin gambles $f \in \mathbb{V}$; this includes all $\mathscr{F}$-measurable gambles and is therefore a very strong result. However, here too, as before, their framework differs substantially from ours and the result requires additional topological assumptions. It is therefore again not immediately clear if, and how, it translates to our context; we leave this for future work. On the other hand, we also have a result by Lopatatzidis et al. [13, Theorem 29] that shows that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) = \overline{\mathrm{E}}_{\mathrm{V}}(f|s)$ for all $s \in \mathscr{X}^*$ and all n-measurable gambles f—and therefore all finitary gambles f in $\mathbb{V}_{\mathrm{fin}}$. Here too, however, the result is not immediately applicable because —as explained directly after Corollary 19—the version of $\overline{\mathrm{E}}_{\mathrm{V}}$ that is used in [13] differs from ours. We therefore give a self-contained proof instead. Nevertheless, our ideas and techniques are essentially the same as the ones that were used to obtain [13, Theorem 29].

We start by showing that $\overline{\mathrm{E}}_{\mathrm{meas}}$ satisfies P3$^=$; the desired equivalence result will then follow as an immediate consequence.

Lemma 20. *The upper expectation* $\overline{\mathrm{E}}_{\mathrm{meas}}$ *satisfies* P3$^=$.

Proof. Consider any $f \in \mathbb{V}_{\mathrm{fin}}$. Let $n \in \mathbb{N}_0$ be such that f is n-measurable—there is at least one such an n because f is finitary. Then, to establish the fact that $\overline{\mathrm{E}}_{\mathrm{meas}}$ satisfies P3$^=$, it suffices to prove that, for all $m \in \mathbb{N}_0$,

$$\text{there is some } p \sim \overline{\mathrm{Q}} \text{ such that } \overline{\mathrm{E}}_{\mathrm{meas}}(\overline{\mathrm{E}}_{\mathrm{meas}}(f|X_{1:m+1})|X_{1:m}) = \overline{\mathrm{E}}_{\mathrm{meas},p}(f|X_{1:m}). \tag{8}$$

Indeed, by definition of $\overline{\mathrm{E}}_{\mathrm{meas}}$, $\overline{\mathrm{E}}_{\mathrm{meas},p}(f|X_{1:m}) \leq \overline{\mathrm{E}}_{\mathrm{meas}}(f|X_{1:m})$ for all $p \sim \overline{\mathrm{Q}}$ and all $m \in \mathbb{N}_0$, and therefore, expression (8) implies that $\overline{\mathrm{E}}_{\mathrm{meas}}(\overline{\mathrm{E}}_{\mathrm{meas}}(f|X_{1:m+1})|X_{1:m}) \leq \overline{\mathrm{E}}_{\mathrm{meas}}(f|X_{1:m})$ for all $m \in \mathbb{N}_0$. The converse inequality then follows from Theorem 18, which says that $\overline{\mathrm{E}}_{\mathrm{meas}}$ satisfies P3.

Let us first show that expression (8) holds for any $m \geq n$. Consider any $\ell \geq n$ and any $x_{1:\ell} \in \mathscr{X}^\ell$. Since f is n-measurable and $\ell \geq n$, f takes a constant value $f(x_{1:n})$ on $\Gamma(x_{1:\ell})$, and therefore, we can write that $f\mathbb{I}_{x_{1:\ell}} = f(x_{1:n})\mathbb{I}_{x_{1:\ell}}$. Hence, for any $p \sim \overline{\mathrm{Q}}$, since $\overline{\mathrm{E}}_{\mathrm{meas},p}$ satisfies P2 due to Theorem 14, we have that

$$\overline{\mathrm{E}}_{\mathrm{meas},p}(f|x_{1:\ell}) \overset{\mathrm{P2}}{=} \overline{\mathrm{E}}_{\mathrm{meas},p}(f\mathbb{I}_{x_{1:\ell}}|x_{1:\ell}) = \overline{\mathrm{E}}_{\mathrm{meas},p}(f(x_{1:n})\mathbb{I}_{x_{1:\ell}}|x_{1:\ell}) \overset{\mathrm{P2}}{=} \overline{\mathrm{E}}_{\mathrm{meas},p}(f(x_{1:n})|x_{1:\ell}) = \mathrm{E}_{\mathrm{meas},p}(f(x_{1:n})|x_{1:\ell}).$$

where the last step follows from Proposition 12 and the fact that $E_{\text{meas},p}(f(x_{1:n})|x_{1:\ell})$ exists because $f(x_{1:n})$ is a constant. Furthermore, since $f(x_{1:n})$ is real-valued [because f is a gamble], M1 implies that $E_{\text{meas},p}(f(x_{1:n})|x_{1:\ell}) = f(x_{1:n})$. Hence, we have that

$$\overline{E}_{\text{meas},p}(f|x_{1:\ell}) = f(x_{1:n}) \text{ for all } p \sim \overline{Q}, \text{ all } \ell \geq n \text{ and all } x_{1:\ell} \in \mathcal{X}^{\ell}. \tag{9}$$

So, since $m \geq n$, the right hand side of Equation (8) is equal to $f(X_{1:n}) = f$, whatever the tree $p \sim \overline{Q}$. Moreover, note that Equation (9) also implies that $\overline{E}_{\text{meas}}(f|x_{1:\ell}) = f(x_{1:n})$ for all $\ell \geq n$ and all $x_{1:\ell} \in \mathcal{X}^{\ell}$. Then, since $m \geq n$, and therefore certainly $m+1 \geq n$, the left hand side of Equation (8) is equal to

$$\overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:m+1})|X_{1:m}) = \overline{E}_{\text{meas}}(f(X_{1:n})|X_{1:m}) = \overline{E}_{\text{meas}}(f|X_{1:m}) = f,$$

where the last step follows once more from the fact that $m \geq n$, and that $\overline{E}_{\text{meas}}(f|x_{1:\ell}) = f(x_{1:n})$ for all $\ell \geq n$ and all $x_{1:\ell} \in \mathcal{X}^{\ell}$. Hence, irrespectively of the tree $p \sim \overline{Q}$, both sides of Equation (8) are equal to f—and therefore equal to each other—for the case that $m \geq n$.

For the case that $m < n$, we will establish expression (8) using a decreasing induction argument in m. The case where $m = n$ will serve as an induction base; that expression (8) holds for this case was just proved above. Now suppose that expression (8) is satisfied with $m = \ell$ where $1 \leq \ell \leq n$, meaning that there is some $p \sim \overline{Q}$ such that $\overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:\ell+1})|X_{1:\ell}) = \overline{E}_{\text{meas},p}(f|X_{1:\ell})$. First note that this implies that $\overline{E}_{\text{meas}}(f|X_{1:\ell}) = \overline{E}_{\text{meas},p}(f|X_{1:\ell})$ because, on the one hand, $\overline{E}_{\text{meas}}(f|X_{1:\ell}) \leq \overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:\ell+1})|X_{1:\ell})$ due to Theorem 18 which says that $\overline{E}_{\text{meas}}$ satisfies P3, and on the other hand, $\overline{E}_{\text{meas},p}(f|X_{1:\ell}) \leq \overline{E}_{\text{meas}}(f|X_{1:\ell})$ due to the definition of $\overline{E}_{\text{meas}}$. Furthermore, because f is a finitary gamble, and therefore an $\mathcal{F}$-measurable gamble, $E_{\text{meas},p}(f|X_{1:\ell})$ exists and therefore, by Proposition 12,

$$\overline{E}_{\text{meas}}(f|X_{1:\ell}) = \overline{E}_{\text{meas},p}(f|X_{1:\ell}) = E_{\text{meas},p}(f|X_{1:\ell}).$$

Moreover, using M1 and M2, and the fact that $\inf f$ and $\sup f$ are both real (because f is a gamble), it can easily be inferred that $\inf f \leq E_{\text{meas},p}(f|X_{1:\ell}) \leq \sup f$. So $\overline{E}_{\text{meas}}(f|X_{1:\ell}) = E_{\text{meas},p}(f|X_{1:\ell})$ is a gamble, and more specifically an ℓ-measurable gamble. Then, fix any $x_{1:\ell-1} \in \mathcal{X}^{\ell-1}$ and note that, since $\overline{E}_{\text{meas}}$ satisfies P1 and P2 by Theorem 18, we can apply Lemma 1 to obtain that

$$\overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:\ell})|x_{1:\ell-1}) = \overline{E}_{\text{meas}}(E_{\text{meas},p}(f|X_{1:\ell})|x_{1:\ell-1}) = \overline{Q}_{x_{1:\ell-1}}(E_{\text{meas},p}(f|x_{1:\ell-1}\cdot)).$$

Then, due to the fact that $\mathbb{P}_{x_{1:\ell-1}}$ is related with $\overline{Q}_{x_{1:\ell-1}}$ through Equation (2), there is some $p'(\cdot|x_{1:\ell-1}) \in \mathbb{P}_{x_{1:\ell-1}}$ such that

$$\overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:\ell})|x_{1:\ell-1}) = \overline{Q}_{x_{1:\ell-1}}(E_{\text{meas},p}(f|x_{1:\ell-1}\cdot)) = \sum_{x_{\ell} \in \mathcal{X}} p'(x_{\ell}|x_{1:\ell-1}) E_{\text{meas},p}(f|x_{1:\ell}). \tag{10}$$

Moreover, since f is n-measurable, we can write that $f = \sum_{z_{1:n} \in \mathcal{X}^n} f(z_{1:n}) \mathbb{I}_{z_{1:n}}$, and therefore, the Lebesgue integral $E_{\text{meas},p}(f|x_{1:\ell})$ for any $x_{\ell} \in \mathcal{X}$, reduces—by definition; see [1, Section 2.6.1]—to the finite sum

$$E_{\text{meas},p}(f|x_{1:\ell}) = \sum_{z_{1:n} \in \mathcal{X}^n} f(z_{1:n}) P_{|x_{1:\ell}}(z_{1:n}) = \sum_{z_{1:\ell} \in \mathcal{X}^{\ell}} \sum_{x_{\ell+1:n} \in \mathcal{X}^{n-\ell}} f(z_{1:\ell} x_{\ell+1:n}) P_{|x_{1:\ell}}(z_{1:\ell} x_{\ell+1:n}),$$

where the last step simply is a convenient change of notation. Due to Equation (6), we have, for all $z_{1:\ell} \in \mathcal{X}^{\ell}$ and all $x_{\ell+1:n} \in \mathcal{X}^{n-\ell}$, that $P_{|x_{1:\ell}}(z_{1:\ell} x_{\ell+1:n}) = 0$ if $z_{1:\ell} \neq x_{1:\ell}$ and $P_{|x_{1:\ell}}(z_{1:\ell} x_{\ell+1:n}) = \prod_{i=\ell}^{n-1} p(x_{i+1}|x_{1:i})$ if $z_{1:\ell} = x_{1:\ell}$. Substituting these expressions in the equality above, gives us

$$E_{\text{meas},p}(f|x_{1:\ell}) = \sum_{x_{\ell+1:n} \in \mathcal{X}^{n-\ell}} f(x_{1:n}) P_{|x_{1:\ell}}(x_{1:n}) = \sum_{x_{\ell+1:n} \in \mathcal{X}^{n-\ell}} f(x_{1:n}) \prod_{i=\ell}^{n-1} p(x_{i+1}|x_{1:i}).$$

Then, plugging this back into Equation (10), we obtain that

$$\overline{E}_{\text{meas}}(\overline{E}_{\text{meas}}(f|X_{1:\ell})|x_{1:\ell-1}) = \sum_{x_{\ell} \in \mathcal{X}} p'(x_{\ell}|x_{1:\ell-1}) \sum_{x_{\ell+1:n} \in \mathcal{X}^{n-\ell}} f(x_{1:n}) \prod_{i=\ell}^{n-1} p(x_{i+1}|x_{1:i})$$

$$= \sum_{x_{\ell:n} \in \mathcal{X}^{n-\ell+1}} f(x_{1:n}) p'(x_{\ell}|x_{1:\ell-1}) \prod_{i=\ell}^{n-1} p(x_{i+1}|x_{1:i}). \tag{11}$$

22

Now let $p^* \sim \overline{Q}$ be any compatible precise probability tree such that $p^*(\cdot|x_{1:\ell-1}) = p'(\cdot|x_{1:\ell-1})$ and $p^*(\cdot|x_{1:i}) = p(\cdot|x_{1:i})$ for all $\ell \leq i \leq n-1$ and all $x_{\ell:i} \in \mathcal{X}^{i-\ell+1}$; it is always possible to choose such a compatible tree p^* because $p \sim \overline{Q}$ and $p'(\cdot|x_{1:\ell-1}) \in \mathbb{P}_{x_{1:\ell-1}}$. Then note that, for any $x_{\ell:n} \in \mathcal{X}^{n-\ell+1}$, the corresponding probability $\mathrm{P}^*_{|x_{1:\ell-1}}(x_{1:n})$, related with p^* through Equation (6), is equal to $p'(x_\ell|x_{1:\ell-1}) \prod_{i=\ell}^{n-1} p(x_{i+1}|x_{1:i})$. Hence, recalling Equation (11), we infer that

$$\overline{\mathrm{E}}_{\mathrm{meas}}(\overline{\mathrm{E}}_{\mathrm{meas}}(f|X_{1:\ell})|x_{1:\ell-1}) = \sum_{x_{\ell:n} \in \mathcal{X}^{n-\ell+1}} f(x_{1:n}) \mathrm{P}^*_{|x_{1:\ell-1}}(x_{1:n}) = \sum_{z_{1:\ell-1} \in \mathcal{X}^{\ell-1}} \sum_{x_{\ell:n} \in \mathcal{X}^{n-\ell+1}} f(z_{1:\ell-1} x_{\ell:n}) \mathrm{P}^*_{|x_{1:\ell-1}}(z_{1:\ell-1} x_{\ell:n})$$
$$= \mathrm{E}_{\mathrm{meas},p^*}(f|x_{1:\ell-1}) = \overline{\mathrm{E}}_{\mathrm{meas},p^*}(f|x_{1:\ell-1}),$$

where the second step follows from the fact that $\mathrm{P}^*_{|x_{1:\ell-1}}(z_{1:\ell-1} x_{\ell:n}) = 0$ if $z_{1:\ell-1} \neq x_{1:\ell-1}$ due to Equation (6), where the penultimate step follows from $f = \sum_{z_{1:n} \in \mathcal{X}^n} f(z_{1:n}) \mathbb{I}_{z_{1:n}}$ and the definition of the Lebesgue integral [1, Section 2.6.1], and where the last step follows from Proposition 12. Hence, since the equation above holds for any $x_{1:\ell-1} \in \mathcal{X}^{\ell-1}$, we have established expression (8) for the case where $m = \ell-1$, concluding our proof of the induction step. □

That $\overline{\mathrm{E}}_{\mathrm{meas}}$, $\overline{\mathrm{E}}_{\mathrm{A}}$ and $\overline{\mathrm{E}}_{\mathrm{V}}$ are all equal on the domain $\mathbb{V}_{\mathrm{fin}}$ now follows straightforwardly by combining Lemma 20 with Corollary 19, Theorem 18 and Proposition 2.

Proposition 21. *We have that* $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) = \overline{\mathrm{E}}_{\mathrm{A}}(f|s) = \overline{\mathrm{E}}_{\mathrm{V}}(f|s)$ *for all* $f \in \mathbb{V}_{\mathrm{fin}}$ *and all* $s \in \mathcal{X}^*$.

Proof. Due to Corollary 19, it suffices to prove that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) \geq \overline{\mathrm{E}}_{\mathrm{A}}(f|s)$ for all $f \in \mathbb{V}_{\mathrm{fin}}$ and all $s \in \mathcal{X}^*$. This can easily be done by combining Theorem 18, Lemma 20 and Proposition 2. Indeed, Theorem 18 says that $\overline{\mathrm{E}}_{\mathrm{meas}}$ is an element of $\overline{\mathbb{E}}_{1-5}(\overline{Q}) \subseteq \overline{\mathbb{E}}_{1-4}(\overline{Q})$, and Lemma 20 says that $\overline{\mathrm{E}}_{\mathrm{meas}}$ satisfies P3$^=$, so Proposition 2 implies that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) \geq \overline{\mathrm{E}}'(f|s)$ for all $f \in \mathbb{V}_{\mathrm{fin}}$, all $s \in \mathcal{X}^*$ and all $\overline{\mathrm{E}}' \in \overline{\mathbb{E}}_{1-4}(\overline{Q})$. In particular, this implies that $\overline{\mathrm{E}}_{\mathrm{meas}}(f|s) \geq \overline{\mathrm{E}}_{\mathrm{A}}(f|s)$ for all $f \in \mathbb{V}_{\mathrm{fin}}$ and all $s \in \mathcal{X}^*$ because, by definition, $\overline{\mathrm{E}}_{\mathrm{A}} \in \overline{\mathbb{E}}_{1-5}(\overline{Q}) \subseteq \overline{\mathbb{E}}_{1-4}(\overline{Q})$. □

Whether the equality with $\overline{\mathrm{E}}_{\mathrm{meas}}$ can be extended to a more general domain, remains an open question that we would like to pursue in our future research. We expect that there is much to learn from the results in [5, Section 9.2], provided that they can be translated to our setting here.

10. Conclusion

We have put forward a small set of simple axioms P1–P5 and argued that they are desirable for a global upper expectation in the context of discrete-time finite-state uncertain processes. We have established the existence of a unique most conservative model under these axioms, and have in addition given sufficient conditions to uniquely characterise this most conservative model. Using this characterisation, we have shown that our most conservative upper expectation coincides with a version of the game-theoretic upper expectation used by Shafer and Vovk, and therefore has particularly powerful mathematical properties, despite the simplicity of our defining axioms.

We also considered an alternative, more traditional measure-theoretic approach, where we defined a global belief model using an upper envelope of upper integrals corresponding to compatible precise probability trees. We have shown that this model satisfies axioms P1–P5 and therefore, that our most conservative model gives guaranteed upper bounds on the value of this measure-theoretic upper expectation. For precise probability trees, both models, and consequently also the game-theoretic model, coincide. In the imprecise case, equality was only established for the domain of finitary gambles, and it remains to be seen whether this can be extended to a more general domain.

Given our current findings, we belief there are a number of reasons why our global model $\overline{\mathrm{E}}_{\mathrm{A}}$ candidates as an excellent choice when it comes to modelling discrete-time finite-state uncertain processes. First and foremost, its definition is based on the axioms P1–P5; a set of properties which, in our opinion, are essential for a global belief model to have. Our model $\overline{\mathrm{E}}_{\mathrm{A}}$ is taken to be the most conservative or, equivalently, the least informative model under these axioms because we do not want to include any further information. In this way, we obtain a definition that is both convincing and interpretationally clear. Moreover, if one desires her global belief model to also have other properties, additional to P1–P5, then our model $\overline{\mathrm{E}}_{\mathrm{A}}$ would still serve as a conservative upper

bound for her model. As a second argument, we recall from Section 8 that our model $\overline{E}_A$ possesses rather strong continuity properties, as well as several fundamental probabilistic laws. Thirdly and lastly, our model $\overline{E}_A$ coincides with the game-theoretic upper expectation $\overline{E}_V$ and dominates the measure-theoretic upper expectation $\overline{E}_{\text{meas}}$. Hence, our model $\overline{E}_A$ plays an important role to anyone: for a person with a game-theoretic background, the equivalent models $\overline{E}_A$ and $\overline{E}_V$ can be used interchangeably and can therefore benefit from each others properties; for a person with a measure-theoretic background, our model $\overline{E}_A$ always gives a conservative upper bound for the model $\overline{E}_{\text{meas}}$.

One counterargument that could be given to the plea above, is whether axiom P5 is really desirable or even justified for a global belief model. Our justification for it is based on the fact that we consider upper expectations of non-finitary variables to be abstract idealisations of upper expectations of finitary variables. Still, this could be seen as a rather artificial and unnecessary assumption, similar to the countable additivity property in measure theory. However, compared to other common continuity properties, e.g. monotone convergence, property P5 is fairly weak since it only applies to sequences of finitary gambles and only imposes an inequality on the value of the global upper expectations. Moreover, that both the game-theoretic and the measure-theoretic model satisfy property P5, can only be seen as additional motivation to adopt it.

Acknowledgements

References

[1] A. N. Shiryaev, Probability-1, Graduate Texts in Mathematics, third ed., Springer, New York, 2016.

[2] A. N. Shiryaev, Probability-2, Graduate Texts in Mathematics, third ed., Springer, New York, 2019.

[3] P. Billingsley, Probability and Measure, Wiley Series in Probability and Statistics, Wiley, 1995.

[4] G. Shafer, V. Vovk, Probability and Finance: It's Only a Game!, Wiley, New York, 2001.

[5] G. Shafer, V. Vovk, Game-Theoretic Foundations for Probability and Finance, Wiley, Hoboken, 2019.

[6] D. Williams, Probability with Martingales, Cambridge mathematical textbooks, Cambridge University Press, 1991.

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

[8] T. Augustin, F. Coolen, G. de Cooman, M. Troffaes, Introduction to Imprecise Probabilities, John Wiley & Sons, Chichester, 2014.

[9] G. de Cooman, M. C. Troffaes, Lower Previsions, Wiley, Chichester, 2014.

[10] P. M. Williams, Notes on conditional previsions, Technical Report, School of Mathematical and Physical Science, University of Sussex, 1975. For a revised journal version, see [15].

[11] M. C. M. Troffaes, G. de Cooman, Lower previsions, John Wiley & Sons, 2014.

[12] G. de Cooman, F. Hermans, Imprecise probability trees: Bridging two theories of imprecise probability, Artificial Intelligence 172 (2008) 1400–1427.

[13] S. Lopatatzidis, Robust Modelling and Optimisation in Stochastic Processes using Imprecise Probabilities, with an Application to Queueing Theory, Ph.D. thesis, Ghent University, 2017.

[14] N. T'Joens, J. De Bock, G. de Cooman, In search of a global belief model for discrete-time uncertain processes, in: Proceedings of the 11th International Symposium on Imprecise Probabilities: Theories and Applications, volume 103, 2019, pp. 377–385.

[15] P. M. Williams, Notes on conditional previsions, International Journal of Approximate Reasoning 44 (2007) 366–383.

[16] H. L. Royden, P. M. Fitzpatrick, Real Analysis, Prentice Hall, New Jersey, 2010.

[17] G. Shafer, V. Vovk, A. Takemura, Lévy's zero–one law in game-theoretic probability, Journal of Theoretical Probability 25 (2012) 1–24.

[18] G. de Cooman, J. De Bock, S. Lopatatzidis, Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems, International Journal of Approximate Reasoning 76 (2016) 18–46.

[19] J. Ville, Étude critique de la notion de collectif, Monographies des probabilités, Gauthier-Villars, Paris, 1939.

[20] N. T'Joens, J. De Bock, G. de Cooman, Game-theoretic upper expectations for discrete-time finite-state uncertain processes, 2020. Submitted for publication; see arXiv:2008.03133 for a preliminary online version.

[21] N. T'Joens, G. de Cooman, J. De Bock, Continuity of the Shafer-Vovk-Ville operator, in: Proceedings of the 9th International Conference on Soft Methods in Probability and Statistics, volume 832, 2018, pp. 200–207.

[22] A. N. Shiryaev, Probability, Graduate Texts in Mathematics, second ed., Springer, New York, 1996.

Appendix A. Basic Measure-Theoretic Concepts

A *measurable space* $(\mathcal{Y}, \mathcal{F})$ is a couple, where $\mathcal{Y}$ is a non-empty set and $\mathcal{F}$ is a σ-algebra on $\mathcal{Y}$. We say that $A \subseteq \mathcal{Y}$ is $\mathcal{F}$-*measurable* if $A \in \mathcal{F}$. Such a set $A \in \mathcal{F}$ is also called an ($\mathcal{F}$-measurable) *event*. We say that an extended real-valued function $f: \mathcal{Y} \to \overline{\mathbb{R}}$ is $\mathcal{F}$-*measurable* if the set $f^{-1}(A) := \{y \in \mathcal{Y}: f(y) \in A\}$ is $\mathcal{F}$-measurable for every $A \in \mathcal{B}(\overline{\mathbb{R}})$. Here, $\mathcal{B}(\overline{\mathbb{R}})$ denotes the Borel algebra on $\overline{\mathbb{R}}$, being the σ-algebra generated by all open—or equivalently, closed—sets in $\overline{\mathbb{R}}$. Recall that we consider the usual order topology on $\overline{\mathbb{R}}$ and therefore, the Borel algebra $\mathcal{B}(\overline{\mathbb{R}})$ can alternatively be generated from the sets $\{x \in \overline{\mathbb{R}}: x \le c\}$ where $c \in \mathbb{R}$; see for instance [1, Section 2.2.2]. Then, as we have done in the main text, we can alternatively characterise the $\mathcal{F}$-measurable functions as those functions $f: \mathcal{Y} \to \overline{\mathbb{R}}$ such that $\{y \in \mathcal{Y}: f(y) \le c\}$ is $\mathcal{F}$-measurable for all $c \in \mathbb{R}$ [3, Theorem 13.1. (i)]. Typically, an $\mathcal{F}$-measurable (extended) real-valued function f is also called a *random variable*.

A *probability space* $(\mathcal{Y}, \mathcal{F}, \mathrm{P})$ is a measurable space $(\mathcal{Y}, \mathcal{F})$ equipped with a σ-additive probability measure P on $\mathcal{F}$. We say that an event $A \in \mathcal{F}$ is P-*null* if $\mathrm{P}(A) = 0$. We will also say that a property about the elements in $\mathcal{Y}$ holds P-*almost surely* (P-a.s.) if there is a P-*null* event $A \in \mathcal{F}$ such that the property holds for all $y \in \mathcal{Y} \setminus A$. Note that the intersection $A \cap B \in \mathcal{F}$ of two P-almost sure events $A, B \in \mathcal{F}$, is itself also P-almost sure. Now, for any probability space $(\mathcal{Y}, \mathcal{F}, \mathrm{P})$ and any extended real-valued function $f: \mathcal{Y} \to \overline{\mathbb{R}}$ that is $\mathcal{F}$-measurable, the (measure-theoretic) expectation $\mathrm{E}(f)$ of f will be defined using the Lebesgue integral [1, 3, 16]: $\mathrm{E}(f) := \int_{\mathcal{Y}} f \, d\mathrm{P}$. As we already mentioned in the main text, this integral is definitely well-defined if f is non-negative and is otherwise only defined if $\min\{\int_{\mathcal{Y}} f^+ d\mathrm{P}, \int_{\mathcal{Y}} f^- d\mathrm{P}\} < +\infty$, with $f^+ = f^{\vee 0}$ and $f^- = -f^{\wedge 0}$, in which case we let $\int_{\mathcal{Y}} f \, d\mathrm{P} := \int_{\mathcal{Y}} f^+ d\mathrm{P} - \int_{\mathcal{Y}} f^- d\mathrm{P}$. When we say that '$\mathrm{E}(f)$ exists', we take this to mean that its defining integral $\int_{\mathcal{Y}} f \, d\mathrm{P}$ is well-defined. Furthermore, an extended real-valued function $f: \mathcal{Y} \to \overline{\mathbb{R}}$ is called P-*integrable* if it is $\mathcal{F}$-measurable and $\int_{\mathcal{Y}} |f| d\mathrm{P} = \int_{\mathcal{Y}} f^+ d\mathrm{P} + \int_{\mathcal{Y}} f^- d\mathrm{P} < +\infty$.

The (measure-theoretic) conditional expectation $\mathrm{E}(f|\mathcal{G})$ of a non-negative $\mathcal{F}$-measurable extended real-valued function f with respect to a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, is any extended real-valued function on $\mathcal{Y}$ that is $\mathcal{G}$-measurable and that satisfies $\int_A f \, d\mathrm{P} = \int_A \mathrm{E}(f|\mathcal{G}) d\mathrm{P}$ or, equivalently, $\int_{\mathcal{Y}} f \mathbf{1}_A d\mathrm{P} = \int_{\mathcal{Y}} \mathrm{E}(f|\mathcal{G}) \mathbf{1}_A d\mathrm{P}$ for all $A \in \mathcal{G}$. If f is $\mathcal{F}$-measurable and non-negative, the conditional expectation $\mathrm{E}(f|\mathcal{G})$ exists and is unique up to a P-null set [1, Section 2.7]. The value of $\mathrm{E}(f|\mathcal{G})$ on a P-null set can be chosen arbitrarily since it will not change the value of the integral $\int_A \mathrm{E}(f|\mathcal{G}) d\mathrm{P} = \int_{\mathcal{Y}} \mathrm{E}(f|\mathcal{G}) \mathbf{1}_A d\mathrm{P}$; see Property M3 below. Similarly to the unconditional case, the measure-theoretic conditional expectation $\mathrm{E}(f|\mathcal{G})$ of a general $\mathcal{F}$-measurable extended real-valued function f with respect to a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is only defined if $\min\{\mathrm{E}(f^+|\mathcal{G}), \mathrm{E}(f^-|\mathcal{G})\} < +\infty$ P-almost surely, and then $\mathrm{E}(f|\mathcal{G}) := \mathrm{E}(f^+|\mathcal{G}) - \mathrm{E}(f^-|\mathcal{G})$ up to a P-null set. If f is P-integrable, such a conditional expectation $\mathrm{E}(f|\mathcal{G})$ always exists; see also [3, Section 34]. We moreover have the following convenient properties.

Lemma 22. *For any probability space* $(\mathcal{Y}, \mathcal{F}, \mathrm{P})$ *and any two extended real-valued functions* f *and* g *that are* $\mathcal{F}$-*measurable, the following properties hold, provided that the considered unconditional or conditional expectations exist.*

M1. $\mathrm{E}(af + b) = a\mathrm{E}(f) + b$ *for all* $a, b \in \mathbb{R}$;

M2. $f \le g \Rightarrow \mathrm{E}(f) \le \mathrm{E}(g)$;

M3. $f = g$ P-almost surely $\Rightarrow \mathrm{E}(f) = \mathrm{E}(g)$;

M4. $\mathrm{E}(f|\mathcal{F}) = f$ P-*almost surely;*

M5. $\mathrm{E}(f|\mathcal{G}^*) = \mathrm{E}(f)$ *where* $\mathcal{G}^* = \{\emptyset, \mathcal{Y}\}$;

M6. $\mathrm{E}(\mathrm{E}(f|\mathcal{G}_2)|\mathcal{G}_1) = \mathrm{E}(f|\mathcal{G}_1)$ P-*almost surely for every two* σ-*algebras* $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$.

M7. $\inf f \le \mathrm{E}(f|\mathcal{G})$ P-*almost surely and* $\mathrm{E}(f|\mathcal{G}) \le \sup f$ P-*almost surely for any* σ-*algebra* $\mathcal{G} \subseteq \mathcal{F}$;

Proof. Let us first show that M5 follows from [1, Section 2.7.4(**E***)]. There, it is stated that $\mathrm{E}(f|\mathcal{G}^*) = \mathrm{E}(f)$ holds P-almost surely. However, since $\mathcal{G}^* = \{\emptyset, \mathcal{Y}\}$ and $\mathrm{E}(f|\mathcal{G}^*)$ is $\mathcal{G}^*$-measurable, we find that $\mathrm{E}(f|\mathcal{G}^*)$ must be constant. So if $\mathrm{E}(f|\mathcal{G}^*) = \mathrm{E}(f)$ holds P-almost surely, it also holds on the entire domain $\mathcal{Y}$.

To prove Property M1, consider [1, Section 2.7.4(**D***)] which in particular states that $\mathrm{E}(af + b|\mathcal{G}^*) = a\mathrm{E}(f|\mathcal{G}^*) + \mathrm{E}(b|\mathcal{G}^*)$ P-almost surely. Since $\mathrm{E}(b|\mathcal{G}^*) = b$ P-almost surely by [1, Section 2.7.4(**A***)], we have that $\mathrm{E}(af + b|\mathcal{G}^*) =$

$a\mathrm{E}(f|\mathscr{G}^*) + b$ P-almost surely, which by M5 implies that $\mathrm{E}(af + b) = a\mathrm{E}(f) + b$ P-almost surely. Since both sides are constants, M1 follows.

Property M2 follows from [1, Section 2.7.4(**B***)], which states that $\mathrm{E}(f|\mathscr{G}^*) \leq \mathrm{E}(g|\mathscr{G}^*)$ P-almost surely if $f \leq g$ P-almost surely. Indeed, if $f \leq g$, then certainly $f \leq g$ P-almost surely and then [1, Section 2.7.4(**B***)] together with M5 implies that $\mathrm{E}(f) \leq \mathrm{E}(g)$ P-almost surely; since both sides are constants, M2 follows.

To prove M3, note that $f = g$ P-almost surely implies that both $f \leq g$ P-almost surely and $f \geq g$ P-almost surely. Then M3 follows by applying the previous reasoning to both of these inequalities. Furthermore, Properties M4 and M6 are taken directly from [1, Section 2.7.4].

To see that property M7 holds, consider [1, Section 2.7.4(**A***)], which in particular says that $\mathrm{E}(\sup f|\mathscr{G}) = \sup f$ P-almost surely. Combining this with mononicity [1, Section 2.7.4(**B***)], and taking into account that the intersection of two P-almost sure events is itself also P-almost sure, we indeed find that $\mathrm{E}(f|\mathscr{G}) \leq \sup f$ P-almost surely. The fact that $\inf f \leq \mathrm{E}(f|\mathscr{G})$ P-almost surely can then easily be obtained using the linearity of $\mathrm{E}(\cdot|\mathscr{G})$ [1, Section 2.7.4(**D***)]. $\square$

Lemma 23. *Consider any probability space* $(\mathscr{Y}, \mathscr{F}, \mathrm{P})$, *any* σ-*algebra* $\mathscr{G} \subseteq \mathscr{F}$ *and any sequence* $\{f_n\}_{n \in \mathbb{N}_0}$ *of extended real-valued functions such that* f_n *is* $\mathscr{F}$-*measurable for all* $n \in \mathbb{N}_0$.

M8. *If* $\{f_n\}_{n \in \mathbb{N}_0}$ *is non-decreasing and there is an* $\mathscr{F}$-*measurable function* f^* *such that* $\mathrm{E}(f^*) > -\infty$ *and* $f_n \geq f^*$ *for all* $n \in \mathbb{N}_0$, *then*

$$\lim_{n \to +\infty} \mathrm{E}(f_n) = \mathrm{E}(f) \text{ where } \lim_{n \to +\infty} f_n = f.$$

M9. *If* $\{f_n\}_{n \in \mathbb{N}_0}$ *is non-increasing and there is an* $\mathscr{F}$-*measurable function* f^* *such that* $\mathrm{E}(f^*) < +\infty$ *and* $f_n \leq f^*$ *for all* $n \in \mathbb{N}_0$, *then*

$$\lim_{n \to +\infty} \mathrm{E}(f_n) = \mathrm{E}(f) \text{ where } \lim_{n \to +\infty} f_n = f.$$

Proof. Both properties can directly be obtained by applying M5 to the properties stated in [1, Theorem 2.7.2] and observing that both sides of the acquired equations will be constants (allowing us to infer that the equations always hold, instead of P-almost surely as is the case for the more general versions in [1, Theorem 2.7.2]). $\square$

A (discrete) *filtration* $\{\mathscr{F}_n\}_{n \in \mathbb{N}_0}$ on a measurable space $(\mathscr{Y}, \mathscr{F})$ is a sequence of increasing σ-algebras in $\mathscr{F}$; so $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \cdots \mathscr{F}$. We will then also use $\mathscr{F}_\infty$ to denote the smallest σ-algebra $\sigma(\cup_{n \in \mathbb{N}_0} \mathscr{F}_n)$ generated by the σ-algebras $\mathscr{F}_n$. We say that $(\mathscr{Y}, \mathscr{F}, \{\mathscr{F}_n\}_{n \in \mathbb{N}_0})$ is a *filtered measurable space* if $(\mathscr{Y}, \mathscr{F})$ is equipped with a filtration $\{\mathscr{F}_n\}_{n \in \mathbb{N}_0}$, and moreover say that $(\mathscr{Y}, \mathscr{F}, \{\mathscr{F}_n\}_{n \in \mathbb{N}_0}, \mathrm{P})$ is a *filtered probability space* if it additionally has a σ-additive measure P on $\mathscr{F}$. A sequence $\{\mathscr{N}_n\}_{n \in \mathbb{N}_0}$ of extended real-valued functions on $\mathscr{Y}$ is called a *measure-theoretic process* in a filtered probability space $(\mathscr{Y}, \mathscr{F}, \{\mathscr{F}_n\}_{n \in \mathbb{N}_0}, \mathrm{P})$ if $\mathscr{N}_n$ is $\mathscr{F}_n$-measurable for all $n \in \mathbb{N}_0$. It is moreover called a *measure-theoretic martingale* if $\mathscr{N}_n$ is real-valued and $\mathrm{E}(\mathscr{N}_{n+1}|\mathscr{F}_n) = \mathscr{N}_n$ P-almost surely for all $n \in \mathbb{N}_0$, where the existence of $\mathrm{E}(\mathscr{N}_{n+1}|\mathscr{F}_n)$ is an implicit condition. The following two results are fundamental in establishing a relation between the measure-theoretic and the game-theoretic framework.

Proposition 24 (Lévy's zero-one law; [2, Theorem 7.4.3])**.** *For any filtered probability space* $(\mathscr{Y}, \mathscr{F}, \{\mathscr{F}_n\}_{n \in \mathbb{N}_0}, \mathrm{P})$ *and any* P*-integrable function* f, *we have that*

$$\lim_{n \to +\infty} \mathrm{E}(f|\mathscr{F}_n) = \mathrm{E}(f|\mathscr{F}_\infty) \text{ P-almost surely}$$

Proposition 25 (Ville's theorem; [4, Proposition 8.14])**.** *For any filtered probability space* $(\mathscr{Y}, \mathscr{F}, \{\mathscr{F}_n\}_{n \in \mathbb{N}_0}, \mathrm{P})$ *and any* $A \in \mathscr{F}_\infty$, *we have that* $\mathrm{P}(A) = 0$ *if and only if there is a non-negative measure-theoretic martingale that converges to* $+\infty$ *on* A.

Appendix B. Proof of Proposition 10

Proposition 10 will be shown to hold by establishing the equivalence between the global measure-theoretic and global game-theoretic (upper) expectation on a particular restricted domain and then applying Theorem 6. Before we relate both global models, however, we first establish the following alternative characterisation for the unique extension $\overline{Q}_s^{\uparrow}$ of the local linear (upper) expectation $\overline{Q}_s$ corresponding to any probability mass function p_s on $\mathscr{X}$:

Lemma 26. *Consider any probability mass function p_s on $\mathscr{X}$ and let $\overline{Q}_s$ be the unique linear expectation on $\mathscr{L}(\mathscr{X})$ corresponding to p_s. Then we have that $\overline{Q}_s^{\uparrow}(f) = \sum_{x \in \mathscr{X}} f(x) p_s(x)$ for all $f \in \overline{\mathscr{L}}(\mathscr{X})$.*

Proof. Let $\overline{E}_s \colon \overline{\mathscr{L}}(\mathscr{X}) \to \overline{\mathbb{R}}$ be defined by $\overline{E}_s(f) := \sum_{x \in \mathscr{X}} f(x) p_s(x)$ for all $f \in \overline{\mathscr{L}}(\mathscr{X})$. We show that $\overline{Q}_s^{\uparrow}$ coincides with $\overline{E}_s$. It is clear that this is the case on the domain $\mathscr{L}(\mathscr{X})$ of all gambles on $\mathscr{X}$; this follows immediately from the fact that $\overline{Q}_s^{\uparrow}$ is an extension of $\overline{Q}_s$ together with the fact that $\overline{E}_s(f) = \sum_{x \in \mathscr{X}} f(x) p_s(x) = \overline{Q}_s(f)$ for all $f \in \mathscr{L}(\mathscr{X})$. To see that both coincide on $\overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$, consider any $f \in \overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$ and observe that

$$\overline{Q}_s^{\uparrow}(f) \overset{\mathrm{C8}}{=} \lim_{c \to +\infty} \overline{Q}_s^{\uparrow}(f^{\wedge c}) = \lim_{c \to +\infty} \sum_{x \in \mathscr{X}} f^{\wedge c}(x) p_s(x) = \sum_{x \in \mathscr{X}} \lim_{c \to +\infty} f^{\wedge c}(x) p_s(x) = \sum_{x \in \mathscr{X}} p_s(x) \lim_{c \to +\infty} f^{\wedge c}(x)$$
$$= \sum_{x \in \mathscr{X}} p_s(x) f(x) = \overline{E}_s(f),$$

where the second equality follows from the fact that $f^{\wedge c} \in \mathscr{L}(\mathscr{X})$ for any $c \in \mathbb{R}$ and, as we have just shown, $\overline{Q}_s^{\uparrow}$ and $\overline{E}_s$ coincide on $\mathscr{L}(\mathscr{X})$; where the third equality follows from the fact that $\mathscr{X}$ is finite and that, for all $x \in \mathscr{X}$, $f^{\wedge c}(x) p_s(x)$ is real and non-decreasing in c; where the fourth equality follows from the fact that $p_s(x) \geq 0$ for all $x \in \mathscr{X}$ and our convention that $0 \cdot (+\infty) = 0$; and where the last equality follows from the definition of $\overline{E}_s$. In an analogous way, we show that $\overline{Q}_s^{\uparrow}$ and $\overline{E}_s$ coincide on their entire domain $\overline{\mathscr{L}}(\mathscr{X})$: consider any $f \in \overline{\mathscr{L}}(\mathscr{X})$ and note that

$$\overline{Q}_s^{\uparrow}(f) \overset{\mathrm{C9}}{=} \lim_{c \to -\infty} \overline{Q}_s^{\uparrow}(f^{\vee c}) = \lim_{c \to -\infty} \sum_{x \in \mathscr{X}} f^{\vee c}(x) p_s(x) = \sum_{x \in \mathscr{X}} \lim_{c \to -\infty} f^{\vee c}(x) p_s(x) = \sum_{x \in \mathscr{X}} p_s(x) \lim_{c \to -\infty} f^{\vee c}(x)$$
$$= \sum_{x \in \mathscr{X}} p_s(x) f(x) = \overline{E}_s(f),$$

where the second equality follows from the fact that $f^{\vee c} \in \overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$ for any $c \in \mathbb{R}$ and, as we have just shown, $\overline{Q}_s^{\uparrow}$ and $\overline{E}_s$ coincide on $\overline{\mathscr{L}}_{\mathrm{b}}(\mathscr{X})$; where the third equality follows from the fact that $\mathscr{X}$ is finite, our convention that $+\infty - \infty = -\infty + \infty = +\infty$ and that, for all $x \in \mathscr{X}$, $f^{\vee c}(x) p_s(x)$ is in $\mathbb{R} \cup \{+\infty\}$ and non-decreasing in c; the fourth equality follows from the fact that $p_s(x) \geq 0$ for all $x \in \mathscr{X}$ and our convention that $0 \cdot (-\infty) = 0$; and the last equality follows from the definition of $\overline{E}_s$. $\qquad \square$

Henceforth, for any precise probability tree p and any $s \in \mathscr{X}^*$, we will always let $\overline{Q}_s^{\uparrow}$ be this unique extension of the linear expectation $\overline{Q}_s$ corresponding to the probability mass function $p(\cdot|s)$. We will then say that a map $\mathscr{M} \colon \mathscr{X}^* \to \overline{\mathbb{R}}$ is a game-theoretic supermartingale *with respect to a precise probability tree p* if $\mathscr{M}$ is a game-theoretic supermartingale with respect to the local upper expectations $\overline{Q}_s^{\uparrow}$.

Furthermore, for any precise probability tree $p \colon s \in \mathscr{X}^* \mapsto p(\cdot|s)$, we let $\mathrm{P} \colon \mathscr{F} \times \mathscr{X}^* \to \mathbb{R}$ be the corresponding conditional probability measure as discussed in Section 9, where $\mathscr{F}$ is the σ-algebra generated by all cylinder events. Recall that, for any $s \in \mathscr{X}^*$, the map $\mathrm{P}(\cdot|s) = \mathrm{P}_{|s}$ is then a probability measure on $\mathscr{F}$. This allows us to apply the concepts and results in Appendix A here, by considering the probability space $(\Omega, \mathscr{F}, \mathrm{P}_{|s})$, for any $s \in \mathscr{X}^*$. For notational convenience, we will let $\mathrm{E}_{|s}(f)$ be the Lebesgue integral $\int_{\Omega} f \, \mathrm{dP}_{|s}$ with respect to the measure $\mathrm{P}_{|s}$ for all $\mathscr{F}$-measurable $f \in \overline{\mathbb{V}}$ and all $s \in \mathscr{X}^*$ such that $\int_{\Omega} f \, \mathrm{dP}_{|s}$ exists. So $\mathrm{E}_{|s}(f)$ is an alternative notation for $\mathrm{E}_{\mathrm{meas},p}(f|s)$. We introduce this notation because it allows us to write $\mathrm{E}_{|s}(f|\mathscr{G})$ to denote a $\mathscr{G}$-measurable function representing the measure-theoretic expectation of f conditional on a σ-algebra $\mathscr{G} \subseteq \mathscr{F}$, as defined in Appendix A. Furthermore, we equip the measurable space $(\Omega, \mathscr{F})$ with the filtration $\{\mathscr{F}_n\}_{n \in \mathbb{N}_0}$ where, for any $n \in \mathbb{N}_0$, $\mathscr{F}_n$ is the σ-algebra generated by all cylinder events $\Gamma(x_{1:n})$ where $x_{1:n} \in \mathscr{X}^n$. Note that, for any $n \in \mathbb{N}_0$, an $\mathscr{F}_n$-measurable function is

then an n-measurable variable because the cylinder events $\Gamma(x_{1:n})$ form the atoms of $\mathscr{F}_n$ and $\mathscr{X}$ is finite. Hence, any measure-theoretic process $\{\mathscr{N}_n\}_{n \in \mathbb{N}_0}$ in $(\Omega, \mathscr{F}, \{\mathscr{F}_n\}_{n \in \mathbb{N}_0})$ is a sequence of n-measurable variables. This allows us to write $\mathscr{N}_n(x_{1:n})$ for any $n \in \mathbb{N}_0$ and any $x_{1:n} \in \mathscr{X}^n$ to mean the constant value of $\mathscr{N}_n$ on all $\omega \in \Gamma(x_{1:n})$.

Lemma 27. *Consider any $x_{1:n} \in \mathscr{X}^*$, any precise probability tree p and let P be the corresponding conditional probability measure. Then, for any non-negative measure-theoretic martingale $\{\mathscr{N}_i\}_{i \in \mathbb{N}_0}$ in the filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_m\}_{m \in \mathbb{N}_0}, \mathrm{P}_{|x_{1:n}})$, there is a non-negative game-theoretic (super)martingale[11] $\mathscr{M}$ with respect to the tree p such that $\liminf \mathscr{M} \geq_{x_{1:n}} \liminf_{i \to +\infty} \mathscr{N}_i$ and moreover $\mathscr{M}(x_{1:n}) = \mathscr{N}_0(\square)$.*

Proof. Fix any non-negative measure-theoretic martingale $\{\mathscr{N}_i\}_{i \in \mathbb{N}_0}$ and let $\mathscr{M} : \mathscr{X}^* \to \overline{\mathbb{R}}$ be defined by

$$
\mathscr{M}(z_{1:m}) := \begin{cases} \mathscr{N}_m(z_{1:m}) & \text{if } n \leq m,\ z_{1:n} = x_{1:n} \text{ and } \mathrm{P}_{|x_{1:n}}(z_{1:m}) > 0; \\ +\infty & \text{if } n \leq m,\ z_{1:n} = x_{1:n} \text{ and } \mathrm{P}_{|x_{1:n}}(z_{1:m}) = 0; \quad \text{for all } z_{1:m} \in \mathscr{X}^*. \\ \mathscr{N}_n(x_{1:n}) & \text{otherwise,} \end{cases}
$$

We show that $\mathscr{M}$ is a non-negative game-theoretic supermartingale such that $\mathscr{M}(x_{1:n}) = \mathscr{N}_0(\square)$ and $\liminf \mathscr{M} \geq_{x_{1:n}} \liminf_{i \to +\infty} \mathscr{N}_i$.

Let us first show that $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}^{\uparrow}_{z_{1:m}}(\mathscr{M}(z_{1:m} \cdot))$ for all $z_{1:m} \in \mathscr{X}^*$ and therefore, that $\mathscr{M}$ is a game-theoretic supermartingale. Recall that, because $\{\mathscr{N}_i\}_{i \in \mathbb{N}_0}$ is a non-negative measure-theoretic martingale, we have that $\{\mathscr{N}_i\}_{i \in \mathbb{N}_0}$ is a sequence of non-negative real i-measurable variables such that $\mathrm{E}_{|x_{1:n}}(\mathscr{N}_{i+1}|\mathscr{F}_i) = \mathscr{N}_i$ $\mathrm{P}_{|x_{1:n}}$-almost surely for all $i \in \mathbb{N}_0$ (note that the considered expectations exist because $\{\mathscr{N}_i\}_{i \in \mathbb{N}_0}$ is non-negative). For all $i \in \mathbb{N}_0$ and all $A \in \mathscr{F}_i$, since $\mathrm{E}_{|x_{1:n}}(\mathscr{N}_{i+1}|\mathscr{F}_i)$ and $\mathscr{N}_i$ only differ on a $\mathrm{P}_{|x_{1:n}}$-null set, we have that $\int_A \mathrm{E}_{|x_{1:n}}(\mathscr{N}_{i+1}|\mathscr{F}_i)\mathrm{dP}_{|x_{1:n}} = \int_A \mathscr{N}_i \mathrm{dP}_{|x_{1:n}}$ because of M3. In particular, this implies that

$$
\int_{\Gamma(z_{1:m})} \mathrm{E}_{|x_{1:n}}(\mathscr{N}_{m+1}|\mathscr{F}_m)\mathrm{dP}_{|x_{1:n}} = \int_{\Gamma(z_{1:m})} \mathscr{N}_m \mathrm{dP}_{|x_{1:n}} \text{ for any } m \in \mathbb{N}_0 \text{ and any } z_{1:m} \in \mathscr{X}^m.
$$

Moreover, since $\Gamma(z_{1:m}) \in \mathscr{F}_m$, it follows from the definition of the measure-theoretic conditional expectation that $\int_{\Gamma(z_{1:m})} \mathscr{N}_{m+1}\mathrm{dP}_{|x_{1:n}} = \int_{\Gamma(z_{1:m})} \mathrm{E}_{|x_{1:n}}(\mathscr{N}_{m+1}|\mathscr{F}_m)\mathrm{dP}_{|x_{1:n}}$. Combining both equalities, we find that

$$
\int_{\Gamma(z_{1:m})} \mathscr{N}_{m+1}\mathrm{dP}_{|x_{1:n}} = \int_{\Gamma(z_{1:m})} \mathscr{N}_m \mathrm{dP}_{|x_{1:n}} = \mathscr{N}_m(z_{1:m}) \int_{\Gamma(z_{1:m})} \mathrm{dP}_{|x_{1:n}} = \mathscr{N}_m(z_{1:m})\mathrm{P}_{|x_{1:n}}(z_{1:m}),
$$

where the second equality follows from M1 and the fact that $\mathscr{N}_m$ is constant and real on the cylinder event $\Gamma(z_{1:m})$. Moreover, because $\mathscr{N}_{m+1}$ is $(m+1)$-measurable and real-valued, the term on the left hand side of the equation above reduces to the finite sum $\sum_{z_{m+1} \in \mathscr{X}} \mathscr{N}_{m+1}(z_{1:m+1})\mathrm{P}_{|x_{1:n}}(z_{1:m+1})$, allowing us to conclude that

$$
\mathscr{N}_m(z_{1:m})\mathrm{P}_{|x_{1:n}}(z_{1:m}) = \sum_{z_{m+1} \in \mathscr{X}} \mathscr{N}_{m+1}(z_{1:m+1})\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) \text{ for any } m \in \mathbb{N}_0 \text{ and any } z_{1:m} \in \mathscr{X}^m. \tag{B.1}
$$

Consider now first any $m \geq n$ and any $z_{1:m} \in \mathscr{X}^m$ such that $z_{1:n} = x_{1:n}$ and $\mathrm{P}_{|x_{1:n}}(z_{1:m}) > 0$. It then follows from the definition of $\mathscr{M}$ that

$$
\begin{aligned}
\mathscr{M}(z_{1:m})\mathrm{P}_{|x_{1:n}}(z_{1:m}) = \mathscr{N}_m(z_{1:m})\mathrm{P}_{|x_{1:n}}(z_{1:m}) &\overset{(\mathrm{B.1})}{=} \sum_{z_{m+1} \in \mathscr{X}} \mathscr{N}_{m+1}(z_{1:m+1})\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) \\
&= \sum_{z_{m+1} \in \mathscr{X}} \mathscr{M}(z_{1:m+1})\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) \\
&= \sum_{z_{m+1} \in \mathscr{X}} \mathscr{M}(z_{1:m+1})\mathrm{P}_{|x_{1:n}}(z_{1:m})p(z_{m+1}|z_{1:m}) \\
&= \mathrm{P}_{|x_{1:n}}(z_{1:m}) \sum_{z_{m+1} \in \mathscr{X}} \mathscr{M}(z_{1:m+1})p(z_{m+1}|z_{1:m}) = \mathrm{P}_{|x_{1:n}}(z_{1:m})\overline{\mathrm{Q}}^{\uparrow}_{z_{1:m}}(\mathscr{M}(z_{1:m} \cdot)),
\end{aligned}
$$

---

[11]As shown in the proof, the inequality $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}^{\uparrow}_{z_{1:m}}(\mathscr{M}(z_{1:m} \cdot))$ is actually an equality, for all $z_{1:m} \in \mathscr{X}^*$. Combining this with the non-negativity of $\mathscr{M}$ and the characterisation of the upper expectations $\overline{\mathrm{Q}}^{\uparrow}_{z_{1:m}}$ that we established in Lemma 26, it can then be deduced that $-\mathscr{M}$ is also a game-theoretic supermartingale. In this case, we can therefore actually call $\mathscr{M}$ a *game-theoretic martingale*.

where the third equality follows because $\mathscr{N}_{m+1}(z_{1:m+1})$ only differs from $\mathscr{M}(z_{1:m+1})$ if $\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) = 0$ [because $n \leq m+1$ and $z_{1:n} = x_{1:n}$] and our convention that $0 \cdot (+\infty) = 0$, where the fourth equality follows from Equation (6) together with the fact that $n < m+1$ and $z_{1:n} = x_{1:n}$, and where the last equality follows from the expression for $\overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}$ that we established in Lemma 26. Recall that $\mathrm{P}_{|x_{1:n}}(z_{1:m}) > 0$, so we can devide both sides by $\mathrm{P}_{|x_{1:n}}(z_{1:m})$ to obtain that $\mathscr{M}(z_{1:m}) = \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$. As a consequence, the condition that $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$ is satisfied for all $z_{1:m} \in \mathscr{X}^*$ such that $m \geq n$, $z_{1:n} = x_{1:n}$ and $\mathrm{P}_{|x_{1:n}}(z_{1:m}) > 0$.

Second, consider any $z_{1:m} \in \mathscr{X}^*$ such that $m \geq n$, $z_{1:n} = x_{1:n}$ and $\mathrm{P}_{|x_{1:n}}(z_{1:m}) = 0$. Then we also have that $\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) = 0$ for any $z_{m+1} \in \mathscr{X}$ because $\Gamma(z_{1:m+1}) \subset \Gamma(z_{1:m})$, and it therefore follows from the definition of $\mathscr{M}$ that $\mathscr{M}(z_{1:m}) = \mathscr{M}(z_{1:m+1}) = +\infty$ for all $z_{m+1} \in \mathscr{X}$. Hence, using Lemma 26, we find that

$$\overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot)) = \sum_{z_{m+1} \in \mathscr{X}} \mathscr{M}(z_{1:m+1}) p(z_{m+1}|z_{1:m}) = +\infty = \mathscr{M}(z_{1:m}),$$

which establishes that indeed $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$. So we have that $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$ for all $z_{1:m} \in \mathscr{X}^*$ such that $m \geq n$ and $z_{1:n} = x_{1:n}$.

Thirdly, consider any other situation $z_{1:m} \in \mathscr{X}^*$. That is, consider any $z_{1:m} \in \mathscr{X}^*$ such that $n > m$ or $z_{1:n} \neq x_{1:n}$. It then follows from the definition of $\mathscr{M}$ that $\mathscr{M}(z_{1:m}) = \mathscr{N}_n(x_{1:n})$ and, as we will now show, that also $\mathscr{M}(z_{1:m+1}) = \mathscr{N}_n(x_{1:n})$ for all $z_{m+1} \in \mathscr{X}$. Our assumption about $z_{1:m}$ tells us that there are two cases: either $n \leq m$ and $z_{1:n} \neq x_{1:n}$, either $n > m$. If $n \leq m$ and $z_{1:n} \neq x_{1:n}$, then also $n \leq m+1$ and therefore, due to the definition of $\mathscr{M}$, we find that $\mathscr{M}(z_{1:m+1}) = \mathscr{N}_n(x_{1:n})$ for all $z_{m+1} \in \mathscr{X}$. If $n > m$, then either $n > m+1$ or $n = m+1$. If $n > m+1$ or, equivalently, $n \not\leq m+1$, then it is once more clear from the definition of $\mathscr{M}$ that $\mathscr{M}(z_{1:m+1}) = \mathscr{N}_n(x_{1:n})$ for all $z_{m+1} \in \mathscr{X}$. If $n = m+1$, then for any $z_{m+1} \in \mathscr{X}$, we either have that $z_{1:m+1} = x_{1:n}$ or that $z_{1:m+1} \neq x_{1:n}$. In the latter case, it is clear, again due to the definition of $\mathscr{M}$, that $\mathscr{M}(z_{1:m+1}) = \mathscr{N}_n(x_{1:n})$. Otherwise, so if $z_{1:m+1} = x_{1:n}$, the definition of $\mathscr{M}$ also implies that $\mathscr{M}(z_{1:m+1}) = \mathscr{M}(x_{1:n}) = \mathscr{N}_n(x_{1:n})$ because $\mathrm{P}_{|x_{1:n}}(x_{1:n}) = 1 > 0$ according to Equation (6). We conclude that, indeed, $\mathscr{M}(z_{1:m}) = \mathscr{N}_n(x_{1:n})$ and $\mathscr{M}(z_{1:m+1}) = \mathscr{N}_n(x_{1:n})$ for all $z_{m+1} \in \mathscr{X}$. It now follows trivially from Lemma 26 that $\mathscr{M}(z_{1:m}) = \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$, and therefore definitely that $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$.

So we have proved that $\mathscr{M}(z_{1:m}) \geq \overline{\mathrm{Q}}_{z_{1:m}}^{\uparrow}(\mathscr{M}(z_{1:m}\cdot))$ for all $z_{1:m} \in \mathscr{X}^*$ and therefore, that $\mathscr{M}$ is indeed a game-theoretic supermartingale with respect to the tree p. Moreover, $\mathscr{M}$ is non-negative because $\{\mathscr{N}_i\}_{i \in \mathbb{N}_0}$ is non-negative. That $\liminf_{i \to +\infty} \mathscr{M}(\omega^i) \geq \liminf_{i \to +\infty} \mathscr{N}_i(\omega)$ holds for all $\omega \in \Gamma(x_{1:n})$, follows immediately from the fact that $\mathscr{M}(\omega^i) \geq \mathscr{N}_i(\omega^i) = \mathscr{N}_i(\omega)$ for any $\omega \in \Gamma(x_{1:n})$ and any $i \geq n$, where the inequality is simply implied by the definition of $\mathscr{M}$ together with the fact that $\omega^n = x_{1:n}$ and $i \geq n$. It still remains to show that $\mathscr{M}(x_{1:n}) = \mathscr{N}_0(\square)$.

We only need to prove that $\mathscr{N}_0(\square) = \mathscr{N}_n(x_{1:n})$, since the desired equality then trivially follows from the definition of $\mathscr{M}$ and the fact that $\mathrm{P}_{|x_{1:n}}(x_{1:n}) = 1 > 0$ due to Equation (6). To do so, consider any $z_{1:m} \in \mathscr{X}^*$ such that $m < n$ and $z_{1:m} = x_{1:m}$. Note that, by Equation (6), we then have that $\mathrm{P}_{|x_{1:n}}(z_{1:m}) = 1$. Moreover, it should also be clear from Equation (6) that $\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) = 1$ if $z_{m+1} = x_{m+1}$ [because then $z_{1:m+1} = x_{1:m+1}$ and $m+1 \leq n$] and that $\mathrm{P}_{|x_{1:n}}(z_{1:m+1}) = 0$ otherwise [because then $z_{1:m+1} \neq x_{1:m+1}$]. Hence, plugging this back into (B.1), we find that $\mathscr{N}_m(z_{1:m}) = \mathscr{N}_{m+1}(z_{1:m} x_{m+1})$. Since this holds for any $z_{1:m} \in \mathscr{X}^*$ such that $m < n$ and $z_{1:m} = x_{1:m}$, it follows that $\mathscr{N}_m(x_{1:m}) = \mathscr{N}_{m+1}(x_{1:m+1})$ for all $m < n$. This clearly implies that $\mathscr{N}_0(\square) = \mathscr{N}_1(x_1) = \cdots = \mathscr{N}_n(x_{1:n})$. $\qquad\square$

In the following proof, we will write, for any two $f, g \in \overline{\mathbb{V}}$, any $s \in \mathscr{X}^*$ and any (unconditional) probability measure P' on $\mathscr{F}$, that $f =_s g$ P-almost surely—and similarly for $\geq_s$ and $\leq_s$—if the event $\{\omega \in \Gamma(s): f(\omega) \neq g(\omega)\}$ is P-null. Note that then $f =_s g$ P-almost surely if $f = g$ P-almost surely.

Proof of Proposition 10. Fix any $\mathscr{F}$-measurable $f' \in \mathbb{V}$ and any $x_{1:n} \in \mathscr{X}^*$. It suffices to show that $\mathrm{E}_{\mathrm{meas},p}(f'|x_{1:n}) = \overline{\mathrm{E}}_{\mathrm{V}}(f'|x_{1:n})$; the desired equality is then automatically implied by Theorem 6. First observe that, because f' is bounded and $\mathscr{F}$-measurable, $\mathrm{E}_{\mathrm{meas},p}(f'|x_{1:n}) = \mathrm{E}_{|x_{1:n}}(f')$ exists. We will now prove that $\mathrm{E}_{\mathrm{meas},p}(f|x_{1:n}) = \overline{\mathrm{E}}_{\mathrm{V}}(f|x_{1:n})$ for the non-negative $\mathscr{F}$-measurable gamble $f := f' - \inf f'$ (the variable f is indeed a gamble because f' is a gamble and therefore $\inf f' \in \mathbb{R}$), which then implies that $\mathrm{E}_{\mathrm{meas},p}(f'|x_{1:n}) = \overline{\mathrm{E}}_{\mathrm{V}}(f'|x_{1:n})$ because $\mathrm{E}_{\mathrm{meas},p}(\cdot|x_{1:n})$ and $\overline{\mathrm{E}}_{\mathrm{V}}$ both satisfy the constant additivity property; see M1 and V4.

We first show that $\overline{\mathrm{E}}_{\mathrm{V}}(f|x_{1:n}) \leq \mathrm{E}_{\mathrm{meas},p}(f|x_{1:n})$. To do so, we will prove that there is some $c \in \mathbb{R}$ such that, for all $\epsilon > 0$, there is a bounded below game-theoretic supermartingale $\mathscr{M}_\epsilon$ with respect to the tree p such that $\mathscr{M}_\epsilon(x_{1:n}) = \mathrm{E}_{\mathrm{meas},p}(f|x_{1:n}) + \epsilon c$ and $\liminf \mathscr{M}_\epsilon \geq_{x_{1:n}} f$. Indeed, the desired inequality then follows immediately from the definition of $\overline{\mathrm{E}}_{\mathrm{V}}$.

Consider the filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_m\}_{m\in\mathbb{N}_0}, P_{|x_{1:n}})$ and the corresponding measure-theoretic expectation $E_{|x_{1:n}} = E_{\mathrm{meas},p}(\cdot|x_{1:n})$. Since f is bounded and $\mathscr{F}$-measurable, it is surely $P_{|x_{1:n}}$-integrable (the Lebesgue integral of a bounded $\mathscr{F}$-measurable function is real-valued), and therefore, by Proposition 24 in Appendix A, we have that

$$\lim_{m\to+\infty} E_{|x_{1:n}}(f|\mathscr{F}_m) = E_{|x_{1:n}}(f|\mathscr{F}_\infty) \quad P_{|x_{1:n}}\text{-almost surely.}$$

Note that $\mathscr{F}_\infty = \sigma(\cup_{m\in\mathbb{N}_0}\mathscr{F}_m) = \mathscr{F}$ is the smallest σ-algebra generated by all cylinder events, which, by M4, implies that $E_{|x_{1:n}}(f|\mathscr{F}_\infty) = f$, $P_{|x_{1:n}}$-almost surely. Hence, since the intersection of two $P_{|x_{1:n}}$-almost sure events is itself also $P_{|x_{1:n}}$-almost sure, we have that $\lim_{m\to+\infty} E_{|x_{1:n}}(f|\mathscr{F}_m) = f$ $P_{|x_{1:n}}$-almost surely. Moreover, note that the conditional expectations $E_{|x_{1:n}}(f|\mathscr{F}_m)$ can be chosen in such a way that the sequence $\{E_{|x_{1:n}}(f|\mathscr{F}_m)\}_{m\in\mathbb{N}_0}$ is a non-negative measure-theoretic martingale in the filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_m\}_{m\in\mathbb{N}_0}, P_{|x_{1:n}})$. Indeed, each $E_{|x_{1:n}}(f|\mathscr{F}_m)$ is real-valued and non-negative $P_{|x_{1:n}}$-almost surely because of M7 and the fact that f is bounded and non-negative. So, since the value of the measure-theoretic conditional expectation can be chosen arbitrarily on a null set, $E_{|x_{1:n}}(f|\mathscr{F}_m)$ can be chosen such that it is real-valued and non-negative everywhere. Moreover, because of M6, we have that $E_{|x_{1:n}}\big(E_{|x_{1:n}}(f|\mathscr{F}_{m+1})|\mathscr{F}_m\big) = E_{|x_{1:n}}(f|\mathscr{F}_m)$ $P_{|x_{1:n}}$-almost surely for all $m\in\mathbb{N}_0$, where $E_{|x_{1:n}}\big(E_{|x_{1:n}}(f|\mathscr{F}_{m+1})|\mathscr{F}_m\big)$ exists because $E_{|x_{1:n}}(f|\mathscr{F}_{m+1})$ is $\mathscr{F}$-measurable and non-negative. So let $\{E_{|x_{1:n}}(f|\mathscr{F}_m)\}_{m\in\mathbb{N}_0}$ be a version of the conditional expectations that forms a non-negative measure-theoretic martingale. We can then use Lemma 27 to infer the existence of a non-negative game-theoretic supermartingale $\mathscr{M}_0$ with respect to the tree p such that $\liminf \mathscr{M}_0 \geq_{x_{1:n}} \liminf_{m\to+\infty} E_{|x_{1:n}}(f|\mathscr{F}_m)$ and $\mathscr{M}_0(x_{1:n}) = E_{|x_{1:n}}(f|\mathscr{F}_0)(\square)$. Then, because $\lim_{m\to+\infty} E_{|x_{1:n}}(f|\mathscr{F}_m) = f$ $P_{|x_{1:n}}$-almost surely and therefore also $\lim_{m\to+\infty} E_{|x_{1:n}}(f|\mathscr{F}_m) =_{x_{1:n}} f$ $P_{|x_{1:n}}$-almost surely, we have that $\liminf \mathscr{M}_0 \geq_{x_{1:n}} f$ $P_{|x_{1:n}}$-almost surely. Moreover, we have that $\mathscr{M}_0(x_{1:n}) = E_{|x_{1:n}}(f|\mathscr{F}_0)(\square) = E_{|x_{1:n}}(f)$, due to the fact that $\mathscr{F}_0 = \{\emptyset, \Omega\}$ and property M5. So, we can conclude that $\mathscr{M}_0$ is a non-negative game-theoretic supermartingale with respect to p such that $\mathscr{M}_0(x_{1:n}) = E_{|x_{1:n}}(f) = E_{\mathrm{meas},p}(f|x_{1:n})$ and $\liminf \mathscr{M}_0 \geq_{x_{1:n}} f$ $P_{|x_{1:n}}$-almost surely.

As our final step towards obtaining $\mathscr{M}_\epsilon$, consider Proposition 25 and note that it ensures that there is a non-negative measure-theoretic supermartingale $\{\mathscr{N}_m\}_{m\in\mathbb{N}_0}$ in $(\Omega, \mathscr{F}, \{\mathscr{F}_m\}_{m\in\mathbb{N}_0}, P_{|x_{1:n}})$ that converges to $+\infty$ on all paths $\omega\in\Gamma(x_{1:n})$ such that $\liminf\mathscr{M}_0(\omega) < f(\omega)$. Indeed, the set of all such paths ω has probability zero because $\liminf\mathscr{M}_0 \geq_{x_{1:n}} f$ $P_{|x_{1:n}}$-almost surely. Let $c \coloneqq \mathscr{N}_0(\square)$, which is real-valued because $\{\mathscr{N}_m\}_{m\in\mathbb{N}_0}$ is a measure-theoretic supermartingale. By Lemma 27, we find that there is a non-negative game-theoretic supermartingale $\mathscr{M}'$ with respect to p such that $\liminf\mathscr{M}' \geq_{x_{1:n}} \liminf_{m\to+\infty}\mathscr{N}_m$ and $\mathscr{M}'(x_{1:n}) = \mathscr{N}_0(\square) = c$. Since $\{\mathscr{N}_m\}_{m\in\mathbb{N}_0}$ converges to $+\infty$ on all paths $\omega\in\Gamma(x_{1:n})$ such that $\liminf\mathscr{M}_0(\omega) < f(\omega)$, it follows that $\mathscr{M}'$ also converges to $+\infty$ on all such paths ω. Consider now any $\epsilon > 0$ and let $\mathscr{M}_\epsilon$ be the process defined by $\mathscr{M}_\epsilon(s) \coloneqq \mathscr{M}_0(s) + \epsilon\mathscr{M}'(s)$ for all $s\in\mathscr{X}^*$. Then $\mathscr{M}_\epsilon$ is clearly non-negative—and therefore bounded below—and it is a game-theoretic supermartingale because of [20, Lemma 12].[12] Furthermore, note that $\liminf\mathscr{M}_\epsilon(\omega) \geq f(\omega)$ for all $\omega\in\Gamma(x_{1:n})$. Indeed, if $\liminf\mathscr{M}_0(\omega) \geq f(\omega)$ for some $\omega\in\Gamma(x_{1:n})$, then also $\liminf\mathscr{M}_\epsilon(\omega) \geq f(\omega)$ because ϵ and $\mathscr{M}'$ are non-negative. If $\liminf\mathscr{M}_0(\omega) < f(\omega)$ for some $\omega\in\Gamma(x_{1:n})$, then $\mathscr{M}'$, and therefore also $\epsilon\mathscr{M}'$, converges to $+\infty$, which, together with the non-negativity of $\mathscr{M}_0$, implies that $\mathscr{M}_\epsilon$ converges to $+\infty$ on ω. Hence, also in this case, we have that $\liminf\mathscr{M}_\epsilon(\omega) \geq f(\omega)$ so we can conclude that $\liminf\mathscr{M}_\epsilon \geq_{x_{1:n}} f$. Moreover, recall that $\mathscr{M}'(x_{1:n}) = c\in\mathbb{R}$ and that $\mathscr{M}_0(x_{1:n}) = E_{|x_{1:n}}(f)$, so we have that $\mathscr{M}_\epsilon(x_{1:n}) = \mathscr{M}_0(x_{1:n}) + \epsilon\mathscr{M}'(x_{1:n}) = E_{|x_{1:n}}(f) + \epsilon c = E_{\mathrm{meas},p}(f|x_{1:n}) + \epsilon c$. Hence, $\mathscr{M}_\epsilon$ satisfies all the desired conditions and we conclude that indeed $\overline{E}_V(f|x_{1:n}) \leq E_{\mathrm{meas},p}(f|x_{1:n})$.

Then we are left to show the remaining inequality $\overline{E}_V(f|x_{1:n}) \geq E_{\mathrm{meas},p}(f|x_{1:n})$. However, this can be easily deduced from the already obtained inequality and the self-conjugacy of $E_{\mathrm{meas},p}$. Indeed, $E_{|x_{1:n}}(-f)$ exists because $-f$ is $\mathscr{F}$-measurable and bounded, so we can apply M1 to find that $E_{\mathrm{meas},p}(f|x_{1:n}) = E_{|x_{1:n}}(f) = -E_{|x_{1:n}}(-f) = -E_{\mathrm{meas},p}(-f|x_{1:n})$. Since we have already shown that $\overline{E}_V(g|x_{1:n}) \leq E_{\mathrm{meas},p}(g|x_{1:n})$ for all $\mathscr{F}$-measurable $g\in\mathbb{V}$, we have in particular that $\overline{E}_V(-f|x_{1:n}) \leq E_{\mathrm{meas},p}(-f|x_{1:n})$, which implies that $E_{\mathrm{meas},p}(f|x_{1:n}) = -E_{\mathrm{meas},p}(-f|x_{1:n}) \leq -\overline{E}_V(-f|x_{1:n}) = \underline{E}_V(f|x_{1:n}) \leq \overline{E}_V(f|x_{1:n})$, where the last inequality follows from V1 and the fact that $\overline{E}_A = \overline{E}_V$—and therefore also $\underline{E}_A = \underline{E}_V$—because of Theorem 6. □

---

[12] Alternatively, instead of using [20, Lemma 12], one could also easily deduce this using the alternative expression for the local models $\overline{Q}_s^\dagger$ that we established in Lemma 26.