Proceedings of Machine Learning Research 147:90–99, 2021 ISIPTA 2021

Average Behaviour of Imprecise Markov Chains: A Single Pointwise Ergodic Theorem for Six Different Models

Jasper De Bock JASPER.DEBOCK@UGENT.BE
Natan T'Joens NATAN.TJOENS@UGENT.BE
Foundations Lab for Imprecise Probabilities, ELIS, Ghent University, Belgium

Abstract

We study the average behaviour of imprecise Markov chains; a generalised type of Markov chain where local probabilities are partially specified, and where structural assumptions such as Markovianity are weakened. In particular, we prove a pointwise ergodic theorem that provides (strictly) almost sure bounds on the long term average of any real function of the state of such an imprecise Markov chain. Compared to an earlier ergodic theorem by De Cooman et al. (2006), our result requires weaker conditions, provides tighter bounds, and applies to six different types of models.

Keywords: imprecise Markov chain, averages, ergodic theorem, weak ergodicity, strictly almost surely, lower and upper expectation

1. Introduction

A Markov chain [5, 10] is a popular and simple type of probabilistic model for describing the uncertain dynamics of a system as it evolves in discrete time steps. If a Markov chain is ergodic, it has the property that the expected value $\mathrm{E}(f(X_i)|X_1 = x)$ of a function $f(X_i)$ that depends on the system's state X_i at time i converges to a limit $\mathrm{E}_\infty(f) := \lim_{i\to+\infty} \mathrm{E}(f(X_i)|X_1 = x)$ that is the same for every value x of the initial state X_1 that we might start in. The resulting limit expectation E_∞ then corresponds to what is known as the limit distribution of a Markov chain.

The importance of this limit distribution and its associated expectation E_∞ is not so much that it describes the uncertainty about X_i for very large i, but rather that it can be used to characterise the average behaviour of X_i over a large time period. In particular, as follows from the so-called pointwise ergodic theorem, the time average $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ of f will almost surely converge to $\mathrm{E}_\infty(f)$ as the time horizon n receeds to infinity, where 'almost surely' means that the probability that it does not happen is zero.

Similar observations can be made for imprecise Markov chains [2, 3, 6, 11, 13, 14]. Loosely speaking, these are Markov chains of which the transition probabilities that parametrise them are partially specified, in the sense that they are only known to belong to some set of probabilities. Since the probabilities are not exactly known, neither are the expectations that are derived from them. A typical inference therefore consists in finding tight bounds on these expectations, called lower and upper expectations.

By analogy with the precise case, an imprecise Markov chain is called ergodic if it has a limit upper (and lower) expectation $\overline{\mathrm{E}}_\infty$ (and $\underline{\mathrm{E}}_\infty$), meaning that $\overline{\mathrm{E}}(f(X_i)|X_1 = x)$ converges to a limit $\overline{\mathrm{E}}_\infty(f)$ that does not depend on x, and similarly for $\underline{\mathrm{E}}_\infty$. Remarkably, ergodic imprecise Markov chains also satisfy a pointwise ergodic theorem [3, Theorem 32]. The time average $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ may not converge now, but will strictly almost surely be eventually contained in the interval $[\underline{\mathrm{E}}_\infty(f), \overline{\mathrm{E}}_\infty(f)]$, in the sense that $\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(X_i)$ and $\liminf_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(X_i)$ both belong to $[\underline{\mathrm{E}}_\infty(f), \overline{\mathrm{E}}_\infty(f)]$. We explain the meaning of 'strictly almost surely' in Section 3, but for now, it suffices to know that for a strictly almost sure event, the upper probability that it does not happen is zero.

We here improve upon this imprecise pointwise ergodic theorem in three ways. First, our result only requires weak ergodicity [13], meaning that $\overline{\mathrm{E}}(\frac{1}{n}\sum_{i=1}^{n} f(X_i)|X_1 = x)$ converges to a limit $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ that does not depend on x, and similarly for $\underline{\mathrm{E}}_{\mathrm{av},\infty}$. This property is implied by ergodicity—in fact, it is strictly weaker. Second, we prove our result for the interval $[\underline{\mathrm{E}}_{\mathrm{av},\infty}(f), \overline{\mathrm{E}}_{\mathrm{av},\infty}(f)]$, which is always included in and sometimes much smaller [13, Example 2] than $[\underline{\mathrm{E}}_\infty(f), \overline{\mathrm{E}}_\infty(f)]$, thus yielding tighter bounds. And third, our pointwise ergodic theorem applies to six different types of imprecise Markov chains, giving it a universal character.

2. Imprecise Markov Chains Unravelled

An imprecise Markov chain [2, 3, 6, 11, 13, 14] is a mathematical model for the uncertain evolution of a system's state in discrete steps. The steps are indexed by the natural numbers $\mathbb{N}$ (excluding zero), typically thought of as points in time. At every time point—or every step—$i \in \mathbb{N}$, the state X_i is uncertain, but known to take values in a fixed state space $\mathscr{X}$. We assume that $\mathscr{X}$ is finite.

To describe a subject's beliefs about the uncertain evolution of the state of a system, an imprecise Markov chain considers, for every $x \in \mathscr{X}$, a (non-empty) set $\mathscr{P}_x$ of probability mass functions on $\mathscr{X}$. From an intuitive point of view, $\mathscr{P}_x$ can be thought of as providing partial information about some 'true' probability mass function p_x that, to every $y \in \mathscr{X}$, assigns a probability $p_x(y)$ that the system will be in state y in the next step given that it is in state x now. The probability mass function p_x—and hence the probability $p_x(y)$—is not necessarily known to our subject though. All that is expressed by the model is that p_x belongs to $\mathscr{P}_x$; this is what makes the model imprecise. Similarly, the subject's beliefs about the first state X_1 are also described by a set of probability mass functions on $\mathscr{X}$, denoted by $\mathscr{P}_\square$. If $\mathscr{P}_\square$ and every $\mathscr{P}_x$, $x \in \mathscr{X}$, each consist of only a single fully specified mass function, then we obtain the well-known special case of a (precise) Markov chain.

But there is more to imprecise Markov chains than the simple intuitive picture that we have painted above. Browsing through the literature on the subject, one will discover numerous different interpretations and characterisations for the sets $\mathscr{P}_\square$ and $\mathscr{P}_x$, and various different methods for turning these local models into a global uncertainty model that describes the complete evolution of the state of the system [2, 6]. The differences between these methods can sometimes be subtle, but they can also lead to fundamentally different inferences and conclusions. We will here distinguish between six approaches; that is, six types of imprecise Markov chains. Rest assured though: for our present purpose of studying the average behaviour of imprecise Markov chains, we will see that it does not matter which of these six is considered. The fact that it indeed doesn't is one of the contributions of this paper, and gives our results a universal character.

For each of the six considered models, we will focus on the corresponding global upper expectation operator. Such an operator $\overline{\mathrm{E}}$ takes two arguments: an extended real function f on the set $\Omega := \mathscr{X}^{\mathbb{N}}$ of all infinite state sequences, and a finite—possibly empty—state sequence $x_{1:n} = (x_1, \cdots, x_n) \in \mathscr{X}^n$; it maps these to a corresponding upper expectation $\overline{\mathrm{E}}(f|x_{1:n}) \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. An infinite state sequence $\omega = (x_1, \ldots, x_n, \ldots) \in \Omega$ is called a path and represents a possible evolution of the system, with $\omega_i := x_i$ the state at time $i \in \mathbb{N}$. Extended real functions on Ω are called variables; we denote the set of all variables by $\mathscr{V}(\Omega) := \overline{\mathbb{R}}^{\Omega}$. A finite state sequence $x_{1:n} \in \mathscr{S} := \cup_{i \in \mathbb{N}_0} \mathscr{X}^i$, with $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, is called a situation; it represents a possible (partial) evolution of the system up to some finite time point n. The empty sequence is called the initial situation, and will also be denoted by $\square := () = x_{1:0}$. If the time point n is of no importance, we will sometimes also use s to denote a generic situation. Situations act as conditional arguments for global upper expectations, in the sense that $\overline{\mathrm{E}}(f|x_{1:n})$ is interpreted as the upper expectation of f conditional on the fact that we observed $X_1 = x_1, \cdots, X_n = x_n$.

Besides upper expectations, one can also consider lower expectations $\underline{\mathrm{E}}$. These are equivalent to upper expectations because they are related to them by conjugacy, in the sense that $\underline{\mathrm{E}}(f|x_{1:n}) = -\overline{\mathrm{E}}(-f|x_{1:n})$ for all $f \in \mathscr{V}(\Omega)$ and $x_{1:n} \in \mathscr{S}$. Unconditional lower and upper expectations correspond to the case $n = 0$; that is, for any $f \in \mathscr{V}(\Omega)$, the unconditional lower expectation of f is given by $\underline{\mathrm{E}}(f) := \underline{\mathrm{E}}(f|\square) = \underline{\mathrm{E}}(f|x_{1:0})$, and similarly for $\overline{\mathrm{E}}(f)$. Lower and upper probabilities, finally, are also obtained as special cases; for any event $A \subseteq \Omega$, its lower and upper probability are defined by $\underline{\mathrm{P}}(A) := \underline{\mathrm{E}}(\mathbb{I}_A)$ and $\overline{\mathrm{P}}(A) := \overline{\mathrm{E}}(\mathbb{I}_A)$, respectively, where $\mathbb{I}_A \in \mathscr{V}(\Omega)$ is the indicator of A, defined for all $\omega \in \Omega$ by $\mathbb{I}_A(\omega) := 1$ if $\omega \in A$ and $\mathbb{I}_A(\omega) := 0$ otherwise.

2.1. Measure-Theoretic Imprecise Markov Chains

Of the six types of imprecise Markov chains that we will consider, the first three are most easily explained in terms of probabilities. For each of them, the local models $\mathscr{P}_\square$ and $\mathscr{P}_x$ are used to define a different set of so-called compatible probability trees. Any such probability tree $p_\bullet$ is a map that associates with each situation $x_{1:n} \in \mathscr{S}$ a probability mass function $p_{x_{1:n}} \in \mathscr{P}_{x_{1:n}}$, where $\mathscr{P}_{x_{1:0}} := \mathscr{P}_\square$ and, for $n \in \mathbb{N}$, $\mathscr{P}_{x_{1:n}} := \mathscr{P}_{x_n}$. Any of these probability trees $p_\bullet$ naturally gives rise to a corresponding conditional probability measure P, simply by positing that $\mathrm{P}(X_{n+1} = y | X_{1:n} = x_{1:n}) := p_{x_{1:n}}(y)$ for all $x_{1:n} \in \mathscr{S}$ and all $y \in \mathscr{X}$, and then subsequently using the conventional extension theorems; see [16, Section 9]. We do not go into further detail here, but do want to point out that this approach is slightly different from the usual measure-theoretic one, in the sense that our notion of a conditional probability measure P defines, for each $x_{1:n} \in \mathscr{S}$, a separate probability measure $\mathrm{P}(\cdot | X_{1:n} = x_{1:n})$.

Different choices for the local mass functions $p_{x_{1:n}}$ lead to different probability trees $p_\bullet$, and therefore different conditional probability measures P. We will use $\mathbb{P}_{\mathrm{ei}}$ to denote the set of all measures obtained in this way; this is known as an imprecise Markov chain under epistemic irrelevance. If we additionally impose that $p_{x_{1:n}}$ only depends on x_n and n—that is, if we impose a Markov assumption on the individual measures—we obtain a subset of $\mathbb{P}_{\mathrm{ei}}$, denoted by $\mathbb{P}_{\mathrm{ci}}$; this is known as an imprecise Markov chain under complete independence.[1] Finally, if we require that $p_{x_{1:n}}$ only depends on x_n—that is, if we impose Markovianity and time-homogeneity on the individual measures—then we obtain yet a smaller set of measures, denoted by $\mathbb{P}_{\mathrm{ri}}$ and referred to as an imprecise Markov chain under repetition independence. A crucial observation is that for each of these three models, the set $\mathscr{P}_{x_{1:n}}$ from which $p_{x_{1:n}}$ is chosen only depends on x_n, but not on earlier states nor on n; this is an *imprecise* version of the Markov (and time-homogeneity) property that justifies why each of these models is called an imprecise *Markov* chain, despite the fact that the individual measures they consist of may not be Markovian.

1. Some authors call this an imprecise Markov chain under strong independence [4]; that name is better suited for the convex hull of $\mathbb{P}_{\mathrm{ci}}$ though [1]. In any case, the difference is not important here since convexifying $\mathbb{P}_{\mathrm{ci}}$ does not affect its lower or upper expectations.

The global upper expectations $\overline{\mathrm{E}}_{\mathrm{ei}}$, $\overline{\mathrm{E}}_{\mathrm{ci}}$ and $\overline{\mathrm{E}}_{\mathrm{ri}}$ that correspond to the three models $\mathbb{P}_{\mathrm{ei}}$, $\mathbb{P}_{\mathrm{ci}}$ and $\mathbb{P}_{\mathrm{ri}}$ are defined as tight upper bounds on the expectations associated with the probability measures in these respective sets,[2] and similarly for lower expectations and lower and upper probabilities. Again, we refer to [16, Section 9] for more details.

2.2. Game-Theoretic Imprecise Markov Chains

The remaining three types of imprecise Markov chains that we will consider are more easily expressed in terms of supermartingales; or to phrase it less technically, in terms of the possible evolutions of a gambler's capital as he gambles in accordance with the subject's local models.

The starting point for these types of imprecise Markov chains are not the local models $\mathscr{P}_\square$ and $\mathscr{P}_x$ themselves, but rather the associated (local) upper expectations. The latter are not defined on variables, but on local gambles, which are real functions on $\mathscr{X}$. We will denote the set of all such gambles by $\mathscr{G}(\mathscr{X})$ and, for all $f \in \mathscr{G}(\mathscr{X})$, let $\|f\| := \max_{x\in\mathscr{X}} |f(x)|$. Then, for any non-empty set of probability mass functions $\mathscr{P}$ on $\mathscr{X}$, the associated (local) upper expectation $\overline{\mathrm{E}}_{\mathscr{P}}$ is defined by

$$\overline{\mathrm{E}}_{\mathscr{P}}(f) := \sup\Big\{\sum_{x\in\mathscr{X}} f(x)p(x) \colon p \in \mathscr{P}\Big\} \text{ for all } f \in \mathscr{G}(\mathscr{X}).$$

One can easily verify that such an upper expectation is coherent, meaning that for all $f,g \in \mathscr{G}(\mathscr{X})$ and real $\lambda \geq 0$:

C1. $\min f \leq \overline{\mathrm{E}}_{\mathscr{P}}(f) \leq \max f$; [bounds]

C2. $\overline{\mathrm{E}}_{\mathscr{P}}(f+g) \leq \overline{\mathrm{E}}_{\mathscr{P}}(f) + \overline{\mathrm{E}}_{\mathscr{P}}(g)$; [subadditivity]

C3. $\overline{\mathrm{E}}_{\mathscr{P}}(\lambda f) = \lambda \overline{\mathrm{E}}_{\mathscr{P}}(f)$. [nonnegative homogeneity]

In the particular case of an imprecise Markov chain, we associate with every situation $x_{1:n} \in \mathscr{S}$ such a local upper expectation $\overline{\mathrm{E}}_{x_{1:n}} := \overline{\mathrm{E}}_{\mathscr{P}_{x_{1:n}}}$ on $\mathscr{G}(\mathscr{X})$. Just like the set $\mathscr{P}_{x_{1:n}}$, this local upper expectation $\overline{\mathrm{E}}_{x_{1:n}}$ only depends on the last state x_n; this is the same imprecise Markov property that we discussed before, and to which imprecise Markov chains owe their name. Alternatively, we could also use local upper expectations $\overline{\mathrm{E}}_{x_{1:n}}$ that are specified directly, without any reference to $\mathscr{P}_{x_{1:n}}$, as long as they are coherent and satisfy the imprecise Markov property. They are then not necessarily interpreted in terms of probabilities, but can be given a direct behavioural interpretation in terms of prices for gambles [3]; for example, $\overline{\mathrm{E}}_\square(f)$ would then be the infimum selling price for the uncertain payoff $f(X_1)$.

2. Following standard measure-theoretic practices, the lower and upper expectation of a variable $f \in \mathscr{V}(\Omega)$ is only defined if f is measurable and Lebesgue integrable for every measure in the considered set. This restriction can be removed though by considering lower and upper Lebesgue integrals; this is also explained in [16]. In any case, for our present purposes, this is not essential since all the inferences that we will consider are measurable and Lebesgue integrable.

Regardless of the interpretation, the concept that turns local upper expectations into global game-theoretic upper expectations is that of a supermartingale. The idea is that a gambler—called Skeptic in this game-theoretic context [9]—will gamble on the values of the subsequent states of the process in such a way that our subject—called Forecaster—*expects* him to lose (or at least not gain) money. As Skeptic gambles, his capital will evolve, and this evolving capital is described by a supermartingale.

In particular, a supermartingale is a real function $\mathscr{M}$ on $\mathscr{S}$ such that Skeptic's capital $\mathscr{M}(x_{1:n})$ in each situation $x_{1:n} \in \mathscr{S}$ is at least as high as his (upper) expected capital at the next time point $n+1$. That is, for any $x_{1:n} \in \mathscr{S}$, a supermartingale $\mathscr{M}$ satisfies

$$\overline{\mathrm{E}}_{x_{1:n}}(\mathscr{M}(x_{1:n},\cdot)) \leq \mathscr{M}(x_{1:n}), \tag{1}$$

where we use $\mathscr{M}(x_{1:n},\cdot)$ to denote the gamble in $\mathscr{G}(\mathscr{X})$ whose value in $x_{n+1} \in \mathscr{X}$ is given by $\mathscr{M}(x_{1:n+1})$. We focus on supermartingales that are bounded below, meaning that there is some real $B > 0$ such that $\mathscr{M}(x_{1:n}) \geq -B$ for all $x_{1:n} \in \mathscr{S}$. This expresses that Skeptic is able to borrow at most a finite amount of capital. A non-negative supermartingale with $\mathscr{M}(\square) = 1$ is called a test supermartingale.

Now, for any variable $f \in \mathscr{V}(\Omega)$, the game-theoretic upper expectation $\overline{\mathrm{E}}_{\mathrm{mbr}}(f)$ is the lowest—or, more correctly, infimum—starting capital $\mathscr{M}(\square)$ of all the bounded below supermartingales $\mathscr{M}$ that eventually hedge f, in the sense that $\liminf \mathscr{M} \geq f$, with $\liminf \mathscr{M} \in \mathscr{V}(\Omega)$ defined by

$$\liminf \mathscr{M}(\omega) := \liminf_{n\to+\infty} \mathscr{M}(\omega_{1:n}) \text{ for all } \omega \in \Omega.$$

More generally, for any situation $x_{1:n} \in \mathscr{S}$, the conditional game-theoretic upper expectation $\overline{\mathrm{E}}_{\mathrm{mbr}}(f|x_{1:n})$ of f is the infimum capital $\mathscr{M}(x_{1:n})$ in the situation $x_{1:n}$ that allows Skeptic to hedge f on all paths that start with $x_{1:n}$ [3, 15]:

$$\begin{aligned}\overline{\mathrm{E}}_{\mathrm{mbr}}(f|x_{1:n}) := \inf\big\{&\mathscr{M}(x_{1:n}) \colon \mathscr{M} \in \mathbb{M}_{\mathrm{mbr}},\\ &(\forall \omega \in \Omega_{x_{1:n}}) \liminf \mathscr{M}(\omega) \geq f(\omega)\big\},\end{aligned}$$

with $\mathbb{M}_{\mathrm{mbr}}$ the set of all super**m**artingales that are **b**ounded below (and **r**eal)—the bold letters clarify our choice of notation—and $\Omega_{x_{1:n}} := \{\omega \in \Omega \colon \omega_{1:n} = x_{1:n}\}$ the set of all paths that start with $x_{1:n}$.

The two other types of game-theoretic upper expectations that we will consider are defined completely analogously; the only difference is that the set of supermartingales $\mathbb{M}_{\mathrm{mbr}}$ is replaced by a slightly different one. For $\overline{\mathrm{E}}_{\mathrm{mbb}}$, we consider the subset $\mathbb{M}_{\mathrm{mbb}}$ of all bounded supermartingales—so super**m**artingales that are, besides **b**ounded below, also **b**ounded above [15]. For $\overline{\mathrm{E}}_{\mathrm{mbe}}$, we replace $\mathbb{M}_{\mathrm{mbr}}$ with the superset $\mathbb{M}_{\mathrm{mbe}}$ of all **e***xtended real* super**m**artingales that are **b**ounded below [9, 15]: extended real-valued functions on $\mathscr{S}$ that satisfy Equation (1) and are bounded below. Care should be taken though because the operator $\overline{\mathrm{E}}_{x_{1:n}}$ in Equation (1) is only defined for local gambles, which are real-valued, whereas $\mathscr{M}(x_{1:n},\cdot)$ is now a function on $\mathscr{X}$ that can also take the value $+\infty$. This can be dealt with by suitable extending the domain of $\overline{\mathrm{E}}_{x_{1:n}}$ using limit arguments [15, 16].

2.3. So How Do They All Relate?

For a fixed choice of the local models $\mathscr{P}_{\square}$ and $\mathscr{P}_x$, we here considered three measure-theoretic global upper expectations ($\overline{\mathrm{E}}_{\mathrm{ri}}$, $\overline{\mathrm{E}}_{\mathrm{ci}}$ and $\overline{\mathrm{E}}_{\mathrm{ei}}$) and three game-theoretic ones ($\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$and $\overline{\mathrm{E}}_{\mathrm{mbb}}$). In that particular order, these models are ranked from most precise—or least conservative—to most imprecise—or most conservative.

Proposition 1 *For all $f\in\mathscr{V}(\Omega)$ and $s\in\mathscr{S}$, we have that*

$$\overline{\mathrm{E}}_{\mathrm{ri}}(f|s)\leq\overline{\mathrm{E}}_{\mathrm{ci}}(f|s)\leq\overline{\mathrm{E}}_{\mathrm{ei}}(f|s)\leq\overline{\mathrm{E}}_{\mathrm{mbe}}(f|s)\leq\overline{\mathrm{E}}_{\mathrm{mbr}}(f|s)\leq\overline{\mathrm{E}}_{\mathrm{mbb}}(f|s).$$

Proof The first two inequalities follow from the fact that $\mathbb{P}_{\mathrm{ri}}\subseteq\mathbb{P}_{\mathrm{ci}}\subseteq\mathbb{P}_{\mathrm{ei}}$. The last two inequalities follow from the fact that $\mathbb{M}_{\mathrm{mbe}}\supseteq\mathbb{M}_{\mathrm{mbr}}\supseteq\mathbb{M}_{\mathrm{mbb}}$. The third inequality is a special case of [16, Corollary 19]. ∎

Due to conjugacy, the opposite inequalities are true for lower expectations. Analogous inequalities hold for lower and upper probabilities: it suffices to apply the results for lower and upper expectations to indicators of events.

For all but the third inequality, the inequalities in Proposition 1 can become strict [4, 7, 8, 15]. For the third inequality, this is—as far as we know—an open question.

That said, for specific types of variables, some of the inequalities in Proposition 1 do turn into equalities. For example, for bounded real variables—that is, $f\in\mathscr{V}(\mathscr{X})$ for which there is some real $B>0$ such that $-B\leq f\leq B$—the three game-theoretic global upper expectations coincide.

Proposition 2 *[15, Proposition 36] For any $f\in\mathscr{V}(\Omega)$ that is bounded and any $s\in\mathscr{S}$, we have that*

$$\overline{\mathrm{E}}_{\mathrm{mbe}}(f|s)=\overline{\mathrm{E}}_{\mathrm{mbr}}(f|s)=\overline{\mathrm{E}}_{\mathrm{mbb}}(f|s).$$

These equalities can be extended to include $\overline{\mathrm{E}}_{\mathrm{ei}}$ if we moreover restrict ourselves to variables that are *finitary*: functions f on Ω that only depend on a finite number of states X_i, meaning that there is some $n\in\mathbb{N}$ and a real function g on $\mathscr{X}^n$ such that $f(\omega)=g(\omega_{1:n})$ for all $\omega\in\Omega$.

Proposition 3 *For any $f\in\mathscr{V}(\Omega)$ that is real and finitary and any $s\in\mathscr{S}$, we have that*

$$\overline{\mathrm{E}}_{\mathrm{ei}}(f|s)=\overline{\mathrm{E}}_{\mathrm{mbe}}(f|s)=\overline{\mathrm{E}}_{\mathrm{mbr}}(f|s)=\overline{\mathrm{E}}_{\mathrm{mbb}}(f|s).$$

Proof The first equality corresponds to [16, Proposition 21]. Since finitary real variables are bounded because of $\mathscr{X}$ is finite, the other two follow from Proposition 2. ∎

Other examples of types of variables for which some of the inequalities in Proposition 1 turn into equalities can be found in, among others, References [4, 6, 7, 12].

3. Almost Sure Events

In (precise) measure-theoretic probability theory, an event $A\subseteq\Omega$ is almost sure if it has probability one, or equivalently, if its complement $A^{\mathrm{c}}:=\Omega\setminus A$ has probability zero. Similarly, in an imprecise probability context, we call an event almost sure if its complement has *upper* probability zero [3, 17]. Since we here consider six global upper expectations, and hence six types of upper probabilities, we distinguish between six notions of almost surely.

Definition 4 *We say that an event $A\subseteq\Omega$ is $\overline{\mathrm{P}}_{\mathrm{ri}}$-almost sure if $\overline{\mathrm{P}}_{\mathrm{ri}}(A^{\mathrm{c}})=0$. We adopt similar definitions for $\overline{\mathrm{P}}_{\mathrm{ci}}$, $\overline{\mathrm{P}}_{\mathrm{ei}}$, $\overline{\mathrm{P}}_{\mathrm{mbe}}$, $\overline{\mathrm{P}}_{\mathrm{mbr}}$ and $\overline{\mathrm{P}}_{\mathrm{mbb}}$.*

Our next result shows how these notions are related.

Corollary 5 *For any event $A\subseteq\Omega$, we have that*

$$\overline{\mathrm{P}}_{\mathrm{mbb}}(A^{\mathrm{c}})=0\Leftrightarrow\overline{\mathrm{P}}_{\mathrm{mbr}}(A^{\mathrm{c}})=0\Leftrightarrow\overline{\mathrm{P}}_{\mathrm{mbe}}(A^{\mathrm{c}})=0\Rightarrow\overline{\mathrm{P}}_{\mathrm{ei}}(A^{\mathrm{c}})=0\Rightarrow\overline{\mathrm{P}}_{\mathrm{ci}}(A^{\mathrm{c}})=0\Rightarrow\overline{\mathrm{P}}_{\mathrm{ri}}(A^{\mathrm{c}})=0.$$

Proof Since $\mathbb{I}_{A^{\mathrm{c}}}$ is a bounded variable, the equivalences follow from Proposition 2. The implications hold because

$$\overline{\mathrm{P}}_{\mathrm{mbe}}(A^{\mathrm{c}})\geq\overline{\mathrm{P}}_{\mathrm{ei}}(A^{\mathrm{c}})\geq\overline{\mathrm{P}}_{\mathrm{ci}}(A^{\mathrm{c}})\geq\overline{\mathrm{P}}_{\mathrm{ri}}(A^{\mathrm{c}})\geq 0,$$

where the last inequality holds because $\mathrm{P}(A^{\mathrm{c}})\geq 0$ for all P in $\mathbb{P}_{\mathrm{ri}}$, and the first three follow from Proposition 1. ∎

So we see that $\overline{\mathrm{P}}_{\mathrm{ri}}$-almost sure is the weakest notion and that the three strongest notions of almost surely are the game-theoretic ones that correspond to $\overline{\mathrm{P}}_{\mathrm{mbe}}$, $\overline{\mathrm{P}}_{\mathrm{mbr}}$ and $\overline{\mathrm{P}}_{\mathrm{mbb}}$. These last three are furthermore equivalent.

Besides these different types of almost sure events, we will also consider strictly almost sure events: events whose complement is deemed so rare that it is possible for Skeptic to start with capital one and, without borrowing, gamble in such a way that he becomes infinitely rich along every path not included in the event (even though our subject expects that Skeptic's capital will not increase) [3, 17].

Definition 6 *We say that an event $A\subseteq\Omega$ is strictly almost sure if there is a (real) test supermartingale $\mathscr{M}$ such that $\lim\mathscr{M}(\omega)=+\infty$ for all $\omega\in A^{\mathrm{c}}$, with $\lim\mathscr{M}(\omega):=\lim_{n\to+\infty}\mathscr{M}(\omega_{1:n})$.*

This strict version of almost surely is stronger than each of the six others we have considered.

Proposition 7 *If an event $A\subseteq\Omega$ is strictly almost sure, then it is also $\overline{\mathrm{P}}_{\mathrm{mbb}}$-almost sure. The same is true for $\overline{\mathrm{P}}_{\mathrm{mbr}}$, $\overline{\mathrm{P}}_{\mathrm{mbe}}$, $\overline{\mathrm{P}}_{\mathrm{ei}}$, $\overline{\mathrm{P}}_{\mathrm{ci}}$ and $\overline{\mathrm{P}}_{\mathrm{ri}}$.*

Proof If A is strictly almost sure, using the terminology of [3], A^{c} is strictly null. Using [3, Proposition 4], this implies that A^{c} is a null event, in the sense that $\overline{\mathrm{P}}_{\mathrm{mbr}}(A^{\mathrm{c}})=0$. The result now follows from Corollary 5. ∎

Furthermore, finite intersections of strictly almost sure events are strictly almost sure themselves.

Lemma 8 *Consider two strictly almost sure events $A, B \subseteq \Omega$. Then $A \cap B$ is strictly almost sure as well.*

Proof Since A and B are strictly almost sure, there are test supermartingales $\mathscr{M}_A$ and $\mathscr{M}_B$ that converge to $+\infty$ on $\Omega \setminus A$ and $\Omega \setminus B$, respectively. It can easily be verified that $\mathscr{M} := \frac{1}{2}(\mathscr{M}_A + \mathscr{M}_B)$ is then a test supermartingale that converges to $+\infty$ on $\Omega \setminus (A \cap B) = (\Omega \setminus A) \cup (\Omega \setminus B)$. Hence, $A \cap B$ is strictly almost sure. ■

4. Weak Ergodicity

The pointwise ergodic theorem that we will present in Section 5 applies to imprecise Markov chains that are weakly ergodic. This property, which is implied by—and hence more generally applicable than—ergodicity, was recently introduced for $\overline{\mathrm{E}}_{\mathrm{ri}}$, $\overline{\mathrm{E}}_{\mathrm{ci}}$ and $\overline{\mathrm{E}}_{\mathrm{ei}}$ [13]. We here extend this concept to $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ and $\overline{\mathrm{E}}_{\mathrm{mbb}}$.

Definition 9 *An imprecise Markov chain with global upper expectation $\overline{\mathrm{E}}_{\mathrm{ri}}$ is weakly ergodic if, for all $f \in \mathscr{G}(\mathscr{X})$, the limit*

$$\overline{\mathrm{E}}_{\mathrm{av},\infty}(f) := \lim_{n\to+\infty} \overline{\mathrm{E}}_{\mathrm{ri}}\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i) \,\middle|\, x\right)$$

exists and is equal for all $x \in \mathscr{X}$. Similar definitions apply for $\overline{\mathrm{E}}_{\mathrm{ci}}$, $\overline{\mathrm{E}}_{\mathrm{ei}}$, $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ and $\overline{\mathrm{E}}_{\mathrm{mbb}}$.

For the global upper expectations $\overline{\mathrm{E}}_{\mathrm{ri}}$, $\overline{\mathrm{E}}_{\mathrm{ci}}$ and $\overline{\mathrm{E}}_{\mathrm{ei}}$, we have recently shown that these notions of weak ergodicity are equivalent, and that the resulting limit values $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ are identical [13]. They depend solely on the local models $\mathscr{P}_x$. In particular, they are completely determined by the upper transition operator $\overline{T}\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{X})$ that, with every gamble $f \in \mathscr{G}(\mathscr{X})$, associates a new gamble $\overline{T}f$ on $\mathscr{X}$ defined by $\overline{T}f(x) := \overline{\mathrm{E}}_{\mathscr{P}_x}(f)$ for all $x \in \mathscr{X}$.

Our next result shows that this remains true if we additionally include $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ and $\overline{\mathrm{E}}_{\mathrm{mbb}}$. It makes use of the concept of a weakly ergodic upper transition operator.

Definition 10 *An upper transition operator $\overline{T}$ is weakly ergodic if, for all $f \in \mathscr{G}(\mathscr{X})$, with*

$$\overline{T}_f\colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{X})\colon h \mapsto \overline{T}_f h := f + \overline{T}h,$$

the limit

$$\overline{\mathrm{E}}_{\mathrm{av},\overline{T}}(f) := \lim_{n\to+\infty} \frac{1}{n}\left[\overline{T}_f^{\,n-1} f\right](x) \tag{2}$$

exists and is equal for all $x \in \mathscr{X}$.

Proposition 11 *An imprecise Markov chain with global upper expectation $\overline{\mathrm{E}}_{\mathrm{ri}}$ is weakly ergodic if and only if the corresponding upper transition operator $\overline{T}$ is weakly ergodic, and if it is, then $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f) = \overline{\mathrm{E}}_{\mathrm{av},\overline{T}}(f)$ for all $f \in \mathscr{G}(\mathscr{X})$. The same holds for $\overline{\mathrm{E}}_{\mathrm{ci}}$, $\overline{\mathrm{E}}_{\mathrm{ei}}$, $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ and $\overline{\mathrm{E}}_{\mathrm{mbb}}$.*

Proof That this result holds for $\overline{\mathrm{E}}_{\mathrm{ci}}$ and $\overline{\mathrm{E}}_{\mathrm{ei}}$ is explained in [13, end of Section 4]. [13, Theorem 14] furthermore shows that the weak ergodicity of $\overline{T}$ is equivalent to $\overline{T}$ being 'top class absorbing'. The result for $\overline{\mathrm{E}}_{\mathrm{ri}}$ therefore follows from [13, Theorem 22]. To see that it also holds for $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ and $\overline{\mathrm{E}}_{\mathrm{mbb}}$, observe that it follows from Proposition 3 that $\overline{\mathrm{E}}_{*}(\frac{1}{n}\sum_{i=1}^{n} f(X_i)|x)$ is equal for all $* \in \{\mathrm{ei}, \mathrm{mbe}, \mathrm{mbr}, \mathrm{mbb}\}$. It therefore does not matter whether Definition 9 is applied to $\overline{\mathrm{E}}_{\mathrm{ei}}$, $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ or $\overline{\mathrm{E}}_{\mathrm{mbb}}$, because the corresponding notion of weak ergodicity, as well as the resulting values of $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$, will be identical. Since we already know that Proposition 11 is true for $\overline{\mathrm{E}}_{\mathrm{ei}}$, it therefore also holds for $\overline{\mathrm{E}}_{\mathrm{mbe}}$, $\overline{\mathrm{E}}_{\mathrm{mbr}}$ and $\overline{\mathrm{E}}_{\mathrm{mbb}}$. ■

So we see that the adopted global upper expectation becomes irrelevant when it comes to weak ergodicity: we can simply say that an imprecise Markov chain is weakly ergodic and then consider the resulting averaged limit upper expectations $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f) = \overline{\mathrm{E}}_{\mathrm{av},\overline{T}}(f)$, regardless of the adopted global model.

In practice, checking if an upper transition operator $\overline{T}$—and hence also an imprecise Markov chain—is weakly ergodic can be done by verifying whether it is top class absorbing [13]. If it is, then $\overline{\mathrm{E}}_{\mathrm{av},\overline{T}}$—and hence also $\overline{\mathrm{E}}_{\mathrm{av},\infty}$—can be computed using Equation (2).

5. A Pointwise Ergodic Theorem

With all terminology in place, we can now finally state the pointwise ergodic theorem that is the subject of this paper. We start with a one-sided version.

Proposition 12 *For a weakly ergodic imprecise Markov chain and any $f \in \mathscr{G}(\mathscr{X})$, strictly almost surely,*

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(X_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f).$$

Due to its technical nature, the proof of this result is deferred to Section 7.

To arrive at a two-sided version, the trick is to apply it to $-f$ as well. This yields a similar result for $\underline{\mathrm{E}}_{\mathrm{av},\infty}(f) := -\overline{\mathrm{E}}_{\mathrm{av},\infty}(-f)$, but with the inequality reversed and lim inf instead of lim sup. Combined with Lemma 8 and Proposition 7, we arrive at the main result of this paper: a two-sided pointwise ergodic theorem for imprecise Markov chains. Its formulation is deceivingly simple, but we hope that its strength and universal character are nevertheless apparent.

Theorem 13 *For a weakly ergodic imprecise Markov chain and any $f \in \mathscr{G}(\mathscr{X})$, strictly almost surely,*

$$\underline{\mathrm{E}}_{\mathrm{av},\infty}(f) \leq \liminf_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(X_i) \leq \limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(X_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f).$$

This result continues to hold if we replace strictly almost surely by $\overline{\mathrm{P}}_{\mathrm{mbb}}$-almost surely, and similarly for $\overline{\mathrm{P}}_{\mathrm{mbr}}$, $\overline{\mathrm{P}}_{\mathrm{mbe}}$, $\overline{\mathrm{P}}_{\mathrm{ei}}$, $\overline{\mathrm{P}}_{\mathrm{ci}}$ and $\overline{\mathrm{P}}_{\mathrm{ri}}$.

Proof By applying Proposition 12 to f, we find that the third inequality of the statement is strictly almost sure. By applying Proposition 12 to $-f$, we find that, strictly almost surely, $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}-f(X_i)\leq\overline{\mathrm{E}}_{\mathrm{av},\infty}(-f)=-\underline{\mathrm{E}}_{\mathrm{av},\infty}(f)$. So the first inequality of the statement is strictly almost sure as well. Since the second inequality is trivially true, it follows from Lemma 8 that the stated three inequalities are (jointly) strictly almost sure. The last part of the theorem now follows directly from Proposition 7. ■

The remainder of this paper is devoted to the proof of Proposition 12. Section 6 presents the two main building blocks on which this proof relies. The proof itself is given at the end of Section 7, preceded by some intuition.

6. Two Crucial Supermartingale Inequalities

Our proof for Proposition 12 essentially consists of two steps. First, we will establish the desired inequality between $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}f(X_i)$ and $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ up to an additive term that is expressed in terms of a supermartingale. The second step consists in showing that, (strictly) almost surely, this additive term can be ignored.

The key result that enables the first of these steps is the proposition below. It uses the weak ergodicity of an imprecise Markov chain to show that for n large enough,

$$\frac{1}{n}\sum_{i=1}^{n}f(X_i)\leq\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)+\varepsilon+\frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(X_{1:i})(X_{i+1}),\quad(3)$$

where $\mathscr{M}$ is a specific supermartingale and $\Delta\mathscr{M}$ is its so-called difference. For any real map $\mathscr{M}$ on $\mathscr{S}$, this difference $\Delta\mathscr{M}$ is the unique map from $\mathscr{S}$ to $\mathscr{G}(\mathscr{X})$ that associates with every $x_{1:n}\in\mathscr{S}$ a local gamble $\Delta\mathscr{M}(x_{1:n})$, defined for all $x_{n+1}\in\mathscr{X}$ by

$$\Delta\mathscr{M}(x_{1:n})(x_{n+1}):=\mathscr{M}(x_{1:n+1})-\mathscr{M}(x_{1:n}).$$

Proposition 14 *Consider a weakly ergodic imprecise Markov chain. Then for all $f\in\mathscr{G}(\mathscr{X})$ and $\varepsilon>0$, there is a real $B>\varepsilon$, a supermartingale $\mathscr{M}$ with $|\Delta\mathscr{M}|\leq B$ and some $N\in\mathbb{N}$ such that, for all $x_{1:n}\in\mathscr{S}$ with $n\geq N$,*

$$\frac{1}{n}\sum_{i=1}^{n}f(x_i)\ \leq\ \overline{\mathrm{E}}_{\mathrm{av},\infty}(f)\ +\ \varepsilon\ +\ \frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(x_{1:i})(x_{i+1}).$$

Our proof for this result uses the following two lemmas, the first of which we borrow from [3] and the second of which is a well-known consequence of coherence.

Lemma 15 *Let $\mathscr{M}$ be a real supermartingale. Then for all $n\in\mathbb{N}_0$ and $x_{1:n}\in\mathscr{X}_{1:n}$:*

$$\mathscr{M}(x_{1:n})\geq\inf_{\omega\in\Omega_{x_{1:n}}}\liminf\mathscr{M}(\omega).$$

Proof Since $\mathscr{M}$ is a (real) supermartingale, $-\mathscr{M}$ is a submartingale in the sense of Reference [3]. The result therefore follows from [3, Lemma 1]. ■

Lemma 16 *A real map $\mathscr{M}$ on $\mathscr{S}$ is a supermartingale if and only if, for all $x_{1:n}\in\mathscr{S}$, $\overline{\mathrm{E}}_{x_{1:n}}(\Delta\mathscr{M}(x_{1:n}))\leq 0$.*

Proof This follows from the definition of a supermartingale and the fact that, for all $x_{1:n}\in\mathscr{S}$,

$$\begin{aligned}\overline{\mathrm{E}}_{x_{1:n}}(\mathscr{M}(x_{1:n},\cdot))&=\overline{\mathrm{E}}_{x_{1:n}}(\mathscr{M}(x_{1:n})+\Delta\mathscr{M}(x_{1:n}))\\&=\mathscr{M}(x_{1:n})+\overline{\mathrm{E}}_{x_{1:n}}(\Delta\mathscr{M}(x_{1:n})),\end{aligned}$$

where the second equality is an instance of the constant additivity of coherent upper expectations [18, 2.6.1(c)]. ■

Proof of Proposition 14. We know from Proposition 11 that $\overline{T}$ is weakly ergodic and that $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)=\overline{\mathrm{E}}_{\mathrm{av},\overline{T}}(f)$. Furthermore, for all $n\in\mathbb{N}$ and $x\in\mathscr{X}$, we know from [13, Lemma 40] that $\min f\leq\frac{1}{n}\big[\overline{T}_f^{n-1}f\big](x)\leq\max f$. Hence, using Definition 10, we see that $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ is a real number.

Next, for all $x\in\mathscr{X}$ and $n\in\mathbb{N}$, let

$$\overline{\mathrm{E}}_{\mathrm{av},n}(f|x):=\overline{\mathrm{E}}_{\mathrm{mbr}}\left(\frac{1}{n}\sum_{i=1}^{n}f(X_i)\,\middle|\,x\right).$$

Definition 9 then tells us that $\lim_{n\to+\infty}\overline{\mathrm{E}}_{\mathrm{av},n}(f|x)=\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ for all $x\in\mathscr{X}$. Therefore, and since $\mathscr{X}$ is finite and $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ is real, there is some $m\in\mathbb{N}$ such that $|\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)-\overline{\mathrm{E}}_{\mathrm{av},m}(f|x)|\leq\varepsilon/4$ for all $x\in\mathscr{X}$. For each $x\in\mathscr{X}$, since $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ is real, this implies that $\overline{\mathrm{E}}_{\mathrm{av},m}(f|x)$ is real as well. Applying the definition of $\overline{\mathrm{E}}_{\mathrm{mbr}}$, this implies that there is, for all $x\in\mathscr{X}$, a real supermartingale $\mathscr{M}_x$ such that $\mathscr{M}_x(x)\leq\overline{\mathrm{E}}_{\mathrm{av},m}(f|x)+\varepsilon/4$ and, for all $\omega\in\Omega_x$, $\liminf\mathscr{M}_x(\omega)\geq\frac{1}{m}\sum_{i=1}^{m}f(\omega_i)$. Hence, for all $x_{1:m+1}\in\mathscr{X}_{1:m+1}$, it follows from Lemma 15 that

$$\begin{aligned}\mathscr{M}_{x_1}(x_{1:m+1})&\geq\inf_{\omega\in\Gamma(x_{1:m+1})}\liminf\mathscr{M}_{x_1}(\omega)\\&\geq\inf_{\omega\in\Gamma(x_{1:m+1})}\frac{1}{m}\sum_{i=1}^{m}f(\omega_i)=\frac{1}{m}\sum_{i=1}^{m}f(x_i),\end{aligned}$$

and therefore also that

$$\begin{aligned}\sum_{i=1}^{m}\Delta\mathscr{M}_{x_1}(x_{1:i})(x_{i+1})&=\mathscr{M}_{x_1}(x_{1:m+1})-\mathscr{M}_{x_1}(x_1)\\&\geq\frac{1}{m}\sum_{i=1}^{m}f(x_i)-\overline{\mathrm{E}}_{\mathrm{av},m}(f|x_1)-\frac{\varepsilon}{4}\\&\geq\frac{1}{m}\sum_{i=1}^{m}f(x_i)-\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)-\frac{\varepsilon}{2}.\quad(4)\end{aligned}$$

Now let $\mathscr{M}$ be the unique real map on $\mathscr{S}$ that satisfies $\mathscr{M}(\Box)=0$, $\Delta\mathscr{M}(\Box)=0$ and

$$\Delta\mathscr{M}(x_{1:m\ell+k}):=m\Delta\mathscr{M}_{x_{m\ell+1}}(x_{m\ell+1:m\ell+k})$$

for all $\ell \in \mathbb{N}_0$, $k \in \{1, \ldots, m\}$ and $x_{1:m\ell+k} \in \mathscr{X}^{m\ell+k}$. Using Lemma 16, we see that this map $\mathscr{M}$ is a supermartingale because $\mathscr{M}_x$ is a supermartingale for each $x \in \mathscr{X}$, and because $\overline{\mathrm{E}}_{\square}(\Delta\mathscr{M}(\square)) = \overline{\mathrm{E}}_{\square}(0) \leq 0$ due to C1. Now let

$$B' := \max\left\{ \|\Delta\mathscr{M}_{z_1}(z_{1:k})\| : k \in \{1, \ldots, m\}, z_{1:k} \in \mathscr{X}_{1:k} \right\},$$

which clearly provides a uniform bound for $\Delta\mathscr{M}$, in the sense that $|\Delta\mathscr{M}| \leq B'$. Then for $B := \max\{B', 2\varepsilon\}$, we have that $B > \varepsilon$ and $|\Delta\mathscr{M}| \leq B$. Furthermore, for any $\tilde{n} \in \mathbb{N}$ and $x_{1:m\tilde{n}+1} \in \mathscr{X}^{m\tilde{n}+1}$, we have that

$$\begin{aligned}
&\frac{1}{m\tilde{n}} \sum_{i=1}^{m\tilde{n}} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) \\
&\quad = \frac{1}{m\tilde{n}} \sum_{\ell=0}^{\tilde{n}-1} \sum_{k=1}^{m} \Delta\mathscr{M}(x_{1:m\ell+k})(x_{m\ell+k+1}) \\
&\quad = \frac{1}{m\tilde{n}} \sum_{\ell=0}^{\tilde{n}-1} \sum_{k=1}^{m} m\Delta\mathscr{M}_{x_{m\ell+1}}(x_{m\ell+1:m\ell+k})(x_{m\ell+k+1}) \\
&\quad = \frac{1}{\tilde{n}} \sum_{\ell=0}^{\tilde{n}-1} \Big(\sum_{k=1}^{m} \Delta\mathscr{M}_{x_{m\ell+1}}(x_{m\ell+1:m\ell+k})(x_{m\ell+k+1}) \Big) \\
&\quad \geq \frac{1}{\tilde{n}} \sum_{\ell=0}^{\tilde{n}-1} \Big(\frac{1}{m} \sum_{k=1}^{m} f(x_{m\ell+k}) - \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) - \frac{\varepsilon}{2} \Big) \\
&\quad = \frac{1}{m\tilde{n}} \sum_{i=1}^{m\tilde{n}} f(x_i) - \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) - \frac{\varepsilon}{2}, \qquad (5)
\end{aligned}$$

using Equation (4) for the inequality. Now let $N \in \mathbb{N}$ be any natural number such that $N\varepsilon \geq 8m\|f\|$, $N\varepsilon \geq 8mB$ and $N > m$, and consider any $n \in \mathbb{N}$ such that $n \geq N$ and any $x_{1:n} \in \mathscr{X}^n$. Let $\tilde{n}$ be the unique $\tilde{n} \in \mathbb{N}$ such that $m\tilde{n} < n \leq m\tilde{n} + m$. Then on the one hand, we find that

$$\begin{aligned}
&\frac{1}{n} \sum_{i=1}^{n} f(x_i) - \frac{1}{m\tilde{n}} \sum_{i=1}^{m\tilde{n}} f(x_i) \\
&\quad = \Big(\frac{1}{n} - \frac{1}{m\tilde{n}} \Big) \sum_{i=1}^{m\tilde{n}} f(x_i) + \frac{1}{n} \sum_{i=m\tilde{n}+1}^{n} f(x_i) \\
&\quad \leq \Big(\frac{1}{m\tilde{n}} - \frac{1}{n} \Big) m\tilde{n} \|f\| + \frac{1}{n}(n - m\tilde{n}) \|f\| \\
&\quad = 2\Big(1 - \frac{m\tilde{n}}{n}\Big) \|f\| = 2\Big(\frac{n - m\tilde{n}}{n}\Big) \|f\| \leq 2\frac{m}{n} \|f\|,
\end{aligned}$$

where the first equality holds because $n > m\tilde{n}$, the first inequality because $\frac{1}{m\tilde{n}} > \frac{1}{n}$ (since $n > m\tilde{n}$) and $-\|f\| \leq f(x_i) \leq \|f\|$, and the second inequality because $n \leq m\tilde{n} + m$. On the other hand, we also find that

$$\begin{aligned}
&\frac{1}{m\tilde{n}} \sum_{i=1}^{m\tilde{n}} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) - \frac{1}{n} \sum_{i=0}^{n-1} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) \\
&\quad = \Big(\frac{1}{m\tilde{n}} - \frac{1}{n} \Big) \sum_{i=1}^{m\tilde{n}} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) \\
&\qquad - \frac{1}{n} \Delta\mathscr{M}(\square)(x_1) - \frac{1}{n} \sum_{i=m\tilde{n}+1}^{n-1} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) \\
&\quad \leq \Big(\frac{1}{m\tilde{n}} - \frac{1}{n} \Big) m\tilde{n}B + \frac{1}{n}(n - m\tilde{n})B \\
&\quad = 2\Big(1 - \frac{m\tilde{n}}{n}\Big)B = 2\Big(\frac{n - m\tilde{n}}{n}\Big)B \leq 2\frac{m}{n}B,
\end{aligned}$$

where the first equality holds because $n - 1 \geq m\tilde{n}$ (since $n > m\tilde{n}$), the first inequality because $\frac{1}{m\tilde{n}} > \frac{1}{n}$ (again since $n > m\tilde{n}$) and $|\Delta\mathscr{M}| \leq B$, and the second inequality because $n \leq m\tilde{n} + m$. If we combine these two inequalities with Equation (5), we find that

$$\begin{aligned}
&\frac{1}{n} \sum_{i=1}^{n} f(x_i) \\
&\leq \frac{1}{m\tilde{n}} \sum_{i=1}^{m\tilde{n}} f(x_i) + 2\frac{m}{n} \|f\| \\
&\leq \frac{1}{m\tilde{n}} \sum_{i=1}^{m\tilde{n}} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) + \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + \frac{\varepsilon}{2} + 2\frac{m}{n} \|f\| \\
&\leq \frac{1}{n} \sum_{i=0}^{n-1} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) + 2\frac{m}{n}B \\
&\qquad\qquad + \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + \frac{\varepsilon}{2} + 2\frac{m}{n} \|f\|,
\end{aligned}$$

which implies that

$$\frac{1}{n} \sum_{i=1}^{n} f(x_i) \leq \frac{1}{n} \sum_{i=0}^{n-1} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) + \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + \varepsilon$$

because $n\varepsilon \geq N\varepsilon \geq 8m\|f\|$ and $n\varepsilon \geq N\varepsilon \geq 8mB$. ■

A crucial feature of this result is that $\mathscr{M}$ has uniformly bounded differences, in the sense that $|\Delta\mathscr{M}| \leq B$ for some $B \in \mathbb{R}$. The reason why this is important is that it allows us to build a new supermartingale that is positive and that becomes exponentially large on all paths where the supermartingale term $\sum_{i=0}^{n-1} \Delta\mathscr{M}(X_{1:i})(X_{i+1})$ in the upper bound of Equation (3) exceeds ε.

Proposition 17 *Consider any real $B > 0$ and $0 < \varepsilon < B$. Let $\mathscr{M}$ be a supermartingale such that $|\Delta\mathscr{M}| \leq B$. Then the function $\mathscr{F}_{\mathscr{M}}$ on $\mathscr{S}$, defined by*

$$\mathscr{F}_{\mathscr{M}}(x_{1:n}) := \prod_{i=0}^{n-1} \Big[1 + \frac{\varepsilon}{2B^2} \Delta\mathscr{M}(x_{1:i})(x_{1+i})\Big]$$

for all $x_{1:n} \in \mathscr{S}$, is a positive supermartingale with $\mathscr{F}_{\mathscr{M}}(\square) = 1$. Furthermore, for all $x_{1:n} \in \mathscr{S}$, we have that $\mathscr{F}_{\mathscr{M}}(x_{1:n}) \leq (3/2)^n$ and

$$\frac{1}{n} \sum_{i=0}^{n-1} \Delta\mathscr{M}(x_{1:i})(x_{i+1}) \geq \varepsilon \Rightarrow \mathscr{F}_{\mathscr{M}}(x_{1:n}) \geq \exp\Big(\frac{n\varepsilon^2}{4B^2}\Big).$$

Proof Since $\mathscr{M}$ is a supermartingale, $-\mathscr{M}$ is a submartingale in the sense of [3]. This result is therefore a particular case of [3, Lemma 9]. That result does not explicitly state the upper bound $\mathscr{F}_{\mathscr{M}}(x_{1:n}) \leq (3/2)^n$ though. To see that this upper bound indeed applies, it suffices to observe that $0 < \varepsilon < B$ and $|\Delta\mathscr{M}| \leq B$ imply that $-\frac{1}{2} \leq \frac{\varepsilon}{2B^2}\Delta\mathscr{M}(x_{1:i})(x_{1+i}) \leq \frac{1}{2}$. ■

Since supermartingales are by definition expected to decrease, it is unlikely for them to become exponentially large. So at least intuitively, we have at this point established that the supermartingale term in Equation (3) is likely to not exceed ε, or equivalently, that it is likely that $\frac{1}{n}\sum_{i=1}^n f(X_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + 2\varepsilon$. To arrive at Proposition 12, three tasks remain: replace $\frac{1}{n}\sum_{i=1}^n f(X_i)$ by $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^n f(X_i)$, remove the 2ε, and formalise our intuition by replacing the informal notion of 'likely' with the formal notion of 'strictly almost surely'.

7. Proving Proposition 12

To obtain an upper bound for $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^n f(X_i)$, we start by taking the limit superior of (both sides of) Equation (3). This yields

$$\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^n f(X_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + \varepsilon + \limsup_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(X_{1:i})(X_{i+1}). \quad (6)$$

Since $\Delta\mathscr{M}$ is uniformly bounded, it follows directly from Proposition 17 that there is a positive supermartingale $\mathscr{F}_{\mathscr{M}}$ with initial capital $\mathscr{F}_{\mathscr{M}}(\square) = 1$ such that $\limsup \mathscr{F}_{\mathscr{M}}$ becomes unbounded on all paths ω where $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(X_{1:i})(X_{i+1})$ exceeds ε. If $\mathscr{F}_{\mathscr{M}}$ would instead converge to $+\infty$ on those paths, then this would mean that, strictly almost surely, $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^n f(X_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + 2\varepsilon$. So it remains to replace $\mathscr{F}_{\mathscr{M}}$ with a supermartingale that converges to $+\infty$ if the inequality fails, and to remove the ε part.

We start by replacing $\mathscr{F}_{\mathscr{M}}$ with a supermartingale that converges to an arbitrarily large number on all paths where the inequality fails. Since $\mathscr{F}_{\mathscr{M}}$ becomes unbounded on those paths, this can be achieved by simply keeping $\mathscr{F}_{\mathscr{M}}$ constant as soon as it exceeds some (large) threshold α.

Corollary 18 *Fix any real $B > 0$, $0 < \varepsilon < B$ and $\alpha > 0$. Let $\mathscr{M}$ be a supermartingale such that $|\Delta\mathscr{M}| \leq B$. Then there is a positive supermartingale $\mathscr{F}_{\mathscr{M},\alpha}$ with $\mathscr{F}_{\mathscr{M},\alpha}(\square) = 1$ such that $\mathscr{F}_{\mathscr{M},\alpha}(x_{1:n}) \leq (3/2)^n$ for all $x_{1:n} \in \mathscr{S}$ and*

$$\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(\omega_{1:i})(\omega_{i+1}) > \varepsilon \Rightarrow \lim_{n\to+\infty}\mathscr{F}_{\mathscr{M},\alpha}(\omega_{1:n}) \geq \alpha,$$

for all $\omega \in \Omega$.

Proof Let $\mathscr{F}_{\mathscr{M}}$ be the positive supermartingale from Proposition 17. Then $\mathscr{F}_{\mathscr{M}}(\square) = 1$ and, for all $x_{1:n} \in \mathscr{S}$, we have that $\mathscr{F}_{\mathscr{M}}(x_{1:n}) \leq (3/2)^n$ and

$$\frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(x_{1:i})(x_{i+1}) \geq \varepsilon \Rightarrow \mathscr{F}_{\mathscr{M}}(x_{1:n}) \geq \exp\left(\frac{n\varepsilon^2}{4B^2}\right). \quad (7)$$

Now let $\mathscr{F}_{\mathscr{M},\alpha}$ be the real map on $\mathscr{S}$ that is equal to $\mathscr{F}_{\mathscr{M}}$ until it exceeds or equals α, at which point it remains constant. That is, for all $x_{1:n} \in \mathscr{S}$, let $\mathscr{F}_{\mathscr{M},\alpha}(x_{1:n}) := \mathscr{F}_{\mathscr{M}}(x_{1:n})$ if $\mathscr{F}_{\mathscr{M}}(x_{1:i}) < \alpha$ for all $i \in \{0,\dots,n-1\}$, and let $\mathscr{F}_{\mathscr{M},\alpha}(x_{1:n}) := \mathscr{F}_{\mathscr{M},\alpha}(x_{1:n-1})$ if $\mathscr{F}_{\mathscr{M}}(x_{1:i}) \geq \alpha$ for some $i \in \{0,\dots,n-1\}$. To see that $\mathscr{F}_{\mathscr{M},\alpha}$ is a supermartingale, observe that for all $x_{1:n} \in \mathscr{S}$, either $\Delta\mathscr{F}_{\mathscr{M},\alpha}(x_{1:n}) = \Delta\mathscr{F}_{\mathscr{M}}(x_{1:n})$ or $\Delta\mathscr{F}_{\mathscr{M},\alpha}(x_{1:n}) = 0$. Since $\mathscr{F}_{\mathscr{M}}$ is a supermartingale and $\overline{\mathrm{E}}_{x_{1:n}}(0) \leq 0$ because of C1, it therefore follows from Lemma 16 that $\mathscr{F}_{\mathscr{M},\alpha}$ is indeed a supermartingale. Next, observe that $\mathscr{F}_{\mathscr{M},\alpha}(\square) = \mathscr{F}_{\mathscr{M}}(\square) = 1$, and that for all $x_{1:n} \in \mathscr{S}$, there is some $m \leq n$ such that $\mathscr{F}_{\mathscr{M},\alpha}(x_{1:n}) = \mathscr{F}_{\mathscr{M}}(x_{1:m}) \leq (3/2)^m \leq (3/2)^n$. Finally, consider any $\omega \in \Omega$ such that $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}(\omega_{1:i})(\omega_{i+1}) > \varepsilon$. Then there are arbitrarily high m for which $\frac{1}{m}\sum_{i=0}^{m-1}\Delta\mathscr{M}(\omega_{1:i})(\omega_{i+1}) \geq \varepsilon$, and therefore, due to Equation (7), arbitrarily high m for which $\mathscr{F}_{\mathscr{M}}(\omega_{1:m}) \geq \exp(\frac{m\varepsilon^2}{4B^2})$. Hence, for any such ω, there is definitely some $m \in \mathbb{N}_0$ for which $\mathscr{F}_{\mathscr{M}}(\omega_{1:m}) \geq \alpha$, implying that $\mathscr{F}_{\mathscr{M},\alpha}(\omega_{1:n}) = \mathscr{F}_{\mathscr{M}}(\omega_{1:m})$ for all $n \geq m$ and therefore, that $\lim_{n\to+\infty}\mathscr{F}_{\mathscr{M},\alpha}(\omega_{1:n}) = \mathscr{F}_{\mathscr{M}}(\omega_{1:m}) \geq \alpha$. ■

Let us now replace Proposition 17 with Corollary 18 in the reasoning below Equation (6). We then end up with a positive supermartingale $\mathscr{F}_{\mathscr{M},\alpha}$ with initial capital 1 that converges and exceeds α on all paths where the inequality $\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=1}^n f(X_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + 2\varepsilon$ does not hold. To prove Proposition 12, the only thing left to do is therefore to replace α by $+\infty$ and ε by zero. As we show in our proof for Proposition 12, this can be achieved simultaneously. The key insight is to realise that the ε in Proposition 14—and hence in all our subsequent reasoning—can be chosen to be arbitrarily small, whereas the α in Corollary 18 can be arbitrarily large. By suitable combining the supermartingales that correspond to increasingly smaller ε and larger α, we finally arrive at a proof for our main result.

Proof of Proposition 12. Fix any $r \in \mathbb{N}$. Then we know from Proposition 14 that there is a real $B > 2^{-r}$, a supermartingale $\mathscr{M}_r$ with $|\Delta\mathscr{M}_r| \leq B$ and some $N \in \mathbb{N}$ such that

$$\frac{1}{n}\sum_{i=1}^n f(x_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + 2^{-r} + \frac{1}{n}\sum_{i=0}^{n-1}\Delta\mathscr{M}_r(x_{1:i})(x_{i+1})$$

for all $n \geq N$ and $x_{1:n} \in \mathscr{X}^n$, and therefore also,

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(\omega_i) \leq \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + 2^{-r} + \limsup_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \Delta\mathscr{M}_r(\omega_{1:i})(\omega_{i+1}) \quad (8)$$

for all $\omega \in \Omega$. Furthermore, due to Corollary 18 (with $\varepsilon = 2^{-r}$ and $\alpha = 2^r$), we know that there is a positive supermartingale $\mathscr{F}_r$ with $\mathscr{F}_r(\square) = 1$, $\mathscr{F}_r(x_{1:n}) \leq (3/2)^n$ for all $x_{1:n} \in \mathscr{S}$ and, for all $\omega \in \Omega$,

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \Delta\mathscr{M}_r(\omega_{1:i})(\omega_{i+1}) > 2^{-r} \Rightarrow \lim_{n\to+\infty} \mathscr{F}_r(\omega_{1:n}) \geq 2^r. \quad (9)$$

Let $\mathscr{F} := \sum_{r\in\mathbb{N}} 2^{-r}\mathscr{F}_r$. This map $\mathscr{F}$ on $\mathscr{S}$ is well-defined and positive because every $\mathscr{F}_r$ is positive. Furthermore, since we know that for all $r \in \mathbb{N}$, $\mathscr{F}_r(\square) = 1$ and $\mathscr{F}_r(x_{1:n}) \leq (3/2)^n$ for all $x_{1:n} \in \mathscr{S}$, and since $\sum_{r\in\mathbb{N}} 2^{-r} = 1$, it follows that also $\mathscr{F}(\square) = 1$ and $\mathscr{F}(x_{1:n}) \leq (3/2)^n$ for all $x_{1:n} \in \mathscr{S}$. Hence, in particular, $\mathscr{F}$ is a positive real function on $\mathscr{S}$ with $\mathscr{F}(\square) = 1$. On the other hand, it is also an extended real supermartingale because of [15, Lemma 12]. So it follows that $\mathscr{F}$ is a positive real supermartingale.

Consider now any path $\omega \in \Omega$ such that

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(\omega_i) > \overline{\mathrm{E}}_{\mathrm{av},\infty}(f). \quad (10)$$

Then clearly, there is some $R \in \mathbb{N}$ such that, for all $r \geq R$,

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(\omega_i) > \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) + 2\cdot 2^{-r},$$

and hence also, using Equation (8),

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \Delta\mathscr{M}_r(\omega_{1:i})(\omega_{i+1}) \geq \limsup_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} f(\omega_i) - \overline{\mathrm{E}}_{\mathrm{av},\infty}(f) - 2^{-r} > 2^{-r}.$$

If we combine this with Equation (9), it follows that $\lim_{n\to+\infty}\mathscr{F}_r(\omega_{1:n}) \geq 2^r$ for all $r \geq R$. Consider now any $m \in \mathbb{N}$. Then

$$\begin{aligned}\liminf_{n\to+\infty} \mathscr{F}(\omega_{1:n}) &= \liminf_{n\to+\infty} \sum_{r\in\mathbb{N}} 2^{-r}\mathscr{F}_r(\omega_{1:n}) \\ &\geq \liminf_{n\to+\infty} \sum_{i=0}^{m-1} 2^{-(R+i)}\mathscr{F}_{(R+i)}(\omega_{1:n}) \\ &= \sum_{i=0}^{m-1} 2^{-(R+i)} \lim_{n\to+\infty} \mathscr{F}_{(R+i)}(\omega_{1:n}) \\ &\geq \sum_{i=0}^{m-1} 2^{-(R+i)}2^{(R+i)} = \sum_{i=0}^{m-1} 1 = m\end{aligned}$$

using the positivity of $\mathscr{F}_r$ for the first inequality. Since $m \in \mathbb{N}$ was arbitrary, it follows that $\lim_{n\to+\infty}\mathscr{F}(\omega_{1:n}) = +\infty$.

We conclude that we have found a positive real supermartingale $\mathscr{F}$ with $\mathscr{F}(\square) = 1$ that converges to $+\infty$ on all paths $\omega \in \Omega$ that satisfy Equation (10). Due to Definition 6, this proves Proposition 12. ■

8. Conclusion and Future Work

When it comes to studying the average behaviour of imprecise Markov chains, the main focus has so far been on ergodic imprecise Markov chains, and the object of interest has always been the limit upper (or lower) expectation $\overline{\mathrm{E}}_\infty$ (or $\underline{\mathrm{E}}_\infty$). With good reason, because, as we explained in the introduction, these objects provide strictly almost sure upper and lower bounds on the long-term time averages of such systems [3, Theorem 32].

The main conclusion of this contribution, however, is that there is now a second and arguably better option, which is to instead focus on weakly ergodic imprecise Markov chains, and to use their averaged limit upper (and lower) expectation $\overline{\mathrm{E}}_{\mathrm{av},\infty}$ (and $\underline{\mathrm{E}}_{\mathrm{av},\infty}$). These too, as we have shown in Theorem 13, provide strictly almost sure upper and lower bounds on the long-term time averages of an imprecise Markov chain. Since our bounds are at least as tight—and sometimes significantly tighter [13, Example 2]—and since the condition of weak ergodicity is more easily satisfied [13], this yields a more powerful approach that is applicable to more models. Remarkably, it does not even matter which type of imprecise Markov chain one considers. Our results apply to measure- and game-theoretic versions alike and, in the measure-theoretic case, for all main notions of independence considered in the literature.

For these reasons, we think that future work on the average behaviour of imprecise Markov chains should focus on weak ergodicity and the corresponding averaged limit upper and lower expectations. The main practical challenge is to compute $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$. As we have seen in Proposition 11, Equation (2) provides a possible method. The scalability of this method remains to be assessed though, and alternative methods for computing $\overline{\mathrm{E}}_{\mathrm{av},\infty}(f)$ would be most welcome. On a more theoretical level, we would like to extend our results to also include a seventh model: the natural extension of the local models $\overline{\mathrm{E}}_{x_{1:n}}$ [18]. The weak continuity properties of that model will make a pointwise ergodic theorem infeasible, we think, but we believe it is possible to prove a finitary version instead, in the style of the weak law of large numbers.

Acknowledgments

The research of Jasper De Bock was partially funded by project number 3GO28919 of the FWO (Research Foundation - Flanders). The research of Natan T'Joens was supported and funded by the Special Research Fund (BOF) of Ghent University (reference number: 356).

References

[1] Fabio G. Cozman. Sets of probability distributions, independence, and convexity. *Synthese*, 186(2):577–600, 2012.

[2] Gert de Cooman, Filip Hermans, and Erik Quaeghebeur. Imprecise Markov chains and their limit behaviour. *Probability in the Engineering and Informational Sciences*, 23(4):597–635, 2009.

[3] Gert de Cooman, Jasper De Bock, and Stavros Lopatatzidis. Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems. *International Journal of Approximate Reasoning*, 76:18–46, 2016.

[4] Filip Hermans and Damjan Škulj. Stochastic Processes. In Thomas Augustin, Frank P.A. Coolen, Gert de Cooman, and Matthias C.M. Troffaes, editors, *Introduction to Imprecise Probabilities*, pages 258–278. Wiley, Chichester, 2014.

[5] John G. Kemeny and J. Laurie Snell. *Finite Markov chains*. Undergraduate Text in Mathematics. Springer, New York, 1976.

[6] Thomas Krak, Natan T'Joens, and Jasper De Bock. Hitting times and probabilities for imprecise markov chains. In *International Symposium on Imprecise Probabilities: Theories and Applications, ISIPTA 2019, Proceedings*, volume 103, pages 265–275. PMLR, 2019.

[7] Stavros Lopatatzidis. *Robust Modelling and Optimisation in Stochastic Processes using Imprecise Probabilities, with an Application to Queueing Theory*. PhD thesis, Ghent University, 2017.

[8] Stavros Lopatatzidis, Jasper De Bock, Gert de Cooman, Stijn De Vuyst, and Joris Walraevens. Robust queueing theory: an initial study using imprecise probabilities. *Queueing Systems*, 82: 75–101, 2016.

[9] G. Shafer and V. Vovk. *Game-Theoretic Foundations for Probability and Finance*. Wiley, Hoboken, New Jersey, 2019.

[10] Albert N. Shiryaev. *Probability*. Graduate Texts in Mathematics. Springer, New York, second edition, 1996.

[11] Damjan Škulj and Robert Hable. Coefficients of ergodicity for Markov chains with uncertain parameters. *Metrika*, 76(1):107–133, 2011.

[12] Natan T'Joens and Jasper De Bock. Global upper expectations for discrete-time stochastic processes: In practice, they are all the same! Accepted for publication in the Proceedings of ISIPTA 2021.

[13] Natan T'Joens and Jasper De Bock. Average behaviour in discrete-time imprecise Markov chains: A study of weak ergodicity. *International Journal of Approximate Reasoning*, 132:181–205, 2021.

[14] Natan T'Joens, Thomas Krak, Jasper De Bock, and Gert De Cooman. A recursive algorithm for computing inferences in imprecise markov chains. In *Proceedings of ECSQARU 2019*, volume 11726 of *Lecture Notes in Computer Science*, pages 455–465, 2019.

[15] Natan T'Joens, Jasper De Bock, and Gert de Cooman. Game-theoretic upper expectations for discrete-time finite-state uncertain processes. arXiv: 2008.03133, 2020.

[16] Natan T'Joens, Jasper De Bock, and Gert De Cooman. A particular upper expectation as global belief model for discrete-time finite-state uncertain processes. *International Journal of Approximate Reasoning*, 131: 30–55, 2021.

[17] Vladimir Vovk and Glenn Shafer. Game-Theoretic Probability. In Thomas Augustin, Frank P.A. Coolen, Gert de Cooman, and Matthias C.M. Troffaes, editors, *Introduction to Imprecise Probabilities*, pages 114–134. Wiley, Chichester, 2014.

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