Randomness and Imprecision: A Discussion of Recent Results

Gert de Cooman GERT.DECOOMAN@UGENT.BE
Jasper De Bock JASPER.DEBOCK@UGENT.BE
Foundations Lab for Imprecise Probabilities, Ghent University, Belgium

Abstract

We discuss our recent work on incorporating imprecision in the field of algorithmic randomness, based on the martingale-theoretic approach of game-theoretic probability. We consider several notions of randomness associated with interval, rather than precise, forecasting systems. We study their properties and argue that there are quite a number of reasons for wanting to do so. First, the richer mathematical structure in this generalisation provides a useful backdrop for a better understanding of precise randomness. Second, randomness associated with non-stationary precise forecasting systems can be captured by a constant but less precise interval forecast: greater model simplicity requires more imprecision. Third, imprecise randomness can't always be explained away as a result of (over)simplification: there are sequences that are random for a constant interval forecast, but never random for any computable (more) precise forecasting system. Incorporating imprecision into randomness therefore allows us to do more than was hitherto possible. Finally, the random sequences for a non-vacuous interval forecast constitute a meagre set, as they do for precise forecasts: imprecise and precise random sequences are equally rare from a topological point of view, and are, in that sense, equally interesting.

Keywords: Martin-Löf randomness, computable randomness, Schnorr randomness, computable stochasticity, imprecise probabilities, game-theoretic probability, interval forecast, supermartingale, computability, meagre set.

1. Introduction

This paper presents an overview of our work on incorporating imprecision into the study of randomness, where we aim at giving a precise mathematical meaning to, and study the mathematical consequences of, associating randomness with interval rather than precise probabilities. We believe it can provide a satisfactory answer to questions raised by a number of researchers [19, 20, 22, 50] about frequentist and 'objective' aspects of interval, or imprecise, probabilities. There are many notions of randomness [1, 4], but we focus here essentially on Martin-Löf, computable, and Schnorr randomness. We refer to the preprint [12] for a much more extensive and detailed version of this paper, with proofs for what we claim below, and to Ref. [13] for a much more limited report on our earlier efforts in this direction.

We consider an infinite sequence $\omega = (z_1, \dots, z_n, \dots)$, whose components z_k are either 0 or 1, and are considered as successive *outcomes* of some experiment. In the literature, the randomness of such a sequence ω is typically associated with a *forecasting system* φ that associates with each finite sequence of outcomes $(x_1, \dots, x_n)$ the (conditional) expectation $\varphi(x_1, \dots, x_n) = E(X_{n+1}|x_1, \dots, x_n)$ for the next, as yet unknown, outcome X_{n+1}. This $\varphi(x_1, \dots, x_n)$ is a (precise) *forecast* for the value of X_{n+1} after observing the values $x_1, \dots, x_n$ of the earlier outcomes $X_1, \dots, X_n$, and can be seen as a fair price for—and therefore a commitment to bet on—the unknown next outcome X_{n+1} after observing the first n outcomes $x_1, \dots, x_n$. The sequence ω is then 'random' when there is no 'allowable' strategy for getting infinitely rich by exploiting the bets made available by the forecasting system φ along the sequence, without borrowing. Betting strategies that are made available by the forecasting system φ are called supermartingales. Which supermartingales are considered 'allowable' differs in various approaches [1, 4, 18, 25, 34], but typically involves some (semi)computability requirement.

This martingale-theoretic, or algorithmic randomness, approach lends itself elegantly to allowing for interval rather than precise forecasts, and therefore to allowing for 'imprecision' in the definition of randomness. As we explain in Section 2, an 'imprecise' forecasting system φ associates with each finite sequence of outcomes $(x_1, \dots, x_n)$ a (conditional) expectation *interval* $\varphi(x_1, \dots, x_n)$ for the next outcome X_{n+1}. The lower bound of this *interval forecast* represents a supremum acceptable buying price, and its upper bound an infimum acceptable selling price, for the next outcome X_{n+1} [2, 44, 49]. This idea allows us to associate supermartingales with an interval forecasting system, and therefore in Section 3 to extend a number of existing notions of randomness to allow for interval, rather than precise, forecasts: we include in particular Martin-Löf, computable, and Schnorr randomness [1, 4, 18, 34]. We discuss interesting properties of these randomness notions in Section 4. In Section 5, we restrict our attention to *stationary* interval forecasts, as an extension of the more classical accounts of randomness, which typically consider a forecasting system with constant forecast $1/2$—corresponding to flipping a fair coin. In the precise case, a given sequence may not be random for any stationary forecast, but as we will see, for interval forecasting there typically is a filter of intervals that a sequence is random for. We show in Section 6 by means of a few examples that this filter may not have a smallest element, and even when it does, this smallest element may be a non-vanishing interval. These examples involve sequences that are random for some computable non-stationary precise forecast, but can't be random for a stationary forecast unless it becomes interval-valued, or imprecise. This might lead to the suspicion that this imprecision is perhaps only an artefact, which results from looking at non-stationary phenomena through an imperfect stationary lens. We continue the argument by showing that this suspicion is unfounded: there are sequences that are random for a stationary interval forecast, but not random for *any computable (more) precise* forecast, be it stationary or not. This serves to corroborate our claim that *there are forms of randomness that are irreducibly imprecise*. Finally, we argue in Section 7 that 'imprecise' randomness is an interesting extension of the existing notions of 'precise' randomness, because it is equally rare: just as for precise stationary forecasts, the set of all sequences that are random for a non-vacuous stationary interval forecast is *meagre*.

2. Preliminaries

We begin by introducing the preliminary notions needed to define and study randomness in the following sections.

2.1. The Forecasting Game

The dynamics of forecasting can be made clear, after the fashion first introduced by Shafer and Vovk [36, 37], by considering a game amongst three players, Forecaster, Sceptic and Reality. It involves a sequence of initially unknown outcomes X_1, X_2, ..., X_n, ... in the set of possible *outcomes* $\{0,1\}$. To stress that they are unknown, we call them *variables*, and use upper-case notation.

Each successive stage $n \in \mathbb{N}$ of the game consists of three steps. Here and in what follows, $\mathbb{N}$ is the set of all natural numbers, without zero, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

In a first step, *Forecaster* specifies an interval $I_n = [\underline{p}_n, \overline{p}_n] \subseteq [0,1]$ for the expectation of the as yet unknown outcome X_n in $\{0,1\}$—or equivalently, for the probability that $X_n = 1$. We interpret this so-called *interval forecast* I_n as a commitment for Forecaster to adopt $\underline{p}_n$ as his *supremum acceptable buying price* and $\overline{p}_n$ as his *infimum acceptable selling price* for the gamble (with reward function) X_n. This means that *Sceptic* can now in a second step take Forecaster up on any (combination) of the following commitments, whose (possibly negative) uncertain pay-offs are expressed in units of a linear utility: (i) for all real $q \leq \underline{p}_n$ and $\alpha \geq 0$, Forecaster is committed to accepting the gamble $\alpha[X_n - q]$, leading to an uncertain reward $-\alpha[X_n - q]$ for Sceptic;[1] and (ii) for all real $r \geq \overline{p}_n$ and $\beta \geq 0$, Forecaster is committed to accepting the gamble $\beta[r - X_n]$, leading to an uncertain reward $-\beta[r - X_n]$ for Sceptic. Finally, in a third step, *Reality* determines the value x_n of X_n in $\{0,1\}$, and the corresponding rewards $-\alpha[x_n - q]$ or $-\beta[r - x_n]$ are paid by Forecaster to Sceptic, who adds them to his current capital.

Elements x of $\{0,1\}$ are called *outcomes*, and elements p of the real unit interval $[0,1]$ will serve as (precise) *forecasts*. We denote by $\mathscr{I}$ the set of non-empty closed subintervals of the real unit interval $[0,1]$. Any element I of $\mathscr{I}$ will serve as an *interval forecast*. We will use the generic notation I for such an interval forecast, and $\underline{p} := \min I$ and $\overline{p} := \max I$ for its lower and upper bounds, respectively. An interval forecast $I = [\underline{p}, \overline{p}]$ is of course *precise* when $\underline{p} = \overline{p} =: p$, and we will then make no distinction between the singleton interval forecast $I = \{p\} \in \mathscr{I}$ and the corresponding precise forecast $p \in [0,1]$.

When Forecaster announces an interval forecast I_n, Sceptic can try and increase her capital by taking a gamble on the unknown outcome X_n. Any such *gamble* can be identified with a map $f_n \colon \{0,1\} \to \mathbb{R}$, and can therefore be represented as a point or vector $(f_n(1), f_n(0))$ in the two-dimensional vector space $\mathbb{R}^2$. $f_n(X_n)$ is then the (possibly negative) increase in Sceptic's capital in stage n of the game, as a function of the outcome variable X_n. Not every gamble $f_n(X_n)$ on the unknown outcome X_n will be available to Sceptic: which gambles she can take is determined by Forecaster's interval forecast I_n. As indicated above, they have the form $f_n(X_n) = -\alpha[X_n - q] - \beta[r - X_n]$, where α and β are non-negative real numbers, $q \leq \underline{p}_n$ and $r \geq \overline{p}_n$. They constitute a closed convex cone $\mathscr{A}_{I_n}$ in $\mathbb{R}^2$.

If we associate with any precise forecast $p \in [0,1]$ the *expectation* E_p, defined by $E_p(f) := pf(1) + (1-p)f(0)$ for any gamble $f \colon \{0,1\} \to \mathbb{R}$, and also consider the so-called *upper expectation* $\overline{E}_I$ associated with an interval forecast $I \in \mathscr{I}$, defined by

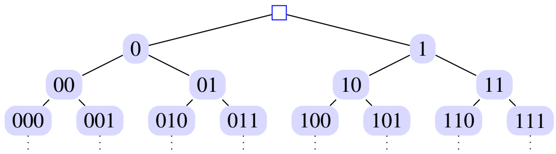
$$\overline{E}_I(f) := \max_{p \in I} E_p(f) = \begin{cases} E_{\overline{p}}(f) & \text{if } f(1) \geq f(0) \\ E_{\underline{p}}(f) & \text{if } f(1) \leq f(0) \end{cases} \quad \text{for any gamble } f \colon \{0,1\} \to \mathbb{R},$$

then *the closed convex cone $\mathscr{A}_{I_n}$ of all gambles $f_n(X_n)$ on the outcome X_n that are available to Sceptic at stage n, after Forecaster announces his interval forecast I_n, is completely determined by the condition* $\overline{E}_{I_n}(f_n) \leq 0$. When Reality then chooses a value x_n for X_n, this results in a (possibly negative) gain in capital $f_n(x_n)$ for Sceptic.

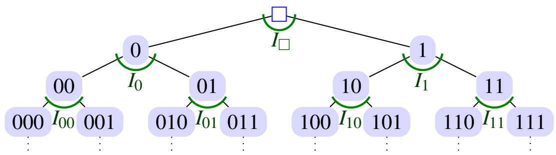
1. Because we allow $q \leq \underline{p}_n$ rather than $q < \underline{p}_n$, we actually see $\underline{p}_n$ as a *maximum* acceptable buying price, rather than a supremum one. We do this because it doesn't affect the conclusions, but it does simplify the mathematics and the discussion somewhat. Similarly for $r \geq \overline{p}_n$.

2.2. The Event Tree and Its Forecasting Systems

We call $(x_1, x_2, \dots, x_n, \dots)$ an outcome sequence, and we collect all possible outcome sequences in the set $\Omega := \{0,1\}^{\mathbb{N}}$. We collect the finite outcome sequences $x_{1:n} := (x_1, \dots, x_n)$ in the set $\mathbb{S} := \{0,1\}^* = \bigcup_{n\in\mathbb{N}_0}\{0,1\}^n$. The finite outcome sequences s in $\mathbb{S}$ and infinite outcome sequences ω in Ω constitute the nodes—also called *situations*—and *paths* in an event tree with unbounded horizon, part of which is depicted below. The empty sequence $x_{1:0} =: \square$ is also called the *initial* situation.



In the repeated game described above, Forecaster will only provide interval forecasts I_n after observing the actual sequence $(x_1, \dots, x_{n-1})$ that Reality has chosen, and the corresponding sequence of gambles $(f_1, \dots, f_{n-1})$ that Sceptic has chosen. This is the essence of so-called prequential forecasting [6, 7, 10]. For the present discussion, it will be more advantageous to consider an alternative, and in some aspects more involved, setting where a forecast I_s is specified in each of the possible situations s in the event tree $\mathbb{S}$; see the figure below.



We can use this idea to extend the notion of a forecasting system in Refs. [8, 48] from precise to interval forecasts.

Definition 1 (Forecasting System)
A forecasting system *is a map $\varphi\colon \mathbb{S} \to \mathscr{I}$, that associates an interval forecast $\varphi(s) \in \mathscr{I}$ with any situation s in the event tree $\mathbb{S}$. With any forecasting system φ we can associate two real processes $\underline{\varphi}$ and $\overline{\varphi}$, defined by $\underline{\varphi}(s) := \min \varphi(s)$ and $\overline{\varphi}(s) := \max \varphi(s)$ for all $s \in \mathbb{S}$. A forecasting system φ is called* precise *if $\underline{\varphi} = \overline{\varphi}$.*

Specifying a forecasting system φ requires that Forecaster should imagine in advance all the moves that Reality (and Sceptic) could make, and that he should devise in advance what forecast $\varphi(s)$ to give in each situation $s \in \mathbb{S}$.

We denote by Φ the set $\mathscr{I}^{\mathbb{S}}$ of all forecasting systems, and use the notation $\varphi \sqsubseteq \varphi^*$ to mean that the forecasting system φ^* is *at least as conservative* as φ, meaning that $\varphi(s) \subseteq \varphi^*(s)$ for all $s \in \mathbb{S}$.

2.3. Imprecise Probability Trees

Since in each situation s the interval forecast $I_s = \varphi(s)$ corresponds to a so-called *local* upper expectation $\overline{E}_{I_s}$, we can use the argumentation in our earlier papers [14, 16, 17] on imprecise stochastic processes to help φ turn the event tree into an *imprecise probability tree*, with an associated *global* upper expectation on paths, and a corresponding notion of 'almost surely' [14, 16, 17, 36, 37, 38, 47].

For any path $\omega \in \Omega$, the initial sequence that consists of its first n elements is a situation in $\{0,1\}^n$, denoted by $\omega_{1:n}$. Its n-th element belongs to $\{0,1\}$ and is denoted by ω_n. As a convention, we let its 0-th element be the *initial* situation $\omega_{1:0} = \omega_0 = \square$.

For any situation $s \in \mathbb{S}$ and path $\omega \in \Omega$, ω *goes through* s if there is some $n \in \mathbb{N}_0$ such that $\omega_{1:n} = s$. We denote by $\Gamma(s)$ the so-called *cylinder set* of all paths $\omega \in \Omega$ that go through s. We write $s \sqsubseteq t$, and say that the situation s *precedes* the situation t, when every path that goes through t also goes through s: $\Gamma(t) \subseteq \Gamma(s)$. We say that the situation s *strictly precedes* the situation t, and write $s \sqsubset t$, when $s \sqsubseteq t$ and $s \neq t$, or equivalently, when $\Gamma(t) \subset \Gamma(s)$.

For any situation $s = (x_1, \dots, x_n) \in \mathbb{S}$, we call $n = |s|$ its depth in the tree. Of course, $|s| \geq |\square| = 0$. Also, for any $x \in \{0,1\}$, we denote by sx the situation $(x_1, \dots, x_n, x)$.

A *process* F is a map defined on $\mathbb{S}$. A *real process* associates a real number $F(s) \in \mathbb{R}$ with every situation $s \in \mathbb{S}$. With any real process F, we can always associate a process ΔF, called the *process difference*. For every $s \in \mathbb{S}$, $\Delta F(s)$ is the gamble on $\{0,1\}$ defined by

$$\Delta F(s)(x) := F(sx) - F(s) \text{ for all } x \in \{0,1\}.$$

The *initial value* of a process F is its value $F(\square)$ in the situation $\square$. We call a real process *non-negative* if it is non-negative in all situations. Similarly, a *positive* real process is (strictly) positive in all situations. We call *test process* any non-negative real process F with $F(\square) = 1$.

We now look at a number of special real processes. In the imprecise probability tree associated with a *given* forecasting system φ, a *supermartingale* M for φ is a real process such that $\overline{E}_{\varphi(s)}(\Delta M(s)) \leq 0$ for all $s \in \mathbb{S}$. In other words, all supermartingale differences have non-positive upper expectation: supermartingales are real processes that Forecaster expects to decrease. We denote the set of all supermartingales for a given forecasting system φ by $\overline{\mathbb{M}}^{\varphi}$—whether a real process is a supermartingale depends of course on the forecasts in the situations.

The supermartingales for φ are effectively all the possible *capital processes* M for a Sceptic who starts with an initial capital $M(\square)$, and in each possible subsequent situation s selects a gamble $f_s = \Delta M(s)$ that is available there because of Forecaster's specification of the interval forecast $I_s = \varphi(s)$: $\overline{E}_{I_s}(f_s) \leq 0$. If Reality chooses the successive outcomes $x_1, \dots, x_n$, then Sceptic will end up in the corresponding situation $s = (x_1, \dots, x_n)$ with a capital

$$M(x_1, \dots, x_n) = M(\square) + \sum_{k=0}^{n-1} \underbrace{\Delta M(x_1, \dots, x_k)(x_{k+1})}_{=f_{(x_1,\dots,x_k)}(x_{k+1})}.$$

We call *test supermartingale* for φ any test process that is also a supermartingale for φ, or in other words, any non-negative supermartingale M for φ with initial value $M(\square) = 1$. It corresponds to Sceptic starting with unit capital and never borrowing. We collect all test supermartingales for φ in the set $\overline{\mathbb{T}}^{\varphi}$.

We also pay attention to a particular way of constructing test supermartingales. Define a *multiplier process* as a map D from $\mathbb{S}$ to *non-negative* gambles on $\{0,1\}$. Given such a multiplier process D, we can construct a test process $D^{\circledcirc}$ by the recursion equation

$$D^{\circledcirc}(sx) := D^{\circledcirc}(s)D(s)(x) \text{ for all } s \in \mathbb{S} \text{ and } x \in \{0,1\},$$

with $D^{\circledcirc}(\square) := 1$. We call $D^{\circledcirc}$ the test process *generated by* the multiplier process D. Any multiplier process D that satisfies the additional condition that $\overline{E}_{\varphi(s)}(D(s)) \leq 1$ for all $s \in \mathbb{S}$, is called a *supermartingale multiplier* for the forecasting system φ. The test process $D^{\circledcirc}$ generated by D is then a test supermartingale for φ.

2.4. Upper Expectations and Null Events

In the context of (imprecise) probability trees, any bounded real-valued map defined on the *sample space* Ω is called a *gamble* on Ω, or also a *global gamble*. An *event* A is a subset of Ω, and its *indicator* $\mathbb{I}_A$ is the gamble on Ω that assumes the value 1 on A and 0 elsewhere.

The supermartingales for a forecasting system φ allow us to associate a *global upper expectation* $\overline{E}^{\varphi}$ with φ:

$$\overline{E}^{\varphi}(g) := \inf\{M(\square) \colon M \in \overline{\mathbb{M}}^{\varphi} \text{ and } \liminf M \geq g\} \quad \text{for all gambles } g \text{ on } \Omega, \tag{1}$$

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

For extensive discussion about why the expression (1) is interesting and useful, we refer to Refs. [14, 17, 36, 37, 39, 40, 41, 43]. For our present purposes, it may suffice to mention that for precise forecasts, it leads to a model that coincides with the one found in measure-theoretic probability theory; see Refs. [36, Chapter 8] and [37, Chapter 9], as well as Ref. [43]. In particular, when all $I_s = \{1/2\}$, it coincides on all measurable global gambles with the usual uniform (Lebesgue) expectation. More generally, for an imprecise forecast $\varphi \in \Phi$, the upper expectation $\overline{E}^{\varphi}$ provides a tight upper bound on the measure-theoretic expectation of every precise forecasting system φ' that is compatible with φ in the sense that $\varphi' \sqsubseteq \varphi$ [39].

For an event $A \subseteq \Omega$, the corresponding *upper probability* is defined by $\overline{P}^{\varphi}(A) := \overline{E}^{\varphi}(\mathbb{I}_A)$. We call an event $A \subseteq \Omega$ *null* for a forecasting system φ if $\overline{P}^{\varphi}(A) = 0$. As usual, any property that holds, except perhaps on a null event, is said to hold *almost surely* for the forecasting system φ. We will then also say that *almost all paths have that property in the imprecise probability tree corresponding to* φ.

2.5. Computability

A *recursive* map $\psi \colon \mathbb{N}_0 \to \mathbb{N}_0$ is a map that can be computed by a Turing machine. By the Church–Turing (hypo)thesis, this is equivalent to the existence of an algorithm that, upon input of a number $n \in \mathbb{N}_0$, outputs the number $\psi(n) \in \mathbb{N}_0$. All notions of computability that we need are based on this notion, and we use the equivalent condition consistently. It is clear that in this definition, we can replace any of the $\mathbb{N}_0$ with any other countable set that is linked with $\mathbb{N}_0$ through a recursive bijection whose inverse is also recursive.

In what follows, we will need a notion of computable real processes, or in other words, computable real-valued maps $F \colon \mathbb{S} \to \mathbb{R}$ defined on the set $\mathbb{S}$ of all situations. Because there is an obvious recursive bijection between $\mathbb{N}_0$ and $\mathbb{S}$, whose inverse is also recursive, we can identify real processes and real sequences, and simply import, *mutatis mutandis*, the definitions for computable real sequences common in the literature [31, Chapter 0, Definition 5].

We call a net of rational numbers $r_{s,n}$ *recursive* if there are three recursive maps a, b, ς from $\mathbb{S} \times \mathbb{N}_0$ to $\mathbb{N}_0$ such that

$$b(s,n) > 0 \text{ and } r_{s,n} = (-1)^{\varsigma(s,n)} \frac{a(s,n)}{b(s,n)} \quad \text{for all } s \in \mathbb{S} \text{ and } n \in \mathbb{N}_0.$$

We call a real process $F \colon \mathbb{S} \to \mathbb{R}$ *computable* if there is a recursive net of rational numbers $r_{s,n}$ and a recursive map $e \colon \mathbb{S} \times \mathbb{N}_0 \to \mathbb{N}_0$ such that

$$n \geq e(s,N) \Rightarrow |r_{s,n} - F(s)| \leq 2^{-N} \quad \text{for all } s \in \mathbb{S} \text{ and } n, N \in \mathbb{N}_0.$$

A forecasting system φ is computable if the processes $\underline{\varphi}$ and $\overline{\varphi}$ are.

A real process F is *lower semicomputable* [34, 26] if it can be approximated from below by a recursive net of rational numbers, meaning that there is some recursive net of rational numbers $r_{s,n}$ such that

(i) $r_{s,n+1} \geq r_{s,n}$ for all $s \in \mathbb{S}$ and $n \in \mathbb{N}_0$;
(ii) $F(s) = \lim_{n\to\infty} r_{s,n}$ for all $s \in \mathbb{S}$.

We say that F is *upper semicomputable* if $-F$ is lower semicomputable. Computability can be related to lower and upper computability: a real process F is computable if and only if it is both lower and upper semicomputable. The set of all (semi)computable processes is countable; see for instance Ref. [48, Lemma 13]. The (semi)computability of multiplier processes is defined similarly, by replacing the domain $\mathbb{S}$ by $\mathbb{S} \times \{0,1\}$.

3. Several Notions of Randomness

We denote by $\mathbb{A}$ any *countable* set of test processes that includes the countable set of all computable positive test processes, which we denote by $\mathbb{A}_{\mathrm{C}}^{+}$. Examples of such sets $\mathbb{A}$ are:

$\mathbb{A}_{\mathrm{C}}^{+}$	all computable positive test processes
$\mathbb{A}_{\mathrm{C}}$	all computable test processes
$\mathbb{A}_{\mathrm{ML}}$	all lower semicomputable test processes
$\mathbb{A}_{\mathrm{ML}}^{\circledcirc}$	all test processes generated by lower semicomputable multiplier processes.

We call such test processes in $\mathbb{A}$ *allowable*. It holds that

$$\mathbb{A}_{\mathrm{C}}^{+} \subseteq \mathbb{A}_{\mathrm{C}} \text{ and } \mathbb{A}_{\mathrm{C}}^{+} \subseteq \mathbb{A}_{\mathrm{ML}}^{\circledcirc} \subseteq \mathbb{A}_{\mathrm{ML}}. \tag{2}$$

The test supermartingales for φ in this set $\mathbb{A}$ are called *allowable test supermartingales*, and collected in the set $\overline{\mathbb{T}}_{\mathbb{A}}^{\varphi} := \mathbb{A} \cap \overline{\mathbb{T}}^{\varphi}$. In particular, $\overline{\mathbb{T}}_{\mathrm{C}}^{\varphi,+} := \mathbb{A}_{\mathrm{C}}^{+} \cap \overline{\mathbb{T}}^{\varphi}$, $\overline{\mathbb{T}}_{\mathrm{C}}^{\varphi} := \mathbb{A}_{\mathrm{C}} \cap \overline{\mathbb{T}}^{\varphi}$, $\overline{\mathbb{T}}_{\mathrm{ML}}^{\varphi} := \mathbb{A}_{\mathrm{ML}} \cap \overline{\mathbb{T}}^{\varphi}$ and $\overline{\mathbb{T}}_{\mathrm{ML}}^{\varphi,\circledcirc} := \mathbb{A}_{\mathrm{ML}}^{\circledcirc} \cap \overline{\mathbb{T}}^{\varphi}$. Hereafter, unless explicitly stated to the contrary, $\mathbb{A}$ is an arbitrary but fixed set of allowable test processes.

We introduce several versions of randomness, each connected with a particular class of test supermartingales.

Definition 2 (Randomness)
Consider any forecasting system $\varphi\colon \mathbb{S} \to \mathscr{I}$ and any path $\omega \in \Omega$. We call ω $\mathbb{A}$-random for φ *if all (allowable) test supermartingales T in $\overline{\mathbb{T}}_{\mathbb{A}}^{\varphi}$ remain bounded above on ω, meaning that $\sup_{n\in\mathbb{N}} T(\omega_{1:n}) < \infty$. We then also say that the forecasting system φ* makes ω $\mathbb{A}$-random.

In other words, $\mathbb{A}$-randomness of a path means that there is no allowable strategy that starts with unit capital and avoids borrowing, and allows Sceptic to increase her capital without bounds by exploiting the bets on the outcomes along the path that are made available to her by Forecaster's specification of the forecasting system φ.

We let $\Phi_{\mathbb{A}}(\omega) := \{\varphi \in \Phi\colon \omega \text{ is } \mathbb{A}\text{-random for } \varphi\}$ denote the set of all forecasting systems that make the path ω $\mathbb{A}$-random. We will also use the special notations $\Phi_{\mathrm{C}}^{+}(\omega)$, $\Phi_{\mathrm{C}}(\omega)$, $\Phi_{\mathrm{ML}}^{\circledcirc}(\omega)$ and $\Phi_{\mathrm{ML}}(\omega)$ in the cases that $\mathbb{A}$ is equal to $\mathbb{A}_{\mathrm{C}}^{+}$, $\mathbb{A}_{\mathrm{C}}$, $\mathbb{A}_{\mathrm{ML}}^{\circledcirc}$ and $\mathbb{A}_{\mathrm{ML}}$, respectively.

When the forecasting system φ is precise and computable, and $\mathbb{A}$ is the set $\mathbb{A}_{\mathrm{ML}}$ of all lower semicomputable test processes, our definition reduces to that of *Martin-Löf randomness* on the Schnorr–Levin (martingale-theoretic) account [1, 4, 34, 35, 48]. We continue to call $\mathbb{A}_{\mathrm{ML}}$-randomness *Martin-Löf randomness* when the forecasting system φ is no longer precise or computable. Because $\mathbb{A}_{\mathrm{ML}}^{\circledcirc}$-randomness is weaker than Martin-Löf randomness, but has a similar flavour, we will also call it *weak Martin-Löf randomness*. When the forecasting system φ is precise and computable, and $\mathbb{A}$ is the set $\mathbb{A}_{\mathrm{C}}$ of all computable test processes, our definition reduces to that of *computable randomness* [1, 4]. We continue to call $\mathbb{A}_{\mathrm{C}}$-randomness *computable randomness* when the forecasting system φ is no longer precise or computable.

We also extend the notion of Schnorr randomness [34, 35] to our present context. To this end, we call a map $\rho\colon \mathbb{N}_0 \to \mathbb{N}_0$ a *growth function* if it is recursive, non-decreasing and unbounded, and call a real-valued map $\mu\colon \mathbb{N}_0 \to \mathbb{R}$ *computably unbounded* if there is some growth function ρ such that $\limsup_{n\to\infty}[\mu(n) - \rho(n)] > 0$.

Definition 3 (Schnorr Randomness)
Consider any forecasting system $\varphi\colon \mathbb{S} \to \mathscr{I}$. We call a path $\omega \in \Omega$ Schnorr random for φ *if no computable test supermartingale $T \in \overline{\mathbb{T}}_{\mathrm{C}}^{\varphi}$ for φ is computably unbounded on ω. We then also say that the forecasting system φ* makes ω Schnorr random, *and we collect all such forecasting systems in the set $\Phi_{\mathrm{S}}(\omega)$.*

4. Properties

The more conservative—imprecise—a forecasting system, the less stringent is the corresponding randomness notion.

Proposition 4 *Let ω be $\mathbb{A}$-random (respectively Schnorr random) for a forecasting system φ. Then ω is also $\mathbb{A}$-random (respectively Schnorr random) for any forecasting system φ^* such that $\varphi \sqsubseteq \varphi^*$.*

The larger a set $\mathbb{A}$ of allowable test processes, the more stringent is the corresponding randomness notion, and the 'fewer' $\mathbb{A}$-random paths there are. And Schnorr randomness is the weakest form of randomness considered here.

Proposition 5 *Consider two sets $\mathbb{A}, \mathbb{A}'$ of allowable test processes such that $\mathbb{A}' \subseteq \mathbb{A}$. If ω is $\mathbb{A}$-random for a forecasting system φ, then ω is also $\mathbb{A}'$-random for φ as well as Schnorr random, so $\Phi_{\mathbb{A}}(\omega) \subseteq \Phi_{\mathbb{A}'}(\omega) \subseteq \Phi_{\mathrm{S}}(\omega)$.*

As a consequence of Equation (2), we can infer from Proposition 5 that

$$\Phi_{\mathrm{ML}}(\omega) \subseteq \Phi_{\mathrm{ML}}^{\circledcirc}(\omega) \subseteq \Phi_{\mathrm{C}}^{+}(\omega) = \Phi_{\mathrm{C}}(\omega) \subseteq \Phi_{\mathrm{S}}(\omega). \tag{3}$$

As a special case, the (computable) *vacuous* forecasting system φ_{v} assigns the vacuous forecast $\varphi_{\mathrm{v}}(s) := [0,1]$ to all situations $s \in \mathbb{S}$. Clearly $\varphi \sqsubseteq \varphi_{\mathrm{v}}$ for all $\varphi \in \Phi$, so φ_{v} is the most conservative forecasting system. It corresponds to Forecaster making no actual commitments. This vacuous forecasting system can be used to conclude that for any path ω there are forecasting systems that make it random.

Proposition 6 *All paths are $\mathbb{A}$-random and Schnorr random for the vacuous forecasting system, so $\varphi_{\mathrm{v}} \in \Phi_{\mathbb{A}}(\omega) \subseteq \Phi_{\mathrm{S}}(\omega)$ for all $\omega \in \Omega$.*

We now turn to a number of important consistency results for the randomness notions we have introduced. We first show that any Forecaster who specifies a forecasting system is consistent in the sense that he believes himself to be well-calibrated: in the imprecise probability tree generated by his own forecasts, almost all paths will be random, so he is 'almost sure' that Sceptic won't be able to become infinitely rich by exploiting his—Forecaster's—forecasts.

Theorem 7 *Consider any forecasting system $\varphi\colon \mathbb{S} \to \mathscr{I}$. Then almost all paths are $\mathbb{A}$-random, and therefore also Schnorr random, for φ in the imprecise probability tree that corresponds to φ.*

This result guarantees in particular that there always are random paths, for any forecasting system, and leads to the following 'converse' to Proposition 6.

Corollary 8 *For any forecasting system φ there is at least one path that is $\mathbb{A}$-random, and therefore also Schnorr random, for φ.*

Theorem 9 below shows that if we concentrate on a *specific* path that is random, then the limsup average gain for Sceptic along *that* path for betting on a fixed gamble $h\colon \{0,1\} \to \mathbb{R}$ with rates provided by Forecaster is non-positive—Forecaster's liminf average loss is then non-negative. In this result, the average can be taken over any recursive selection of situations. To formalise this, we call any process that assumes values in $\{0,1\}$ a *selection process*. For any $k \in \mathbb{N}_0$, the situation $\omega_{1:k}$ is then selected on the path ω only if $S(\omega_{1:k}) = 1$.

Theorem 9 (Average Gains: Selection Processes) *Consider any computable forecasting system $\varphi\colon \mathbb{S} \to \mathscr{I}$, any path $\omega \in \Omega$ that is $\mathbb{A}$-random for φ, and the corresponding sequence $(I_1,\dots,I_n,\dots)$ of interval forecasts $I_n := \varphi(\omega_{1:n-1})$ for the path ω. If $S\colon \mathbb{S} \to \{0,1\}$ is a recursive selection process such that $\lim_{n\to\infty}\sum_{k=0}^{n} S(\omega_{1:k}) = \infty$, then*

$$\liminf_{n\to\infty} \frac{\sum_{k=0}^{n-1} S(\omega_{1:k})\left[h(\omega_{k+1}) - \underline{E}_{I_{k+1}}(h)\right]}{\sum_{k=0}^{n-1} S(\omega_{1:k})} \geq 0$$

for any gamble h on $\{0,1\}$.

For Schnorr randomness we can only prove a weaker result, involving the simpler notion of a *selection function* $\sigma\colon \mathbb{N} \to \{0,1\}$: at any 'time point' $k \in \mathbb{N}$, the outcome ω_k is selected along the path ω only if $\sigma(k) = 1$.

Theorem 10 (Average Gains: Selection Functions) *Consider any computable forecasting system $\varphi\colon \mathbb{S} \to \mathscr{I}$, any path $\omega \in \Omega$ that is $\mathbb{A}$-random for φ, and the corresponding sequence $(I_1,\dots,I_n,\dots)$ of interval forecasts $I_n := \varphi(\omega_{1:n-1})$ for the path ω. If σ is a recursive selection function such that $\lim_{n\to\infty}\sum_{k=1}^{n}\sigma(k) = \infty$, then*

$$\liminf_{n\to\infty} \frac{\sum_{k=1}^{n} \sigma(k)\left[h(\omega_k) - \underline{E}_{I_k}(h)\right]}{\sum_{k=1}^{n}\sigma(k)} \geq 0$$

for any gamble h on $\{0,1\}$. The same conclusion continues to hold when ω is Schnorr random for φ.

5. Stationary Forecasting Systems

We now turn to the special case where the interval forecasts $I \in \mathscr{I}$ are constant, and don't depend on the already observed outcomes. This leads to a generalisation of the classical case of randomness associated with a fair coin, which corresponds to $I = \{1/2\}$. For any interval $I \in \mathscr{I}$, we denote by $\gamma_I\colon \mathbb{S} \to \mathscr{I}$ the corresponding so-called *stationary* forecasting system that assigns the same interval forecast I to all situations: $\gamma_I(s) := I$ for all $s \in \mathbb{S}$.

In order to investigate the mathematical properties of imprecise randomness, we associate, with any path ω, the collection of all interval forecasts that make ω $\mathbb{A}$-random: $\mathscr{I}_{\mathbb{A}}(\omega) := \{I \in \mathscr{I} \colon \gamma_I \in \Phi_{\mathbb{A}}(\omega)\}$. We use the special notations $\mathscr{I}_{\mathrm{C}}^{+}(\omega)$, $\mathscr{I}_{\mathrm{C}}(\omega)$, $\mathscr{I}_{\mathrm{ML}}^{\circledcirc}(\omega)$ and $\mathscr{I}_{\mathrm{ML}}(\omega)$ in the cases that $\mathbb{A}$ is equal to $\mathbb{A}_{\mathrm{C}}^{+}$, $\mathbb{A}_{\mathrm{C}}$, $\mathbb{A}_{\mathrm{ML}}^{\circledcirc}$ and $\mathbb{A}_{\mathrm{ML}}$, respectively. Similarly, $\mathscr{I}_{\mathrm{S}}(\omega) := \{I \in \mathscr{I} \colon \gamma_I \in \Phi_{\mathrm{S}}(\omega)\}$.

5.1. Computable Stochasticity

We begin our study of the randomness associated with stationary forecasting systems by considering the behaviour of relative frequencies along random paths. Theorem 9 implies the consistency property in Corollary 11 below, which is a counterpart in our more general context of the notion of *computable stochasticity* or *Church randomness* in the precise fair-coin case where $I = \{1/2\}$ [1]. Interestingly, this corollary doesn't impose any computability requirements on the interval forecast I.

Computable stochasticity, or Church randomness, goes back to Alonzo Church's account of randomness [5]. He required of a random path ω that for any recursive selection process S such that $\sum_{k=0}^{n} S(\omega_{1:k}) \to \infty$,

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} S(\omega_{1:k})\omega_{k+1}}{\sum_{k=0}^{n-1} S(\omega_{1:k})} = \frac{1}{2}.$$

In other words, the relative frequencies of the ones—the successes—in the outcomes that S selects along the random path ω should converge to the constant probability $1/2$ of a success. It is well-known that all paths that are computably random—and therefore also all Martin-Löf random paths—for a stationary forecast $I = \{1/2\}$ are also Church random; see for instance Refs. [1, 51].

Our generalised notions of randomness no longer imply such convergence, but we're still able to conclude that the limits inferior and superior of the relative frequencies of the successes in the selected outcomes of a random path must lie in the forecast interval.

Corollary 11 (Church Randomness) *For any path $\omega \in \Omega$, any constant interval forecast $[\underline{p},\overline{p}] \in \mathscr{I}_{\mathbb{A}}(\omega)$ that makes ω $\mathbb{A}$-random, and any recursive selection process $S\colon \mathbb{S} \to \{0,1\}$ such that $\sum_{k=0}^{n} S(\omega_{1:k}) \to \infty$:*

$$\underline{p} \leq \liminf_{n\to\infty} \frac{\sum_{k=0}^{n-1} S(\omega_{1:k})\omega_{k+1}}{\sum_{k=0}^{n-1} S(\omega_{1:k})} \leq \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} S(\omega_{1:k})\omega_{k+1}}{\sum_{k=0}^{n-1} S(\omega_{1:k})} \leq \overline{p}.$$

That Corollary 11 needn't hold for Schnorr randomness, is in accordance with the fact that, in the particular fair-coin case where $I = \{1/2\}$, Schnorr randomness was shown by Wang [51] not to imply computable stochasticity either [51]. We can prove a weaker result for paths that are (only) Schnorr random, now based on Theorem 10.

Corollary 12 (Weak Church Randomness) *For any path $\omega \in \Omega$, any constant interval forecast $[\underline{p}, \overline{p}] \in \mathscr{I}_{\mathbb{A}}(\omega)$ that makes ω $\mathbb{A}$-random, and any recursive selection function σ such that $\lim_{n\to\infty} \sum_{k=0}^{n} \sigma(k) = \infty$:*

$$\underline{p} \leq \liminf_{n\to\infty} \frac{\sum_{k=1}^{n} \sigma(k)\omega_k}{\sum_{k=1}^{n} \sigma(k)} \leq \limsup_{n\to\infty} \frac{\sum_{k=1}^{n} \sigma(k)\omega_k}{\sum_{k=1}^{n} \sigma(k)} \leq \overline{p}.$$

The same conclusion continues to hold when the interval forecast $[\underline{p}, \overline{p}]$ makes ω Schnorr random.

If we were to strengthen the requirements on the selection processes S in Theorem 9 and Corollary 11 from 'being recursive' to 'being recursive and displaying recursive behaviour on the path ω under consideration', then the corresponding (weaker) computable stochasticity result would still hold for all Schnorr random paths. This is essentially what we do in Theorem 10 and Corollary 12. Any criticism of Schnorr randomness along the lines of Wang's argument [51] will therefore have to include an argumentation for why such a strengthening of the requirements on the selection processes isn't reasonable, or undesirable, or alternatively, why selection processes rather than selection functions appear in the requirements.

5.2. The Structure of the Interval Forecasts That Make a Path Random

It is guaranteed by Proposition 5 and Equation (3) that $\mathscr{I}_{\mathbb{A}}(\omega) \subseteq \mathscr{I}_{\mathrm{S}}(\omega)$ and

$$\mathscr{I}_{\mathrm{ML}}(\omega) \subseteq \mathscr{I}_{\mathrm{ML}}^{\circledcirc}(\omega) \subseteq \mathscr{I}_{\mathrm{C}}(\omega) = \mathscr{I}_{\mathrm{C}}^{+}(\omega) \subseteq \mathscr{I}_{\mathrm{S}}(\omega). \quad (4)$$

Most of our efforts here will be devoted to investigating the mathematical structure of these sets of interval forecasts.

As immediate consequences of the results in Section 4, we find that all these sets of interval forecasts associated with a random path are non-empty and increasing.

Proposition 13 (Non-emptiness) *For all $\omega \in \Omega$, $[0,1] \in \mathscr{I}_{\mathbb{A}}(\omega) \subseteq \mathscr{I}_{\mathrm{S}}(\omega)$, so any sequence of outcomes ω has at least one stationary forecast that makes it $\mathbb{A}$-random and therefore also Schnorr random.*

Proposition 14 (Increasingness) *For all $\omega \in \Omega$ and any $I, J \in \mathscr{I}$:*
(i) *if $I \in \mathscr{I}_{\mathbb{A}}(\omega)$ and $I \subseteq J$, then $J \in \mathscr{I}_{\mathbb{A}}(\omega)$;*
(ii) *if $I \in \mathscr{I}_{\mathrm{S}}(\omega)$ and $I \subseteq J$, then $J \in \mathscr{I}_{\mathrm{S}}(\omega)$.*

Corollary 12 allows us to derive the following consistency result: any collection of interval forecasts that make some path random must have a non-empty intersection.

Proposition 15 *For any $\omega \in \Omega$, $\mathscr{I}_{\mathbb{A}}(\omega)$ and $\mathscr{I}_{\mathrm{S}}(\omega)$ have the intersection property: any of their subsets has a non-empty intersection. In fact,*

$$\left[\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \omega_k, \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \omega_k\right] \subseteq \bigcap \mathscr{I}_{\mathbb{A}}(\omega) \subseteq \bigcap \mathscr{I}_{\mathrm{S}}(\omega). \quad (5)$$

Proposition 16 below guarantees, together with Proposition 14, that $\mathscr{I}_{\mathrm{C}}(\omega)$, $\mathscr{I}_{\mathrm{ML}}^{\circledcirc}(\omega)$ and $\mathscr{I}_{\mathrm{S}}(\omega)$ are *set filters*: increasing sets that are closed under finite intersections. We have no proof for a corresponding result for Martin-Löf randomness: it is an open problem whether the set of constant interval forecasts $\mathscr{I}_{\mathrm{ML}}(\omega)$ is closed under finite intersections, and therefore a set filter.

Proposition 16 *For any $\omega \in \Omega$, the sets of interval forecasts $\mathscr{I}_{\mathrm{ML}}^{\circledcirc}(\omega)$, $\mathscr{I}_{\mathrm{C}}(\omega)$ and $\mathscr{I}_{\mathrm{S}}(\omega)$ are closed under finite intersections.*

In these specific cases, any interval forecast that includes the non-empty closed intervals $\bigcap \mathscr{I}_{\mathbb{A}}(\omega) =: [\underline{p}_{\mathbb{A}}(\omega), \overline{p}_{\mathbb{A}}(\omega)]$ and $\bigcap \mathscr{I}_{\mathrm{S}}(\omega) =: [\underline{p}_{\mathrm{S}}(\omega), \overline{p}_{\mathrm{S}}(\omega)]$ strictly on both sides will make the path ω $\mathbb{A}$-random, respectively Schnorr random. We will see that it may depend on the case at hand whether the interval forecasts $[\underline{p}_{\mathbb{A}}(\omega), \overline{p}_{\mathbb{A}}(\omega)]$ and $[\underline{p}_{\mathrm{S}}(\omega), \overline{p}_{\mathrm{S}}(\omega)]$ themselves do the job: in the following sections, we will come across a number of examples where they do, and another example where they don't.

5.3. Examples at the Extreme Ends

We conclude the discussion in this section with a few immediate examples of possible sets of interval forecasts.

For any precise forecast $p \in [0,1]$, there always are paths ω that are $\mathbb{A}$-random, and at least as many that are Schnorr random, for the *precise* stationary forecasting system γ_p; see Corollary 8. A constant interval forecast I will make any such path ω $\mathbb{A}$-random if and only if it contains the precise forecast p: $\mathscr{I}_{\mathbb{A}}(\omega) = \{I \in \mathscr{I} : p \in I\}$; and similarly for Schnorr random paths.

On the other hand, any *recursive* path with infinitely many zeroes and ones will only be random for the vacuous interval forecast.

Proposition 17 *If a path $\omega \in \Omega$ is recursive and has infinitely many zeroes and infinitely many ones, then the only interval forecast that makes ω $\mathbb{A}$-random, or Schnorr random, is the vacuous one: $\mathscr{I}_{\mathbb{A}}(\omega) = \mathscr{I}_{\mathrm{S}}(\omega) = \{[0,1]\}$.*

The examples in the next section will show that, in between these extremes of total imprecision and maximal precision, there lies an uncharted realm of paths whose unpredictability is 'similar' to that of the ones traditionally called 'random', but for which $0 < \underline{p}_{\mathbb{A}}(\omega) < \overline{p}_{\mathbb{A}}(\omega) < 1$, and similarly, $0 < \underline{p}_{\mathrm{S}}(\omega) < \overline{p}_{\mathrm{S}}(\omega) < 1$.

6. Why We Claim That Some Types of Randomness Are Irreducibly Imprecise

We have learnt from our work on imprecise Markov chains [11, 15, 17, 24, 42] that we can often compute tight bounds on expectations in non-stationary precise Markov chains very efficiently by replacing them with stationary but imprecise versions. Similarly, in statistical modelling, when learning from data sampled from a distribution with a varying (non-stationary) parameter, it is quite a challenge to estimate the time sequence of its values, but we may be more successful in learning about its (stationary) interval range. Such ideas also lie behind the proposal by Fierens *et al.* [20] of a frequentist interpretation for imprecise probability models, based on non-stationarity.

Here, we explore this idea in the context of our study of imprecise randomness, and illustrate in a number of interesting examples that randomness associated with non-stationary precise forecasting systems can be captured by a stationary forecasting system, which must then be less precise: we gain simplicity of representation by going from a non-stationary to a stationary one, but we *must* then pay for it by losing precision.

We start with a simple example to introduce the basic idea. Fix any p,q in $[0,1]$ with $p<q$, and any path ω that is $\mathbb{A}$-random for the forecasting system $\varphi_{p,q}$, defined by

$$\varphi_{p,q}(s) := \begin{cases} p & \text{if } |s| \text{ is odd} \\ q & \text{if } |s| \text{ is even} \end{cases} \qquad \text{for all } s \in \mathbb{S}.$$

Corollary 8 guarantees that there is at least one such path. Then $\mathscr{I}_{\mathbb{A}}(\omega) = \mathscr{I}_{\mathrm{S}}(\omega) = \{I \in \mathscr{I} \colon [p,q] \subseteq I\}$.

We next look at sequences that are 'nearly' random for the constant precise forecast $1/2$, but not quite. Consider the following sequence $\{p_n\}_{n\in\mathbb{N}_0}$ of precise forecasts:

$$p_n := \frac{1}{2} + (-1)^n \delta_n \text{ with } \delta_n := \sqrt{\frac{8}{n+33}} \quad \text{for all } n \in \mathbb{N}_0.$$

We see that $p_n \to 1/2$ and that $p_n \in (0,1)$ for all $n \in \mathbb{N}_0$. Focus on an arbitrary but fixed path ω that is $\mathbb{A}_{\mathrm{ML}}$-random for the computable precise forecasting system $\varphi_{\sim 1/2}$ with

$$\varphi_{\sim 1/2}(s) := p_{|s|} \text{ for all } s \in \mathbb{S}.$$

There is at least one such path, by Corollary 8. Then for all $\mathbb{A}$ such that $\mathbb{A}_{\mathrm{C}}^{+} \subseteq \mathbb{A} \subseteq \mathbb{A}_{\mathrm{ML}}$:

$$\mathscr{I}_{\mathbb{A}}(\omega) = \mathscr{I}_{\mathrm{S}}(\omega) = \left\{[\underline{p},\overline{p}] \in \mathscr{I} \colon \underline{p} < 1/2 < \overline{p}\right\}.$$

These two examples indicate that randomness associated with a non-stationary precise forecasting system can also be 'described' as randomness for a simpler, stationary but then necessarily imprecise, forecasting system. They might lead us to suspect that all stationary imprecise forms of randomness could be 'explained away' as such simpler representations of non-stationary but precise forms of randomness. This would imply that the imprecision in the stationary forecasts isn't essential, and can always be dismissed as a necessary consequence of using a stationary representation that isn't powerful enough to allow for the ideal, precise but non-stationary, representation.

We will now argue that this suspicion is misguided, and in fact wrong when we focus on computable forecasting systems: we outline in the theorem below that there are paths that are random for a (computable) stationary interval forecasting system but never for any computable precise forecasting system, be it stationary or not. This serves to corroborate our claim that there is randomness that is irreducibly imprecise, as its imprecision can't be explained away as an effect of oversimplification. The imprecision involved is furthermore non-negligible, and can be made arbitrarily large, because besides excluding the possibility of randomness of such paths for precise computable forecasting systems, we also show they can't be random for any computable forecasting system whose highest imprecision is smaller than that of the original, stationary one.

Theorem 18 (Irreducible Imprecision)
Consider any set of allowable test processes $\mathbb{A}$, and any interval forecast $[\underline{p},\overline{p}] \in \mathscr{I}$. Then there is path $\omega \in \Omega$ that is $\mathbb{A}$-random—and therefore also Schnorr random—for the stationary interval forecast $[\underline{p},\overline{p}]$, but that is never Schnorr random—and therefore never $\mathbb{A}$-random—for any computable forecasting system φ whose highest imprecision is smaller than that of $[\underline{p},\overline{p}]$, in the specific sense that $\sup_{s\in\mathbb{S}}\left[\overline{\varphi}(s) - \underline{\varphi}(s)\right] < \overline{p} - \underline{p}$.

For an example showing that the computability condition in this result can't be dropped, and a discussion on the theoretical and practical relevance of this condition, we refer to recent work by Persiau and us [30].

7. Random Sequences Are Topologically Rare

Theorem 7 tells us that the set of all random paths for a forecasting system has lower probability one—since its complement has upper probability zero—so there are many such random paths in a '*measure-theoretic*' sense. But we will see presently that, in a *topological* sense, random paths are also 'few', in the sense that they typically constitute only a meagre set. This is a known result for precise randomness, that was, as far as we can judge, first formulated in the context of a much more encompassing discussion on the nature of randomness by Muchnik, Semenov and Uspensky [28], who showed that the set of all paths that correspond to a precise stationary forecast is meagre. It also appeared in a related form in the wake of discussions [3, 9, 29, 32, 33] of Philip Dawid's calibration papers [6, 8], and was foreshadowed by some of Terry Fine's results [21].

Here is the essence of Muchnik, Semenov and Uspensky's argument [28]. They call a path ω *lawful* if there is some algorithm that, given as input any situation s on the path ω, outputs a non-trivial finite set $R(s)$ of situations $t \sqsupset s$ such that one of these 'extensions' t is also on the path—meaning that $\omega \in \Gamma(t)$. By 'non-trivial', they mean that $R(s)$ is restrictive: it actually eliminates possible extensions. They then go on to show that the set of all lawful paths is meagre, and finally, that random paths, because they satisfy the law of large numbers, are lawful.

We now show that we can extend this argument to imprecise stationary forecasts. First of all, let us give a definition of lawfulness that makes the formulation above more precise. A *partial* function on a domain D is a function that need not be defined on all elements of D.

Definition 19 (Lawfulness [28, Definition 2.1]) *We call* algorithm *any recursive (partial) function R from $\mathbb{S}$ to the collection of finite subsets of $\mathbb{S}$. A path $\omega \in \Omega$ is called* lawful *for an algorithm R if for all $m \in \mathbb{N}_0$:*

(i) *R is defined in the situation $\omega_{1:m}$;*
(ii) *$R(\omega_{1:m})$ is a non-empty finite subset of $\mathbb{S}$ such that $\omega_{1:m} \sqsubset t$ for all $t \in R(\omega_{1:m})$;*
(iii) *$R(\omega_{1:m})$ is* non-trivial*: $\bigcup_{t \in R(\omega_{1:m})} \Gamma(t) \subset \Gamma(\omega_{1:m})$;*
(iv) *there is some $t \in R(\omega_{1:m})$ such that $\omega \in \Gamma(t)$.*

A path $\omega \in \Omega$ is called lawful *if it is lawful for some algorithm R. A path that isn't lawful is called* lawless.

A set of paths $A \subseteq \Omega$ is *nowhere dense in* Ω [28] if for every $s \in \mathbb{S}$, there is some $t \in \mathbb{S}$ such that $s \sqsubseteq t$ and $A \cap \Gamma(t) = \emptyset$. A set of paths $B \subseteq \Omega$ is then called *meagre*, or *first category*, it it is a countable union of nowhere dense sets. We rely on the following central result in Ref. [28].

Theorem 20 ([28, Corollary 2.3]) *Any subset of Ω containing only lawful paths is meagre.*

To see that a set of random paths is meagre, it therefore suffices to prove that these random paths are all lawful. This turns out to be not too difficult, because relative frequencies along lawless paths behave very 'wildly'.

Proposition 21 *Let $\omega \in \Omega$ be a lawless path. Then*

$$\liminf_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \omega_k = 0 \text{ and } \limsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \omega_k = 1.$$

So, in order to derive our result, it now suffices to consider that relative frequencies along random paths can't behave so wildly, because they are constrained by our 'weak computable stochasticity' result in Corollary 12. *Random paths are typically lawful.*

Theorem 22 *Let $I = [\underline{p}, \overline{p}] \in \mathscr{I}$ be any closed subinterval of $[0,1]$ strictly included in $[0,1]$, so $\underline{p} > 0$ or $\overline{p} < 1$. Then the set of all paths that are $\mathbb{A}$-random for the stationary forecasting system γ_I is meagre. Similarly, the set of all Schnorr random paths for γ_I is meagre.*

We see that the important distinction for random paths lies not between precise and imprecise stationary forecasts, but rather between vacuous and non-vacuous ones: for any non-vacuous stationary forecast, the set of random paths is meagre, whereas for the vacuous forecast, all paths are random, and therefore the corresponding set of random paths is *co-meagre*—the complement of a meagre set.

It also suggests that the paths that are random for non-vacuous interval forecasts are 'equally rare' as those that are random for precise forecasts, which, we believe, only adds to their mathematical interest.

8. Conclusion

There have been a number of papers [19, 20, 22, 50] that aim to introduce imprecision for probabilities that have a physical, or frequentist, interpretation. The present paper wants to continue that tradition.

We believe the work described here is a first systematic attempt at reconciling imprecision with the study of algorithmic randomness along the lines of von Mises [46], Church [5], Kolmogorov [23], Ville [45], Martin-Löf [27], Levin [25] and Schnorr [34, 35]. Our results indicate that this is possible and interesting.

Besides the sequences that are random for precise forecasts, new realms of sequences arise that are random only for interval forecasts, and have interesting properties. Topologically speaking, they are as rare as their precise counterparts, as they also constitute meagre sets. Our examples show that incorporating imprecision into the study of randomness allows for a richer mathematical structure to arise, and our treatment allows us to better understand, as special cases, the existing results in the precise limit.

On the one hand, 'imprecise randomness' arises as a useful stationary model simplification when dealing with non-stationarity. But, we have also shown that it has a more fundamental role, as there are sequences that are random for a given computable interval forecast, but not for any computable (more) precise forecast.

All this leads us to the conviction that there is more to randomness than the classical account for precise forecasts seems to suggest.

Acknowledgments

We are grateful to Philip Dawid and Volodya Vovk for their inspiring and helpful comments and guidance. Gert's research and travel were partly funded through project number G012512N of the Research Foundation – Flanders (FWO). Jasper was a Post-Doctoral Fellow of the FWO when much of this research was being done, and he wishes to acknowledge its financial support. In more recent years, Gert and Jasper's work was also supported by H2020-MSCA-ITN-2016 UTOPIAE, grant agreement 722734.

Author Contributions

As with most of our joint work, there is no telling, after a while, which of us had what idea, or did what, exactly. We have both contributed equally to this paper. But since a paper must have a first author, we decided it should be the one who took the first significant steps: Gert, in this case.

References

[1] Klaus Ambos-Spies and Antonín Kucera. Randomness in computability theory. *Contemporary Mathematics*, 257:1–14, 2000.

[2] Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes, editors. *Introduction to Imprecise Probabilities*. John Wiley & Sons, 2014.

[3] Gordon Belot. Failure of calibration is typical. *Statistics and Probability Letters*, 83(10):2316–2318, 2013.

[4] Laurent Bienvenu, Glenn Shafer, and Alexander Shen. On the history of martingales in the study of randomness. *Electronic Journal for History of Probability and Statistics*, 5, 2009.

[5] Alonzo Church. On the concept of a random sequence. *Bulletin of the American Mathematical Society*, 46(2): 130–136, 1940.

[6] A. Philip Dawid. The well-calibrated Bayesian. *Journal Of The American Statistical Association*, 77(379): 605–610, 1982.

[7] A. Philip Dawid. Statistical theory: The prequential approach. *Journal of the Royal Statistical Society, Series A*, 147:278–292, 1984.

[8] A. Philip Dawid. Calibration-based empirical probability. *Annals Of Statistics*, 13(4):1251–1274, 1985.

[9] A. Philip Dawid. Self-calibrating priors do not exist: Comment. *Journal of the American Statistical Association*, 80(390):340–341, 1985.

[10] A. Philip Dawid and Vladimir G. Vovk. Prequential probability: principles and properties. *Bernoulli*, 5: 125–162, 1999.

[11] Jasper De Bock, Alexander Erreygers, and Thomas Krak. Sum-product laws and efficient algorithms for imprecise Markov chains. Submitted for publication in the Proceedings of UAI 2021.

[12] Gert de Cooman and Jasper De Bock. Randomness is inherently imprecise, 2021. ArXiv: 2103.00071 [math.PR].

[13] Gert de Cooman and Jasper De Bock. Computable randomness is inherently imprecise. In Alessandro Antonucci, Giorgio Corani, Inés Couso, and Sébastien Destercke, editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 133–144. PMLR, 10–14 July 2017.

[14] Gert de Cooman and Filip Hermans. Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence*, 172(11):1400–1427, 2008.

[15] 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.

[16] Gert de Cooman, Jasper De Bock, and Stavros Lopatatzidis. A pointwise ergodic theorem for imprecise Markov chains. In Thomas Augustin, Serena Doria, Enrique Miranda, and Erik Quaeghebeur, editors, *ISIPTA '15 – Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 107–116. Aracne Editrice, 2015.

[17] 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.

[18] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic Randomness and Complexity*. Springer, 2010.

[19] Pablo I. Fierens. An extension of chaotic probability models to real-valued variables. *International Journal of Approximate Reasoning*, 50(4):627–641, 2009.

[20] Pablo I. Fierens, Leandro C. Rego, and Terrence L. Fine. A frequentist understanding of sets of measures. *Journal of Statistical Planning and Inference*, 139(6): 1879–1892, 2009.

[21] Terrence L. Fine. On the apparent convergence of relative frequency and its implications. *IEEE Transactions on Information Theory*, 16(3):251–257, 1970.

[22] Igor I. Gorban. *The Statistical Stability Phenomenon*. Springer, 2016.

[23] Andrei N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:1–7, 1965.

[24] Thomas Krak, Jasper De Bock, and Arno Siebes. Imprecise continuous-time Markov chains. *International Journal of Approximate Reasoning*, 88:452–528, 2017.

[25] Leonid A. Levin. On the notion of a random sequence. *Soviet Math. Doklady*, 14(5):1413–1416, 1973.

[26] Ming Li and Paul M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, 1993.

[27] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9(6):602–619, 1966.

[28] Andrei A. Muchnik, Alexei L. Semenov, and Vladimir A. Uspensky. Mathematical metaphysics of randomness. *Theoretical Computer Science*, 207(2): 263–317, 1998.

[29] David Oakes. Self-calibrating priors do not exist. *Journal of the American Statistical Association*, 80 (390):339, 1985.

[30] Floris Persiau, Jasper De Bock, and Gert de Cooman. A remarkable equivalence between non-stationary precise and stationary imprecise uncertainty models in computable randomness. In *Proceedings of ISIPTA 2021*, Granada, Spain, 2021. Accepted for publication.

[31] Marian Boykan Pour-El and Jonathan Ian Richards. *Computability in Analysis and Physics*. Springer, 1989.

[32] Mark J. Schervish. Calibration-based empirical probability – discussion. *The Annals of Statistics*, 13(4): 1274–1282, 1985.

[33] Mark J. Schervish. Self-calibrating priors do not exist: Comment. *Journal of the American Statistical Association*, 80(390):341–342, 1985.

[34] Claus Peter Schnorr. *Zufälligkeit und Wahrscheinlichkeit: Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*. Springer, 1971.

[35] Claus Peter Schnorr. Process complexity and effective random tests. *Journal of Computer and System Sciences*, 7(4):376–388, 1973.

[36] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, 2001.

[37] Glenn Shafer and Vladimir Vovk. *Game-Theoretic Foundations for Probability and Finance*. Wiley, 2019.

[38] Glenn Shafer, Vladimir Vovk, and Akimichi Takemura. Lévy's zero–one law in game-theoretic probability. *Journal of Theoretical Probability*, 25:1–24, 2012.

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

[40] Natan T'Joens, Jasper De Bock, and Gert de Cooman. Continuity properties of game-theoretic upper expectations, 2019. ArXiv: 1902.09406 [math.PR].

[41] Natan T'Joens, Jasper De Bock, and Gert de Cooman. In search of a global belief model for discrete-time uncertain processes. *Proceedings of Machine Learning Research*, 103:377–385, 2019.

[42] Natan T'Joens, Thomas Krak, Jasper De Bock, and Gert de Cooman. A recursive algorithm for computing inferences in imprecise Markov chains. In *ECSQARU 2019: Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 455–465. Springer, 2019.

[43] 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.

[44] Matthias C. M. Troffaes and Gert de Cooman. *Lower Previsions*. Wiley, 2014.

[45] J. Ville. *Étude critique de la notion de collectif*. Gauthier-Villars, 1939.

[46] Richard von Mises. *Probability, Statistics and Truth*. Dover, second revised edition edition, 1981.

[47] 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*. Wiley, 2014.

[48] Vladimir G. Vovk and Alexander Shen. Prequential randomness and probability. *Theoretical Computer Science*, 411(29–30):2632–2646, 2010.

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

[50] Peter Walley and Terrence L. Fine. Towards a frequentist theory of upper and lower probability. *Annals of Statistics*, 10:741–761, 1982.

[51] Yongge Wang. *Randomness and Complexity*. PhD thesis, Naturwissenschaftlich-Mathematischen Gesamtfakultät, Ruprecht-Karls-Universität, Heidelberg, 1996.