



HAL
open science

Applications of Effective Probability Theory to Martin-Löf Randomness

Mathieu Hoyrup, Cristobal Rojas

► **To cite this version:**

Mathieu Hoyrup, Cristobal Rojas. Applications of Effective Probability Theory to Martin-Löf Randomness. 36th International Colloquium on Automata, Languages and Programming - ICALP 2009, Jul 2009, Rhodes, Greece. pp.549-561, 10.1007/978-3-642-02927-1 . hal-00425560v1

HAL Id: hal-00425560

<https://hal.science/hal-00425560v1>

Submitted on 22 Oct 2009 (v1), last revised 5 Nov 2009 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Applications of Effective Probability Theory to Martin-Löf Randomness

Mathieu Hoyrup¹ and Cristóbal Rojas²

¹ LORIA - 615, rue du jardin botanique, BP 239

54506 Vandœuvre-lès-Nancy, FRANCE, hoyrup@loria.fr

² Institut de Mathématiques de Luminy, Campus de Luminy, Case 907

13288 Marseille Cedex 9, FRANCE, rojas@iml.univ-mrs.fr

Abstract. We pursue the study of the framework of *layerwise computability* introduced in a preceding paper and give three applications.

(i) We prove a general version of Birkhoff's ergodic theorem for random points, where the transformation and the observable are supposed to be *effectively measurable* instead of *computable*. This result significantly improves V'yugin and Nandakumar's ones. (ii) We provide a general framework for deriving sharper theorems for random points, sensitive to the speed of convergence. This offers a systematic approach to obtain results in the spirit of Davie's ones. (iii) Proving an effective version of Prokhorov theorem, we positively answer a question recently raised by Fouché: can random Brownian paths reach any random number? All this shows that *layerwise computability* is a powerful framework to study Martin-Löf randomness, with a wide range of applications.

1 Introduction

Algorithmic randomness emerged as an early achievement of Kolmogorov's program to base probability theory on the theory of computing. Yet a framework allowing the combination of these two theories is still lacking: for instance, computable analysis is mainly concerned with effective versions of topological notions, and not probabilistic/measure-theoretic ones. For this reason, the study of algorithmic randomness has not reached its expected range of application: general probability theory. Let us recall the main contributions of algorithmic randomness to probability theory developed so far.

Theorems for random points. The main novelty brought by algorithmic randomness is that probabilistic laws can be strengthened in principle, holding at every random point and not only with probability one. Classical examples can be found in [1–3] for instance. When proving this kind of result the key hypothesis is the *computability* of the random variables involved. However, it is well-known that computability notions are the effective versions of topological ones (the computable functions are precisely the effectively continuous ones, the semi-decidable sets are precisely the effectively open sets, and so on). Hence the computability assumption on random variables is (i) inappropriate in principle,

as probability theory is grounded on measure theory and not on topology; (ii) a priori too strong, as in the classical setting only properties as measurability, integrability are required. This leads to the following:

Problem 1. Theorems for random points should hold for “effectively measurable” objects and not only computable ones.

This problem has already been independently investigated in [4, 5] where ergodic theorems for random points are proved for different types of “almost everywhere computable” functions. These works are, however, still far from catching the effective version of measurable functions. For instance in Birkhoff’s ergodic theorem, nothing can be said about the mean sojourn time of algorithmically random points in fractal sets having effective constructions, as the *Smith-Volterra-Cantor* (or *fat Cantor*) set $A \subseteq [0, 1]$, which is homeomorphic to the Cantor set and has Lebesgue measure $\frac{1}{2}$.

Information given by the randomness degree. A further contribution of algorithmic randomness to probability theory consists in making use of the “randomness degree” of a random point x to get additional information about the way x satisfies a given probabilistic law. For instance in [6], the speed of convergence in the Strong Law of Large Numbers is computed from the *compressibility coefficient*, or *deficiency of randomness* of each random sequence. This kind of result gives a much sharper insight into probabilistic phenomena and, we believe, new tools are needed in order to make this approach systematic and applicable on abstract spaces:

Problem 2. Having a general framework to get sharper theorems for random points, using the information given by the randomness degree.

Layerwise computability. In [7], working in the context of computable probability spaces (to which Martin-Löf randomness has been recently extended, see [8, 9]), effective versions of measure-theoretic notions were examined and another contribution of algorithmic randomness to probability theory was developed: the setting of a new framework for computability adapted to the probabilistic context. This was achieved by making a fundamental use of the existence of a universal Martin-Löf test to endow the space with what we call the *Martin-Löf layering*. In this new framework, which we call **layerwise computability**, the *layerwise* versions of virtually all computability notions can be naturally defined. The contributions of this setting can be summarized in the following principle, supported by the main results in [7]:

Correspondence Principle (CP). Under effectivity assumptions, measure-theoretic notions correspond exactly to *layerwise* versions of topological ones.

Intuitively, this gives evidence that the *layering structure* grasps a large part of the probabilistic phenomena: each probabilistic notion, that by nature intimately depends on the underlying measure μ , can be expressed without referring

to μ but only to its imprint on the space, captured by the *layering*. In this paper, elaborating on [7] and developing layerwise computability further, we give solutions to Problems 1 and 2. The CP is at the core of these solutions, that we briefly present now.

Solution to Problem 1. We prove general versions of theorems for random points and *effectively measurable* random variables, in particular Birkhoff's ergodic theorem. This is a significant improvement of [4, 5] as it implies in particular a positive result for the Smith-Volterra-Cantor set. To prove these results we develop tools allowing to adapt the existent techniques (used in the *computable* context) to the *layerwise computable* context. Then, the results for *effectively measurable* objects follow from the CP. This strategy is very general and applicable in a wide range of situations.

Solution to Problem 2. As a further illustration of the CP we prove that under effectivity assumptions, *almost everywhere* convergence corresponds to the *layerwise* version of *uniform* convergence. This result gives evidence that the *layering* encodes information from which sharper results can be stated, providing a systematic approach to obtain results in the spirit of [6]. In particular, we use it to compute the speed of convergence of random points in both the Strong Law of Large Number and the Ergodic Theorem, in their general versions. The explicit connection between our framework and [6] is also given.

Our framework also enables us to give a simple answer to a question raised in [10] for algorithmically random Brownian motion (see Sect. 5.3).

In Sect. 2 we recall the background on computable probability spaces and Martin-Löf randomness and prove the effective version of a Prokhorov's result. In Sect. 3 we set the framework for *layerwise computability* and state the results relating it to effective measurability. In Sect. 4 we study the convergence of random variables from the effective point of view. We finish in Section 5 by applying all this machinery to obtain the general results announced above, giving solutions to Problems 1 and 2.

2 Preliminaries

2.1 Computable probability spaces

We work on the well-studied computable metric spaces (see [11]).

Definition 1. A **computable metric space** is a triple (X, d, \mathcal{S}) where:

1. (X, d) is a separable metric space,
2. $\mathcal{S} = \{s_i : i \in \mathbb{N}\}$ is a countable dense subset of X with a fixed numbering,
3. $d(s_i, s_j)$ are uniformly computable real numbers.

\mathcal{S} is called the set of **ideal points**. If $x \in X$ and $r > 0$, the metric ball $B(x, r)$ is defined as $\{y \in X : d(x, y) < r\}$. The set $\mathcal{B} := \{B(s, q) : s \in \mathcal{S}, q \in \mathbb{Q}, q > 0\}$ of **ideal balls** has a canonical numbering $\mathcal{B} = \{B_i : i \in \mathbb{N}\}$. An **effectively open set** is an open set U such that there is a r.e. set $E \subseteq \mathbb{N}$ with $U = \bigcup_{i \in E} B_i$.

A compact set K is **effectively compact** if the set $\{\langle i_1, \dots, i_n \rangle : K \subseteq B_{i_1} \cup \dots \cup B_{i_n}\} \subseteq \mathbb{N}$ is r.e. Let $K \subset X$. A set V is **effectively open in K** if there is an effective open set U such that $V \cap K = U \cap K$. A set V is **decidable in K** if V and $X \setminus V$ are effectively open in K . A function $f : X \rightarrow Y$ is **computable on K** if the preimages of effectively open sets are effectively open in K , in a uniform way. A real function $f : X \rightarrow [-\infty, +\infty]$ is **lower semi-computable** if the sets $f^{-1}(q_i, +\infty)$ are uniformly effectively open, it is **upper semi-computable** if $-f$ is lower semi-computable. Any object that has some effectivity can be naturally encoded into a (possibly more than one) integer, called its *Gödel number*.

Remark 1. Let K be effectively compact. It is not difficult to see that the complement $X \setminus K$ is an effective open set, uniformly in K , and that if U is an effective open set, then $K \setminus U$ is effectively compact, uniformly in U, K .

Several approaches to the computability of Borel probability measures have been proposed and happen to give the same notion, which can then be considered as a robust one.

Definition 2 (from [9, 12, 13]). Let (X, d, \mathcal{S}) be a computable metric space. A Borel probability measure μ on X is **computable** if $\mu(B_{i_1} \cup \dots \cup B_{i_n})$ are lower semi-computable, uniformly in i_1, \dots, i_n .

A **computable probability space** is a pair (X, μ) where X is a computable metric space and μ is a computable Borel probability measure on X .

Algorithmic randomness. Martin-Löf randomness was first defined in [1] on the space of infinite symbolic sequences. Its generalization to abstract spaces has been investigated in [8, 9, 14, 15]. We follow the approaches [8, 9] developed on any computable probability space (X, μ) .

Definition 3. A **Martin-Löf test (ML-test)** V is a sequence of uniformly effective open sets V_n such that $\mu(V_n) < 2^{-n}$. A point x **passes a ML-test V** if $x \notin \bigcap_n V_n$. A point is **Martin-Löf random (ML-random)** if it passes all ML-tests. The set of ML-random points is denoted by ML_μ .

Theorem 1 (adapted from [1]). Every computable probability space (X, μ) admits a universal Martin-Löf test, i.e. a ML-test U such that for all $x \in X$, x is ML-random $\iff x$ passes the test U . Moreover, for each ML-test V there is a constant c (computable from any Gödel number of V) such that $V_{n+c} \subseteq U_n$ for all n .

From now and beyond, we fix a particular universal ML-test U . One can assume w.l.o.g. that $U_{n+1} \subseteq U_n$. When the underlying space is complete, even if it is unbounded the finite character of probability measures makes the probabilistic phenomena concentrate in a small region. This is formally expressed by Prokhorov's theorem: on a complete separable metric every Borel probability measure is *tight*. We prove its effective version:

Theorem 2 (Effective Prokhorov theorem). On a complete computable metric space, every computable Borel probability measure is effectively tight: the sets $K_n := X \setminus U_n$ are uniformly effective compact sets and $\mu(K_n) > 1 - 2^{-n}$.

Effective measurability. The following is an adaptation of [7] to complete spaces.

Definition 4. A set A is **effectively μ -measurable** if there are uniformly effective compact sets C_n and open sets U_n such that $C_n \subseteq A \subseteq U_n$ and $\mu(U_n \setminus C_n) < 2^{-n}$. A function $f : X \rightarrow Y$ is **effectively μ -measurable** if there is a basis $\hat{\mathcal{B}} = \{\hat{B}_1, \hat{B}_2, \dots\}$ of Y effectively equivalent to \mathcal{B} such that $f^{-1}(\hat{B}_i)$ are uniformly effectively μ -measurable.

In the original definition the sets C_n are complements of effective open sets. When the space is complete, requiring C_n to be effectively compact gives the same notion, using effective tightness (Thm. 2) and Rmk. 1. Any effective open set whose μ -measure is computable, as the complement of the fat Cantor set for the Lebesgue measure, is effectively μ -measurable. Conversely, the measure of any effectively μ -measurable set is computable. Decidable sets and computable functions are always effectively μ -measurable, whatever the computable measure μ may be. A set is effectively μ -measurable if and only if so is its indicator function.

3 Layerwise computability

Now we enter in the main novelty of this article. With effective versions of measure-theoretic notions at our disposal, we can hope to solve Problem 1. However the notions developed so far are difficult to handle and rather heavy, see Def. 4. It was demonstrated in [7] that algorithmic randomness and the universal test offer an alternative elegant way of handling effective measurability notions.

Let (X, μ) be a *complete* computable probability space. It comes with a canonical universal ML-test U_n , with $U_{n+1} \subseteq U_n$. Hence the set of ML-random points is *layered* by an increasing sequence of uniformly effective compact sets: $\text{ML} = \bigcup_n K_n$ (Thm. 2).

Definition 5 (Martin-Löf Layering). Let (X, μ) be a computable probability space. We call the sequence $(K_n)_{n \in \mathbb{N}}$ the **Martin-Löf layering** of the space.

Definition 6 (Layerwise computability notions).

1. A set $A \subseteq X$ is **layerwise semi-decidable** if for all n , A is effectively open on K_n , uniformly in n , i.e. there are uniformly effective open sets U_n such that $A \cap K_n = U_n \cap K_n$,
2. A set $A \subseteq X$ is **layerwise decidable** if for all n , A is decidable on K_n , uniformly in n , i.e. both A and $X \setminus A$ are layerwise semi-decidable,
3. A function $f : (X, \mu) \rightarrow Y$ is **layerwise computable** if for all n , f is computable on K_n , uniformly in n , i.e. $f^{-1}(B_i)$ are uniformly layerwise semi-decidable (B_i are the canonical ideal open balls).

More generally, every computability (or *effective topological*) notion has its layerwise counterpart. For instance, the layerwise counterpart of effective uniform convergence (i.e., in the sup norm) of functions will be examined in Sect.

4. In general, layerwise computability is not stable under composition, simply because layerwise computable functions may not preserve randomness. This can be overcome under measure-preservation:

Proposition 1 (from [7]). *Let (X, μ) be a computable probability space, $T : X \rightarrow X$ a layerwise computable function which preserves μ (i.e. $\mu \circ T^{-1} = \mu$) and $f : X \rightarrow Y$ a layerwise computable function. Then:*

1. T preserves ML-randomness. Moreover there is a constant c such that $T(K_n) \subseteq K_{n+c}$ for all n .
2. $f \circ T$ is layerwise computable, uniformly in f and T .

3.1 Relation with effective measurability

In [7], we prove the following equivalences:

Theorem 3. *Let $A \subseteq X$ be a set and $f : X \rightarrow Y$ a function.*

1. A is effectively μ -measurable $\iff A$ is layerwise decidable,
2. $f : X \rightarrow Y$ is effectively μ -measurable $\iff f$ is layerwise computable.

Therefore, under effectivity assumptions measure-theoretical notions are the layerwise versions of topological ones. Observe that the latter are expressed without referring to μ but only to the Martin-Löf layering. In other words, the layering catches the essential part of the probabilistic features. This is the first illustration of the Correspondence Principle (see Introduction).

3.2 Layerwise tests

We now state the theorem which will allow to solve Problem 1: making theorems on random points hold under effective measurability assumptions. The surprising point is that weakening randomness tests to their layerwise versions leave them close to plain tests and does not spawn higher-order tests. This will enable us to strengthen many existing results.

Definition 7. *A layerwise Martin-Löf test A is a sequence of uniformly layerwise semi-decidable sets A_n such that $\mu(A_n) < 2^{-n}$. A layerwise integrable test is a layerwise lower semi-computable function $t : X \rightarrow [0, +\infty]$ such that $\int t d\mu < \infty$.*

Theorem 4. *Let U be a layerwise semi-decidable set, A a layerwise ML-test and t a layerwise integrable test.*

1. If $\mu(U) = 1$ then $ML_\mu \subseteq U$.
2. If x is ML-random, then $x \notin \bigcap_n A_n$. Moreover, there is a constant c (computable from a Gödel number of the sequence A) such that $A_{n+c} \cap K_n = \emptyset$ for all n .
3. If x is ML-random, then $t(x) < \infty$. Moreover, there is a constant c such that $t < 2^{n+c}$ on K_n .

4 Convergence of random variables

In [3] the following result for convergence of random variables on random points is stated: if computable functions converge almost everywhere in an effective way then they converge on ML-random points. Here we improve this in several ways:

- using layerwise tests, we weaken the hypothesis: the functions are now assumed to be effectively measurable only, which gives a solution to Problem 1,
- using the layering, we get information about the speed of convergence on random points, providing a solution to Problem 2,
- under effectivity assumptions, we get a characterization of a probabilistic notion (namely, almost everywhere convergence) as the layerwise version of a topological one (namely, uniform convergence), which further illustrate the Correspondence Principle, beyond Theorem 3.
- we give other results for random points under different types of assumptions on the convergence of the sequence.

Observe that what follows works on any computable probability space (algorithmic randomness was only developed on the Cantor space when [3] was written). Let $f_i : X \rightarrow \mathbb{R}$ be a sequence of random variables and f another random variable (expected to be the limit of f_i). Let $D_n(\delta) := \{x : \exists i \geq n, |f_i - f| > \delta\}$. It is a standard observation that the sequence f_i converge almost everywhere to f if and only if the measure of the sets $D_n(\delta)$ tends to zero, for each δ . This motivates the following:

Definition 8. *Functions f_n converge effectively almost everywhere (effectively a.e.) if $\mu(D_n(\delta))$ converge to 0, effectively from δ . In other words there is a computable function $n(\delta, \varepsilon)$ such that $\mu(D_{n(\delta, \varepsilon)}(\delta)) < \varepsilon$.*

As already said, V'yugin [3] proved that if f_n are uniformly *computable* functions that converge effectively a.e. then they converge at each ML-random point. Actually, the result also holds when the functions f_n are uniformly *effectively μ -measurable* and we can even go further. Let us first introduce the layerwise version of effective convergence for the uniform norm:

Definition 9. *Functions f_i converge layerwise effectively uniformly to f if for each k , the restrictions of f_i to K_k converge to the restriction of f to K_k for the uniform norm, effectively from k . In other words, there is a computable function $n(\delta, k)$ such that $\|f_i - f\|_{K_k} := \sup_{K_k} |f_i - f| \leq \delta$ for all $i \geq n(\delta, k)$.*

In the same way that uniform convergence implies pointwise convergence, such functions converge on each ML random point.

Proposition 2. *If f_i are uniformly layerwise computable functions that converge layerwise effectively uniformly to f then f is layerwise computable.*

As said above, effective a.e. convergence implies layerwise effective uniform convergence. Actually this is a characterization, which provides another illustration of the Correspondence Principle:

Theorem 5. *Let f_n be uniformly effectively μ -measurable functions. Then f_n converge effectively a.e. if and only if f_n converge layerwise effectively uniformly.*

At the same time, this result gives evidence that layerwise computability is a solution to Problems 1 and 2: this convergence for random points holds for effectively μ -measurable functions and not only computable ones, and the speed of convergence can be computed from the layer a random point belongs to.

Corollary 1. *If f_i are uniformly effectively μ -measurable functions that converge effectively a.e. to f , then f is effectively μ -measurable.*

The simplicity of this proof shall be very general, as soon as the Correspondence Principle holds: a result about effective measure-theoretical notions can be split into two parts (i) the layerwise version of the corresponding result for effective topological (i.e. computability) notions (as Prop. 2) and (ii) the characterizations of effective measure-theoretical notions as layerwise topological ones (as Thm. 3 and 5).

Non-effective convergence. When the convergence a.e. is not effective we can still say something concerning random points.

Theorem 6. *Let f_n, f be uniformly layerwise computable functions, and c some (not necessarily computable) constant.*

- *If f_n converge a.e. to a constant c then $\liminf f_n(x) \leq c \leq \limsup f_n(x)$ for all $x \in ML$.*
- *If f_n converge a.e. to a layerwise computable function f , then $\liminf f_n(x) \leq f(x) \leq \limsup f_n(x)$ for all $x \in ML$.*

5 Applications

5.1 Ergodic theorems for effectively measurable functions

We now apply the tools developed so far to solve Problem 1 for Poincaré recurrence theorem and Birkhoff's ergodic theorem. The version of the latter theorem for random points has been proved by V'yugin [3] (i) on the Cantor space $X = \{0, 1\}^{\mathbb{N}}$ and (ii) for a computable transformation $T : X \rightarrow X$ and a computable observable $f : X \rightarrow \mathbb{R}$. Point (i) is not a real restriction as the proof for general spaces remains unchanged. However the condition (ii) is an unnatural and *a priori* too strong restriction, as explained in the introduction. Moreover, on general spaces this restriction becomes much more important since the theorem cannot be applied to indicators of sets anymore (in connected spaces they cannot be computable as they are not continuous).

In [4], Birkhoff's ergodic theorem for random points is extended to include functions having some discontinuities at computable points. A further step is given in [5] where the result is proved to hold for the indicator functions of every (not necessarily constructive) set of continuity (i.e., a set whose boundary has

measure zero). Yet, nothing can be said about some natural sets having effective constructions, like the **Smith-Volterra-Cantor set** (or *fat Cantor set*) whose Lebesgue measure is $\frac{1}{2}$ but has empty interior, and hence is not a continuity set.

We now give a definite solution by proving the ergodic theorems for *effectively measurable* functions. In particular, indicator functions of sets like the fat Cantor set fall in this class.

Theorem 7 (Poincaré recurrence theorem for random points). *Let (X, μ) be a computable probability space, $T : X \rightarrow X$ an effectively μ -measurable measure-preserving map and A a layerwise semi-decidable set with positive measure. Every ML-random point $x \in A$ falls infinitely often in A by iteration of T . If the system is moreover ergodic, then every ML-random point falls infinitely often in A by iteration of T .*

Theorem 8 (Ergodic theorem for random points). *Let (X, μ) be a computable probability space, $T : X \rightarrow X$ an effectively μ -measurable measure-preserving map and $f \in L^1(X, \mu)$ be an effectively μ -measurable observable. Then:*

- (i) *For every ML-random point x , the limit $\bar{f}(x) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x)$ exists.*
- (ii) *If the system is moreover ergodic, then $\bar{f}(x) = \int f d\mu$ for every ML-random x .*

5.2 Layerwise computable speed of convergence on random points

In this section we show how the framework developed so far provides a solution to Problem 2. Let us first recall some results established by Davie [6].

Davie's results. To state them some background is needed first. On the Cantor space, implicitly endowed with the uniform measure λ , the *compressibility coefficient* or *deficiency of randomness* of a sequence ω is defined as $d_\lambda(\omega) = \sup_n \{n - H(\omega_{1:n})\}$ where $H(w)$ is the *Kolmogorov-Chaitin complexity* of the finite word w . A fundamental result from algorithmic randomness and information theory is that a sequence is ML-random w.r.t. λ if and only if $d_\lambda(\omega)$ is finite. Davie defines $K^c := \{\omega : d_\lambda(\omega) \leq c\}$ and proves:

Theorem 9 (Davie, 2001). *If A_i are uniformly effective open sets such that $\sum_i \mu(A_i)$ is a finite computable real number, then there is a computable function $n(c)$ such that for all $\omega \in K^c$ and all $m > n(c)$, $\omega \notin A_m$.*

Theorem 10 (Davie, 2001). *There is a computable function $n(c, \varepsilon)$ such that for all $\omega \in K^c$ and all $n > n(c, \varepsilon)$, $\left| \frac{S_n(\omega)}{n} - \frac{1}{2} \right| < \varepsilon$ where $S_n(\omega)$ is the number of ones in the prefix of ω of length n .*

The equivalence between the paradigm of effective measure theory (Martin-Löf's approach) and the paradigm of compressibility (Chaitin's approach) is a strong non-trivial result, partly based on the technical *coding theorem*. Davie's results follow this line as they relate the compressibility coefficient of a sequence to the way the sequence satisfies a probability law, and thus their proofs consist in a finer use of the coding theorem. In our framework, we stay on the side of effective measure theory. In this way, the relation between the layer a random point belongs to and the way it satisfy laws is much simpler to derive, as it is essentially already contained in the existing proofs. This provides a solution to Problem 2. At the same time, as layerwise computability provides a solution to Problem 1 too, our results hold for effectively measurable sets/functions.

As an illustration, we first state here the refined version of classical results in algorithmic randomness due to Solovay. The proofs are straightforward combinations of the usual proofs together with Thm. 4. Note that the first one is the generalization of Thm. 9 due to Davie.

Proposition 3 (Borel-Cantelli 1). *There is a computable function $n(c, p)$ such that if A_n are uniformly layerwise semi-decidable sets such that $\alpha := \sum_n \mu(A_n)$ is finite and computable, then there is a constant c , computable from a Gödel number of the sequence A_n and α , such that if $x \in K_p$ then $x \notin A_n$ for all $n \geq n(c, p)$.*

We can also get a weaker result when the sum is not computable.

Proposition 4 (Borel-Cantelli 2). *Let A_i be uniformly layerwise semi-decidable sets such that $\sum_i \mu(A_i) < \infty$. There is c , computable from a description of the sequence A_i , such that every x in K_n falls in the A_i 's at most 2^{n+c} times.*

We can now easily prove:

Theorem 11 ((Very) Strong Law of Large Numbers.) *Let $X_i : (X, \mu) \rightarrow \mathbb{R}$ be i.i.d. effectively μ -measurable random variables such that $\int |X_i|^4 d\mu < +\infty$. Let $S_n := X_0 + \dots + X_{n-1}$. Hence, there is a computable function $n(c, \varepsilon)$ such that if $x \in K_c$ then for all $n > n(c, \varepsilon)$, $\left| \frac{S_n(x)}{n} - \int X_0 d\mu \right| < \varepsilon$.*

Effective convergence in Birkhoff's theorem The convergence of the Birkhoff averages is not effective in general. In [3], on the Cantor space V'yugin builds a computable probability measure which is invariant under the shift transformation, and such that the convergence of the averages of $\mathbf{1}_{[1]}$ is not effective. This measure is an infinite combination of ergodic measures and it is still an open question if a computable ergodic measure could be built for which the convergence is not effective.

However, in [16] it is shown that for a class of ergodic systems, the convergence in Birkhoff theorem is effective. Let us recall that a system is ergodic if and only if for any two integrable functions f and g , the quantity $\gamma_n(f, g) := \left| \frac{1}{n} \sum_{i < n} \int f \circ T^i \cdot g d\mu - \int f d\mu \int g d\mu \right|$ goes to 0. A system is said to be **ln²-ergodic for (f, g)** , if there is a constant $c_{f,g} > 0$ such that $\gamma_n(f, g) \leq \frac{c_{f,g}}{(\ln(n))^2}$ for all $n \geq 2$.

Theorem 12 (Ergodic theorem for random points). *Let (X, μ) be a computable probability space, $T : X \rightarrow X$ be an effectively measurable measure-preserving map and $f \in L^1(X, \mu)$ be an effectively measurable function. If T is \ln^2 -ergodic for (f, f) , then there is a computable function $n(c, \varepsilon)$ such that if $x \in K_c$ then for all $n > n(c, \varepsilon)$, $|\frac{1}{n} \sum_{i < n} f(x) \circ T^i(x) - \int f d\mu| < \varepsilon$.*

Relation between K_n and K^n . Let X be the Cantor space endowed with a computable Borel probability measure μ . The version of the *compressibility coefficient* or *deficiency of randomness* adapted to μ is $d_\mu(\omega) := \sup_n \{-\log \mu[\omega_{1:n}] - H(\omega_{1:n})\}$. This function is known to be the logarithm of a universal integrable μ -test, which means that for every integrable μ -test t there is a constant a such that $\log t \leq a + d_\mu$. On the other hand, every computable probability space admits a universal integrable test t_μ (see [8, 9]). Generalizing Davie, let us define $K^c := \{x : t_\mu(x) \leq 2^c\}$. As $ML_\mu = \bigcup_c K^c$, the sequence $(K^c)_{c \in \mathbb{N}}$ can be used as an alternative layering and underlying alternative versions of Def. 6 and 9. Actually, this would lead to the same notions. Indeed, using classical results from algorithmic randomness and information theory (see [17, 18]), it can be proved that there is a constant c such that $K_n \subseteq K^{n+c}$ and $K^n \subseteq K_{n+2 \log n+c}$ for all n . Hence K^n are also uniformly effective compact sets and all layerwise computability notions relative to K^n are equivalent to the notions relative to K_n .

5.3 An application to Brownian motion

The study of Brownian motion from the algorithmic randomness point of view is carried out in [10, 19]. Algorithmically random paths, called *complex oscillations* as they are defined in terms of Kolmogorov-Chaitin complexity, are the Martin-Löf random points of the computable probability space $(\mathcal{C}([0, 1]), W)$, where $\mathcal{C}([0, 1])$ is the space of continuous functions $x : [0, 1] \rightarrow \mathbb{R}$ with the uniform norm and W is the Wiener probability measure. In [19] it is proved that if $t \in [0, 1]$ is computable and x is a complex oscillation then $x(t)$ is not computable. At the end of [10] the following question is raised: can it be lower semi-computable?

We say that $y \in \mathbb{R}$ is λ -ML-random if $y = n + z$ where $n \in \mathbb{Z}$ and $z \in [0, 1]$ is ML-random w.r.t. the Lebesgue measure λ on $[0, 1]$. As noticed in [10], it is a corollary of [20] that $x(t)$ is actually λ -ML-random. But then can it be a Chaitin's Ω (which are lower semi-computable λ -ML-random reals)? The compactness of the layers (Thm. 2) enables us to give a positive answer. Indeed, Prop. 1 can be reinforced using Thm. 2:

Proposition 5. *Let (X, μ) and (Y, ν) be computable probability spaces such that X is complete. Let $T : X \rightarrow Y$ be a layerwise computable function which maps μ to ν . Then $T(ML_\mu) = ML_\nu$, i.e. T preserves randomness but it is also onto.*

Now, given a computable $t \in (0, 1]$, the function $T_t : \mathcal{C}([0, 1]) \rightarrow \mathbb{R}$ mapping x to $x(t)$ is computable. It pushes the Wiener measure W to a gaussian measure G . As G has bounded density w.r.t. the uniform measure and vice versa, ML_G is exactly the set of λ -ML-random reals. Hence,

Corollary 2. *Let x be a complex oscillation. For each computable $t \in (0, 1]$, $x(t)$ is λ -ML-random. Moreover, given any λ -ML-random y and any non-zero computable t , there exists a complex oscillation x such that $x(t) = y$.*

References

1. Martin-Löf, P.: The definition of random sequences. *Information and Control* **9**(6) (1966) 602–619
2. Vovk, V.G.: The law of the iterated logarithm for random kolmogorov, or chaotic, sequences. *Theory of Probability and Applications* **32** (1987) 413–425
3. V’yugin, V.V.: Effective convergence in probability and an ergodic theorem for individual random sequences. *SIAM Theory of Probability and Its Applications* **42**(1) (1997) 39–50
4. Nandakumar, S.: An effective ergodic theorem and some applications. In: *STOC ’08: Proceedings of the 40th annual ACM symposium on Theory of computing*, New York, NY, USA, ACM (2008) 39–44
5. Galatolo, S., Hoyrup, M., Rojas, C.: Effective symbolic dynamics, random points, statistical behavior, complexity and entropy. *Inf. Comput.* (2007) In press.
6. Davie, G.: The Borel-Cantelli lemmas, probability laws and Kolmogorov complexity. *Annals of Probability* **29**(4) (2001) 1426–1434
7. Hoyrup, M., Rojas, C.: An application of Martin-Löf randomness to effective probability theory. In: *LNCS. Proceedings of CiE’09*.
8. Gács, P.: Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science* **341** (2005) 91–137
9. Hoyrup, M., Rojas, C.: Computability of probability measures and Martin-Löf randomness over metric spaces. *Inf. and Comput.* (2009) In press.
10. Fouché, W.L.: Dynamics of a generic brownian motion: Recursive aspects. *Theor. Comput. Sci.* **394**(3) (2008) 175–186
11. Weihrauch, K.: *Computable Analysis*. Springer, Berlin (2000)
12. Edalat, A.: When Scott is weak on the top. *Mathematical Structures in Computer Science* **7**(5) (1997) 401–417
13. Schröder, M.: Admissible representations of probability measures. *Electronic Notes in Theoretical Computer Science* **167** (2007) 61–78
14. Zvonkin, A., Levin, L.: The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematics Surveys* **256** (1970) 83–124
15. Hertling, P., Weihrauch, K.: Random elements in effective topological spaces with measure. *Information and Computation* **181**(1) (2003) 32–56
16. Galatolo, S., Hoyrup, M., Rojas, C.: A constructive Borel-Cantelli lemma. Constructing orbits with required statistical properties. *Theor. Comput. Sci.* (2007) In press.
17. Li, M., Vitanyi, P.M.B.: *An Introduction to Kolmogorov Complexity and Its Applications*. Springer-Verlag, Berlin (1993)
18. Gács, P.: Lecture notes on descriptonal complexity and randomness. <http://www.cs.bu.edu/faculty/gacs/papers/ait-notes.pdf>
19. Fouché, W.L.: Arithmetical representations of brownian motion i. *J. Symb. Log.* **65**(1) (2000) 421–442
20. Kjos-Hanssen, B., Nerode, A.: The law of the iterated logarithm for algorithmically random brownian motion. In: *LFCS*. (2007) 310–317