Journal of Logic & Analysis 14:2 (2022) 1–14
ISSN 1759-9008

Computability of convergence rates in the Ergodic Theorem for Martin-Löf random points

GEORGE DAVIE

Abstract: In this paper we look at the convergence rates for the ergodic averages in the pointwise ergodic theorem for computable ergodic transformations on Cantor space. While, for example, these rates are layerwise computable for Martin-Löf random points and effectively open sets with measure a computable real, they are also layerwise computable for an arbitrary interval. For the shift operator, however, there are effectively open sets for which there are *no* effective rates, in particular, not layerwise computable ones. We also show that, when the measure of the effectively open set is any real α, the convergence rates are computable in α and the layers relative to α.

2020 Mathematics Subject Classification 28D05 (primary); 68Q30 (secondary)

Keywords: Layerwise Computable; Martin-Löf random; Ergodic Theorem; Effectively open set; Effective convergence rates.

1 Introduction

Probability laws that hold for almost all points can often be proven to hold for all Martin-Löf random points, a set of measure 1. For example, the Laws of Large Numbers, and the Law of the Iterated Logarithm, Vovk [17].

Further, under certain conditions one can use the randomness degree (the *compressibility coefficient* or *layer*) of a point to find the rate at which it satisfies the probability law, see Davie [5] and Hoyrup and Rojas [9].

In the first theme above, it was independently shown in the two papers Bienvenu, Day, Hoyrup, Mezhirov and Shen [3] and Franklin, Greenberg, Miller and Ng [6] that the ergodic average of every Martin-Löf random point will equal the measure of the corresponding set for all effectively open/closed sets for computable ergodic transformations in computable measure spaces. In the terminology of [6], every Martin-Löf random point is *Birkhoff* for this context.

In general, the convergence *rates* of the ergodic averages are, however, not computable.

Published: September 2022 https://doi.org/10.4115/jla.2022.14.2

While the results of Avigad, Gerhardy and Towsner [2] and those of Hoyrup and Rojas [9] imply that the convergence rates are layerwise computable when the measure of the effectively open set is a computable number, the computability of the measure of the set is not necessary. In fact, rates of convergence of the ergodic averages are layerwise computable for *any* interval.

We also construct, for the classical ergodic operator on Cantor space, the shift operator, an effectively open set for which convergence rates are not layerwise computable (in fact not even effective), and deduce from this a sufficient condition for layerwise computability of rates.

Lastly we show (using the result from [2] again) that for the measure of the effectively open set any real α, the convergence rates are computable in α and the layers relative to α.

2 Definitions and notation

Our ergodic setting is Cantor space with the normal product measure. We denote an infinite sequence by ω. Our ergodic transformations T will be computable unless noted otherwise. A function f on infinite binary sequences is computable if there is an algorithm that, given access to an infinite binary sequence ω, will, on input n, output the first n digits of $f(\omega)$. Of course, the algorithm uses only finitely many digits of ω for each such computation.

We work throughout with *effectively open sets*. An effectively open set O is a union of a computably enumerable set of basic open sets, in our case intervals. Note that, if we are interested in the membership of such sets, then we do not at any finite stage, have the complete set (and hence the characteristic function 1_O). At each stage n of enumeration, we have only an approximation O_n to the set and 1_n to the characteristic function.

Martin-Löf random points are defined via Martin-Löf *tests* [14]:

Definition 2.1 A Martin-Löf test U is a sequence of uniformly effective open sets U_n such that $\mu(U_n) < 2^{-n}$. A point ω passes the test U if $\omega \notin \bigcap_n U_n$. A point is Martin-Löf random if it passes all Martin-Löf-tests.

We will assume that the test is *nested*, that is $U_{n+1} \subseteq U_n$. Call the first $k \in N$ for which $\omega \notin U_k$, the *actual* layer of ω. Any $l > k, l \in N$ is a *valid layer* for ω.

Martin-Löf random points have many equivalent definitions, including the property that no betting strategy can succeed against the points when considered as infinite binary sequences and also that the points, when considered as such sequences, have incompressible initial segments.

There exist universal Martin-Löf tests which are such that, if a sequence passes it, then it passes all Martin-Löf tests. We will work with an *optimal* Martin-Löf test, which is an even stronger notion. A Martin-Löf test is optimal if, for any other Martin-Löf test V there exists a $c \in N$ such that $V_{c+n} \subset U_n$ for all $n \in N$.

Algorithmic randomness has been generalised to *computable probability spaces*, see Weihrauch [18] for the context of computable analysis, and Gàcs [1] and Hoyrup and Rojas [8, 10] for the generalisation of Martin-Löf randomness to computable probability spaces. We will mostly stay in Cantor space in this paper.

2.1 Layerwise computable and decidable

Layerwise computability is a weakening of the standard notion of computability which has profound links with measure theory and topology. The seminal papers in this area are Hoyrup and Rojas [8, 9]. For an excellent recent survey, see Hoyrup [11].

Definition 2.2 A function f defined on the random infinite binary sequences is layerwise computable if it is computable in the pair (ω, l), where l is any valid layer for ω.[1]

That is, we need a layer l as an extra input alongside ω. The function f gives the same output on (ω, l) for l any valid layer. For example, if the actual layer of ω is 1, then the function is defined on each of $(\omega, k), k \geq l$ with the same output.

Definition 2.3 A set is layerwise decidable if the membership is decidable given a pair (ω, l).

No claim is made for behaviour on an input (ω, k) where either ω is not random, or k is not a valid layer for ω.

Note also that a layerwise computable function can be "made computable" on as large a measure set as we want. Hence if we want, for example, a layerwise function to act like

[1] There is also a notion of *exactly* layerwise computable/decidable in which the actual layer must be given, see Hölzl [7].

a normal computable function on more than measure $1 - 2^{-k}$ points, we can just give all input points the same layer l for some effective $l > k$.

A typical example of layerwise computability is the following. Consider an enumeration of an effectively open set O. It is in general not decidable whether a given point will appear in the enumeration or not. When the measure of the set is a computable number however, the decision becomes layerwise decidable, see Davie [5].

This principle is very useful to us and we state it as a lemma:

Lemma 2.4 *Let O be an effectively open set with measure a computable real. Upper bounds for the stage of appearance of Martin-Löf random points in O are then layerwise computable; consequently, their membership of O is layerwise decidable. Conversely, if the membership is layerwise computable then the measure is a computable real.*

In the more general context of computable measure spaces, the counterpart of Lemma 2.4 is that a set is so-called *effectively μ-measurable* if and only if it is layerwise decidable. Another fundamental result in this context is that a function is layerwise computable if and only if it is so-called *effectively measurable*, Hoyrup and Rojas [8].

3 Hitting times

As an illustration of the concept of layerwise computability we look at the following theorem of Kučera [12], which played a central role in both papers Bienvenu et al [3], and Franklin et al [6]:

Theorem 3.1 (Kučera) *If A is an effectively closed set of Cantor space with measure greater than 0, then every Martin-Löf random sequence ω must have a tail in A.*

In other words, under the shift operator, there will be a finite *hitting time* for (shifts of) ω to enter A.

Kučera in fact proved the converse too, that is, that having a tail in every such set A characterizes the Martin-Löf random points.

When the measure of the set A is a computable real we have the following layerwise computable version:

Proposition 3.2 *Let an effectively closed set A have measure a computable real, then finding the first tail of an infinite binary sequence ω that hits A is layerwise computable.*

That is, hitting times for the set A under the shift operator are layerwise computable.

Proof Consider the complement of A which is an effectively open set O with computable measure, we find the first tail that *misses* O. By Hoyrup and Rojas [9] the layer can change by at most a (computable) constant every time we take the next tail. We can thus use Lemma 2.4 for each consecutive tail. For each tail we thus have a computable upper bound on how long we have to wait for it to appear in the enumeration of the effectively open set. For some tail this stage of the enumeration will be reached before the tail appears in the enumeration. We can then conclude that this tail will not appear and this is therefore the first tail which misses O. □

Note that $\omega \in O$ if and only if the hitting time of ω is 0. Hence, if we could decide hitting times layerwise computably, then we could decide membership of O layerwise computably, and by Lemma 2.4, $\mu(O)$ would be a computable real. (Thanks to the referee of a previous submission for pointing this out.) Hence we have:

Corollary 3.3 *Hitting times for effectively open sets are layerwise computable if and only if the set has measure a computable real.*

It is shown in Pauly, Fouché and Davie [15] that if we could decide hitting times, we could construct random finite sequences of arbitrary length.

4 Ergodic systems

Definition 4.1 A dynamical system is *ergodic* if for every set S and almost every point ω, the ergodic average $\frac{1}{n}\sum_{i=0}^{n-1} 1_S(T^i\omega)$ converges as follows:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} 1_S(T^i\omega) = \mu(S)$$

Definition 4.2 Write $A_n(\omega) = \frac{1}{n}\sum_{i=0}^{n-1} 1_S(T^i\omega)$.

Following Franklin et al [6], call a point ω which satisfies the above condition for a set S a *Birkhoff point* for S.

Kučera's theorem is used in both [3] and [6] to show that:

Theorem 4.3 (Bienvenu, Day, Hoyrup, Mezhirov, Shen / Franklin, Greenberg, Miller and Ng) *Let T be a computable ergodic measure-preserving transformation. Then every Martin-Löf point is a Birkhoff point for every effectively open/closed set with measure > 0.*

4.1 Convergence rates for ergodic systems

Let T be a computable ergodic transformation on Cantor space. To compute the convergence rate for a point ω and set S in the Ergodic Theorem is to have a computable function which, on inputs ω and ε, will output a number of iterations of the ergodic function T after which the ergodic average of the point ω will not vary more than ε from $\mu(S)$, the measure of the set.

For the convergence rates to be *layerwise* computable, we need not only the point ω as input, but also a valid *layer* of ω to find the number of iterations of T needed.

4.2 Layerwise computable for O with computable measure

We have the following two important theorems, the first from Hoyrup and Rojas [8]:

Theorem 4.4 (Hoyrup, Rojas) *If f_n are uniformly layerwise computable functions and f_n converges effectively almost everywhere, then the convergence rates for f_n for random sequences are layerwise computable.*

We also have the very powerful result of Avigad, Gerhardy and Townser [2]:

Theorem 4.5 *Let X be a separable metric space and T be an ergodic measure-preserving transformation. Then, for any f in $L^2(X)$, the function $n(\varepsilon)$ such that for every $k \geq n(\varepsilon)$*

$$\mu\Big(\{\omega \mid \max_{n(\varepsilon)\leq m\leq k} |A_m(\omega) - A_{n(\varepsilon)}(\omega)| > \varepsilon\}\Big) < \varepsilon$$

is computable in f and T.

If we set $f_n = A_n$, the ergodic average after n steps, then Theorem 4.5 says that f_n converges effectively almost everywhere. Then by Theorem 4.4 the convergence rates for $f_n = A_n$ are layerwise computable:

Proposition 4.6 *Let T be a computable ergodic transformation on Cantor space. If an effective open set O has measure a computable real, then the rate of convergence of $A_n(\omega)$ to $\mu(O)$ in the pointwise Ergodic Theorem for Martin-Löf random points is layerwise computable. That is, there is a layerwise computable $n(\varepsilon)$ such that for all $k > n(\varepsilon)$:*

$$|A_k(\omega) - \mu(A)| < \varepsilon$$

A referee has informed us that Proposition 4.6 has been noted before and in fact holds for all effectively measure-preserving transformations (or equivalently, layerwise computable transformations)—not just for computable ones—and for all L^1-computable observables, not only indicator sets. For good measure, such a generalisation should also be for any computable probability space.

5 The converse fails

Is it perhaps the case that, analogous to Lemma 2.4, the convergence rates are layerwise computable only if the measure of the set O has computable measure? That is, does the converse of Proposition 4.6 hold?

This is not the case; in fact, convergence is layerwise computable for any interval. The next result also shows that the layerwise computability of the convergence rates for effectively open sets has less to do with the measure than the *packing* of the effectively open set.

Theorem 5.1 *The convergence of the ergodic averages is layerwise computable (and hence effective) for any interval.*

Proof We do the proof for the interval I having left endpoint 0. The general case is similar. Given n and $\varepsilon = 1/2^k$, we must find a stage i after which

$$|A_m(\omega) - \mu(I)| \le \frac{1}{2^k}$$

for all ω in layer n. Divide the unit interval into 2^{k+2} dyadic intervals, each of length $1/2^{k+2}$. Since each of these intervals I_k is computable, Proposition 4.6 allows us to find, for each I_k, a stage after which, for all ω in layer n:

$$|A_m(\omega) - \mu(I_k)| \le \frac{\varepsilon}{2^{3k+3}} = \frac{2^{-k}}{2^{3k+3}} = \frac{1}{2^{2k+3}}$$

Take the maximum off these stages. At this stage then the piece I' of I which is entirely contained in such intervals has:

$$|A_m(\omega) - \mu(I')| \le 2^{k+2} \times \frac{1}{2^{2k+3}} = \frac{1}{2^{(k+1)}}$$

The remaining piece I'' of I is also contained in an interval of length $1/2^{(k+2)}$. Hence on this piece, the average can be no less than 0 (clearly) and no more than:

$$\frac{1}{2^{k+2}} + \frac{1}{2^{2k+3}}$$

Since these pieces are disjoint we add the averages to get that:

$$|A_m(\omega) - \mu(I)| \leq \frac{1}{2^k} \qquad \square$$

Ergodic averages for intervals are effective because of the neat "packing" of the set. It is clear that we can, in particular, transform any effectively open set O into an interval by just adding, for each interval enumerated in O, an interval of equal measure adjacent to our rightmost endpoint at that stage, to form in the limit the interval $\big[0, \mu(O)\big)$.

Since the "bad" intervals above (for which we are forced to use the entire interval length as approximation) are those intervals that are only partly contained in I, we see that we can approximate the averages effectively as long as there are not too many of these bad intervals:

Corollary 5.2 *Let $J \subset [0,1]$ and let k_n be the number of dyadic intervals of length 2^{-n} which only partially overlap the set J. Then the convergence rates are layerwise computable (and hence effective) if $k_n \times 2^{-n} \to 0$ in an effective way.*

Hence, if we can, given ε, find an $n(\varepsilon)$ such that dividing into 2^n intervals and finding enough iterations of T to get the average on these intervals close, then the ergodic averages are layerwise computable and hence effective.

In the following section we will construct an example where the rates are far from computable.

6 Convergence is not effective for the shift operator

Layerwise computability (equivalently, uniform layerwise convergence) implies effective convergence (Hoyrup and Rojas [9]), so the next theorem shows in particular that there are no layerwise computable rates here. The proof uses ideas in the proof of the well known result of Krengel [13], that there are no general rates of convergence for the ergodic theorems. In particular, we use the idea of Rokhlin's Lemma [16] to "spread" the unit interval out into many small sets $A, T^{-1}A, \ldots, T^{-n}A$ with large union.

Theorem 6.1 *There is an effectively open set O for which the convergence of the ergodic averages under the shift operator T is not effective.*

Proof Recall that the shift operator T, "shifts" an infinite binary sequence to the left; eg, $T(010111..) = 10111\ldots$ This is well known to be a measure-preserving ergodic transformation.

Chaitin's halting probability, Ω (see Chaitin [4]), is the measure of an enumeration of halting programs in a dovetailing of the running of all programs on the empty input. A program p that halts (and is then enumerated) is seen as the interval $0.p$ and measure $1/2^{|p|}$ is then added to the total measure so far, which approaches Ω. Since the set of programs is prefix free, $\Omega < 1$.

To ensure that $\mu(O)$ is bounded far below 1, we add a fixed initial segment of two zeroes to each halting program to ensure that the total measure of our effectively open set is less than $1/4$. We will use this modified halting probability along with its associated effectively open set as a "measure provider". That is, we will not enumerate the modified interval $0.00p$ itself into our set O but will add many small intervals, adding to the same measure $1/2^{|p|+2}$.

(1) Dovetail the running of all programs. Let p be the ith program which halts in the dovetailing.
(2) When p halts, we have measure $1/2^{|p|+2}$ available to add new intervals to our set O, as follows.
(3) Form a string of 0's as long as the number of steps l ran in the dovetailing until p halted. That is, form 0^l.
(4) See this string of 0's as the interval adjacent to 0 in the unit interval. (By ergodicity of the shift operator, this interval will, in the limit, be hit on average its measure, $1/2^l$, by almost all binary sequences.)
(5) As the ith part of our set O we now enumerate the interval 0^l and a set of its inverse images under T. The first three intervals to be enumerated are then

$$0^l, 0(0^l) \text{ and } 1(0^l).$$

Continue taking inverse images of these, namely

$$00(0^l), 10(0^l), 10(0^l), 11(0^l).$$

We thus obtain one interval of measure 2^{-l}, two of measure 2^{-l-1}, four of measure 2^{-l-2} and so on. Call the nth set of intervals consisting of 2^n intervals of total measure 2^{-l} set I_n.
(6) Continue to enumerate these inverse images until we have exhausted the measure of $0.00p$. If for example $|p| = k-2$ then we have measure $1/2^k$ available and

since each set I_n has measure 2^{-l} we can enumerate in O at least[2] s of the sets I_n where $s \times 2^{-l} = 2^{-k}$, hence at least 2^{l-k} sets I_n. Note that the set I_n consists of binary strings w for which $T^n(w) \in 0^l$; that is, binary sequences which hit 0^l after n steps.

(7) Make a note of when half of the measure is used up, that is, the I_n we are at when this happens. Note that $n \geq 2^{l-k-1}$. Call the union of the I_n's enumerated *after* this point, that is, intervals from length $l+n$ and up, set H, with $\mu(H) = 2^{-k-1}$. Now, for each of the strings w in H it is the case that $T^m(w)$ is in C for $0 \leq m \leq 2^{l-k-1}$. That is, while $\mu(O) < 1/4$, the ergodic average for at least measure 2^{-k-1} of sequences is 1 for $0 \leq m \leq 2^{l-k-1}$. Recall that l is the total runtime so far, which makes 2^{l-k-1} a very late stage under T.

Now, if there was an algorithm which, given k, could find bounds after which fewer than 2^{-k-1} of sequences will have ergodic average 1, we would have the runtime l_k for the longest running program p of length $k-2$. Being able to do this for each k solves the Halting Problem. This gives us a contradiction. □

7 Relativised randomness

In the example above we did not have effective convergence, since we were enumerating very long intervals of large total measure, very late. Having access to Ω would have helped of course, since then we could see how close we are to the total measure Ω at each stage of the enumeration. Theorem 4.5 states that access to Ω is enough for the convergence rates to be effective. We would like to prove a layerwise version of this.

To do this we must define *relative* layerwise computability. This will use the standard notion of relative randomness, with α seen as an oracle:

Definition 7.1 A Martin-Löf test relative to the binary sequence α is a sequence of sets $U_i : i \in N$ uniformly c.e. in α with $\mu(U_i) < 2^{-i}$. A point ω passes the test U if $\omega \notin \bigcap_n U_n$. A point is Martin-Löf random relative to α if it passes all Martin-Löf tests relative to α.

We will call such tests *relative Martin-Löf tests*.

[2]**Note:** Why "at least"? The intervals enumerated above will contain some overlap, eg the interval $0(0^l) \subset 0^l$. We do not need to use any measure up on $0(0^l)$ but only on new intervals that are disjoint from all previous chosen ones.

Changing from a standard Martin-Löf test to a relative Martin-Löf test will generally change the layers of some points, and classify some previously random sequences as non-random. For example, the sequences computable in α, ie Turing reducible to α, will—even if random before—now be computable (in α).

Note that there are more relative Martin-Löf tests than normal Martin-Löf tests since we have access to an oracle which can enable us to build sets which are effectively open relative to the oracle, but not really effectively open. Hence, a sequence which is Martin-Löf random relative to any number is also Martin-Löf random.

Definition 7.2 A function is layerwise computable in α if it is computable given, as extra inputs, a valid layer relative to α and (an oracle for) α.

In Davie [5] a few generalisations of Lemma 2.4 are listed which all fail, including a seemingly natural oracle form. We note that in fact, the natural oracle form of Lemma 2.4 is the following natural relativisation, which holds by adapting the proof in the obvious way:

Lemma 7.3 *Let O be an effectively open set with measure a real α. Upper bounds for the stage of appearance of Martin-Löf random points relative to α in O are then relatively layerwise computable; consequently, their membership of O is relatively layerwise decidable.*

7.1 Relativised effective Borel–Cantelli Lemmas

As a digression we note that Lemma 7.3 allows us to formulate relativised versions of Theorem 2 in [5].[3]

Theorem 7.4 (Borel–Cantelli Lemmas: effective relative forms)

(i) *Let a sequence of events A_i form a uniform sequence of effectively open sets relative to a real α such that $\sum_i \mu(A_i) = \alpha$, where α is finite. Then k such that $\omega \notin A_n$ for all $n \geq k$ is relatively layerwise computable.*

(ii) *If (a) the events A_n are mutually independent and (b) $\sum_i \mu(A_i)$ diverges, then, given m, a k such that at least one of the events A_n $(m < n < k)$ occur is relatively layerwise computable.*

[3] See also Hoyrup and Rojas [8] for generalisations of the effective Borel–Cantelli lemmas in [5].

7.2 Layerwise computable in $\mu(O)$

By Lemma 7.3 we then have the following relative layerwise computable form for the ergodic theorem for any effectively open set:

Theorem 7.5 *The rate of convergence of $A_n(\omega)$ to $\mu(O)$ is layerwise computable in $\mu(O)$.*

In other words, for any effectively open set, we do indeed have rates for the convergence of each element of a layer to the ergodic average. Only in this case, the layer is relative to the measure of the set we are considering. This allows some standard Martin-Löf random sequences to not have relatively layerwise convergence rates (those that are no longer random with respect to α). Note however, that Theorem 4.3 assures us that these points will be Birkhoff whatever the measure of the set.

So this also means that the only points which have a chance of not being Birkhoff points for these sets are the points which are not random with respect to the measure of the set. That is, for a point not to satisfy the convergence in Birkhoff's theorem it must be intimately related to the measure of the set.

8 2–random sequences

We have seen that we get layerwise computability in $\mu(O)$ for the convergence rates, by Theorem 7.5. We now show that for a set of measure 1, the rates are layerwise computable in Ω, whatever O is.

The set of 2–random binary sequences is the set of binary sequences which are Martin-Löf random with respect to $\alpha = \Omega$.[4] By Theorem 7.5, the rates for the 2–random sequences will be governed by the particular $\mu(O)$; but, since the measure of every effectively open set is computable in Ω, the rates will also be layerwise computable in Ω. Hence:

Theorem 8.1 *For 2–random ω and any effective open set O, the convergence rates of $A_n(\omega)$ to $\mu(O)$ in the pointwise ergodic theorem are layerwise computable in Ω.*

[4] Recall that Ω encodes the halting probability for our Turing machine, hence encodes the Halting Problem.

9 Acknowledgement

I would like to thank the referee for remarks and suggestions which improved the paper substantially. I would also like to thank the referees to a previous submission of this paper for helpful remarks.

References

[1] **P Gács**, *Uniform test of algorithmic randomness over a general space*, Theoretical Computer Science, 341(1-3) (2005) 91–137; https://doi.org/10.1016/j.tcs.2005.03.054

[2] **J Avigad**, **P Gerhardy**, **H Towsner**, *Local stability of ergodic averages*, Transactions of the American Mathematical Society 362, Number 1 (2010) 261–288; https://doi.org/10.1090/s0002-9947-09-04814-4

[3] **L Bienvenu**, **A Day**, **M Hoyrup**, **I Mezhirov**, **A Shen**, *A constructive version of Birkhoff's ergodic theorem for Martin-Lof random points*, Information and Computation 210 (2012) 21–30; https://doi.org/10.1016/j.ic.2011.10.006

[4] **G J Chaitin**, *Algorithmic Information Theory*, Cambridge Univ. Press (1987); https://doi.org/10.1017/cbo9780511608858

[5] **G Davie**, *Recursive events in random sequences*, Archive for Mathematical Logic 40, 629–638 (2001); https://doi.org/10.1007/s001530100075

[6] **J Franklin**, **N Greenberg**, **J Miller**, **K Ng**, (2012). *Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets*, Proceedings of the American Mathematical Society, 140(10), 3623–3628; https://doi.org/10.1090/s0002-9939-2012-11179-7

[7] **R Hölzl**, **P Shafer**, *Universality, optimality, and randomness deficiency*, Annals of Pure and Applied Logic Volume 166, Issue 10 (2015) 1049–1069; https://doi.org/10.1016/j.apal.2015.05.006

[8] **M Hoyrup**, **C Rojas**, *Applications of Effective Probability Theory to Martin-Löf Randomness*, ICALP 2009. Lecture Notes in Computer Science, vol 5555. Springer, Berlin, Heidelberg (2009); https://doi.org/10.1007/978-3-642-02927-1_46

[9] **M Hoyrup**, **C Rojas**, *An Application of Martin-Löf Randomness to Effective Probability Theory*, In: Ambos-Spies K., Löwe B., Merkle W. (eds) Mathematical Theory and Computational Practice. CiE Lecture Notes in Computer Science, vol 5635. Springer, Berlin, Heidelberg (2009) 260–269; https://doi.org/10.1007/978-3-642-03073-4_27

[10] **M Hoyrup**, **C Rojas**, *Computability of probability measures and Martin-Löf randomness over metric spaces*, Information and Computation, 207(7) (2009) 830–847; https://doi.org/10.1016/j.ic.2008.12.009

[11] **M Hoyrup**, *Algorithmic randomness and layerwise computability*, in Algorithmic Randomness: Progress and Prospects, edited by Johanna N. Y. Franklin and Christopher P. Porter, Lecture Notes in Logic, Cambridge University Press, Cambridge (2020) 115–133; https://doi.org/10.1017/9781108781718.005

[12] **A Kučera**, *Measure, Π^0_1-classes and complete extensions of PA*, in Ebbinghaus, HD., Müller, G.H., Sacks, G.E. (eds) Recursion Theory Week, Lecture Notes in Mathematics, vol 1141, Springer, Berlin, Heidelberg (1985); https://doi.org/10.1007/BFB0076224

[13] **U Krengel**, *On the speed of convergence in the ergodic theorem*, Monatsh. Math., 86:3-6, (1978); https://doi.org/10.1007/BF01300052

[14] **P Martin-Löf**, *The definition of random sequences*, Information and Control, 9(6):602-619 (1966); https://doi.org/10.1016/s0019-9958(66)80018-9

[15] **A Pauly**, **W Fouché**, **G Davie**, *Weihrauch-completeness for layerwise computability*, Logical Methods in Computer Science 14(2) (2018); https://doi.org/10.23638/LMCS-14(2:11)2018

[16] **A Rokhlin**, *A "general" measure-preserving transformation is not mixing*, Doklady Akademii Nauk SSSR (N.S.) 60:349-351 (1948)

[17] **V G Vovk**, *The law of the iterated logarithm for random Kolmogorov, or chaotic, sequences*, Theory of Probability and Applications, 32:413–425 (1988); https://doi.org/10.1137/1132061

[18] **K Weihrauch**, Computable Analysis: An Introduction. Springer, Berlin (2000); https://doi.org/10.1007/978-3-642-56999-9

Department of Decision Sciences, School of Economic and Financial Sciences, University of South Africa, South-Africa

davieg@unisa.ac.za

Received: 10 September 2019 Revised: 18 July 2022