ALGORITHMIC RANDOMNESS, MARTINGALES AND DIFFERENTIABILITY

JASON RUTE

Abstract. In this paper, a number of almost-everywhere convergence theorems are looked at using computable analysis and algorithmic randomness. These include various martingale convergence theorems and almost-everywhere differentiability theorems. General conditions are given for when the rate of convergence is computable and for when convergence takes place on the Schnorr random points. Examples are provided to show that these almost-everywhere convergence theorems characterize Schnorr randomness.

Contents

1. Introduction

The subjects of analysis and probability contain many convergence theorems of the following form.

A.E. Convergence Theorem. *If a sequence of functions $(f_n)_{n\in\mathbb{N}}$ satisfies some property P, then (f_n) converges to some integrable function f almost everywhere as $n \to \infty$. (Alternatively, $(f_r)_{r>0}$ converges to f as $r \to 0$.)*

Consider the following closely related examples.

Example 1.1. (Lebesgue differentiation theorem) If $g\colon [0,1] \to \mathbb{R}$ is integrable, then $\frac{1}{2r}\int_{x-r}^{x+r} g(y)\,dy \to g(x)$ for almost every x as $r \to 0$.

Example 1.2. (Lebesgue's theorem) If $f\colon [0,1] \to \mathbb{R}$ is a function of bounded variation function, then f is differentiable almost everywhere. (In this case, $f_r(x) = \frac{g(x+r)-g(x-r)}{2r}$.)

Example 1.3. (Doob's martingale convergence theorem) If (M_n) is a martingale and $\|M_n\|_{L^1} < \infty$, then M_n converges almost everywhere to an integrable function.

Example 1.4. (Ergodic theorem) If g is integrable, and T is a measure preserving transformation, then $\frac{1}{n}\sum_{k<n} g(T^k(x))$ converges almost everywhere. If T is ergodic, then $\frac{1}{n}\sum_{k<n} g(T^k(x)) \to \int g(x)\;dx$ for almost every x as $n \to \infty$.

For all the above theorems, it is natural to ask the following computability questions:

Question 1. *Is the rate of convergence effective (in the parameters of the theorem)?*

It is well known what it means for a sequence of functions to converge effectively in normed spaces like L^1 and L^2. A similar characterization can be given for almost everywhere convergence: a sequence of functions (f_n) converges to f with a effective rate of almost everywhere convergence if given $\varepsilon > 0$ and $\delta > 0$, we can compute some $m \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq m$ and all x except on a set of size less than δ.

Some a.e. convergence theorems have effective rates of convergence. For example, Avigad, Gerhardy, and Towsner [2] showed that the rate of almost everywhere convergence in the ergodic theorem is computable from T and g when T is ergodic. I will show a similar result for the Lebesgue differentiation theorem.

However, not all the theorems have computable rates of convergence. This is the case for Lebesgue's theorem, Doob's martingale convergence theorem, and the ergodic theorem (in the nonergodic setting). However, when certain additional conditions are assumed, one can then compute a rate of convergence.

Question 2. *If the rate of convergence is not effective, what are additional conditions that guarantee an effective rate of convergence.*

For example, Avigad et al. [2] showed that the rate of convergence in the ergodic theorem is computable from g, T and the limit g^*. (Note, it is not trivial to compute the rate of convergence from the limit of a series. For example, it is easy to construct a computable sequence of constant functions which converge to 0, but do not do so effectively.) In the L^2-case, Avigad et al. [2] showed the rate of convergence is computable from g, T, and the L^2-norm of g.

In this paper I will give similar results for Lebesgue's theorem, Doob's martingale convergence theorem, and others. All the results follow the pattern in this observation.[1]

Observation 1. *For most a.e. convergence theorems, a rate of almost everywhere convergence is computable from the sequence* (f_n)*, the limit* f*, and the bounds* $\inf_n \|f_n\|_{L^1}$, $\sup_n \|f_n\|_{L^1}$.

In many cases, such as in the ergodic theorem, $\inf_n \|f_n\|_{L^1}$ and $\sup_n \|f_n\|_{L^1}$ are computable from the sequence (f_n) and the limit f, and therefore they are not explicitly needed. In other cases, such as the Lebesgue differentiation theorem, all three extra conditions are naturally computable from the parameters of the theorem (which is why the rate of convergence in the Lebesgue differentiation theorem is computable without additional assumptions). Further, if we work in L^2 instead of L^1, we do not need the limit f, just its L^2-norm $\|f\|_{L^2}$.

Question 3. *At which points does the sequence converge (under various computability conditions)?*

For example, if we consider the Lebesgue differentiation theorem, we can ask at which x does $\frac{1}{2r}\int_{x-r}^{x+r} f(y)\,dy$ converge for all f computable in the L^1-norm. Notice the set of such x is measure one, since there are only countably-many f computable in the L^1-norm.

This question was first asked by Pathak [38] using the tools of algorithmic randomness. Algorithmic randomness classifies measure-one sets of points that behave randomly with respect to "computable tests". Pathak showed that convergence happens on all Martin-Löf random x. She left it as an open question whether this

[1]It is important to note that Observation 1 is not itself a theorem or metatheorem. Indeed, there are (contrived) cases where it fails to hold—let f_n be some computable sequence of constant functions converging to zero with a noncomputable rate of convergence.

could be strengthened to a larger class of points. In this paper, I will show that it can be strengthened to Schnorr randomness, and that this is the best possible. In other words, the Lebesgue differentiation theorem characterizes Schnorr randomness. This same result was independently and concurrently discovered by Pathak, Rojas, and Simpson [39].

Similar investigations have been made into randomness and the ergodic theorem [49, 36, 24, 20, 15, 3, 16], randomness and Lebesgue's theorem [9, 7, 17], and randomness and martingale convergence [46]. In this paper, I expand on these results, specifically looking at Schnorr randomness. Indeed, I ask this converse to Question 3.

Question 4. *Which conditions guarantee convergence on Schnorr randoms?*

It turns out the answers to Questions 1 and 2 provide an answer, when using this informal observation—which will be made formal in Lemma 3.19.

Observation 2. *Effective a.e. convergence implies convergence on Schnorr randoms.*

This will allow us to "kill two birds with one stone", by focusing on questions in computable analysis (Questions 1 and 2), we can answer questions in algorithmic randomness (Questions 3 and 4) for free. However, to show that one cannot strengthen Schnorr randomness to a larger class of random points, we will need an example for each theorem showing that if x is not Schnorr random, then there are computable parameters for which convergence does not happen on x. I provide a number of such examples.

1.1. **Summary of results.** The results of this paper are diverse and the paper is organized by the tools and lemmas needed to prove the theorems. Table 1 is a summary of all the known convergence theorems which characterize Schnorr randomness. The first column is a short description of the convergence theorem. The second column is a reference to the result showing that the sequence in question converges on all Schnorr randoms. The third column is a reference to the result showing that if a point is not Schnorr random, then there exists such a sequence which fails to converge on that point. If a cell is blank, that direction is subsumed by a stronger result in another row. A "?" means this direction is still an open question (and so that row may not really be a characterization of Schnorr randomness). Some of the results are due to others, or were independently discovered. I provide footnotes in these cases.

1.2. **A comment on the martingale results.** A significant portion of this paper concerns martingales. Informally, martingales are formalizations of gambling strategies—a martingale (M_n) is a sequence of random variables representing the capital of a gambling strategy at time n. They are widely used in probability theory and analysis, as well as in algorithmic randomness. In algorithmic randomness, martingales can be used to characterizes a number of randomness notions, including Schnorr randomness, computable randomness, and Martin-Löf randomness (see [11, 37]). However, there is a difference between how martingales are treated in algorithmic randomness and how they are used in probability theory and analysis.

For one, in algorithmic randomness, the martingales are (usually) gambling strategies on coin flips. Such martingales, which I will call dyadic martingales, are a specific instance of the more general martingales used in probability theory.

Table 1: Characterizations of Schnorr randomness by a.e. convergence theorems. (See Section 1.1 for an explanation of this table.)

Convergence of martingales: $M_n \to M_\infty$		
(M_n) is L^1-comp.; M_∞ is L^1-comp.; $\sup_n \|M_n\|_{L^1}$ is computable	Thm. 7.10	
(M_n) is uniformly integrable, L^1-comp.; M_∞ is L^1-computable	Lem. 6.5	
(M_n) is nonnegative, singular ($M_\infty = 0$), L^1-computable	Lem. 7.4	
(M_n) is nonnegative, singular ($M_\infty = 0$), computable		Thm. 12.9
(M_n) is L^2-computable; $\sup_n \|M_n\|_{L^2} = \|M_\infty\|_{L^2}$ is computable	Cor. 6.8	
(M_n) is nonneg., unif. int., dyadic, computable; $\|M_\infty\|_{L^2}$ is computable		Thm. 12.7
Convergence of super/submartingales: $M_n \to M_\infty$		
(M_n) is L^1-comp., super/submart.; $\lim_n \|M_n\|_{L^1}$ is comp.; M_∞ is L^1-comp.	?[2]	
(M_n) is nonnegative, L^1-comp., supermart.; M_∞ is L^1-comp.	Thm. 8.1	
(M_n) is nonnegative, L^1-comp., supermart.; $M_\infty = 0$	Lem. 7.4	Thm. 12.9
(M_n) is nonneg., L^1-comp., submart.; M_∞ is L^1-comp.; $\sup_n \|M_n\|_{L^1}$ is comp.	Thm. 8.5	Thm. 12.6
Convergence of reverse martingales: $M_{-n} \to M_{-\infty}$		
(M_{-n}) is L^1-computable; $M_{-\infty}$ is L^1-computable	Thm. 11.2	
(M_{-n}) is L^2-computable; $\|M_{-\infty}\|_{L^2}$ is L^2-computable	Thm. 11.2	
(M_{-n}) is bounded, a.e. computable; M_∞ is computable constant		Cor. 12.17
Lebesgue differentiation theorem: $\int_{B(x,r)} f(y)\,dy/\lambda(B(x,r)) \to f(x)$		
f is L^1-computable	Thm. 4.10, [39][3]	Thm. 12.3
f is L^2-computable		
f is bounded, L^1-computable		[39, 17]
Lebesgue density theorem: $\lambda(A \cap B(x,r))/\lambda(B(x,r)) \to \mathbf{1}_A(x)$		
A is effectively measurable	Cor. 4.16	[39, 17][4]
A is effectively closed; $\lambda(A)$ is computable	Cor. 4.16	?
Differentiability of functions f (with derivative Df and total variation $V(f)$)		
f is comp. on dense set, bounded var.; Df is L^1-comp.; $V(f)$ is comp.	Thm. 9.19	
f is comp., bounded variation; Df is L^1-comp.; $V(f)$ is comp.	Cor. 9.20	
f is absolutely continuous; Df is L^1-comp.	Cor. 4.18, [17][5]	
f is computable, absolutely continuous; $\|Df\|_{L^2}$ is computable	Cor. 6.9	
f is increasing, effectively absolutely continuous, comp.; $\|Df\|_{L^2}$ is comp.		Cor. 12.5[6]
f is increasing, Lipschitz, effectively absolutely continuous		[17]
f is computable, increasing, singular ($Df = 0$ a.e.)		Cor. 12.15
f is comp. on dense set, increasing, singular, only contains jumps		Cor. 12.12

[2]See Problem 8.6.

[3]This was independently discovered by this author and by Pathak, Rojas, and Simpson [39].

[4]While neither paper makes this explicit, the example function f they each give for the Lebesgue differentiation theorem is $0, 1$ valued and therefore the indicator function of some effectively measurable set A.

[5]This result is a direct corollary of the effective Lebesgue differentiation theorem (Thm. 4.10, [39]) that was noticed by this author and Freer, Kjos-Hanssen, Nies, and Stephan [17].

[6]This also follows from the Lipschitz result of Freer, Kjos-Hanssen, Nies, and Stephan [17] in the next line.

$\nu(B(x,r))/\lambda(B(x,r) \to \frac{d\nu}{d\lambda}(x)$ for signed measures ν		
ν is computable; $\frac{d\nu}{d\lambda}$ is L^1-comp.; $\|\nu\|_{TV}$ is comp.	Thm. 9.12	
ν is absolutely continuous, computable, positive; $\frac{d\nu}{d\lambda}$ is L^1-comp.	Cor. 4.21	Cor. 12.4
ν is continuous, singular ($\frac{d\nu}{d\lambda} = 0$ a.e.), computable, positive		Cor. 12.14
ν is atomic, singular ($\frac{d\nu}{d\lambda} = 0$ a.e.), computable, positive		Cor. 12.11

Ergodic theorem: $\frac{1}{n}\sum_{k<n} f \circ T^n \to f^*$		
f is L^1-comp.; T is effectively measurable; f^* is L^1-comp.	Thm. 10.2[7]	
f is L^2-comp.; T is effectively measurable; $\|f^*\|_{L^2}$ is comp.	Thm. 10.2	
f is a.e. comp.; T is a.e. comp., ergodic		[20]

Monotone convergence thm: Convergence of (f_n) increasing		
(f_n) is L^1-comp.; $\|f_n\|_{L^1}$ is computable	Prop. 8.2	
(f_n) is L^2-comp.; $\|f_n\|_{L^2}$ is computable	Prop. 8.2	
(f_n) is computable; $\|f_n\|_{L^2}$ is computable		Thm. 12.2[8]

$\varphi_n \to f$ for test functions, dyadic averages, trigonometric polynomials		
φ_n is fast Cauchy sequence of test functions in L^1 (or in measure)	Prop. 3.18, [39][9]	
φ_n is fast Cauchy sequence of test functions converging to 0 in L^2		Thm. 12.18
$f^{(n)}$ are dyadic averages; f is L^1-computable	Prop. 4.6	Thm. 12.6
$\sigma_n(f)$ are trig. polynomials (from Fejér kernel); f is $L^1(\mathbb{T} \to \mathbb{C})$-comp.	Cor. 4.23	?

SLLN, de Finetti's thm: Convergence of $\frac{1}{n}\sum_{k<n} X_k$ for integrable random variables (X_n)		
(X_n) is L^1-comp., exchangeable (de Finetti's theorem)	Cor. 11.11	
(X_n) is L^1-computable, i.i.d. (strong law of large numbers)	Cor. 11.7	Cor. 12.17

[7] Theorem 10.2 is a summary of the results from [2, 20, 21, 39] with a few gaps filled in.

[8] This is closely related to the Schnorr integral tests of Miyabe [33].

[9] This is also a theorem of Pathak, Rojas, and Simpson [39] where the test functions are rational polynomials and convergence is in L^1.

By considering, this larger class of martingales, we can ask new questions that could not be asked of dyadic martingales.

Another difference is that algorithmic randomness is more concerned with success than convergence. We say that a martingale (M_n) succeeds on x if $\limsup_n M_n(x) = \infty$, that is the strategy M_n wins arbitrarily large amounts of money on x. Of the three most common randomness notions—Schnorr randomness, computable randomness, and Martin-Löf randomness—only computable randomness has a well-known characterization in terms of martingale convergence instead of martingale success.[10]

For example, consider the following three characterization of Schnorr randomness. The characterization in (2) is the classical martingale characterization of Schnorr randomness and (3) and (4) are new characterizations which follow from results in this paper (Theorem 7.10, Corollary 6.8, and Theorem12.6). (Note, in (3) and (4) we could replace convergence with success and the characterization would still hold.)

Example 1.5. Recall, a computable dyadic martingale is a computable function $M\colon 2^{<\omega} \to \mathbb{R}$ such that $\frac{1}{2}M(\sigma 0) + \frac{1}{2}M(\sigma 1) = M(\sigma)$. Use the notation $M_n(x) = M(x \upharpoonright n)$. The following are equivalent.

(1) $x \in 2^{\mathbb{N}}$ is Schnorr random (on the fair-coin measure).
(2) (Classical) For all nonnegative computable dyadic martingales (M_n) and all computable, nondecreasing, unbounded functions $h\colon \mathbb{N} \to \mathbb{N}$, we have that $M_n(x) \leq h(n)$ for all but finitely-many n.
(3) (New) For all nonnegative computable dyadic martingales (M_n) such that $\lim_n M_n$ is L^1-computable, we have that $M_n(x)$ converges.
(4) (New) For all nonnegative, computable dyadic martingales (M_n) such that $\sup_n \|M_n\|_{L^2}$ is computable, we have that $M_n(x)$ converges.

However, the results in this paper go far beyond giving a new dyadic martingale characterization of Schnorr randomness. Not only does algorithmic randomness provide us with new tools to study computable analysis; computable analysis also provides us new tools to study algorithmic randomness. Martingales are one such tool. This paper makes significant use of martingales to prove results. One particular type of martingale not previously used in algorithmic randomness is the backwards martingale. To demonstrate their usefulness in algorithmic randomness, I use backwards martingales to prove a new variation of Kučera's theorem in Corollary 11.4: for every Schnorr random $x \in 2^{\mathbb{N}}$ and for every closed set C of positive computable measure, C contains some y which equals x, except that finitely many bits are permuted.

1.3. **A comment on measurable functions in computable analysis.** There is an inherent challenge when working with measurable functions in a computable

[10]There are not-widely-known published results, which when combined, give a martingale-convergence characterization of Martin-Löf randomness. Takahashi [46] showed that Doob's upcrossing inequality implies that computable martingales (that is, martingales in the more general sense of probability theory where (M_n) is a computable sequence of computable functions) converge on Martin-Löf randoms. Edward Dean [personal communication], independently, showed that layerwise-computable martingales converge on Martin-Löf randoms. Merkle, Mihalović, and Slaman [32] gave an example of a computable martingale (in the more general sense of probability theory) which only converges on Martin-Löf randoms.

setting. Measurable functions are not continuous and therefore it is difficult to describe them as maps in a computable manner. Moreover, a single function is best thought of as an equivalence class (under a.e. equivalence). It is challenging to talk about the value $f(x)$ when f is an equivalence class (an important issue when asking about which points an a.e. convergence theorem holds!).

Some authors have taken the easy approach and restricted their attention to computable functions or a.e. computable functions. However, in this paper, I will try to express the theorems in full generality. In order to do this, I will need a clear theory of effectively measurable functions. The space of measurable functions (modulo a.e. equivalence) is naturally described as a computable metric space under a suitable metric which characterizes convergence in measure. The class of effectively measurable functions includes the real-valued functions computable in the L^1-norm as well as other functions (which may not even be integrable or real-valued).

In order to talk about the valuation of functions, each effectively measurable function f will have a representative $\widetilde{f}$. This representative is well-defined (and well-behaved) on Schnorr random points. This representative approach is adapted from Pathak [38] (and is also used in Pathak, Rojas, and Simpson [39]). The same ideas are implicit in the reverse mathematics of the dominated convergence theorem [57, 1].

Other computable approaches to measurable sets and functions include [4, 42, 30, 19, 55, 54, 14, 23, 33]. These approaches are essentially the same as either the metric space approach or the representative approach used in this paper. The biggest difference is that some representative approaches—e.g. layerwise computability [23]—only define f on the Martin-Löf random points. However, it is possible to uniquely extend each such f to the Schnorr random points.

My hope is that Section 3 (on effectively measurable functions) not only serves the needs of this paper, but is of use to other researchers in the field.

1.4. **Outline of the paper.** In Section 2, I give background on computable analysis and Schnorr randomness.

In Section 3, I present a theory of measurable functions, integrable functions, and measurable sets. This also includes the important Lemma 3.19 that effective a.e. convergence implies convergence on Schnorr randoms. Most of the proofs have been moved to Appendix A.

In Section 4, I prove an effective version of the Lebesgue differentiation theorem and discuss many of its corollaries. The proof relies on Kolmogorov's inequality for dyadic martingales.

In Section 5, I give a computable presentation of martingale theory, which will be needed for most of the rest of this paper.

In Section 6, I prove an effective version of the Lévy 0-1 law, which is a simpler version of Doob's martingale convergence theorem and an analog to the Lebesgue differentiation theorem.

In Section 7, I prove an effective version of the martingale convergence theorem. I also give another version for square integrable martingales.

In Section 8, I prove an effective version of the submartingale and supermartingale convergence theorems. I also give another version for square integrable martingales.

In Section 9, I return to differentiability, using effective martingale convergence to prove more differentiability results that extend the Lebesgue differentiation theorem and Lebesgue's theorem.

In Section 10, I survey some results in ergodic theory, filling in gaps in the published literature.

In Section 11, I discuss backwards martingales and some of their applications, including a variation of Kučera's theorem, the strong law of large numbers, and de Finetti's theorem. I also, compare them with ergodic averages.

I intend to follow up this paper with a sequel, exploring martingale convergence and differentiability when the limit is not computable. Such cases characterize computable randomness, Martin-Löf randomness, and weak-2 randomness.

1.5. **Acknowledgments.** I would like to thank André Nies, Jeremy Avigad, and Bjørn Kjos-Hanssen for many helpful corrections on earlier drafts of this paper. I would also like to thank Laurent Bienvenu, Johanna Franklin, Mathieu Hoyrup, Kenshi Miyabe, Noopur Pathak, Cristóbal Rojas, and Stephen Simpson for helpful discussions on parts of this work.

2. Background

In this section I give the necessary background in computable analysis, effective measure theory, effective probability theory, and Schnorr randomness.

2.1. **Notation.** Let $2^{<\omega}$ be the space of finite binary strings, $2^{\mathbb{N}}$ be the space of infinite binary strings, $\varnothing_{string}$ be the empty string, $\sigma \prec \tau$ and $\sigma \prec x$ mean σ is a proper initial segment of $\tau \in 2^{<\omega}$ or $x \in 2^{\mathbb{N}}$, $[\sigma] = \{x \in 2^{\mathbb{N}} \mid \sigma \prec x\}$. Also for $\sigma \in 2^{<\omega}$ (or $x \in 2^{\mathbb{N}}$), let $\sigma(n)$ be the nth digit of σ (where $\sigma(0)$ is the "0th" digit) and $\sigma \upharpoonright n = \sigma(0) \cdots \sigma(n-1)$. A set of strings $\{\sigma_0, \sigma_1, \ldots\}$ is prefix free if the no string in the set is a prefix of another (equivalently, the collection $\{[\sigma_0], [\sigma_1], \ldots\}$ is pair-wise disjoint).

2.2. **Computable analysis.** Here I present some basics of computable analysis. For additional information on the basics see Pour El and Richards [40], Weihrauch [51], or Brattka et al. [6]. I assume the reader has some familiarity with basic computability theory on $\mathbb{N}$, $2^{\mathbb{N}}$, and $\mathbb{N}^{\mathbb{N}}$ as in [45]. It would also help to have some familiarity with the theory of computation on the reals.

Definition 2.1. Fix an enumeration of the rationals $\mathbb{Q} = \{q_i\}_{i \in \mathbb{N}}$ (such that addition and multiplication are computable). A real $x \in \mathbb{R}$ is COMPUTABLE if there is a computable function $h \colon \mathbb{N} \to \mathbb{N}$ such that for all $m > n$, we have $|q_{h(m)} - q_{h(n)}| \leq 2^{-n}$ and $x = \lim_{n\to\infty} q_{h(n)}$.

This can be generalized to an arbitrary complete metric space.

Definition 2.2. A COMPUTABLE (POLISH) METRIC SPACE is a triple $\mathbb{X} = (X, d, S)$ such that

(1) X is a complete metric space with metric $d \colon X \times X \to [0, \infty)$.
(2) $S = \{a_i\}_{i \in \mathbb{N}}$ is a countable dense subset of X (the SIMPLE POINTS of $\mathbb{X}$) .
(3) The distance $d(a_i, a_j)$ is computable uniformly from i and j.

A point $x \in X$ is said to be COMPUTABLE if there is a computable function $h \colon \mathbb{N} \to \mathbb{N}$ such that for all $m > n$, we have $d(a_{h(m)}, a_{h(n)}) \leq 2^{-n}$ and $x = \lim_{n\to\infty} a_{h(n)}$. The sequence $(a_{h(m)})$ is the CAUCHY-NAME for x.

Example 2.3. For the differentiability results, I will be using two spaces. The first is the unit cube $[0,1]^d$ with the usual Euclidean distance. The second is the unit torus $\mathbb{T}^d := (\mathbb{R}/\mathbb{Z})^d$, which will be identified as the half open unit cube $[0,1)^d$ with the Euclidean metric that wraps around each edge, i.e. given $x = (x_1, \ldots, x_d), y = (y_1, \ldots, y_d) \in [0,1)^d$,

$$d(x,y) = \left(\sum_{i=1}^{d} \Big(\min\{|x_i - y_i|, 1 - |x_i - y_i|\} \Big)^2 \right)^{1/2}.$$

The simple points of $\mathbb{T}^d$ and $[0,1]^d$ are taken to be vectors with rational components. A little thought reveals that a vector $x \in [0,1]^d$ (or $x \in \mathbb{T}^d$) is computable if and only if each coordinate is a computable real.

On a computable metric space $\mathbb{X} = (X, S, d)$, the BASIC OPEN BALLS are sets of the form $B(a,r) = \{x \in X \mid d(x,a) < r\}$ where $a \in S$ and $r > 0$ is rational. The Σ^0_1 sets (EFFECTIVELY OPEN SETS) are computable unions of basic open balls; Π^0_1 sets (EFFECTIVELY CLOSED SETS) are the complements of Σ^0_1 sets. A function $f \colon \mathbb{X} \to \mathbb{R}$ is COMPUTABLE (or EFFECTIVELY CONTINUOUS) if for each Σ^0_1 set U in $\mathbb{R}$, the set $f^{-1}(U)$ is Σ^0_1 in $\mathbb{X}$ (uniformly in U), or equivalently, there is an algorithm which sends every Cauchy-name of x to a Cauchy-name of $f(x)$ (see [51]). A function $f \colon \mathbb{X} \to [0,\infty]$ is LOWER SEMICOMPUTABLE if it is the supremum of a computable sequence of computable functions $f_n \colon \mathbb{X} \to [0,\infty)$.

A real x is said to be LOWER (UPPER) SEMICOMPUTABLE if it is the supremum (resp. infimum) of a computable sequence of rationals.

Definition 2.4. If $\mathbb{X} = (X, d, S)$ is a computable metric space, then a Borel measure μ is a COMPUTABLE MEASURE on $\mathbb{X}$ if $\int g\, d\mu$ is computable uniformly from g for all computable $g \colon \mathbb{X} \to [0,1]$. A COMPUTABLE PROBABILITY SPACE is a pair $(\mathbb{X}, \mu)$ where $\mathbb{X}$ is a computable metric space, μ is a computable measure on $\mathbb{X}$, and $\mu(X) = 1$.

There are a number of other equivalent definitions of computable measure, including the following characterization.

Proposition 2.5 ([44, 25])**.** *A measure μ on a computable metric space $\mathbb{X} = (X, d, S)$ is computable if and only if the value $\mu(X)$ is computable, and for each effectively open set $U \subseteq \mathbb{X}$, the measure $\mu(U)$ is lower semicomputable uniformly from U.*

Moreover, the computable probability measures on $\mathbb{X}$ are exactly the computable points in the space of probability measures under the Prokhorov metric.

I will often blur the distinction between a metric space—or a probability space—with its set of points, e.g. writing $x \in \mathbb{X}$ or $x \in (\mathbb{X}, \mu)$ to mean that $x \in X$ where $\mathbb{X} = (X, d, S)$.

Example 2.6. The manifolds $[0,1]^d$ and $\mathbb{T}^d$ can be endowed with the LEBESGUE MEASURE λ. (The Lebesgue measure on $\mathbb{T}^d$ is just the Lebesgue measure on $[0,1)^d$.) Both are computable probability measures, and further $(\mathbb{T}^d, \lambda)$ is translation-invariant. Similarly, on $2^{\mathbb{N}}$ let λ be the FAIR-COIN MEASURE, i.e. the measure such that $\lambda([\sigma]) = 2^{-|\sigma|}$.

Definition 2.7. Let $\mathbb{X} = (X, S, d)$ be a computable metric space.

(1) $(\mathbb{X}, +, \cdot)$ is a COMPUTABLE (TOPOLOGICAL) VECTOR SPACE if $\mathbb{X}$ is a vector space and with computable vector addition $+$ and scalar multiplication $\cdot$ operations.
(2) $(\mathbb{X}, \|\cdot\|, +, \cdot)$ is a COMPUTABLE BANACH SPACE if $(\mathbb{X}, +, \cdot)$ is a computable vector space and the metric d comes from a computable norm $\|\cdot\|$.
(3) $(\mathbb{X}, \|\cdot\|, \langle\rangle, +, \cdot,)$ is a COMPUTABLE HILBERT SPACE if $(\mathbb{X}, \|\cdot\| +, \cdot)$ is a computable Banach space with computable inner product $\langle\cdot,\cdot\rangle$.
(4) $(\mathbb{X}, \wedge, \vee)$ is a COMPUTABLE (TOPOLOGICAL) LATTICE if $\mathbb{X}$ is a lattice with computable meet $\wedge$ and join $\vee$ operations.
(5) $(\mathbb{X}, \wedge, \vee, \neg, \bot, \top)$ is a COMPUTABLE (TOPOLOGICAL) BOOLEAN ALGEBRA if $\mathbb{X}$ is a Boolean algebra with computable meet $\wedge$, join $\vee$, and complement $\neg$ operations and computable bottom $\bot$ and top $\top$ elements.

Remark 2.8. There are a number of natural Banach spaces that are not computable, for example the space of signed Borel measures on $[0, 1]$. This is because they have no countable dense subset. However, we may still represent these spaces using a weaker topology as will be done in Section 9.1.

2.3. **Schnorr randomness.**

Definition 2.9. Let $(\mathbb{X}, \mu)$ be a computable probability space. A SCHNORR TEST (U_n) is a computable sequence of effectively open sets U_n such that $\mu(U_n) \leq 2^{-n}$ for all n and $\mu(U_n)$ is uniformly computable in n. For any $x \in \mathbb{X}$, say x is COVERED by (U_n) if $x \in \bigcap_n U_n$. Say $x \in \mathbb{X}$ is SCHNORR RANDOM if x is not covered by any Schnorr test.

Remark 2.10. We may assume a Schnorr test (U_n) is decreasing by taking an intersection. Similarly, we may also replace 2^{-n} by any computable sequence that decreases to 0 by taking a subsequence (see [11, 37]).

Example 2.11. Let $y_1, \ldots, y_d \in [0, 1]$ (resp. $\mathbb{T}$). For each $1 \leq i \leq d$, let x_i be some binary expansion of y_i. It is easy to see that $(y_1, \ldots, y_d)$ is Schnorr random on $([0, 1]^d, \lambda)$ (resp. on $(\mathbb{T}^d, \lambda)$) if and only if $x_1 \oplus \cdots \oplus x_n \in 2^{\mathbb{N}}$ is Schnorr random on $(2^{\mathbb{N}}, \lambda)$. (Recall, $x_1 \oplus x_2$ is the join operation on $2^{\mathbb{N}}$ defined by $(x_1 \oplus x_2)(2n) := x_1(n)$ and $(x_1 \oplus x_2)(2n + 1) := x_2(n)$.)

3. FUNCTIONS AND CONVERGENCE IN MEASURE THEORY

This section provides background on measurable functions and convergence. It is quite important to the results in this paper. (For example, the frequently used Lemma 3.19 is the only fact the reader will need to know about Schnorr randomness in Sections 4 through 11.)

As mentioned in the introduction, there is a need for two approaches to working with measurable functions (and sets).[11]

(1) Use *equivalence classes* of almost-everywhere equivalent objects.
(2) Use *specific functions and sets* that are defined and unique up to some *specific* measure-one set (which will turn out to be the set of Schnorr random points).

Table 2 compares the two approaches (in the setting of L^1-computable functions).

[11] A third approach may come to mind: use Borel measurable functions and sets, ignoring a.e. equivalence. The difficulty with this approach is that even effectively open sets may not have a computable measure. The situation becomes more complex as one moves up the Borel hierarchy.

Equivalence classes	Specific functions
f an L^1-limit of fast Cauchy sequences	$\widetilde{f}$ a pointwise limit of fast Cauchy sequences
f unique a.e.	$\widetilde{f}$ unique on Schnorr randoms
f computable in the L^1-norm	$\widetilde{f}$ "computable" on Schnorr randoms

TABLE 2. The two approaches to the computability of L^1 functions.

Besides giving definitions and basic facts, the main result of this section is Lemma 3.19, that a.e. convergence implies convergence on Schnorr randoms. (This fact has been hinted at in some of the work on convergence for Schnorr randoms, including Pathak, Rojas, and Simpson [39]. It was also known to Hoyrup and Rojas [personal communication] independently of this author.)

3.1. **Integrable functions, measurable functions, and measurable sets.** Let us start with real-valued functions on the space $(2^{\mathbb{N}}, \lambda)$.

Proposition 3.1. *On $(2^{\mathbb{N}}, \lambda)$ the following hold.*

(1) *(Functions) Consider the following spaces (of a.e. equivalence classes $[f]_\sim$) of Borel measurable functions. Let the* TEST FUNCTIONS $\mathcal{T}$ *be those of the form*

$$\varphi = \sum_{i=0}^{k-1} c_i \mathbf{1}_{[\sigma_i]} \quad (\sigma_0, \ldots, \sigma_{k-1} \in 2^{<\omega}; c_{0,} \ldots, c_{k-1} \in \mathbb{Q}). \tag{3.1}$$

Also consider the lattice given by

$$f \wedge g = \min(f, g) \quad \text{and} \quad f \vee g = \max(f, g).$$

(a) *The* MEASURABLE FUNCTIONS $L^0(2^{\mathbb{N}}, \lambda)$ *with the metric*[12]

$$d_{meas}(f, g) = \int \min\{|f - g|, 1\}\, d\lambda$$

form a computable metric space, a computable vector space, and a computable lattice

$$(L^0(2^{\mathbb{N}}, \lambda), \mathcal{T}, d_{meas}, +, \cdot, \min, \max).$$

(b) *The* INTEGRABLE FUNCTIONS $L^1(2^{\mathbb{N}}, \lambda)$ *with norm*

$$\|f\|_{L^1} = \int |f|\, d\lambda$$

form a computable Banach space and a computable lattice

$$(L^1(2^{\mathbb{N}}, \lambda), \mathcal{T}, \|\cdot\|_{L^1}, +, \cdot, \min, \max).$$

(c) *The* SQUARE INTEGRABLE FUNCTIONS $L^2(2^{\mathbb{N}}, \lambda)$ *with inner product and norm*

$$\langle f, g\rangle = \int f, g\, d\lambda, \quad \|f\|_{L^2} = \left(\int |f|^2\, d\lambda\right)^{1/2}$$

[12] As we shall see, this metric characterizes convergence in measure. It is equivalent to the Ky-Fan metric $d_{\mathrm{KF}}(f, g) := \inf\{\varepsilon > 0 \mid \mu(\{x \mid |f - g| \geq \varepsilon) \leq \varepsilon\}$. (Indeed, $(L^0(2^{\mathbb{N}}, \lambda), \mathcal{T}, d_{KF}, +, \cdot, \min, \max)$ is also a computable metric space, a computable vector space, and a computable lattice with the same computable points as $(L^0(2^{\mathbb{N}}, \lambda), \mathcal{T}, d_{meas})$.)

form a computable Hilbert space and a computable lattice

$$(L^2(2^{\mathbb{N}},\lambda),\mathcal{T},\|\cdot\|_{L^2},\langle\rangle_{L^2},+,\cdot,\min,\max).$$

(2) *(Set spaces) Consider the following space (of a.e. equivalence classes $[A]_\sim$) of Borel measurable sets. Let the* TEST SETS $\mathcal{T}$ *be those of the form*

$$C=\bigcup_{i=0}^{k-1}[\sigma_i]\quad(\textit{prefix-free }\sigma_0,\ldots,\sigma_{k-1}\in 2^{<\omega}).$$

(d) *The* MEASURABLE SETS $\mathcal{B}(2^{\mathbb{N}},\lambda)$ *with metric*

$$d(A,B)=\lambda(A\triangle B)$$

form a computable metric space and a computable Boolean algebra

$$(\mathcal{B}(2^{\mathbb{N}},\lambda),\mathcal{T},d,\cup,\cap,\cdot^c,\varnothing,2^{\mathbb{N}}).$$

Proof. straightforward. □

Definition 3.2. The computable points of each of the above spaces are, respectively, called the EFFECTIVELY MEASURABLE FUNCTIONS (L^0_{comp}), the L^1-COMPUTABLE FUNCTIONS (L^1_{comp}), the L^2-COMPUTABLE FUNCTIONS (L^2_{comp}), and the EFFECTIVELY MEASURABLE sets.

We may also consider measurable functions taking values in other computable metric spaces $\mathbb{Y}=(Y,S,d_{\mathbb{Y}})$.

Proposition 3.3. *Let $\mathbb{Y}=(Y,S,d_{\mathbb{Y}})$ be a computable metric space. The space of measurable functions from $(2^{\mathbb{N}},\lambda)$ to $\mathbb{Y}=(Y,S,d_{\mathbb{Y}})$ is a computable metric space under the metric*

$$d_{meas}(f,g)=\int\min\{d_{\mathbb{Y}},1\}\,d\lambda$$

and TEST FUNCTIONS *of the form*

$$\varphi(x)=c_i\mathbf{1}_{[\sigma_i]}\quad\textit{when } x\in[\sigma_i]\quad(\textit{prefix-free }\sigma_0,\ldots,\sigma_{k-1}\in 2^{<\omega};c_0,\ldots,c_{k-1}\in S).$$

The computable points in this space are called EFFECTIVELY MEASURABLE FUNCTIONS*.*

Proof. straightforward. □

Remark 3.4. The space of measurable sets and the space of $0,1$-valued measurable functions (Proposition 3.3 with $\mathbb{Y}=\{0,1\}$) are the same space. (More specifically, the map $A\mapsto\mathbf{1}_A$ is a bijective isometry where test sets are mapped to test functions.)

The above definitions extend to any computable probability space $(\mathbb{X},\mu)$. The only thing that changes is the choice of test functions. This requires a technical lemma.

Lemma 3.5 (Bosserhoff [5], Hoyrup and Rojas [25])**.** *For any computable metric space $\mathbb{X}=(X,\mathcal{S},d)$ with computable probability measure μ, there is a computable sequence of pairs $\{(a_i,r_i)\}_{i\in\mathbb{N}}$ ($a_i\in\mathcal{S}$, $r_i\in\mathbb{R}$) representing a family of balls $Basis(\mathbb{X},\mu)=\{B(a_i,r_i)\}_{i\in\mathbb{N}}$.*

(1) *Each $B(a_i,r_i)$ has a μ-null boundary. (Hence $\mu(B(a_i,r_i))$ is computable uniformly from i.)*

(2) *Basis*$(\mathbb{X},\mu)$ *is an effective basis of* $\mathbb{X}$*, i.e. for every effectively open set* U*, there is a computable sequence* (i_k) *of indices computable uniformly from (each name for)* U *such that* $U = \bigcup_{k=0}^{\infty} B(a_{i_k}, r_{i_k})$*.*

Since the choice of basis is not unique, let *Basis*$(\mathbb{X},\mu)$ denote a fixed choice of basis for each space $(\mathbb{X},\mu)$.

Definition 3.6. Say that $C \subseteq \mathbb{X}$ is a CELL of *Basis*$(\mathbb{X},\mu)$ if $C = A_1 \cap \ldots \cap A_\ell \cap B_1^c \cap \ldots \cap B_k^c$ for $A_1, \ldots A_\ell, B_1, \ldots, B_k \in$ *Basis*$(\mathbb{X},\mu)$. (Notice, using the enumeration of *Basis*$(\mathbb{X},\mu)$ that each cell is coded by some $\sigma \in 2^{<\omega}$.)

Proposition 3.7. *The measure of each cell of Basis*$(\mathbb{X},\mu)$ *is computable from its code* σ*.*

Proof. See Appendix A.2. □

Definition 3.8. On $(\mathbb{X},\mu)$, the spaces of real-valued functions $L^0(\mathbb{X},\mu)$, $L^1(\mathbb{X},\mu)$, $L^2(\mathbb{X},\mu)$ as well as the space of measurable sets and the space of $\mathbb{Y}$-valued measurable functions are defined as before, except that cylinder sets $[\sigma_i]$ are replaced with cells C_i of *Basis*$(\mathbb{X},\mu)$. Replace the requirement that $\{\sigma_0, \ldots, \sigma_{k-1}\}$ is prefix-free with the requirement that $\{C_0, \ldots, C_{k-1}\}$ is pairwise-disjoint.

Remark 3.9. For the real-valued computable metric spaces L^0, L^1, L^2, a number of other test functions have been used in the literature. The resulting computable metric spaces are equivalent.

(1) On $(\mathbb{X},\mu)$: functions as in Definition 3.8 (for any choice of *Basis*$(\mathbb{X},\mu)$).
(2) On $(\mathbb{X},\mu)$: any computable family $\mathcal{A} = \{\varphi_n\}_{n\in\mathbb{N}}$ of bounded computable functions, such that $\varphi_n \colon \mathbb{X} \to \mathbb{R}$ is computable uniformly in n, there is a bound C_n uniformly computable in n such that $\|\varphi_n(x)\|_\infty \le C_n$ for all n, and $\{\varphi_n\}$ is dense in $L^1(\mathbb{X},\mu)$.
(3) On $(\mathbb{X},\mu)$: the set $\mathcal{E}$ of bounded computable Lipschitz functions in [19, Section 2].
(4) On effectively compact $(\mathbb{X},\mu)$: any computable family $\mathcal{A} = \{\varphi_n\}_{n\in\mathbb{N}}$ of computable functions which is dense in $C(\mathbb{X})$.
(5) On effectively compact $(\mathbb{X},\mu)$: the "polynomials" in [56, 58] closed under pointwise multiplication and addition. (This family is dense in $C(\mathbb{X})$ by the Stone-Weierstrass theorem).
(6) On $(2^{\mathbb{N}},\lambda)$: the test functions in equation (3.1).
(7) On $([0,1]^d,\lambda)$: polynomials with rational coefficients.
(8) On $([0,1]^d,\lambda)$, $(\mathbb{T}^d,\lambda)$: dyadic functions of the form

$$\varphi = \sum_{i=0}^{k-1} c_i \mathbf{1}_{D_i} \quad (c_i \in \mathbb{Q}, D_i \text{ is a dyadic set})$$

where the DYADIC SETS are those of the form

$$\left[\frac{i_1}{2^n}, \frac{i_1+1}{2^n}\right) \times \cdots \times \left[\frac{i_d}{2^n}, \frac{i_d+1}{2^n}\right).$$

3.2. **Effective modes of convergence.** In measure theoretic probability, there are various modes of convergence for measurable functions. I have already mentioned convergence in the L^1-, L^2-norms and the metric d_{meas}.

Definition 3.10. Let (f_i) be a sequence of measurable $\mathbb{Y}$-valued functions and f a measurable $\mathbb{Y}$-valued function.

(1) The sequence f_i converges to f ALMOST UNIFORMLY if there is a RATE OF ALMOST-UNIFORM CONVERGENCE $n(\varepsilon_1, \varepsilon_2)$ such that for all $\varepsilon_1, \varepsilon_2 > 0$,

$$\mu\left(\left\{x \,\middle|\, \sup_{i \geq n(\varepsilon_1,\varepsilon_2)} d_{\mathbb{Y}}(f_i(x), f(x)) > \varepsilon_1\right\}\right) \leq \varepsilon_2.$$

(2) The sequence f_i converges to f IN MEASURE if there is a RATE OF CONVERGENCE IN MEASURE $n(\varepsilon_1, \varepsilon_2)$ such that for all $\varepsilon_1, \varepsilon_2 > 0$,

$$\forall i \geq n(\varepsilon_1, \varepsilon_2)\ \mu\left(\{x \mid d_{\mathbb{Y}}(f_i(x), f(x)) > \varepsilon_1\}\right) \leq \varepsilon_2.$$

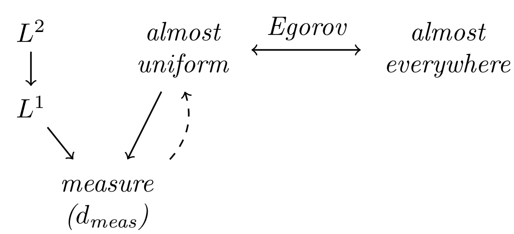
These definition can be extended to continuously-indexed sequences (i.e. functions $(f_r)_{r>0}$ with $r \to 0$) in the usual manner.

Fact 3.11. *Convergence in measure is the same as convergence in the metric d_{meas}.*

Proof. I give an effective version in Proposition 3.15. (For a similar proof with the Ky-Fan metric, see [12].) □

Fact 3.12 (Egorov's theorem, see [12])**.** *On a probability space, almost uniform convergence and almost everywhere convergence are the same (assuming (f_n) is a discretely-indexed sequence of measurable functions taking values in a complete separable metric space).*

Fact 3.13 (Modes of convergence, see [12])**.** *On a probability space, the following implications (and their transitive closures) hold between the modes of convergence. (Note, L^2 and L^1 only apply to real-valued functions. The dotted arrow represents convergence on some subsequence.)*

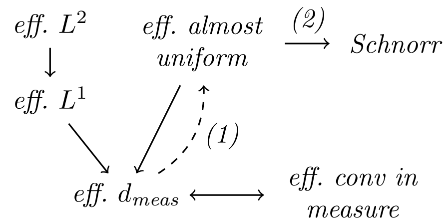


The goal of this section is to give the effective analog of the above chart.

Definition 3.14. Let (f_i) and f be uniformly effectively measurable. Then $f_i \to f$ EFFECTIVELY ALMOST UNIFORMLY, EFFECTIVELY IN MEASURE, or EFFECTIVELY IN d_{meas} if the respective rate of convergence is computable.

Further if real-valued (f_i) and f are uniformly L^1-computable (resp. L^2-computable), then $f_i \to f$ EFFECTIVELY IN THE L^1-NORM (resp. EFFECTIVELY IN THE L^2-NORM) if the corresponding rate of convergence is computable.

Proposition 3.15 (Modes of effective convergence)**.** *On a computable probability space $(\mathbb{X}, \mu)$, the following implications are effective—in that a rate of convergence for the latter is computable from the former. (L^1 and L^2 only apply to real-valued functions.)*



(1) *The dotted arrow represents that if $f_i \to f$ with a geometric rate of convergence in the metric d_{meas}, e.g. $\forall j \geq i\ d_{meas}(f_j, f) \leq 2^{-i}$, then $f_i \to f$ effectively almost uniformly.*
(2) *For the arrow going to "Schnorr", see Lemma 3.19 below.*

Proof. See Appendix A.3. □

Rather than use the term "effectively almost uniformly", we will use the more common term EFFECTIVELY ALMOST EVERYWHERE (or EFFECTIVELY A.E.). This is justified by Egorov's theorem (Fact 3.12).

The following limit properties are also useful.

Proposition 3.16. *Let (f_n) and f be uniformly effectively measurable real-valued functions.*

(1) *If $f_n \to f$ effectively a.e.. and $g_n \to g$ effectively a.e.., then $f_n + g_n \to f + g$ effectively a.e..*
(2) *If $f^j_n \to f^j$ effectively a.e.. ($j \in \{0, \ldots, k-1\}$), and g is computable with a uniform modulus of continuity, then $g(f^0_n, \ldots, f^{k-1}_n) \to g(f^0, \ldots, f^{k-1})$ effectively a.e..*
(3) *(Squeeze theorem) Assume $f_n \leq g_n \leq h_n$ a.e. and that $f_n \to g$ effectively a.e.. and $h_n \to g$ effectively a.e.., then $g_n \to g$ effectively a.e.*

Further, in all cases the rates of convergence for the latter are computable from the former (in (2) use the modulus of continuity for g). Indeed, we do not need to assume the functions are effectively measurable, just that the rates of convergence are computable. The same results hold for continuous convergence, e.g. $f_r \to f$ as $r \to 0$.

3.3. **Convergence on Schnorr randoms.** Now we define representatives for each (equivalence class of an) effectively measurable function. The proofs are in Appendix A.4.

Recall that Cauchy-names are computable sequences of test functions with a geometric rate of convergence.

Definition 3.17. Let $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ be effectively measurable with Cauchy-name (φ_n) in the metric d_{meas}. Define

$$\widetilde{f}(x) = \begin{cases} \lim_{n\to\infty} \varphi_n(x) & \text{if the limit exists} \\ \textit{undefined} & \text{otherwise} \end{cases}.$$

If A is an effectively measurable set (and therefore $\mathbf{1}_A\colon (\mathbb{X}, \mu) \to \{0, 1\}$ is effectively measurable), then define $\widetilde{A}$ as

$$x \in \widetilde{A} \quad \Leftrightarrow \quad \widetilde{\mathbf{1}_A}(x) = 1.$$

These definitions are justified as follows. Similar versions of this proposition are in Pathak [38] (L^1-computable functions and Martin-Löf randomness) and Pathak, Rojas, and Simpson [39] (L^1-computable functions and Schnorr randomness).

Proposition 3.18. *Suppose $f\colon (\mathbb{X},\mu) \to \mathbb{Y}$ is effectively measurable with Cauchy-name (φ_n) (in the metric d_{meas}, L^1-norm, or L^2-norm).*

1. *(Existence) The limit $\lim_{n\to\infty} \varphi_n(x)$ exists on all Schnorr randoms x.*
2. *(Uniqueness) Given another Cauchy-name (ψ_n) for f,*

$$\lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} \psi_n(x) \quad (\text{on Schnorr random } x).$$

In Theorem 12.19, I show that Schnorr randomness is the best possible for the previous theorem.

This next lemma is quite useful and for much of the paper is the only fact about Schnorr randomness needed.

Lemma 3.19 (Convergence Lemma)**.** *Suppose that (f_k) and f are uniformly effectively measurable. If*

$$f_k \to f \quad (\text{effectively a.e.})$$

then

$$\widetilde{f}_k(x) \to \widetilde{f}(x) \quad (\text{for all Schnorr random } x).$$

3.4. **Properties of effectively measurable functions.** The proofs are in Appendix A.5.

Proposition 3.20. *The following implications hold for real-valued functions (and all the computations are uniform).*

1. *$f \in L^2_{comp} \Rightarrow f \in L^1_{comp} \Rightarrow f \in L^0_{comp}$. (The converses do not hold in general.)*
2. *If $0 \le f \le 1$, then $f \in L^2_{comp} \Leftrightarrow f \in L^1_{comp} \Leftrightarrow f \in L^0_{comp}$.*
3. *$f \in L^1_{comp} \Leftrightarrow (f \in L^0_{comp}$ and $\|f\|_{L^1}$ is computable).*
4. *$f \in L^2_{comp} \Leftrightarrow (f \in L^0_{comp}$ and $\|f\|_{L^2}$ is computable).*
5. *If $f \in L^1_{comp}$ then $\int f\,d\mu$ is computable.*
6. *If B is effectively measurable, then $\mu(B)$ is computable.*
7. *If $0 \le g \le 1$, $g \in L^1_{comp}$, and $f \in L^1_{comp}$, then $g \cdot f \in L^1_{comp}$.*

Proposition 3.21 (Effective Lusin's theorem)**.** *Given an effectively measurable $f\colon (\mathbb{X},\mu) \to \mathbb{Y}$, and some rational $\varepsilon \ge 0$, there are an effectively closed set K of computable measure $\mu(K) \ge 1-\varepsilon$ and a computable function $g\colon K \to \mathbb{Y}$ such that $g = \widetilde{f} \upharpoonright K$ on Schnorr randoms. (Further, g and K are computable uniformly from ε and any name for f.) Moreover, if $\mathbb{Y} = \mathbb{R}$, then $g\colon K \to \mathbb{Y}$ can be extended (uniformly from its name) to a total computable function $g\colon \mathbb{X} \to \mathbb{Y}$ such that $g = \widetilde{f} \upharpoonright K$ on Schnorr randoms.*

Proposition 3.22 (Effective inner/outer regularity)**.** *Given $A \subseteq (\mathbb{X},\mu)$ effectively measurable, and some rational $\varepsilon > 0$, there are an effectively open set U and an effectively closed set C both of computable measure such that $C \subseteq \widetilde{A} \subseteq U$ for Schnorr randoms, and such that $\mu(U) - \mu(C) \le \varepsilon$. (The sets U, C and their measures $\mu(U), \mu(C)$ are uniformly computable from ε and any name for A.)*

Remark 3.23. The effectively closed sets K and C in the last two propositions can be made to be EFFECTIVELY COMPACT IN THE STRONGER SENSE (that is $K = f(2^{\mathbb{N}})$ for some total computable map $f\colon 2^{\mathbb{N}} \to \mathbb{X}$). This is not needed in this paper and is left as an exercise for the reader.

This next result is the converse to the effective Lusin's theorem and shows that the representative functions of this paper are the same as the Schnorr layerwise computable functions of Miyabe [33], which are an extension of the layerwise computable functions of Hoyrup and Rojas [23]. Miyabe [33], proved the corresponding result for L^1-computable functions.

Proposition 3.24 (Schnorr layerwise computability)**.** *Consider a (pointwise-defined) measurable function $f\colon \mathbb{X} \to \mathbb{Y}$ that is* SCHNORR LAYERWISE COMPUTABLE, *that is, there is a computable sequence (C_n) of effectively closed sets of computable measure $\mu(C_n) \le 2^{-n}$, such that $f \upharpoonright C_n$ is computable on C_n uniformly in n. Then there is an effectively measurable $g\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ such that $\widetilde{g} = f$ on Schnorr randoms.*

In this next proposition, an ALMOST-EVERYWHERE COMPUTABLE FUNCTION $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ is a partial computable function whose domain is measure one. (Here I mean "domain" to mean the points x for which the underlying computation computes a name for $f(x)$ from a name for x. To avoid ambiguity, I could alternately define an almost-everywhere computable function as a function $f\colon A \subseteq \mathbb{X} \to \mathbb{Y}$ which is computable on a measure-one Π^0_2 set A. See [41] for more discussion.)

Proposition 3.25 (Examples of effectively measurable functions and sets)**.** *All of these functions $f\colon \mathbb{X} \to \mathbb{Y}$ and sets $A \subseteq \mathbb{X}$ are effectively measurable, and $\widetilde{f} = f$ and $\widetilde{A} = A$ on Schnorr randoms.*

1. *Test functions and test sets as in Propositions 3.1 and 3.3 and in Definition 3.8.*
2. *Computable functions and decidable sets (i.e., computable 0,1-valued functions).*
3. *Almost-everywhere computable functions $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ and almost-everywhere decidable sets (i.e., almost everywhere computable 0,1-valued functions).*
4. *Nonnegative lower semicomputable functions $f\colon \mathbb{X} \to \mathbb{R}$ with a computable integral, effectively open sets $U \subseteq \mathbb{X}$ of computable measure, and effectively closed sets $C \subseteq \mathbb{X}$ of computable measure.*

Recall that for a measurable function $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$, the PUSH-FORWARD MEASURE of μ along f (denoted $\mu_* f$) is the measure on $\mathbb{Y}$ defined by $\int \varphi \, d\mu_* f = \int \varphi \circ f \, d\mu$ for bounded computable φ.

Proposition 3.26 (Push-forward measures)**.** *If $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ is effectively measurable, then the push-forward measure $(\mathbb{Y}, \mu_* f)$ is a computable probability space (uniformly from $(\mathbb{X}, \mu)$, $\mathbb{Y}$, and f).*

Proposition 3.27 (Preservation of Schnorr randomness)**.** *If $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ is effectively measurable and x is Schnorr random, then $\widetilde{f}(x)$ is Schnorr random on $(\mathbb{Y}, \mu_* f)$.*

Proposition 3.28 (Composition and tuples)**.**

1. *(Composition) Given $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ and $g\colon (\mathbb{Y}, \mu_* f) \to \mathbb{Z}$ effectively measurable, the composition $g \circ f$ is effectively measurable (uniformly from f*

and g) and

$$\widetilde{f \circ g} = \widetilde{f} \circ \widetilde{g} \quad \text{(on Schnorr randoms).}$$

(2) *(Tuples) Given $f_n\colon (\mathbb{X},\mu) \to \mathbb{Y}_n$ effectively measurable (uniformly in n), the tuples*

$$(f_0,\dots,f_{k-1})\colon (\mathbb{X},\mu) \to \mathbb{Y}_0 \times \cdots \times \mathbb{Y}_{k-1}$$

and

$$(f_n)_{n\in\mathbb{N}}\colon (\mathbb{X},\mu) \to \prod_{n\in\mathbb{N}} \mathbb{Y}_n$$

are effectively measurable (uniformly from (f_n)) and

$$\widetilde{(f_0,\dots,f_{k-1})} = (\widetilde{f}_0,\dots,\widetilde{f}_{k-1}) \quad \text{and} \quad \widetilde{(f_i)_{i\in\mathbb{N}}} = (\widetilde{f}_i)_{i\in\mathbb{N}} \quad \text{(on Schnorr randoms).}$$

These two combinations, along with the results about computable functions in Proposition 3.25, can be used to prove a number of useful facts.

Proposition 3.29 (Combinations of measurable functions)**.**

(1) *(Computable pointwise operations). All computable pointwise operations, including vector, lattice, and Boolean algebra operations preserve effective measurability. Moreover, given $f,g\colon (\mathbb{X},\mu) \to \mathbb{R}$ and $A,B \subseteq (\mathbb{X},\mu)$ effectively measurable, we have*

$$\widetilde{f+g} = \widetilde{f} + \widetilde{g}, \quad \widetilde{af} = a\widetilde{f}, \quad \widetilde{f\cdot g} = \widetilde{f}\cdot\widetilde{g}$$

$$\widetilde{\min(f,g)} = \min(\widetilde{f},\widetilde{g}), \quad \widetilde{\max(f,g)} = \max(\widetilde{f},\widetilde{g}), \quad \widetilde{|f|} = \left|\widetilde{f}\right|$$

$$\widetilde{A\cup B} = \widetilde{A}\cup\widetilde{B}, \quad \widetilde{A\cap B} = \widetilde{A}\cap\widetilde{B}, \quad \widetilde{A^c} = \widetilde{A}^c, \quad \widetilde{\mathbb{X}} = \mathbb{X}, \quad \widetilde{\varnothing} = \varnothing$$

on Schnorr randoms, and

$$f \le g \text{ a.e.} \quad \Leftrightarrow \quad \widetilde{f} \le \widetilde{g} \text{ (on Schnorr randoms)}$$

$$A \subseteq B \text{ a.e.} \quad \Leftrightarrow \quad \widetilde{A} \subseteq \widetilde{B} \text{ (on Schnorr randoms).}$$

(2) *(Inverse image) Given $f\colon (\mathbb{X},\mu) \to \mathbb{Y}$ and $B \subseteq (\mathbb{Y},\mu_* f)$ effectively measurable then $f^{-1}(B)$ is effectively measurable and $\widetilde{f^{-1}(B)} = \widetilde{f}^{-1}(\widetilde{B})$ on Schnorr randoms.*

(3) *(Rotations) Given $f\colon (\mathbb{T}^d,\lambda) \to \mathbb{R}$ effectively measurable, and a computable vector $t \in \mathbb{T}^d$, then $h(x) := f(x-t)$ is effectively measurable and $\widetilde{h}(x) = \widetilde{f}(x-t)$ on Schnorr randoms.*

(4) *(Indicator functions) Given $A \subseteq (\mathbb{X},\mu)$, A is effectively measurable if and only if $\mathbf{1}_A\colon (\mathbb{X},\mu) \to \mathbb{R}$ is effectively measurable (equivalently, L^1-computable by Proposition 3.20 (2)) and $x \in \widetilde{A}$ if and only if $\widetilde{\mathbf{1}_A}(x) = 1$ on Schnorr randoms. (Notice the codomain of $\mathbf{1}_A$ is $\mathbb{R}$ here rather than $\{0,1\}$ as in Definition 3.17.)*

Proposition 3.30. *The following implications hold for real-valued functions (and all the computations are uniform).*

(1) *If $f \in L^1_{comp}$ and A is effectively measurable, then $\int_A f\,d\mu$ is computable.*

(2) *If $\mathbb{X}$ is effectively compact (see [35])—as is $[0,1]^d$, $\mathbb{T}^d$, and $2^{\mathbb{N}}$—and $g\colon \mathbb{X} \to \mathbb{R}$ is computable, then g is L^1-computable (since it has computable bounds).*

(3) *If* $f\colon (\mathbb{X},\mu) \to \mathbb{Y}$ *is effectively measurable and* $g \in L^1_{comp}(\mathbb{Y},\mu_* f)$ *(resp.* $L^2_{comp}(\mathbb{Y},\mu_* f)$*), then* $g \circ f \in L^1_{comp}(\mathbb{X},\mu)$ *(resp.* $L^2_{comp}(\mathbb{X},\mu)$*).*

Proposition 3.31. *Given a measurable map* $f\colon (\mathbb{X},\mu) \to \mathbb{Y}$*, the following are equivalent.*

(1) f *is effectively measurable.*
(2) *The push-forward measure* $(\mathbb{Y},\mu_* f)$ *is computable and one (or all) of the following "pull-back" maps are computable:*
  (a) *(*L^1 *functions)* $g \in L^1(\mathbb{Y},\mu_* f) \;\mapsto\; g \circ f \in L^1(\mathbb{Y},\mu_* f)$*.*
  (b) *(*L^2 *functions)* $g \in L^2(\mathbb{Y},\mu_* f) \;\mapsto\; g \circ f \in L^2(\mathbb{Y},\mu_* f)$*.*
  (c) *(Measurable sets)* $B \subseteq (\mathbb{Y},\mu_* f) \;\mapsto\; f^{-1}(B) \subseteq (\mathbb{X},\mu)$*.*

4. Differentiability

In this section I present effective versions of the Lebesgue differentiation theorem and its corollaries.

4.1. **The dyadic Lebesgue differentiation theorem.**

Before considering the full Lebesgue differentiation theorem, let us consider the simpler dyadic version on the fair-coin measure $(2^{\mathbb{N}},\lambda)$. This will contain most of the work for the version on $[0,1]$.

Fact 4.1 (Dyadic Lebesgue differentiation theorem)**.** *Given* $f \in L^1(2^{\mathbb{N}},\lambda)$*,*

$$\frac{\int_{[x\upharpoonright k]} |f - f(x)|\, d\lambda}{\lambda([x \upharpoonright k])} \xrightarrow[k\to\infty]{} 0 \quad (\lambda\text{-}a.e.\ x \in 2^{\mathbb{N}}).$$

In particular

$$\frac{\int_{[x\upharpoonright k]} f\, d\lambda}{\lambda([x \upharpoonright k])} \xrightarrow[k\to\infty]{} f(x) \quad (\lambda\text{-}a.e.\ x \in 2^{\mathbb{N}}).$$

As a helpful notation, I will write

$$f^{(k)}(x) := \frac{\int_{[x\upharpoonright k]} f\, d\lambda}{\lambda([x \upharpoonright k])}.$$

Notice $f^{(k)} \in L^1(2^{\mathbb{N}},\lambda)$ and that $f^{(k)}$ is constant on each cylinder set $[\sigma]$ where $\sigma \in 2^{k'}$ $(k' \geq k)$. Further, we can use $f^{(k)}$ to approximate f in the L^1-norm as follows.

Fact 4.2 (Lebesgue approximation theorem)**.** *Given* $f \in L^1(2^{\mathbb{N}},\lambda)$*,*

$$f^{(k)} \xrightarrow[k\to\infty]{L^1} f.$$

As we will see, Facts 4.1 and 4.2 are both instances of the more general Lévy 0-1 law (Fact 6.2).

Proposition 4.3 (Effective Lebesgue approximation theorem)**.** *Suppose we are given* $f \in L^1_{comp}(2^{\mathbb{N}},\lambda)$*. Then*

$$f^{(k)} \xrightarrow[k\to\infty]{L^1} f \quad (\textit{effectively}).$$

Proof. We compute the rate of convergence $k(\varepsilon)$. Pick a rational $\varepsilon > 0$. Let φ be a simple function approximating f such that $\|f-\varphi\|_{L^1} \leq \varepsilon/2$. By the definition of simple function, there is some k' such that φ is constant on all cylinder sets $[\sigma]$ where $\sigma \in 2^k$ $(k \geq k')$. In particular, $\varphi^{(k)} = \varphi$ $(k \geq k')$. Let $k(\varepsilon) = k'$. Then for $k \geq k(\varepsilon)$,

$$\begin{aligned}
\left\| f - f^{(k)} \right\|_{L^1} &\leq \|f-\varphi\|_{L^1} + \left\| \varphi^{(k)} - f^{(k)} \right\|_{L^1} \\
&= \|f-\varphi\|_{L^1} + \sum_{\sigma \in 2^k} \left| \frac{\int_{[\sigma]} \varphi - f \, d\lambda}{\lambda([\sigma])} \right| \\
&\leq \|f-\varphi\|_{L^1} + \sum_{\sigma \in 2^k} \frac{\int_{[\sigma]} |\varphi - f| \, d\lambda}{\lambda([\sigma])} \\
&= 2\,\|f-\varphi\|_{L^1} \leq \varepsilon.
\end{aligned}$$

□

Recall the following dyadic version of Kolmogorov's inequality.

Fact 4.4 (Dyadic Kolmogorov's inequality, see [11])**.** *Let $M\colon 2^{<\omega} \to [0,\infty)$ be a nonnegative* DYADIC MARTINGALE *on the $(2^{\mathbb{N}}, \lambda)$, that is $\frac{1}{2}M(\sigma 0) + \frac{1}{2}M(\sigma 1) = M(\sigma)$ for all $\sigma \in 2^{<\omega}$. Then for all $\varepsilon > 0$*

$$\lambda\left(\left\{ x \in 2^{\mathbb{N}} \,\middle|\, \sup_{k \geq 0} M(x \upharpoonright k) \geq \varepsilon \right\}\right) \leq \frac{M(\varnothing_{string})}{\varepsilon}.$$

As a special case we have the following.

Lemma 4.5. *Given nonnegative $f \in L^1(2^{\mathbb{N}}, \lambda)$,*

$$\lambda\left(\left\{ x \in 2^{\mathbb{N}} \,\middle|\, \sup_{k \geq 0} f^{(k)}(x) \geq \varepsilon \right\}\right) \leq \frac{\|f\|_{L^1}}{\varepsilon}.$$

Proof. Let $M(\sigma) = \int_{[\sigma]} f \, d\lambda / \lambda([\sigma])$. This is a nonnegative dyadic martingale since f is nonnegative. Apply Kolmogorov's inequality noting that $f^{(k)}(x) = M(x \upharpoonright k)$ and $\|f\|_{L^1} = \int f \, d\lambda = M(\varnothing_{string})$.
Now we have the effective version of Proposition 4.1. □

Proposition 4.6 (Effective dyadic Lebesgue differentiation theorem)**.** *Given $f \in L^1_{comp}(2^{\mathbb{N}}, \lambda)$, let*

$$g_k(x) := |f - f(x)|^{(k)}(x) = \frac{\int_{[x \upharpoonright k]} |f(y) - f(x)| \, d\lambda(y)}{\lambda([x \upharpoonright k])}.$$

Then $g_k \to 0$ a.e. as $k \to \infty$ with an effective rate $k(\delta, \varepsilon)$ of a.e. convergence. Hence $f^{(k)} \to f$ effectively a.e. as $k \to \infty$.

Further,

$$\frac{\int_{[x \upharpoonright k]} |f(y) - \widetilde{f}(x)| \, d\lambda(y)}{\lambda([x \upharpoonright k])} \xrightarrow[k \to \infty]{} 0 \quad (\textit{on Schnorr random } x).$$

Hence, $f^{(k)}(x) \to \widetilde{f}(x)$ on Schnorr randoms x as $k \to \infty$.

Proof. Pick $\delta > 0$ and $\varepsilon > 0$. By Proposition 4.3, from f we can effectively find some $k' \in \mathbb{N}$ such that $\left\| f - f^{(k')} \right\|_{L^1} \leq \frac{\delta\varepsilon}{4}$. Let $k(\delta, \varepsilon) = k'$. Then for any $k \geq k'$ and all $x \in 2^{\mathbb{N}}$ we have

$$\begin{aligned}
0 \leq g_k(x) &= \frac{\int_{[x\restriction k]} |f(y) - f(x)| \, d\lambda(y)}{\lambda([x \restriction k])} \\
&\leq \frac{\int_{[x\restriction k]} \left| f(y) - f^{(k')}(y) \right| \, d\lambda(y)}{\lambda([x \restriction k])} + \frac{\int_{[x\restriction k]} \left| f^{(k')}(y) - f^{(k')}(x) \right| \, d\lambda(y)}{\lambda([x \restriction k])} \\
&\qquad + \frac{\int_{[x\restriction k]} \left| f^{(k')}(x) - f(x) \right| \, d\lambda(y)}{\lambda([x \restriction k])} \qquad (4.1) \\
&= \left| f - f^{(k')} \right|^{(k)}(x) + 0 + \left| f^{(k')}(x) - f(x) \right|.
\end{aligned}$$

To bound the last line, use Lemma 4.5 for the first term,

$$\lambda\left(\left\{ x \,\middle|\, \sup_{k \geq k'} \left| f - f^{(k')} \right|^{(k)}(x) \geq \frac{\varepsilon}{2} \right\} \right) \leq \frac{2 \left\| f - f^{(k')} \right\|_{L^1}}{\varepsilon},$$

and use Markov's inequality (Fact A.2) for the last term,

$$\lambda\left(\left\{ x \,\middle|\, \sup_{k \geq k'} |f^{(k')}(x) - f(x)| \geq \frac{\varepsilon}{2} \right\} \right) \leq \frac{2 \left\| f - f^{(k')} \right\|_{L^1}}{\varepsilon}.$$

Putting them together (see Fact A.1), we have

$$\lambda\left(\left\{ x \,\middle|\, \sup_{k \geq k'} g_k(x) \geq \varepsilon \right\} \right) \leq 4 \frac{\left\| f - f^{(k')} \right\|_{L^1}}{\varepsilon} \leq \delta.$$

Since $g_k \geq 0$, this shows that $g_k(x) \to 0$ effectively a.e. Moreover,

$$\left| f^{(k)}(x) - f(x) \right| = \left| (f - f(x))^{(k)}(x) \right| \leq |f - f(x)|^{(k)}(x) = g(x) \to 0.$$

Hence, $f^{(k)} \to f$ effectively a.e.

Now let us show convergence on Schnorr randoms. Since $g_k \to 0$ and $f^{(k)}(x) \to f(x)$, both effectively a.e., we have, by Lemma 3.19, that $\widetilde{g}_k(x) \to 0$ and $\widetilde{f^{(k)}}(x) \to \widetilde{f}(x)$ on Schnorr randoms x. Notice that $g_k(x) = h_k \circ (f(x), x)$ where $h_k(a, b) = |f - a|^{(k)}(b)$. Further, both h_k and $f^{(k)}$ are computable functions uniformly in k (since f is L^1-computable). By the results in Section 3.4 we have on Schnorr randoms x that as $k \to \infty$,

$$\left| f - \widetilde{f}(x) \right|^{(k)}(x) = h_k\left(\widetilde{f}(x), x \right) = \widetilde{g}_k(x) \to 0.$$

and

$$f^{(k)}(x) = \widetilde{f^{(k)}}(x) \to \widetilde{f}(x).$$

(In both equations, the first instance of f acts as equivalence class and does not require the tilde.) □

4.2. **The Lebesgue differentiation theorem.** Now I wish to prove an effective version of the Lebesgue differentiation theorem. To simplify the geometry I will use the unit torus $\mathbb{T}^d$ (identified with $[0,1)^d$) and the Lebesgue measure λ. The argument for $[0,1]^d$ is similar. First, recall the Lebesgue differentiation theorem. Here $A_r f(x)$ is the average of f over the ball $B(x,r)$,

$$A_r f(x) = \frac{\int_{B(x,r)} f(y)\,dy}{\lambda(B(x,r))}.$$

Fact 4.7 (Lebesgue differentiation theorem, see [48]). *Given an integrable function f on $(\mathbb{T}^d, \lambda)$,*

$$A_r|f - f(x)|\,(x) = \frac{\int_{B(x,r)} |f(y) - f(x)|\,dy}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 0 \quad (\lambda\text{-a.e. } x \in \mathbb{T}^d). \tag{4.2}$$

In particular,

$$A_r f(x) \xrightarrow[r\to 0]{} f(x) \quad (\lambda\text{-a.e. } x \in \mathbb{T}^d).$$

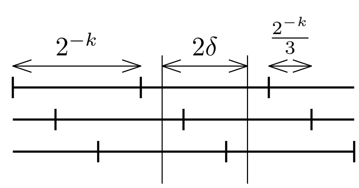
The points x for which the limit (4.2) holds are the LEBESGUE POINTS of f.

If, instead of averaging over balls, we averaged over dyadic sets, the Lebesgue differentiation theorem would be the dyadic Lebesgue differentiation theorem of Fact 4.1. However, the full Lebesgue differentiation theorem is a geometric theorem. The theorem concerns the *simultaneous* convergence of overlapping balls (or cubes). Moreover, if the balls or cubes were replaced by, say, ellipses or rectangles of arbitrary aspect ratio, the theorem would not hold. The main idea behind any proof of the Lebesgue differentiation theorem is to restrict one's attention to a disjoint set of cubes (or balls). The classical proof does this through Vitali's covering lemma (see [48]). Here I use an alternate method of Morayne and Solecki [34], which uses martingale theory and a useful geometric lemma.

If $t = (t_1, \dots, t_d) \in \mathbb{T}^d$ and $Q \subseteq \mathbb{T}^d$, define $t + Q = \{t + x \mid x \in Q\}$, i.e. Q rotated in each ith coordinate by t_i. Let $\mathcal{B}_k$ denote the set of dyadic cubes of measure $\left(2^{-k}\right)^d$. Define $\mathcal{B}_k^t = \{t + Q \mid Q \in \mathcal{B}_k\}$, i.e. translate the dyadic cubes by the vector $t \in \mathbb{T}^d$. Let $I_k^t(x)$ be the unique element of $\mathcal{B}_k^t$ that contains x. The next fact and lemma show that it is enough to consider convergence along dyadic cubes and finitely many shifts.

Fact 4.8 (Morayne and Solecki [34, Lemma 2]). *Let $x \in \mathbb{T}^d$. Consider a cube $Q = x + (-\delta, \delta)^d$ such that $0 < \delta < 2^{-k}/3$. Then $Q \subseteq \bigcup_{t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d} I_k^t(x)$.*

Proof sketch. The main idea is that any interval of length 2δ where $\delta < 2^{-k}/3$ must either be contained in a dyadic interval of length 2^{-k}, or in a dyadic interval shifted by $2^{-k}/3$ in either direction as this picture shows.



Then notice a dyadic interval of length 2^{-k} shifted by $2^{-k}/3$ is also a (different) dyadic interval of length 2^{-k} shifted by $1/3$. □

Lemma 4.9. *Let $x \in \mathbb{T}^d$ and $f \in L^1(\mathbb{T}^d, \lambda)$ (such that f is pointwise defined at x). Then the following are equivalent.*

(1) $A_r|f - f(x)|(x) \xrightarrow[r\to 0]{} 0$ *(i.e., x is a Lebesgue point of f).*
(2) $\frac{1}{\lambda(Q_\delta(x))}\int_{Q_\delta(x)} |f(y) - f(x)|\,dy \xrightarrow[\delta\to 0]{} 0$ *for $Q_\delta(x) = x + (-\delta, \delta)^d$.*
(3) $\frac{1}{\lambda(Q_i)}\int_{Q_i} |f(y) - f(x)|\,dy \xrightarrow[i\to\infty]{} 0$ *for any sequence of cubes $Q_0 \supseteq Q_1 \supseteq \dots$ where $\bigcap_i Q_i = \{x\}$ (the sequence need not be computable).*
(4) $\frac{1}{\lambda(I_k^t(x))}\int_{I_k^t(x)} |f(y) - f(x)|\,dy \xrightarrow[k\to\infty]{} 0$ *for all $t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d$.*

(1) *through* (4) *also hold when $\mathbb{T}^d$ is replaced by $[0,1]^d$.*

Proof. We will show (4) implies (2) implies (1). The other equivalences are standard results that follow similarly. Their proofs are left to the reader.

For (2) implies (1), pick $r > 0$ and let $\delta = r$. Then $\lambda(B(x,r)) = \lambda(Q_\delta)/C$ for some constant C depending only on the dimension d, and

$$A_r|f - f(x)|(x) = \frac{\int_{B(x,r)} |f(y) - f(x)|\,dy}{\lambda(B(x,r))} \leq C \cdot \frac{\int_{Q_\delta} |f(y) - f(x)|\,dy}{\lambda(Q_\delta)}. \tag{4.3}$$

For (2) implies (1), pick $\delta > 0$ and let k be such that $2^{-k}/3 > \delta \geq 2^{-k-1}/3$. Then $\lambda(Q_\delta(x)) \geq \lambda(I_k^t(x))/3^d$ for all $t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d$. Therefore, by Lemma 4.8,

$$\frac{\int_{Q_\delta(x)} |f(y) - f(x)|\,dy}{\lambda(Q_\delta(x))} \leq \frac{\sum_{t\in\{-\frac{1}{3},0,\frac{1}{3}\}^d} \int_{I_k^t} |f(y) - f(x)|\,dy}{\lambda(Q_\delta(x))} \tag{4.4}$$
$$\leq \sum_{t\in\{-\frac{1}{3},0,\frac{1}{3}\}^d} \frac{\int_{I_k^t(x)} |f(y) - f(x)|\,dy}{\lambda\left(I_k^t(x)\right)/3^d} = 3^d \cdot \sum_{t\in\{-\frac{1}{3},0,\frac{1}{3}\}^d} \frac{\int_{I_k^t(x)} |f(y) - f(x)|\,dy}{\lambda\left(I_k^t(x)\right)}. \quad \square$$

Theorem 4.10 (Effective Lebesgue differentiation theorem)**.** *Given $f \in L^1_{comp}(\mathbb{T}^d, \lambda)$,*

$$A_r|f - f(x)|(x) = \frac{\int_{B(x,r)} |f(y) - f(x)|\,dy}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 0 \quad (\lambda\text{-a.e. } x \in \mathbb{T}^d)$$

with an effective rate of a.e. convergence $r(\delta, \varepsilon)$. Hence $A_r f \xrightarrow[r\to 0]{} f$ effectively a.e.

Further,

$$A_r|f - \widetilde{f}(x)|(x) \xrightarrow[r\to 0]{} 0 \quad (\text{on Schnorr random } x).$$

Hence, all Schnorr randoms are Lebesgue points of $\widetilde{f}$ and $A_r f(x) \xrightarrow[r\to 0]{} \widetilde{f}(x)$ on Schnorr randoms x. These statements also hold when $\mathbb{T}^d$ is replaced by $[0,1]^d$.

Proof. Combining inequalities (4.3) and (4.4) in the proof of Lemma 4.9 we have for $2^{-k}/3 > r \geq 2^{-(k+1)}/3$ that

$$0 \leq A_r|f - f(x)|(x) \leq C \cdot \sum_{t\in\{-\frac{1}{3},0,\frac{1}{3}\}^d} \frac{\int_{I_k^t(x)} |f(y) - f(x)|\,dy}{\lambda\left(I_k^t(x)\right)}$$

for some constant C depending only on the dimension of d. Using Proposition 4.6, with $f(y - t)$ in place of $f(y)$, we have that

$$\frac{\int_{I_k^t(x)} |f(y) - f(x)|\,dy}{\lambda\left(I_k^t(x)\right)} \xrightarrow[k\to\infty]{} 0$$

with an effective rate of a.e. convergence for each $t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d$. Hence, by the squeeze theorem (Proposition 3.16 (3))— $A_r|f - f(x)|(x) \xrightarrow[r\to 0]{} 0$ with an effective rate of a.e. convergence. The result for Schnorr randomness follows by a similar argument. (Note that if $h(y) := f(y-t)$ for a computable t, then $\widetilde{h}(y) = \widetilde{f}(y-t)$ by Proposition 3.29)

For $[0,1]^d$, just use the same argument (as for $\mathbb{T}^d$), but also adjust for the error near the boundary (which is straightforward, although somewhat tedious). □

Remark 4.11. Even though it was shown that $A_r f \xrightarrow[r\to 0]{} f$ effectively a.e., one can not directly apply Lemma 3.19 to show that $A_r f(x) \xrightarrow[r\to 0]{} \widetilde{f}(x)$ for all Schnorr randoms x (as in the proof of Proposition 4.6); Lemma 3.19 is only for discretely-indexed sequences. (A continuously-indexed version of Lemma 3.19 is possible, but it would need additional conditions on the sequence of functions (f_r).)

Remark 4.12. Setting aside computational concerns, this proof of the Lebesgue differentiation theorem is very similar to the standard proof. The key differences are that this proof uses Lemma 4.9 to handle the geometric concerns while the standard proof uses the Vitali covering lemma, and we use Kolmogorov's inequality to show convergence, while the standard proof uses the Hardy-Littlewood maximal lemma. The effective proof of Pathak et al. [39] indeed follows the usual proof. For another method to handle the geometry see Brattka et al. [7].

4.3. **Corollaries to the Lebesgue differentiation theorem.** From the effective Lebesgue differentiation theorem (Theorem 4.10), we have the following corollaries. Note that all of these have "dyadic" versions on $2^{\mathbb{N}}$ as well.

Let A be a measurable set on and $x \in [0,1]^d$. We say x is a POINT OF DENSITY of A if

$$\frac{\lambda(A \cap B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 1.$$

Then we have the following well-known corollary to the Lebesgue differentiation theorem.

Fact 4.13 (Lebesgue density theorem)**.** *Let A be a measurable set. Almost every $x \in A$ is a point of density.*

Corollary 4.14 (Effective Lebesgue density theorem)**.** *Let A be an effectively measurable set in $[0,1]^d$. Every Schnorr random in $\widetilde{A}$ is a point of density.*

Proof. Assume x is in $\widetilde{A}$ and is Schnorr random. By Definition 3.17, $\widetilde{\mathbf{1}}_A(x) = 1$. The rest follows from the Lebesgue differentiation theorem (Theorem 4.10) applied to $\mathbf{1}_A$. □

For the next application of the Lebesgue density theorem, if A, B are subsets of $\mathbb{R}$ then denote $A + B := \{x + y \mid x \in A, y \in B\}$, and similarly for $A - B$.

Fact 4.15 (Steinhaus, see [47])**.** *Let A and B be measurable subsets of $\mathbb{R}$ with positive Lebesgue measure and let x and y be points of density of A and B, respectively. Then $A + B$ contains an open neighborhood around $x + y$. Therefore, if A has positive Lebesgue measure, then $A - A$ contains an open neighborhood around 0.*

Corollary 4.16. *Let $A, B \subseteq [0,1]^d$ be effectively measurable with positive measure. If $x \in \widetilde{A}$ and $y \in \widetilde{B}$ are Schnorr randoms, then there is an open neighborhood in $\widetilde{A} + \widetilde{B}$ around $x + y$.*

Proof. By the effective Lebesgue density theorem (Corollary 4.16), x and y are points of density. Apply Steinhaus' theorem (Fact 4.15). □

A function $h\colon [0,1] \to \mathbb{R}$ is said to be ABSOLUTELY CONTINUOUS if it is of the form $F(x) = \int_0^x f(y)\,dy + F(0)$ for some integrable function f. It is clear that absolute continuity implies continuity. We have yet another corollary to the Lebesgue differentiation theorem.

Fact 4.17 (Lebesgue, see [48])**.** *An absolutely continuous function F is differentiable a.e. with derivative $\frac{d}{dx}F = f$ a.e.*

We say F is EFFECTIVELY ABSOLUTELY CONTINUOUS if the derivative f is L^1-computable. (This is equivalent to being a computable point in the Banach space $(AC[0,1], \|\cdot\|_{AC})$ where $\|F\|_{AC} = |f(0)| + \|f\|_{BV}$. See [17].) If F is effectively absolutely continuous, then it is computable (by the computability of integration). However, not every computable and absolutely continuous function is effectively absolutely continuous. This follows from this next corollary combined with the example of Brattka, Miller, and Nies [7] of a computable absolutely continuous function which is only differentiable on Martin-Löf randoms (which are a proper subset of the Schnorr randoms).

Corollary 4.18. *Assume $z \in [0,1]$ is Schnorr random and F is effectively absolutely continuous, hence $F(x) = \int_0^x f(y)\,dy + F(0)$ for all x for some L^1-computable f. Then F is differentiable at z with derivative $\frac{d}{dx}F|_{x=z} = \widetilde{f}(z)$.*

Proof. It suffices to show

$$\frac{F(z+t_i) - F(z)}{t_i} = \frac{\int_z^{z+t_i} f(y)\,dy}{t_i} \xrightarrow[i\to\infty]{} \widetilde{f}(z)$$

for any decreasing sequence $t_i \to 0^+$ (and the same for any increasing sequence $t_i \to 0^-$). Letting $Q_i = [z, t_i]$, this becomes

$$\frac{\int_{Q_i} f(y)\,dy}{\lambda(Q_i)} \xrightarrow[i\to\infty]{} \widetilde{f}(z),$$

which follows from the stronger result

$$\frac{\int_{Q_i} |f(y) - \widetilde{f}(z)|\,dy}{\lambda(Q_i)} \xrightarrow[i\to\infty]{} 0.$$

By item (3) in Lemma 4.9, this is equivalent to z being a Lebesgue point of $\widetilde{f}$—which z is by the effective Lebesgue differentiation theorem (Theorem 4.10). □

Variations of Corollary 4.18 are given in Corollary 6.9, Theorem 9.19, and Corollary 9.20. Further, in Section 12, I will give an example showing that Corollary 4.18 characterizes Schnorr randomness.

Related to absolutely continuous functions is the following theorem about Radon-Nikodym derivatives.

Fact 4.19 (Radon-Nikodym, see [48])**.** *Let μ be a probability measure on $[0,1]^d$. If μ is absolutely continuous with respect to λ (i.e. $\lambda(A) = 0$ implies $\mu(A) = 0$ for all Borel-measurable A), then there is a λ-a.e. unique integrable function $\frac{d\mu}{d\lambda}$, called* *the* RADON-NIKODYM DERIVATIVE *or* DENSITY, *such that for all Borel-measurable sets A,*

$$\mu(A) = \int_A \frac{d\nu}{d\lambda}(x)\,dx.$$

Fact 4.20 (See [48]). *Let μ be a probability measure on $[0,1]^d$ that is absolutely continuous with respect to λ. Then*

$$\frac{\mu(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \frac{d\mu}{d\lambda}(x) \quad (\lambda\text{-a.e. } x).$$

Given a computable measure μ, absolutely continuous with respect to λ, say that μ is COMPUTABLY NORMABLE relative to λ if and only if $\frac{d\mu}{d\lambda} \in L^1_{comp}(\lambda)$. (See [27, 26] for an equivalent characterization of computably normable using norms.)

Corollary 4.21. *Let μ be a computable probability measure on $[0,1]^d$ that is absolutely continuous with respect to λ, and computably normable relative to λ. Then*

$$\frac{\mu(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \widetilde{\frac{d\mu}{d\lambda}}(x) \quad (\text{on } \lambda\text{-Schnorr random } x).$$

Proof. Since μ is computably normable relative to λ, we have $\frac{d\mu}{d\lambda} \in L^1_{comp}(\lambda)$. So then

$$\frac{\mu(B(x,r))}{\lambda(B(x,r))} = \frac{\int_{B(x,r)} \frac{d\mu}{d\lambda}(x)\,dx}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \widetilde{\frac{d\mu}{d\lambda}}(x)$$

on Schnorr randoms x by the effective Lebesgue differentiation theorem (Theorem 4.10). □

An extension of Corollary 4.21 to signed measures is given in Theorem 9.12. In Section 12, I will give an example showing that Corollary 4.21 characterizes Schnorr randomness.

I end this section with an application to effective harmonic analysis. Rescale $\mathbb{T}$ to be $[0,2\pi)$ and here i will denote $\sqrt{-1}$. Let $f \in L^1(\mathbb{T} \to \mathbb{C})$ be a complex-valued integrable function on $\mathbb{T}$. Let $\{\hat{f}(j)\}_{j\in\mathbb{Z}}$ be the complex-valued FOURIER COEFFICIENTS of f, that is

$$\hat{f}(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cdot e^{-ijt}dt.$$

Then f can be approximated by the following complex-valued trigonometric polynomials $\sigma_n(f)$ (arising from the Fejér kernel)

$$\sigma_n(f)(x) = \frac{1}{n+1}\sum_{k=0}^{n}\sum_{j=-k}^{k}\hat{f}(j)e^{ijx}.$$

We have the following theorem of Lebesgue. (Note the definition of Lebesgue point naturally extends to complex-valued functions.)

Fact 4.22 (Lebesgue, see [29]). *If $x \in \mathbb{T}$ is a Lebesgue point of f, then $\sigma_n(f)(x) \to f(x)$ as $n \to \infty$.*

To given an effective version, note that f is computable in $L^1(\mathbb{T} \to \mathbb{C})$ (with a suitable choice of test functions) if and only if both its real and imaginary parts are computable in $L^1(\mathbb{T} \to \mathbb{R})$. It is worth noting that $\hat{f}(j)$ is computable in $\mathbb{C}$ uniformly from f and j (use the facts in Proposition 3.20 and that e^{-ijt} is bounded and computable), and that $\sigma_n(f)$ is a computable complex-valued function uniformly in f and n.

Corollary 4.23. *If $f \in L^1_{comp}(\mathbb{T} \to \mathbb{C})$ and x is Schnorr random, then $\sigma_n(f)(x) \to \widetilde{f}(x)$ as $n \to \infty$.*

Proof. By the effective Lebesgue differentiation theorem (Theorem 4.10) x is a Lebesgue point of both the real and imaginary parts of f. Therefore x is also a Lebesgue point of f. The rest of the corollary follows from Fact 4.22. □

5. Martingales in computable analysis

The remainder of this paper is devoted to the effective convergence properties of martingales and applications thereof. This section develops the theory of martingales in computable analysis.

So far, we have only used dyadic martingales on $2^{\mathbb{N}}$, i.e. functions $M\colon 2^{<\omega} \to \mathbb{R}$ that satisfy $\frac{1}{2}M(\sigma 0)+\frac{1}{2}M(\sigma 1) = M(\sigma)$. As motivation, one may represent a dyadic martingale as a sequence of functions, $M_n(x) = M(x \upharpoonright n)$ for $x \in 2^{\mathbb{N}}$. This alternate notation is the common one used in probability theory and it allows for a much more general class of martingales. We will define what it means for a martingale in this more general sense to be computable.

Throughout this section, fix an arbitrary computable probability space $(\mathbb{X}, \mu)$.

5.1. **Conditional expectation.** An important concept in probability theory is that of conditional expectation. Recall that a σ-algebra is a collection of sets closed under complement, countable intersection and countable union. The collection $\mathcal{B}$ of Borel sets is a σ-algebra. We will only consider sub-σ-algebras of $\mathcal{B}$, and we will only consider them up to μ-a.e. equivalence. (Two σ-algebras $\mathcal{F}, \mathcal{G}$ are μ-A.E. EQUIVALENT if every $A \in \mathcal{F}$ is μ-a.e. equivalent to some $B \in \mathcal{G}$, and vice versa. For example, a σ-algebra with only measure 0 and measure 1 sets is equivalent to the trivial σ-algebra $\{\varnothing, \mathbb{X}\}$.) Hence every σ-algebra should be understood as a collection of equivalence classes of measurable sets.

An important type of σ-algebra is one generated by a finite partition $\mathcal{P} = \{Q_0, \dots, Q_{k-1}\}$ of $\mathbb{X}$ (i.e. $\bigcup_{i=0}^{k-1} Q_i = \mathbb{X}$ μ-a.e.). Given such a finite partition $\mathcal{P}$, and given $f \in L^1(\mathbb{X}, \mu)$, the CONDITIONAL EXPECTATION $\mathbb{E}[f \mid \mathcal{P}] \in L^1(\mathbb{X}, \mu)$ is defined by

$$\mathbb{E}[f \mid \mathcal{P}] := \sum_{i=0}^{k-1} \frac{\int_{Q_i} f\, dx}{\mu(Q_i)} \cdot \mathbf{1}_{Q_i}.$$

We may leave $\mathbb{E}[f \mid \mathcal{P}](x)$ undefined when $x \in Q_i$ and $\mu(Q_i) = 0$, as we only wish to define $\mathbb{E}[f \mid \mathcal{P}]$ as an a.e. equivalence class. Notice that $\mathbb{E}[f \mid \mathcal{P}]$ is a step function constant on each Q_i, and so I will sometimes abuse notation and write $\mathbb{E}[f \mid \mathcal{P}](Q_i) := \frac{1}{\mu(Q_i)} \int_{Q_i} f\, d\mu$ where convenient. Below, and throughout the paper, $\widetilde{\mathbb{E}}[f \mid \mathcal{P}](x)$ will mean $\widetilde{g}$ where $g = \mathbb{E}[f \mid \mathcal{P}]$.

Proposition 5.1. *Let $\mathcal{P} = \{Q_0, \dots, Q_{k-1}\}$ be a finite partition of $\mathbb{X}$ into effectively measurable sets, and let f be an L^1-computable function. Then the following hold.*

1. *$\mathbb{E}[f \mid \mathcal{P}]$ is L^1-computable uniformly from (the names for) f and $\mathcal{P}$.*
2. *The value $\mathbb{E}[f \mid \mathcal{P}](Q_i)$ is computable from f and Q_i.*
3. *$\widetilde{\mathbb{E}}[f \mid \mathcal{P}](x) = \mathbb{E}[f \mid \mathcal{P}](Q_i)$ assuming $x \in \widetilde{Q}_i$ and x is Schnorr random.*

Proof. Items (1) and (2) are straightforward. For (3), assume x is Schnorr random and $x \in \widetilde{Q}_i$. Then, $\mu(Q_i) > 0$. By Definition 3.17, $\widetilde{\mathbf{1}_{Q_i}}(x) = 1$. Moreover, $\widetilde{\mathbf{1}_{Q_j}}(x) = 0$ for $j \neq i$ (since, by Proposition 3.29, $\widetilde{Q}_i \cap \widetilde{Q}_j = \widetilde{Q_i \cap Q_j} = \widetilde{\varnothing} = \varnothing$). Then, by Proposition 3.29, we have

$$\widetilde{\mathbb{E}}[f \mid \mathcal{P}]\,(x) = \sum_{j=0}^{k-1} \frac{\int_{Q_j} f\,dx}{\mu\,(Q_j)} \cdot \widetilde{\mathbf{1}_{Q_j}}(x) = \frac{\int_{Q_i} f\,dx}{\mu\,(Q_i)} = \mathbb{E}\,[f \mid \mathcal{P}]\,(Q_i). \qquad \square$$

The definition of conditional expectation can be extended to any σ-algebra. The CONDITION EXPECTATION $\mathbb{E}[f \mid \mathcal{F}]$ is the a.e. unique function $\mathbb{E}[f \mid \mathcal{F}] \in L^1(\mathbb{X}, \mu)$ such that $\int_A \mathbb{E}[f \mid \mathcal{F}](x)\,d\mu(x) = \int_A f\,d\mu$ for all measurable $A \in \mathcal{F}$. (Alternately, $\mathbb{E}[f \mid \mathcal{F}]$ can be defined directly using the Radon-Nikodym derivative.) If $\mathcal{F}$ is the σ-algebra generated by a partition $\mathcal{P}$, then $\mathbb{E}[f \mid \mathcal{F}] = \mathbb{E}[f \mid \mathcal{P}]$ μ-a.e. The following facts about conditional expectation will be used quite often (sometimes without reference).

Fact 5.2 (See [13, 53])**.** *Assume $f, g, f_n \in L^1(\mathbb{X}, \mu)$, and $\mathcal{F}$, $\mathcal{F}_1$, $\mathcal{F}_2$ are σ-algebras.*

(1) *$\mathbb{E}\,[f \mid \mathcal{F}]$ is $\mathcal{F}$-measurable.*
(2) *$\int \mathbb{E}\,[f \mid \mathcal{F}]\,(x)\,dx = \int f(x)\,dx$.*
(3) *If f is $\mathcal{F}$-measurable, then $\mathbb{E}\,[f \mid \mathcal{F}] = f$ a.e.*
(4) *(Tower property) If $\mathcal{F}_1 \subseteq \mathcal{F}_2$ (as σ-algebras), then $\mathbb{E}\,[\mathbb{E}\,[f \mid \mathcal{F}_2] \mid \mathcal{F}_1] = \mathbb{E}\,[f \mid \mathcal{F}_1]$ a.e.*
(5) *If $\int |g(x)f(x)|\,dx < \infty$ and g is $\mathcal{F}$-measurable, then $\mathbb{E}\,[gf \mid \mathcal{F}] = g \cdot \mathbb{E}\,[f \mid \mathcal{F}]$ a.e.*
(6) *(Linearity) $\mathbb{E}\,[af + g \mid \mathcal{F}] = a\mathbb{E}\,[f \mid \mathcal{F}] + \mathbb{E}\,[g \mid \mathcal{F}]$ a.e.*
(7) *If $f \leq g$ a.e., then $\mathbb{E}\,[f \mid \mathcal{F}] \leq \mathbb{E}\,[g \mid \mathcal{F}]$ a.e.*
(8) *(Conditional Jensen's inequality) $|\mathbb{E}\,[f \mid \mathcal{F}]\,| \leq \mathbb{E}\,[|f| \mid \mathcal{F}]$ a.e. (or replace $|\cdot|$ with any convex function).*
(9) *If $\mathcal{F}_1 \subseteq \mathcal{F}_2$ (as σ-algebras), then $\|\mathbb{E}\,[f \mid \mathcal{F}_1]\|_{L^1} \leq \|\mathbb{E}\,[f \mid \mathcal{F}_2]\|_{L^1} \leq \|f\|_{L^1}$ (also for the L^2-norm).*
(10) *(Conditional Fatou's lemma) $\mathbb{E}[\limsup_{n\to\infty} f_n \mid \mathcal{F}] \geq \limsup_{n\to\infty} \mathbb{E}[f_n \mid \mathcal{F}]$ if there is some $g \in L^1$ such that $f_n \geq g$ for all n.*

5.2. **L^1-computable martingales.** A FILTRATION $(\mathcal{F}_k)$ is a chain of σ-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$. We say a filtration $(\mathcal{F}_k)$ CONVERGES to the σ-algebra $\mathcal{F}_\infty$, written $\mathcal{F}_k \uparrow \mathcal{F}_\infty$, when $\mathcal{F}_\infty = \sigma\left(\bigcup_k \mathcal{F}_k\right)$. One example of a filtration is a chain of increasingly fine partitions. The only filtration we will use by name is the filtration generated by the chain of partitions $(\mathcal{B}_k)$ where, on $2^{\mathbb{N}}$, $\mathcal{B}_k = \{[\tau] \mid |\tau| = k\}$, and on $\mathbb{T}^d$ or $[0,1]^d$, $\mathcal{B}_k$ is the set of dyadic cubes with side length 2^{-k}. It is clear that $\mathcal{B}_k \uparrow \mathcal{B}$, where $\mathcal{B}$ is the Borel σ-algebra.

A MARTINGALE adapted to a filtration $(\mathcal{F}_k)$ is a sequence of integrable functions (M_k) such that M_k is $\mathcal{F}_k$-measurable and

$$\mathbb{E}\,[M_{k+1} \mid \mathcal{F}_k] = M_k \text{ a.e.} \tag{5.1}$$

Assuming the filtration $(\mathcal{F}_k)$ is given by a sequence of partitions $(\mathcal{P}_k)$, then M_k is constant on all $Q \in \mathcal{P}_k$. We then may write $M(Q)$ for $M_k(x)$ where $x \in Q$.

Example 5.3. Every dyadic martingale $M\colon 2^\omega \to \mathbb{R}$ is equivalent to a martingale (M_k) on $(2^{\mathbb{N}}, \lambda)$ with respect to the filtration $(\mathcal{B}_k)$, and vice versa, under the translation $M_k(x) = M(x \upharpoonright k)$. It is easy to see condition (5.1) is equivalent to

$$M(\sigma 0)\mu(\sigma 0) + M(\sigma 1)\mu(\sigma 1) = M(\sigma)\mu(\sigma).$$

In algorithmic randomness, it is customary to assume the martingales are nonnegative. We do not make that assumption here.

Martingales are useful for their well-behaved convergence properties. Also, they have a natural interpretation in terms of gambling. In general, $\mathcal{F}_k$ is the information known to the gambler at time k, and M_k is the capital of the gambler at time k following a betting strategy given by M.

It is not necessary to refer to a specific filtration when talking about martingales. Any martingale (M_k) is also a martingale with respect to the filtration $(\mathcal{F}_k)$ where

$$\mathcal{F}_k = \sigma(M_0, \ldots, M_k) = \sigma\left(\bigcup_{i=0}^{k} \left\{M_i^{-1}(A) \mid A \in \mathcal{B}\right\}\right)$$

i.e. the minimal σ-algebra with respect to which $M_0, \ldots, M_k$ are measurable. (In the definition of $\sigma(M_0, \ldots, M_k)$, it is sufficient to replace $\mathcal{B}$ with any countable generator of $\mathcal{B}$.) Hence (M_k) is a martingale (with respect to some filtration) if and only if $\mathbb{E}[M_{k+1} \mid M_0, \ldots, M_k] = M_k$ (where $\mathbb{E}[M_{k+1} \mid M_0, \ldots, M_k]$ is defined as $\mathbb{E}[M_{k+1} \mid \sigma(M_0, \ldots, M_k)]$).

We say a martingale (M_k) is an L^1-COMPUTABLE MARTINGALE if (M_k) is a computable sequence of L^1-computable functions.

Last, we mention the general form of Kolmogorov's inequality (compared with Fact 4.4) which extends Markov's inequality (Fact A.2). We will use it quite often.

Fact 5.4 (Kolmogorov's inequality, see [53])**.** *For a martingale (M_k), and $n, m \in \mathbb{N}$,*

$$\mu\left(\left\{x \in \mathbb{X} \,\middle|\, \max_{k \in [n,m]} |M_k(x)| \geq \varepsilon\right\}\right) \leq \frac{\|M_m\|_{L^1}}{\varepsilon}.$$

6. The Lévy 0-1 law and uniformly integrable martingales

6.1. **Some martingale convergence theorems.** Assume in this section that $(\mathbb{X}, \mu)$ is a computable probability space. Consider the following class of martingales.

Example 6.1. If $f \in L^1(\mathbb{X}, \mu)$ and $(\mathcal{F}_k)$ is a filtration, then $\mathbb{E}[f \mid \mathcal{F}_k]$ is a martingale on $(\mathcal{F}_k)$ by Fact 5.2 (4). In the case that $\mathbb{X} = 2^{\mathbb{N}}, \mathbb{T}^d, [0,1]^d$, then the sequence $f^{(k)}$ from the Section 4.1 is equal to $\mathbb{E}[f \mid \mathcal{B}_k]$.

Fact 6.2 (Lévy 0-1 law, see [13, 53])**.** *Given a filtration $(\mathcal{F}_k)$ such that $\mathcal{F}_k \uparrow \mathcal{F}_\infty$ and $f \in L^1$, then*

$$\mathbb{E}[f \mid \mathcal{F}_k] \xrightarrow[k \to \infty]{} \mathbb{E}[f \mid \mathcal{F}_\infty] \quad (L^1 \text{ and a.e.}).$$

Therefore, if f is $\mathcal{F}_\infty$-measurable, then $\mathbb{E}[f \mid \mathcal{F}_\infty] = f$ a.e. and

$$\mathbb{E}[f \mid \mathcal{F}_k] \xrightarrow[k \to \infty]{} f \quad (L^1 \text{ and a.e.}).$$

In this section I give an effective version of the Lévy 0-1 law.

Theorem 6.3 (Effective Lévy 0-1 law)**.** *Let $(\mathcal{F}_k)$ be any filtration with limit $\mathcal{F}_\infty$. Assume $f \in L^1_{comp}$, $\mathbb{E}[f \mid \mathcal{F}_k]$ is L^1-computable uniformly in k, and $\mathbb{E}[f \mid \mathcal{F}_\infty] \in L^1_{comp}$. Then*

$$\mathbb{E}\left[f \mid \mathcal{F}_k\right] \xrightarrow[k\to\infty]{} \mathbb{E}\left[f \mid \mathcal{F}_\infty\right] \quad (\text{effectively } L^1 \text{ and effectively a.e.}).$$

Hence, by Lemma 3.19,

$$\widetilde{\mathbb{E}}\left[f \mid \mathcal{F}_k\right](z) \xrightarrow[k\to\infty]{} \widetilde{\mathbb{E}}\left[f \mid \mathcal{F}_\infty\right](z) \quad (\text{on Schnorr random } z).$$

To prove this theorem, we will rely on the following characterization of martingales which converge in the L^1-norm. A martingale (M_k) is called UNIFORMLY-INTEGRABLE if it satisfies either of the following equivalent conditions.

Fact 6.4 (see [13])**.** *If (M_k) is a martingale on the filtration $(\mathcal{F}_k)$ the following are equivalent.*

(1) *(M_k) converges in the L^1-norm.*
(2) *There exists $f \in L^1$ such that $M_k = \mathbb{E}\left[f \mid \mathcal{F}_k\right]$ a.e. for all k.*
(3) *The sequence of functions (M_k) is uniformly integrable, i.e.*

$$\sup_k \int_{\left\{x\in\mathbb{X} \,\middle|\, |M_k(x)|>C\right\}} |M_k| \, d\mu \xrightarrow[C\to\infty]{} 0.$$

(Condition 3 is will not be used in this paper.) By the Lévy 0-1 law, every uniformly-integrable martingale has a limit. By Fact 6.4, the effective Lévy 0-1 law (Theorem 6.3) follows from the next lemma.

Lemma 6.5. *Assume (M_k) is a uniformly-integrable, L^1-computable martingale with limit $M_\infty \in L^1_{comp}$. Then*

$$M_k \xrightarrow[k\to\infty]{} M_\infty \quad (\text{effectively } L^1 \text{and effectively a.e.}).$$

Hence, by Lemma 3.19, $\widetilde{M}_k(z) \xrightarrow[k\to\infty]{} \widetilde{M}_\infty(z)$ for Schnorr randoms z.

Proof. Since we know that, $M_k \xrightarrow[k\to\infty]{L^1} M_\infty$ and since M_∞, M_k are uniformly L^1-computable, we can find a subsequence $\left(M_{k_j}\right)$ such that for all $j \geq i$ we have $\left\|M_{k_j} - M_{k_i}\right\|_{L^1} \leq 2^{-i}$. The subsequence converges effectively in L^1 and a.e. (Proposition 3.15).

First, we show convergence in the L^1-norm. Fix i. Notice that $N_k := (M_k - M_{k_i})$ is a martingale for $k \geq k_i$. (This is easy to verify using conditional expectation facts (Facts 5.2) and the fact that M_{k_i} is $\mathcal{F}_{k_i}$-measurable.) The L^1-norm of the martingale (N_k) is nondecreasing (Facts 5.2) and hence for any $j \geq i$,

$$\max_{k\in[k_i,k_j]} \left\|M_k - M_{k_i}\right\|_{L^1} \leq \left\|M_{k_j} - M_{k_i}\right\|_{L^1} \leq 2^{-i}. \tag{6.1}$$

Since i and j are arbitrary, this shows (M_n) is effectively Cauchy in the L^1-norm.

To show effective a.e. convergence, again fix i and use Kolmogorov's inequality (Fact 5.4) on the martingale $N_k := (M_k - M_{k_i})$ to get

$$\mu\left(\left\{x \,\middle|\, \max_{k\in[k_j,k_i]} |M_k(x) - M_{k_i}(x)| \geq 2^{-i/2}\right\}\right) \leq \frac{\left\|M_{k_j} - M_{k_i}\right\|_{L^1}}{2^{-i/2}} \leq 2^{-i/2}. \tag{6.2}$$

Since i and j are arbitrary, this shows (M_n) is effectively a.e. Cauchy. □

Remark 6.6. Notice in the case that $\mathbb{X} = 2^{\mathbb{N}}, \mathbb{T}^d, [0,1]^d$ and $M_k = f^{(k)}$ (as in Section 4.1), then Lemma 6.5 follows from the effective Lebesgue approximation theorem (Proposition 4.3) (L^1 convergence) and the effective dyadic Lebesgue differentiation theorem (Proposition 4.6) (a.e. convergence).

If the martingale is L^2-computable and L^2-BOUNDED, i.e. $\sup_k \|M_k\|_{L^2} < \infty$, then it is sufficient to know the L^2-bound instead of the limit. (This is not true of the L^1 case.)

Fact 6.7 (See [13]). *Assume (M_k) is an L^2-bounded martingale. Then (M_k) is uniformly-integrable, has a limit M_∞ in the L^2-norm (and L^1-norm), and $\sup_k \|M_k\|_{L^2} = \|M_\infty\|_{L^2}$.*

Corollary 6.8. *Assume (M_k) is an L^2-computable martingale with limit M_∞ and with computable L^2-bound $b = \sup_k \|M_k\|_{L^2} = \|M_\infty\|_{L^2}$. Then*

$$M_k \xrightarrow[k\to\infty]{} M_\infty \quad (\textit{eff. } L^2, \textit{eff. } L^1, \textit{eff. a.e.}).$$

Therefore, M_∞ is L^2- and L^1-computable (uniformly from (M_k) and b), and $\widetilde{M}_k(z) \xrightarrow[k\to\infty]{} \widetilde{M}_\infty(z)$ for Schnorr randoms z.

Proof. The space of L^2-functions is a Hilbert space and the conditional expectation $f \mapsto \mathbb{E}[f \mid \mathcal{F}]$ is a projection onto the space of $\mathcal{F}$-measurable functions [13]. Therefore, by the Pythagorean theorem, for $k \geq j$,

$$\|M_k - M_j\|_{L^2}^2 = \|M_k\|_{L^2}^2 - \|M_j\|_{L^2}^2 \leq b^2 - \|M_j\|_{L^2}^2 .$$

Since the L^2-bound b is finite and computable, this implies effective convergence in the L^2-norm and hence in the L^1-norm as well. Hence the limit is L^1- and L^2-computable (uniformly from (M_k) and b). Since (M_k) converges in L^1, the martingale is uniformly-integrable (Fact 6.4). The rest follows from Lemma 6.5. □

This gives the following variation of Corollary 4.18.

Corollary 6.9. *Let $F\colon [0,1] \to \mathbb{R}$ be a computable function which is also absolutely-continuous with derivative $f = \frac{d}{dx}F$. Assume that $\|f\|_{L^2}$ is computable. Then f is L^2-computable (uniformly from F and $\|f\|_{L^2}$), F is effectively absolutely continuous, and F is differentiable on Schnorr randoms.*

Proof. For any non-dyadic real $x \in [0,1]$, let $x \upharpoonright n$ denote the binary expansion of x truncated at the nth bit and let $0.x \upharpoonright n$ denote the corresponding dyadic rational. Then

$$\frac{d}{dx}F(x) = \lim_{n\to\infty} \frac{F(2^{-n} + 0.x \upharpoonright n) - F(0.x \upharpoonright n)}{2^{-n}}.$$

The term under the limit is an L^2-computable martingale as follows. If f is the derivative of F, then

$$\frac{F(2^{-n} + 0.(x \upharpoonright n)) - F(0.(x \upharpoonright n))}{2^{-n}} = \frac{\int_{[x\upharpoonright n]} f \, d\lambda}{2^{-n}} = f^{(n)}(x)$$

where $f^{(n)}(x)$ is the martingale defined in Section 4.1 (see Example 6.1). Each $f^{(n)}$ is L^2-computable from F and n since it is a test function. We know $f^{(n)} \xrightarrow[n\to\infty]{L^1} f$ (Fact 4.2). Since $\|f\|_{L^2}$ is computable, by Corollary 6.8, the derivative f is L^2-computable and F is effectively absolutely continuous. The rest follows by Corollary 4.18. □

In Section 12, I will give examples showing that the theorems of this section characterize Schnorr randomness.

7. More martingale convergence results

7.1. **Martingale convergence results.** A martingale (M_k) is said to be L^1-BOUNDED if $\sup_k \|M_k\|_{L^1} < \infty$.

The Lévy 0-1 Law above is a special case of the following theorem.

Fact 7.1 (Doob's martingale convergence theorem, see [13, 53])**.** *If (M_k) is an L^1-bounded martingale, then M_k converges pointwise a.e. and in measure to an integrable function.*

Example 7.2. If a martingale is uniformly-integrable or nonnegative then it is L^1-bounded. Indeed, given a uniformly-integrable martingale (M_k), there is some $f \in L^1$ such that $M_k = \mathbb{E}[f \mid \mathcal{F}_k]$ (Fact 6.4) and $\|\mathbb{E}[f \mid \mathcal{F}_k]\|_{L^1} \leq \|f\|_{L^1}$ (Facts 5.2). For a nonnegative martingale (M_k), we have (using Facts 5.2) that

$$\|M_k\|_{L^1} = \int M_k \, d\mu = \int \mathbb{E}[M_k \mid \mathcal{F}_0] \, d\mu = \int M_0 \, d\mu = \|M_0\|_{L^1}.$$

While martingale convergence in general is not effective, it can be under certain circumstances. We have already seen the case when the martingale is uniformly-integrable.

Unlike uniform integrability, being merely L^1-bounded only implies pointwise convergence, not convergence in the L^1-norm.

Example 7.3. Consider a doubling strategy, whereby the gambler bets all his capital on at each stage until he loses. The limit of his capital is almost-surely zero, but the martingale is nonnegative, so the L^1-norm stays constant and does not converge in the L^1-norm.

Now I consider the case when (M_k) is a nonnegative singular supermartingale. A SUPERMARTINGALE (M_k) is an ADAPTED PROCESS, i.e. M_k is $\mathcal{F}_k$-measurable such that $\mathbb{E}[M_{k+1} \mid \mathcal{F}_k] \leq M_k$ for all k. (A SUBMARTINGALE (M_k) is the same except $\mathbb{E}[M_{k+1} \mid \mathcal{F}_k] \geq M_k$.) Notice, every martingale is a supermartingale (and submartingale). A supermartingale (M_k) is SINGULAR if $M_k(x) \xrightarrow[k\to\infty]{} 0$ a.e.

Lemma 7.4. *Let M be a nonnegative L^1-computable singular supermartingale. Then $M_k \xrightarrow[k\to\infty]{} 0$ effectively a.e., and hence (by Lemma 3.19) $\widetilde{M}_k(z) \xrightarrow[k\to\infty]{} 0$ for all Schnorr randoms z.*

Proof. By Fact 7.1, $M_k \xrightarrow[k\to\infty]{} 0$ in measure. Hence we can effectively find a subsequence (k_i) such that (M_{k_i}) converges rapidly the metric d_{meas} (Fact 3.11), namely

$$d_{meas}(M_{k_i}, 0) = \|\min\{|M_{k_i}|, 1\}\|_{L^1} < 2^{-(i+1)}.$$

Fix i. Since M_{k_i} is nonnegative, it follows by Markov's inequality (Fact A.2) that

$$\mu\Big(\underbrace{\{x \mid 0 \leq M_{k_i}(x) < 1\}}_{=:C_i}\Big) \leq 1 - 2^{-(i+1)}. \tag{7.1}$$

The set C_i in $\sigma(M_{k_i})$, and hence $C_i \in \mathcal{F}_{k_i}$ for any filtration $(\mathcal{F}_k)$ to which (M_k) is adapted. For $k > k_i$ let $N_k := \mathbf{1}_{C_i} M_k$. The following calculation shows that $(N_k)_{k \geq k_i}$ is still a supermartingale adapted to $(\mathcal{F}_k)$:

$$\mathbb{E}\left[\mathbf{1}_{C_i} M_{k+1} \mid \mathcal{F}_k\right] = \mathbf{1}_{C_j} \mathbb{E}\left[M_{k+1} \mid \mathcal{F}_k\right] \leq \mathbf{1}_{C_j} M_k \text{ a.e.}$$

(Intuitively what makes N_k a supermartingale is that on C_i, the process (N_k) behaves as the supermartingale (M_k), and on the complement of C_i, the process (N_k) is the constant zero supermartingale.) The L^1-norms of nonnegative supermartingales decrease, and therefore for all $k \geq k_j$,

$$\|\mathbf{1}_{C_i} M_k\|_{L^1} \leq \|\mathbf{1}_{C_i} M_{k_i}\|_{L^1} \leq \|\min(M_{k_i}, 1)\|_{L^1} \leq 2^{-(i+1)}.$$

Kolmogorov's inequality (Fact 5.4) also holds for nonnegative supermartingales, and therefore for $j > i$

$$\mu\left(\left\{x \,\middle|\, \max_{k \in [k_i, k_j]} \mathbf{1}_{C_i}(x) M_k(x) \geq 2^{-(i+1)/2}\right\}\right) \leq \frac{2^{-(i+1)}}{2^{-(i+1)/2}} \leq 2^{-(i+1)/2}.$$

Call this set A_i. Then

$$\mu\left(\left\{x \,\middle|\, \max_{k \in [k_i, k_j]} M_k(x) \geq 2^{-(i+1)/2}\right\}\right) \leq \mu(A_i) + (1 - \mu(C_i)) \leq 2^{-i/2}.$$

As i and j are arbitrary, $M_k \to 0$ effectively a.e. □

Our goal, however, is to show any martingale converges effectively a.e. if the L^1-bound and the limit are known. To prove this, I will use two complimentary martingale decompositions. In this next decomposition, M_k^+ denotes the nonnegative part of the martingale decomposition, whereas $[M_k]^+$ will mean $\max(M_k, 0)$—and similarly for M_k^- and $[M_k]^-$. (Whereas (M_k^+) is a martingale, $([M_k]^+)$ is only a submartingale.) Also, for a martingale $N = (N_k)$, denote $\|N\|_{M^1} = \sup_k \|N_k\|_{L^1}$.

Fact 7.5 (Krickeberg Decomposition, see [8, Chapter V, Section 4])**.** *Let (M_k) be an L^1-bounded martingale with respect to the filtration $(\mathcal{F}_k)$. Then there are two nonnegative martingales (M_k^+) and (M_k^-) such that such that $M_k = M_k^+ - M_k^-$ a.e. for all k, and $\|M\|_{M^1} = \|M^+\|_{M^1} + \|M^-\|_{M^1} = \left\|M_k^+\right\|_{L^1} + \left\|M_k^-\right\|_{L^1}$ for all k. Further, this decomposition is a.e. unique; $M_k^+ = \sup_n \mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]$ a.e.; $M_k^- = \sup_n \mathbb{E}[[M_n]^- \mid \mathcal{F}_k]$ a.e.; $\lim_{k\to\infty} M_k^+ = [\lim_k M_k]^+$ a.e.; and $\lim_{k\to\infty} M_k^- = [\lim_k M_k]^-$ a.e.*

Let (M_k) be an L^1-bounded martingale with respect to the filtration $(\mathcal{F}_k)$ and let $M_\infty = \lim_n M_n$. Then there is a uniformly-integrable martingale (M_k^{ui}) and a singular martingale (M_k^s) such that $M_k = M_k^s + M_k^{ui}$ a.e. for all k. Further, this decomposition is a.e. unique; $M_k^{ui} = \mathbb{E}[M_\infty \mid \mathcal{F}_k]$ a.e.; $M_k^s = \mathbb{E}[M_k - M_\infty \mid \mathcal{F}_k]$ a.e.; and $\|M_k\|_{M^1} = \|M_k^s\|_{M^1} + \left\|M_k^{ui}\right\|_{M^1}$.

Remark 7.6. To make the decompositions computable, we need the filtration to be computable. The filtration $(\mathcal{F}_k)$ can be represented by the sequence of operators $f \mapsto \mathbb{E}[f \mid \mathcal{F}_k]$ from L^1 to L^1. Say that $(\mathcal{F}_k)$ is COMPUTABLE if $f \mapsto \mathbb{E}[f \mid \mathcal{F}_k]$ is a computable operator from L^1 to L^1 uniformly in k. If $(\mathcal{P}_k)$ is a computable chain of computable partitions, where $\mathcal{P}_{k+1}$ is a refinement of $\mathcal{P}_k$, then the corresponding filtration is computable. Assuming the filtration $(\mathcal{F}_k)$ is computable, the above decompositions are computable using the L^1-bound $\|M\|_{M^1}$ and the limit M_∞, respectively, as follows.

Proposition 7.7 (Effective Krickeberg decomposition)**.** *Let (M_k) be an L^1-computable martingale with respect to a computable filtration $(\mathcal{F}_k)$. Then the Krickeberg decomposition (M_k^+), (M_k^-) is computable from (M_k), $(\mathcal{F}_k)$, and the L^1-bound $\|M\|_{M^1}$. Further, the limits $\lim_{k\to\infty} M_k^+ = [M_\infty]^+$ and $\lim_{k\to\infty} M_k^- = [M_\infty]^-$ are L^1-computable from the limit M_∞.*

Proof. We wish to compute $M_k^+ = \sup_n \mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]$ and $M_k^- = \sup_n \mathbb{E}[[M_n]^- \mid \mathcal{F}_k]$. Note that $\mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]$ is L^1-computable from n, k, and $(\mathcal{F}_k)$, since the filtration is computable. To show each supremum is L^1-computable, fix $\varepsilon > 0$ and k. Then choose $n > k$ such that

$$\left\|\mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]\right\|_{L^1} + \left\|\mathbb{E}[[M_n]^- \mid \mathcal{F}_k]\right\|_{L^1} > \|M\|_{M^1} - \varepsilon$$

Since $M_k^+ \geq \mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]$ and $M_k^- \geq \mathbb{E}[[M_n]^- \mid \mathcal{F}_k]$ for all n, we have

$$\begin{aligned}
&\left\|M_k^+ - \mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]\right\|_{L^1} + \left\|M_k^- - \mathbb{E}[[M_n]^- \mid \mathcal{F}_k]\right\|_{L^1} \\
&= \|M_k^+\|_{L^1} + \|M_k^-\|_{L^1} - \left(\left\|\mathbb{E}[[M_n]^+ \mid \mathcal{F}_k]\right\|_{L^1} + \left\|\mathbb{E}[[M_n]^- \mid \mathcal{F}_k]\right\|_{L^1}\right) \\
&\leq \|M\|_{M^1} - \left\|\mathbb{E}[[M_n]^+ \mid \mathcal{F}_k] + \mathbb{E}[[M_n]^- \mid \mathcal{F}_k]\right\|_{L^1} \leq \varepsilon.
\end{aligned}$$

Hence M_k^+ and M_k^- are L^1-computable uniformly in k.

To compute the limits, just use that fact that $[M_\infty]^+$ and $[M_\infty]^-$ are L^1-computable from M_∞. ☐

Proposition 7.8 (Effective Uniformly Integrable/Singular Decomposition)**.** *Let (M_k) be an L^1-computable martingale with respect to a computable filtration $(\mathcal{F}_k)$. Then the decomposition (M_k^{ui}), (M_k^s) is computable from (M_k), $(\mathcal{F}_k)$, and the limit M_∞. Further, the L^1-bound $\|M^s\|_{M^1} = \|M\|_{M^1} - \left\|M^{ui}\right\|_{M^1} = \|M\|_{M^1} - \|M_\infty\|_{L^1}$ is computable from $\|M\|_{M^1}$.*

Proof. Since the filtration is computable, $M_k^{ui} = \mathbb{E}[M_\infty \mid \mathcal{F}_k]$ is computable in the L^1-norm uniformly from M_∞, k, and $(\mathcal{F}_k)$. Then $M_k^s = M_k - M_k^{ui}$ is computable in the L^1-norm. To compute $\|M^s\|_{M^1}$ just use that $\left\|M^{ui}\right\|_{M^1} = \|M_\infty\|_{L^1}$ is computable. ☐

In the martingale convergence results so far, there have been no computability requirements on the filtration $(\mathcal{F}_k)$. We can continue to work without specifying the computability of the filtration. The trick is to approximate M by a different martingale whose filtration is given by a chain of partitions.

Proposition 7.9. *Let M be an L^1-computable martingale (resp. supermartingale, submartingale). There is a computable martingale (resp. supermartingale, submartingale) N adapted to a computable chain of computable partitions $(\mathcal{P}_k)$ such that for all k, $\mathcal{P}_k \subseteq \sigma(M_0, \ldots, M_k)$ and $\|N_k - M_k\|_{L^1} \leq 2^{-k}$. If M is nonnegative, then so is N. Further, if M is a martingale or nonnegative submartingale, then $\sup_n \|M_n\|_{L^1} = \sup_n \|N_n\|_{L^1}$.*

Proof. The main idea is to take each σ-algebra in the canonical filtration $\mathcal{F}_k = \sigma(M_0, \ldots, M_k)$ and approximate it with a finite sub-σ-algebra, i.e. a partition $\mathcal{P}_k \subseteq \mathcal{F}_k$.

For each k, let $T_k \colon (\mathbb{X}, \mu) \to \mathbb{R}^{k+1}$ be the map $T_k = (M_0, \ldots, M_k)$. Recall that $\sigma(M_0, \ldots, M_k) = \sigma(T_k) = \sigma(\{T_k^{-1}(B) \mid B \in \mathcal{C}\})$ where $\sigma(\mathcal{C})$ is generates the Borel sigma algebra on the push forward measure space $(\mathbb{R}^{k+1}, \mu_* T_k)$. Recall that $\mu_* T_k$ is computable (Proposition 3.26) and therefore we can take $\mathcal{C} =$

$Basis(\mathbb{R}^{k+1}, \mu_* T_k)$ as in Lemma 3.5. Let $\{B_i^k\}_i$ be a computable enumeration of $Basis(\mathbb{R}^{k+1}, \mu_* T_k)$. Then by Proposition 3.20, $\{T_k^{-1}(B_i^k)\}_{i,k}$ is a computable double sequence of effectively measurable sets which generates $\sigma(M_0, \ldots, M_k)$. That is, if $\mathcal{Q}_i^k = \{T_k^{-1}(B_0^k), \ldots, T_k^{-1}(B_{i-1}^k)\}$, then $\sigma(\mathcal{Q}_i^k) \underset{i\to\infty}{\uparrow} \sigma(M_0, \ldots, M_k)$.

By the Lévy 0-1 law (Fact 6.2), $\mathbb{E}[M_k \mid \mathcal{Q}_i^k] \xrightarrow[i\to\infty]{L^1} M_k$. Since each $\mathbb{E}[M_k \mid \mathcal{Q}_i^k]$ is L^1-computable from i and k, find some i_k such that $\left\|\mathbb{E}[M_k \mid \mathcal{Q}_{i_k}^k] - M_k\right\|_{L^1} \leq 2^{-k}$. Define $\mathcal{P}_k = \mathcal{Q}_{i_k}^k$ and $N_k = \mathbb{E}[M_k \mid \mathcal{P}_k]$.

If M is a supermartingale, then N is as well. Indeed, by two applications of the tower property (Facts 5.2),

$$\begin{aligned}
\mathbb{E}[N_{k+1} \mid \mathcal{P}_k] &= \mathbb{E}[\mathbb{E}[M_{k+1} \mid \mathcal{P}_{k+1}] \mid \mathcal{P}_k] && \text{(definition of } N_{k+1}\text{)} \\
&= \mathbb{E}[M_{k+1} \mid \mathcal{P}_k] && \text{(tower property)} \\
&= \mathbb{E}[\mathbb{E}[M_{k+1} \mid \sigma(M_0, \ldots, M_k)] \mid \mathcal{P}_k] && \text{(tower property)} \\
&\leq \mathbb{E}[M_k \mid \mathcal{P}_k] && (M \text{ is a supermartingale}) \\
&= N_k && \text{(definition of } N_k\text{).}
\end{aligned}$$

If M is a martingale, or submartingale, the same argument works.

In general, $\|N_k\|_{L^1} = \|\mathbb{E}[M_k \mid \mathcal{P}_k]\|_{L^1} \leq \|M_k\|_{L^1}$ which is just a property of conditional expectation (Facts 5.2). Moreover, $|\|M_k\|_{L^1} - \|N_k\|_{L^1}| \leq \|M_k - N_k\|_{L^1}$. If M is a martingale or nonnegative submartingale, then $\|M_k\|_k$ is increasing and hence $\sup_n \|N_k\|_k = \sup_n \|M_k\|_k$. □

Theorem 7.10. *Let M be an L^1-computable martingale with computable L^1-bound $\|M\|_{M^1}$ and L^1-computable limit M_∞. Then $M_k \xrightarrow[k\to\infty]{} M_\infty$ effectively a.e., and hence, by Lemma 3.19, $\widetilde{M}_k(z) \xrightarrow[k\to\infty]{} \widetilde{M}_\infty(z)$ for all Schnorr randoms z.*

Proof. Let N be as in Proposition 7.9. Since $\|N_k - M_k\|_{L^1} \leq 2^{-k}$ for all k, $(N_k - M_k) \xrightarrow[k\to\infty]{} 0$ effectively a.e. It follows that $M_k \xrightarrow[k\to\infty]{} M_\infty$ effectively a.e. if and only if $N_k \xrightarrow[k\to\infty]{} M_\infty$ effectively a.e.

Since N is a martingale with respect to a computable sequence of partitions, N is effectively decomposable (Proposition 7.8) into a uniformly integrable part N^{ui} and a singular part N^s. We know $N_k^{ui} \xrightarrow[k\to\infty]{} M_\infty$ converges effectively a.e. by the effective Lévy 0-1 law (Theorem 6.3).

Since $\|M\|_{M^1}$ is computable, then so is $\|N^s\|_{M^1}$. Therefore, N^s can be effectively decomposed (Proposition 7.7) into two nonnegative L^1-computable singular martingales N^{s+} and N^{s-}. By Lemma 7.4, $N_k^{s+} \xrightarrow[k\to\infty]{} 0$ and $N_k^{s-} \xrightarrow[k\to\infty]{} 0$ effectively a.e.

Putting this all together we have that $N_k = N_k^{ui} + N_k^{s+} - N_k^{s-} \xrightarrow[k\to\infty]{} M_\infty$ effectively a.e. □

In Section 12, I show that Lemma 9.6 (and hence Theorem 7.10) characterizes Schnorr randomness.

8. Submartingales and supermartingales

Recall from the previous section, a sequence (X_k) of integrable functions is a submartingale (resp. supermartingale) adapted to a filtration $(\mathcal{F}_n)$ if X_k is $\mathcal{F}_k$-measurable for all k, and $\mathbb{E}[X_{k+1} \mid \mathcal{F}_k] \geq X_k$ (resp. $\mathbb{E}[X_{k+1} \mid \mathcal{F}_k] \leq X_k$) for all k.

It can be show that L^1-computable, nonnegative submartingales and supermartingales converge effectively a.e. when their L^1-bounds and limits are known. The proofs are different for each.

Theorem 8.1. *Let (X_n) be a nonnegative L^1-computable supermartingale whose limit X_∞ is L^1-computable. Then $X_n \xrightarrow[n\to\infty]{} X_\infty$ effectively a.e. and, by Lemma 3.19, $\widetilde{X}_n(x) \xrightarrow[n\to\infty]{} \widetilde{X}_\infty(x)$ on Schnorr randoms x. (Instead assuming X_n is nonnegative, we may assume that $X_n \geq Z$ for some integrable function Z.)*

Proof. As in the proof of Theorem 7.10, we may use Proposition 7.9 to assume, without loss of generality, that X_n is adapted to a computable filtration. By the fact that (X_n) is a supermartingale, the fact that X_n is nonnegative (or bounded from below by an integrable function Z), and the conditional Fatou's theorem (Facts 5.2), we have

$$X_n \geq \liminf_k \mathbb{E}[X_k \mid \mathcal{F}_n] \geq \mathbb{E}[X_\infty \mid \mathcal{F}_n].$$

Then we have a nonnegative, L^1-computable supermartingale $Y_n = X_n - \mathbb{E}[X_\infty \mid \mathcal{F}_n]$ which converges to 0 a.e. But Y_n converges to 0 effectively a.e. by Lemma 7.4. Also $\mathbb{E}[X_\infty \mid \mathcal{F}_n]$ converges effectively a.e. by the effective Lévy 0-1 law. Putting them together completes the proof. □

For the submartingale case, I first use an effective version of the monotone convergence theorem.

Proposition 8.2 (Effective monotone convergence theorem)**.** *Assume f_n is an nondecreasing sequence of L^1-computable functions. Also assume $\sup_n \|f_n\|_{L^1}$ is finite and computable. Then $f_n \to \sup_n f_n$ effectively in the L^1-norm and effectively a.e. By Lemma 3.19, $\widetilde{f}_n \to \widetilde{\sup_n f_n}$ (or equivalently $\widetilde{\sup_n f_n} = \sup_n \widetilde{f}_n$) on Schnorr randoms.*

Proof. Find a subsequence (n_k) such that $\left(\sup_n \|f_n\|_{L^1}\right) - \|f_{n_k}\|_{L^1} \leq 2^{-k}$. Fix k. By monotonicity, $\|f_n - f_{n_k}\|_{L^1} \leq 2^{-k}$ for all $n \geq n_k$. Also, by monotonicity, Markov's inequality, and the monotone convergence theorem,

$$\begin{aligned}\mu\left(\left\{\sup_n |f_n - f_{n_k}| > 2^{-k/2}\right\}\right) &= \mu\left(\left\{\left(\sup_n f_n\right) - f_{n_k} > 2^{-k/2}\right\}\right)\\ &\leq \frac{\|\sup_n f_n - f_{n_k}\|_{L^1}}{2^{-k/2}}\\ &= \frac{\sup \|f_n\|_{L^1} - \|f_{n_k}\|_{L^1}}{2^{-k/2}} \leq 2^{-k/2}.\end{aligned}$$

Since k is arbitrary, this gives effective convergence in L^1 and effective a.e. convergence. □

I also use an effective version of Doob's decomposition theorem.

Fact 8.3 (Doob decomposition, see [53]). *Let (X_n) be a submartingale with respect to $(\mathcal{F}_n)$. Then there is a martingale (M_n) with respect to $(\mathcal{F}_n)$ and a* PREDICTABLE PROCESS *A_n (i.e. A_{n+1} is $\mathcal{F}_n$ measurable) such that $A_0 = 0$ and $X_n = M_n + A_n$. Moreover, this decomposition is a.e. unique; $A_{n+1} - A_n = \mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n]$; and A_n is nondecreasing.*

Proposition 8.4 (Effective Doob Decomposition). *If (X_n) is an L^1-computable submartingale and $(\mathcal{F}_n)$ is a computable filtration, then the Doob decomposition is effective.*

Proof. It is enough that $\mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n]$ is L^1-computable from the parameters. □

Theorem 8.5. *Let (X_n) be a nonnegative, L^1-computable submartingale such that the L^1-bound $\sup_n \|X_n\|_{L^1}$ is computable and the limit X_∞ is L^1-computable. Then $X_n \xrightarrow[n\to\infty]{} X_\infty$ effectively a.e. and $\widetilde{X}_n(x) \xrightarrow[n\to\infty]{} \widetilde{X}_\infty(x)$ on Schnorr randoms x.*

Proof. With out loss of generality, the filtration $(\mathcal{F}_n)$ is one of partitions (Proposition 7.9; the same argument holds for submartingales). Then decompose $X_n = M_n + A_n$ as in the effective Doob decomposition (Proposition 8.4). Notice that $0 \le A_n \le X_n$ using induction on the formula for A_n, hence both (M_n) and (A_n) are nonnegative. Recall, also that $\|M_n\|_{L^1}$ is nondecreasing in n since (M_n) is a martingale, and $\|A_n\|_{L^1}$ is nondecreasing since (A_n) is nondecreasing. Hence

$$\sup_n \|X_n\|_{L^1} = \sup_n (\|M_n\|_{L^1} + \|A_n\|_{L^1}) = \left(\sup_n \|M_n\|_{L^1}\right) + \left(\sup_n \|A_n\|_{L^1}\right)$$

Since each term is lower semicomputable and $\sup_n \|X_n\|_{L^1}$ is computable, both $\sup_n \|M_n\|_{L^1}$ and $\sup_n \|A_n\|_{L^1}$ are computable.

Moreover, let $X_\infty, M_\infty, A_\infty$ be the limits of $(X_n), (M_n), (A_n)$, respectively. Clearly, $X_\infty = M_\infty + A_\infty$. Notice X_∞ is L^1-computable by assumption, and A_∞ is L^1-computable by the effective monotone convergence theorem (Proposition 8.2). Hence M_∞ is L^1-computable. Therefore, the convergence of (M_n) and (A_n) is effective a.e. using the effective convergence theorem for martingales (Theorem 7.10) and the effective monotone convergence theorem (Proposition 8.2). Convergence on Schnorr randoms follows similarly. □

In Section 12, I will show these theorems characterize Schnorr randomness.

These theorems are both require a lower bound and are not as general as they could be. We leave the following open problem.

Problem 8.6. Let (X_n) be a nonnegative, L^1-computable submartingale (or supermartingale) such that the L^1-bounds $\sup_n \|X_n\|_{L^1}$ and $\inf_n \|X_n\|_{L^1}$ are computable and the limit X_∞ is L^1-computable. Does (X_n) converge to X_n effectively a.e.? What if $\|X_n\|_{L^1}$ is computable? What if the rate of convergence of $\|X_n\|_{L^1}$ is computable?

9. More differentiability results

In this section we will explore some more differentiability-type results. The results follow from Sections 6 and 7. In some cases, we only sketch the details.

9.1. **Signed measures and Radon-Nikodym derivatives.** Signed measures are (informally) measures that may assign positive or negative mass to sets. A signed measure ν has a total variation norm $\|\nu\|_{TV}$ that represents the sum of both the positive and negative mass. If μ is a positive measure on $[0,1]^d$ (i.e. a measure that gives nonnegative mass to every set), then $\|\mu\|_{TV} = \mu([0,1]^d)$. We will only consider finite signed measures, i.e. where $\|\mu\|_{TV} < \infty$. The (finite) signed measures can be characterized by the Riesz representation theorem as follows. We will use this as our definition of SIGNED MEASURE.

Fact 9.1 (Riesz representation theorem, see [48])**.** *There is a one-to-one correspondence between (finite) signed measures ν on $[0,1]^d$ and bounded linear functionals $T\colon \mathcal{C}([0,1]^d) \to \mathbb{R}$, namely each T is the integration map $f \mapsto \int f\, d\nu$ of a signed measure ν. Further, $\|\nu\|_{TV}$ is equal to the operator norm $\|T\| := \sup_{f\in\mathcal{C}([0,1])} |T(f)|/\|f\|_\infty$.*

Definition 9.2. A signed measure ν is said to be COMPUTABLE if the corresponding functional T_ν is computable (i.e. $\int f\, d\nu$ is computable uniformly from f).[13]

Remark 9.3. If T_ν is positive (i.e. $T_\nu(f) \geq 0$ when $f \geq 0$), then ν is a POSITIVE MEASURE and $\|\nu\|_{TV} = T_\nu(\mathbf{1}_{[0,1]^d})$, which is computable from T_ν. A little thought reveals that the positive, computable signed measures are precisely the computable measures of Definition 2.4. Similarly, the positive, computable signed measures with norm one are precisely the computable probability measures.

Recall that λ denotes the Lebesgue measure. In this next fact, which extends Fact 4.20, ν-a.e. means outside a measurable set C such that $\nu(B) = 0$ for all measurable $B \subseteq C$.

Fact 9.4 (Radon-Nikodym theorem and decomposition, see [48])**.** *Given a signed measure ν on $[0,1]^d$, there is a λ-a.e. unique, λ-integrable function f and a ν-a.e. unique, λ-null set D such that for all measurable sets A,*

$$\nu(A) = \int_A f\, d\lambda + \nu(A \cap D).$$

The function f is the RADON-NIKODYM DERIVATIVE $d\nu/d\lambda$.

Fact 9.5 (See [48])**.** *Let ν be a signed measure on $[0,1]^d$. Then*

$$\frac{\nu(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \frac{d\nu}{d\lambda}(x) \quad (\lambda\text{-a.e. } x).$$

When ν is a nonnegative absolutely continuous measure, Fact 9.5 is equivalent to Fact 4.20, which is a version of the Lebesgue differentiation theorem (Fact 4.7). An effective version of Fact 9.5 will be given in Theorem 9.12, but first consider the "singular" case where $d\nu/d\lambda = 0$.

[13] In general, the norm $\|\nu\|_{TV}$ is only lower semicomputable, so the space of signed measures is not a computable Banach space. The representation I am using implicitly uses the weak-$*$ topology (or topology of pointwise convergence) on the space of bounded linear functionals of $\mathcal{C}([0,1])$. That is the minimal topology for which each T_ν is continuous. The unit ball in this topology is metrizable and one could alternately use this fact to classify the computable signed measures as the computable points in the corresponding computable metric space.

Lemma 9.6. *If μ is a positive measure on $[0,1]^d$ such that $d\mu/d\lambda = 0$ then*

$$\frac{\mu(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 0$$

effectively a.e. and for all λ-Schnorr randoms x.

Proof sketch. Without loss of generality we may work on $(\mathbb{T}^d, \lambda)$. By modifying the argument in Lemma 4.9, it is enough to show on λ-Schnorr randoms x that

$$\frac{\mu(I_k^t(x))}{\lambda(I_k^t(x))} \xrightarrow[k\to\infty]{} 0$$

for all $t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d$. However, $\mu(I_k^t(x))$ may not be computable, which happens when the boundary of the cube $I_k^t(x)$ has positive mass. To handle this, replace $I_k^t(x)$ by $I_k^{t+s}(x)$ (that is the dyadic cube shifted by $t+s$ that contains x) for some computable vector $s \in [0,1]^d$, such that $\mu(I_k^{t+s}(x))$ is computable for all $k \in \mathbb{N}$ and all $t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d$. One can show, by a diagonalization argument, that there is such an s.

Fix such an s. It is enough to show that

$$\frac{\mu(I_k^{t+s}(x))}{\lambda(I_k^{t+s}(x))} \xrightarrow[k\to\infty]{} 0.$$

We have $M_k^t(x) := \mu(I_k^{t+s}(x))/\lambda(I_k^{t+s}(x))$ is a nonnegative, singular, L^1-computable martingale for each $t \in \{-\frac{1}{3}, 0, \frac{1}{3}\}^d$. The statement of the lemma follows from Lemma 7.4. □

To effectivize Fact 9.5 in its full generality, I will use two decompositions, which are analogies to the martingale decompositions in Section 7.

Fact 9.7 (Lebesgue decomposition, see [12])**.** *Given a signed measure ν on $[0,1]^d$, there is a unique decomposition of ν into two signed measures ν_{ac} and ν_s (the absolutely continuous part and the singular part, respectively) such that $\nu = \nu_{ac} + \nu_s$; $\nu_{ac}(A) = \int_A (d\nu_{ac}/d\lambda)\, d\lambda$; and $d\nu_s/d\lambda = 0$. Further, $\|\nu\|_{TV} = \|\nu_{ac}\|_{TV} + \|\nu_s\|_{TV}$; if f and D are as in the Radon-Nikodym theorem (Fact 9.4), then $d\nu_{ac}/d\lambda = f$ and for all measurable A,*

$$\nu_{as}(A) = \int_A f\, d\lambda \quad \textit{and} \quad \nu_s(A) = \nu(A \cap D).$$

Recall, the notation $[f]^+ = \max\{f, 0\}$ and $[f]^- = \max\{-f, 0\}$.

Fact 9.8 (Jordan decomposition, see [12])**.** *Given a signed measure ν on $[0,1]^d$, there is a unique decomposition of ν into two signed measures ν^+ and ν^- such that $\nu = \nu^+ - \nu^-$, $\|\nu\|_{TV} = \|\nu^+\|_{TV} + \|\nu^-\|_{TV}$. Further $d\nu^+/d\lambda = [d\nu/d\lambda]^+$, $d\nu^-/d\lambda = [d\nu/d\lambda]^-$. .*

Denote $|\nu| = \nu^+ + \nu^-$. The Jordan decomposition is related to the Hahn Decomposition.

Fact 9.9 (Hahn decomposition, see [12])**.** *Given a signed measure ν on $[0,1]^d$, there is a unique partition of $[0,1]^d$ into measurable sets N, P such that $\nu^+(A) = \nu(A \cap P)$ and $\nu^- = \nu(A \cap N)$.*

Here are effective versions of the Lebesgue and Jordan decompositions.

Proposition 9.10 (Effective Lebesgue decomposition)**.** *Let ν be a computable signed measure on $[0,1]^d$ such that $d\nu/d\lambda$ is L^1-computable. Then the Lebesgue decomposition ν_{ac}, ν_s is computable. Further, $\|\nu_{ac}\|_{TV}$ and $\|\nu_s\|_{TV}$ are computable from $\|\nu\|_{TV}$.*

Proof. It is easy to see that ν_{ac}, defined by $\nu_{ac}(A) = \int_A f\,d\lambda$, is a computable signed measure where f is the L^1-computable Radon-Nikodym derivative. Then define $\nu_s := \nu - \nu_{ac}$.

Notice $\|\nu_{ac}\|_{TV} = \|f\|_{L^1}$, so $\|\nu_s\|_{TV} = \|\nu\|_{TV} - \|\nu_{ac}\|_{TV}$ is computable when $\|\nu\|_{TV}$ is computable. □

Let ν be a computable signed measure on $[0,1]^d$ such that $\|\nu\|_{TV}$ is computable. Then the Lebesgue decomposition ν^+, ν^- is computable. Further, if $d\nu/d\lambda$ is L^1-computable, then the Radon-Nikodym derivatives $d\nu^+/d\lambda = [d\nu/d\lambda]^+$ and $d\nu^-/d\lambda = [d\nu/d\lambda]^-$ are L^1-computable. (Further, P and N are effectively measurable in the probability measure $|\nu|/\|\nu\|_{TV}$.)

Proof. The proof is very similar to Proposition 7.7. Using the total variation of ν, the Riesz representation, and the fact that computable functions are dense in $\mathcal{C}([0,1]^d)$, we can effectively find a computable function $f\colon [0,1]^d \to [-1,1]$ such that $\|\nu\|_{TV} - \int f\,d\nu \le \varepsilon$ for any ε. This function approximates the Hahn decomposition $\mathbf{1}_P - \mathbf{1}_N$. Notice for any computable $\varphi\colon [0,1]^d \to [0,1]$, we have by nonnegativity,

$$\int \varphi\,d\nu^+ \ge \int \varphi\cdot[f]^+ d\nu^+ \ge \int \varphi\cdot[f]^+ d\nu^+ - \int \varphi\cdot[f]^+ d\nu^- = \int \varphi\cdot[f]^+\,d\nu$$

and similarly $\int \varphi\,d\nu^- \ge \int -\varphi\cdot[f]^- d\nu$. Then we have

$$\begin{aligned}
&\left|\int \varphi\,d\nu^+ - \int \varphi\cdot[f]^+\,d\nu\right| + \left|\int \varphi\,d\nu^- - \int -\varphi\cdot[f]^-\,d\nu\right| \\
&= \int \varphi\cdot(1-f)\,d\nu^+ - \int \varphi\cdot(1+f)\,d\nu^+ \\
&\le \int (1-f)\,d\nu^+ - \int (1+f)\,d\nu^+ \quad (-1 \le f \le 1) \\
&= \int d|\nu| - \int f\,d\nu \le \varepsilon.
\end{aligned}$$

Hence ν^+ and ν^- are computable from $\|\nu\|_{TV}$ and ν. Moreover, this shows that $\mathbf{1}_P - \mathbf{1}_N$ is L^1-computable in $|\nu|/\|\nu\|_{TV}$ and therefore P and N are effectively measurable.

If $d\nu/d\lambda$ is L^1-computable, then so are $[d\nu/d\lambda]^+$ and $[d\nu/d\lambda]^-$ (Proposition 3.1). □

Theorem 9.12. *If ν is a computable signed measure such that $\|\nu\|_{TV}$ is computable and $d\nu/d\lambda$ is L^1-computable, then*

$$\frac{\nu(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \frac{d\nu}{d\lambda} \quad (\textit{effectively a.e.})$$

and

$$\frac{\nu(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \widetilde{\frac{d\nu}{d\lambda}}(x) \quad (\textit{on Schnorr random } x).$$

Proof. By the effective decompositions (Propositions 9.10 and 9.11) decompose ν into $\nu = \nu_{ac}^+ + \nu_{ac}^- + \nu_s^+ + \nu_s^-$ (the order of the decompositions does not matter). Then

$$\frac{\nu_s^+(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 0 \quad \text{and} \quad \frac{\nu_s^-(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 0 \quad (\lambda\text{-a.e.})$$

and

$$\frac{\nu_{ac}^+(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \left[\frac{d\nu}{d\lambda}\right]^+ \quad \text{and} \quad \frac{\nu_{ac}^-(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} \left[\frac{d\nu}{d\lambda}\right]^- \quad (\lambda\text{-a.e.}).$$

Apply Lemma 9.6 and Corollary 4.21 respectively. □

Remark 9.13. An alternate proof would be to prove the following stronger version of Fact 9.5. Since signed measures form a vector space, denote $a \cdot \nu$ for the signed measure given by scaling ν by $a \in \mathbb{R}$. Also by $|\nu|$ we mean the positive measure $\nu^+ + \nu^-$. One can show that

$$\frac{\left|\nu - \frac{d\nu}{d\lambda}(x)\cdot\lambda\right|(B(x,r))}{\lambda(B(x,r))} \xrightarrow[r\to 0]{} 0 \quad (\lambda\text{-a.e.}).$$

We could decompose this effectively into $\left|\nu - \frac{d\nu}{d\lambda}(x)\cdot\lambda\right| = \left|\nu_{ac} - \frac{d\nu}{d\lambda}(x)\cdot\lambda\right| + \nu_s^+ + \nu_s^-$. The first term can be handled by the same proof as the effective Lebesgue differentiation theorem (Theorem 4.10), and the last terms can be handled using Lemma 9.6.

In Section 12, I give some examples which show the theorems of this section characterize Schnorr randomness.

9.2. **Functions of bounded variation.** A function $f\colon [0,1] \to \mathbb{R}$ is of BOUNDED VARIATION if there is some bound b such that for all finite sequences $0 = a_0 \le a_1 \le \ldots \le a_k = 1$ we have

$$\sum_{i<k} |f(a_{i+1}) - f(a_k)| \le b.$$

The smallest such b is the TOTAL VARIATION (NORM) of f and is written $V(f)$. We have the following fact.

Fact 9.14 (See [12])**.** *Every function on* $[0,1]$ *of bounded variation is differentiable almost-everywhere, and the derivative is integrable.*

Since every absolutely continuous function is of bounded variation, Fact 9.14 implies Fact 4.17.

There are a number of approaches to represent functions of bounded variation and their differentiability using computable analysis. The simplest approach is to only consider computable functions of bounded variation [7]. However, not all bounded variation functions are continuous.

The most general approach is to consider functions defined on a computably enumerable, countable, dense subset of $[0,1]$. Then instead of differentiability we will consider pseudo-differentiability. This approach has been used in both constructive mathematics [4, 9] and computable analysis [31, 7, 28].

Definition 9.15. Let $\{a_n\}_{n\in\mathbb{N}}$ be a uniformly computable dense sequence of distinct reals in $[0,1]$ with $a_0 = 0$ and $a_1 = 1$. Let $f\colon \{a_n\}_{n\in\mathbb{N}} \to \mathbb{R}$ be a function. We say f is COMPUTABLE if $f(a_n)$ is uniformly computable from n. Define the

TOTAL VARIATION of f as follows where the supremum is over finite sequences $a_{n_0} < \ldots < a_{n_k}$ in $\{a_n\}_{n\in\mathbb{N}}$.

$$V(f) = \sup_{a_{n_0} < \ldots < a_{n_k}} \sum_{i<k} |f(a_{i+1}) - f(a_i)|$$

Let $x \in (0,1)$. Then define the PSEUDO-DERIVATIVE of f at x as

$$\check{D}f\,(x) = \lim_{|b-a|\to 0} \frac{f(b) - f(a)}{b-a} \tag{9.1}$$

where the limit is over all $a, b \in \{a_n\}_{n\in\mathbb{N}}$ such that $a < x < b$. Say f is PSEUDO-DIFFERENTIABLE at x if the limit converges.

Proposition 9.16. *All functions f as in Definition 9.15 such that $V(f) < \infty$ are pseudo-differentiable for a.e. $x \in (0,1)$, and the derivative is an integrable function.*

Proof. Just extend f to a total bounded variation function g by setting

$$g(x) = \begin{cases} f(x) & x \in \{a_n\}_{n\in\mathbb{N}} \\ \lim\limits_{a\to x^-} f(x) \quad (a \in \{a_n\}_{n\in\mathbb{N}}) & \text{otherwise} \end{cases}.$$

(This limit exists since $V(f) < \infty$.) Then apply Fact 9.14. □

Consider these examples of functions of bounded variation.

Example 9.17. Assume $g\colon [0,1] \to \mathbb{R}$ is a computable (and hence continuous) function of bounded variation. Assume $\{a_n\}_{n\in\mathbb{N}}$ is as in Definition 9.15. Let $f = g \restriction \{a_n\}_{n\in\mathbb{N}}$ (i.e. the restriction of g to $\{a_n\}_{n\in\mathbb{N}}$). Then $f\colon \{a_n\}_{n\in\mathbb{N}} \to \mathbb{R}$ is computable (as in Definition 9.15) and of bounded variation. Moreover, $V(f) = V(g)$ and the derivative $\frac{d}{dx}g$ is equal to $\check{D}f$ for all $x \in (0,1)$.

Conversely, assume $f\colon \{a_n\}_{n\in\mathbb{N}} \to \mathbb{R}$ is a computable function of bounded variation with a continuous extension g and that $V(f)$ is computable. Assume that f can be extended to a continuous function $g\colon [0,1] \to \mathbb{R}$ (i.e. $f = g \restriction \{a_n\}_{n\in\mathbb{N}}$). Then g is a computable function (uniformly computable from $V(f)$ and f). (The modulus of continuity is computable from the variation. See Lu and Weihrauch [31].)

Remark 9.18. One could also consider L^1-computable functions of bounded variation, as well as functions of the form $f(x) = \nu([0,x])$ for some computable signed measure ν. However, it requires some care to work with these types of functions and I will not do so here.

Theorem 9.19. *Let $f\colon \{a_n\}_{n\in\mathbb{N}} \to \mathbb{R}$ be computable (as in Definition 9.15). Assume $V(f)$ is computable (and hence finite) and the derivative $F := \check{D}f$ is L^1-computable. Then f is pseudo-differentiable on all Schnorr randoms. Further*

$$\frac{f(b) - f(a)}{b-a} \xrightarrow[(b-a)\to 0]{} F \quad (a, b \in \{a_n\}_{n\in\mathbb{N}},\ a < x < b).$$

converges effectively a.e., and $\check{D}f\,(x) = \widetilde{F}(x)$ on Schnorr randoms x.

Proof sketch. Follow the arguments of Section 9.1. Replace the norm $\|\nu\|_{TV}$ with the total variation norm $V(f)$; positive measures with increasing functions; the Radon-Nikodym derivative with the pseudo-derivative; absolutely continuous measures with absolutely continuous functions; singular measures with functions of derivative zero; and the Lebesgue/Jordan decompositions with their corresponding versions for functions of bounded variation. See [28] for an effective version of the Jordan decomposition for functions of bounded variation. □

Corollary 9.20. *Let $g\colon [0,1] \to \mathbb{R}$ be a computable function of bounded variation. Assume $V(g)$ is computable and the derivative $G := \frac{d}{dx} g$ is L^1-computable. Then g is differentiable on all Schnorr randoms. Further the derivative converges effectively a.e. to G, and $\frac{d}{dx} g|_{x=z} = \widetilde{G}(z)$ on Schnorr randoms z.*

Proof. Use Theorem 9.19 and Example 9.17. □

In Section 12, I give some examples which show the theorems of this section characterize Schnorr randomness.

10. The ergodic theorem

There has been a great deal of interest in the effectivity of the ergodic theorems, both in terms of rates of convergence and randomness. In this section, I briefly summarize the results for Schnorr randomness.

Fact 10.1 (Ergodic theorems, see [50]). *Let $(\mathbb{X}, \mu)$ be a probability space. Let*

$$A_n f = \frac{1}{n^d} \sum_{i_0 < n} \cdots \sum_{i_{d-1} < n} f \circ T_0^{i_0} \circ \cdots \circ T_{d-1}^{i_{d-1}}$$

where f is integrable and $T_0, \ldots, T_{d-1}$ *are commuting, measure-preserving maps $T_j\colon (\mathbb{X}, \mu) \to (\mathbb{X}, \mu)$. Let $\mathcal{I}nv(T_0, \ldots, T_{d-1})$ be the σ-algebra of sets invariant under $T_0, \ldots, T_{d-1}$. Then $A_n f \to f^* := \mathbb{E}[f \mid \mathcal{I}nv(T_0, \ldots, T_{d-1})]$ a.e. and in the L^1-norm. If f is L^2, then convergence is in the L^2-norm as well. If $\mathcal{I}nv(T_0, \ldots, T_{d-1})$ is* TRIVIAL *($\mathbb{E}[f \mid \mathcal{I}nv(T_0, \ldots, T_{d-1})] = \int f \, d\mu$ for all f), then the system is said to be* ERGODIC *and $A_n f \to f^* = \int f \, d\mu$.*

This next theorem is a combination of results from a number of authors. I use techniques from this paper to fill in a few gaps not explicitly in the literature.

Theorem 10.2 (Effective ergodic theorems). *Let $(\mathbb{X}, \mu)$ be a computable probability space. Let f* be integrable and $T_0, \ldots, T_{d-1}$ *be commuting measure-preserving maps $T_j\colon (\mathbb{X}, \mu) \to (\mathbb{X}, \mu)$ (not necessarily effectively measurable).*

(1) (a) If $A_n f$ is L^1-computable uniformly in n and the limit f^* is L^1-computable, then $A_n f \to f^*$ both effectively in L^1 and effectively a.e. Hence $\widetilde{A_n f}(z) \to \widetilde{f^*}(z)$ on Schnorr randoms z.
    (b) If $A_n f$ is L^2-computable uniformly in n and $\|f^*\|_{L^2}$ is computable, then f^* is L^2-computable and $A_n f \to f^*$ effectively in the L^2-norm, the L^1-norm, and effectively a.e. Hence $\widetilde{A_n f}(z) \to \widetilde{f^*}(z)$ on Schnorr randoms z.
(2) In particular, assume $T_0, \ldots, T_{d-1}$ are effectively measurable and f is L^1-computable. Then $A_n f$ is L^1-computable uniformly in n and

$$\widetilde{A_n f} = \frac{1}{n^d} \sum_{i_0 < n} \cdots \sum_{i_{d-1} < n} \widetilde{f} \circ \widetilde{T}_0^{i_0} \circ \cdots \circ \widetilde{T}_{d-1}^{i_{d-1}}.$$

Further assume the system is ergodic, or more generally, $\mathbb{E}[\,\cdot \mid \mathcal{I}nv(T_0,\dots,T_{d-1})]$ is a computable operator on $L^1 \to L^1$.[15] Then f^* is L^1-computable and the results in (1) hold. The same holds of L^2 in place of L^1.

Proof. (1)(a): The first sentence follows from Avigad, Gerhardy, Towsner [2] and Galatolo, Hoyrup, Rojas [21] in the case that there is a single measure preserving map T, the system is ergodic, f is L^1-computable, and T is effectively measurable[16]. The proof of Galatolo et al. can be generalized in the following ways:

- In the proof, it is only necessary to know $f, f\circ T, f\circ T^2, \dots$. Therefore, one may just assume that $A_n f$ is L^1-computable uniformly in n (without explicit knowledge of T).
- The proof also holds in the non-ergodic case by replacing $\int f\,d\mu$ with the L^1-computable limit f^* [personal communication with Hoyrup and Rojas].
- The proof also holds for multiple $T_0,\dots,T_{d-1}$, by a straight-forward modification to multiple dimensions.

The part about Schnorr randomness follows from Lemma 3.19. (For certain ergodic systems this was proved in Gács, Hoyrup, Rojas [20, 39] and in Pathak, Rojas, Simpson [39].)

(1)(b): The first sentence follows from Avigad, Gerhardy, Towsner [2] in the case that there is a single measure preserving map T, the system is ergodic, f is L^1-computable, and T is effectively measurable.

In the general case of multiple $T_0,\dots,T_{d-1}$, assume we know the value of $\|f^*\|_{L^2}$. First we will show that f^* is L^2-computable. Since $A_k f \xrightarrow{L^2} f^*$, for $\varepsilon > 0$, we can effectively choose k such that $|\|A_k f\|_{L^2} - \|f^*\|_{L^2}| \le \varepsilon$. Let $g = A_k f$. It holds that

$$\lim_n A_n g = \lim_n A_{nk} f = f^*.$$

Therefore, $f^* = \mathbb{E}[g \mid \mathcal{I}nv(T_0,\dots,T_{d-1})]$ is a projection of g in the L^2-norm. By the Pythagorean theorem,

$$\|g - f^*\|_{L^2} = \|g\|_{L^2} - \|f^*\|_{L^2} \le \varepsilon.$$

Hence f^* is L^2-computable.

The rest follows from the proof of Galatolo, Hoyrup, Rojas [21]; besides the generalizations listed in part (a), their proof also holds for L^2 in place of L^1.

(2): This easily follows from the results in Section 3.4. □

This next corollary is a generalization of Kučera's theorem. A one dimensional, Martin-Löf random version of this next corollary can be found in Bienvenu, Day, Hoyrup, Mezhirov, and Shen [3]. My proof is basically the same as theirs.

---

[15] Hoyrup [22] mentions that for $\mathbb{E}[\,\cdot \mid \mathcal{I}nv(T_0,\dots,T_{d-1})]$ is a computable operator on $L^1 \to L^1$ (or $L^2 \to L^2$) if and only if the ergodic decomposition $x \mapsto \mu_x$ is effectively measurable. (Actually, Hoyrup says that $x \mapsto \mu_x$ is "layerwise computable", but this can be replaced with "effectively measurable". Also while his proof is only for the shift map, the proof extends to effectively measurable $T_0,\dots,T_{d-1}$.)

[16] For Avigad et al. [2] the measure preserving map T is "computable" if the corresponding operator $f \mapsto f\circ T$ is a computable from L^2 to L^2. By Proposition 3.31, this is the same as effectively measurable.

While the Galatolo et al. result is for a.e. computable T, the proof works for effectively measurable T by the fact that if f is L^1- or L^2-computable, then so is $f\circ T$ (uniformly from f and T) (Proposition 3.30).

Corollary 10.3. *Assume $T_0, \dots, T_{d-1}$ are effectively measurable, commuting, measure preserving maps on a computable probability space $(\mathbb{X}, \mu)$, the system is ergodic, and A is an effectively measurable set. Then for all Schnorr randoms x, there are infinitely-many tuples $(k_0, \dots, k_{d-1})$ such that $\widetilde{T}_0^{k_0} \cdots \widetilde{T}_{d-1}^{k_d}(x) \in \widetilde{A}$.*

Proof. In the case of a single map T, by Theorem 10.2, $\frac{1}{n}\sum_{k<n} \mathbf{1}_{\widetilde{A}}(\widetilde{T}^k(x)) \to \mu(A) > 0$. Hence, there are infinitely many k such that $\widetilde{T}^k(x) \in \widetilde{A}$. The multiple map version is the same. □

Corollary 10.4 (Kučera's theorem for Schnorr randomness)**.** *If $C \subseteq 2^{\mathbb{N}}$ is a closed set of positive measure and $x \in 2^{\mathbb{N}}$ is Schnorr random, then some tail of x is in C.*

Proof. In the previous result, let T be the left shift map ($T(0x) = T(1x) = x$) and let $A = C$. □

11. Backwards martingales and their applications

In this section, I discuss backwards martingales. Unlike "forward martingales" and ergodic averages, backwards martingales have not before been used before in algorithmic randomness. However, like forward martingales and ergodic averages, they are a powerful tool.

The definition of martingale can be extended to any linearly ordered (or partially ordered) index set I. Namely, $(\mathcal{F}_i)_{i\in I}$ is a FILTRATION if $\mathcal{F}_i \subseteq \mathcal{F}_j$ for any $i \le j$, and $(M_i)_{i\in I}$ is a MARTINGALE adapted to $(\mathcal{F}_i)_{i\in I}$ if each M_i is $\mathcal{F}_i$-measurable and $\mathbb{E}[M_j \mid \mathcal{F}_i] = M_i$ for any $i \le j$. If the index set I is the nonpositive integers, then we say M is a BACKWARDS (or REVERSE) MARTINGALE, often written (M_{-k}) to denote that the martingale is backwards. As opposed to "forward martingales", backwards martingales always converge a.e. and in the L^1-norm.

Fact 11.1 (See [13])**.** *Let (M_{-k}) be a backwards martingale adapted to the filtration $(\mathcal{F}_{-k})$ and let $\mathcal{F}_{-\infty} = \bigcap_k \mathcal{F}_{-k}$. Then $M_{-k} \to M_{-\infty} = \mathbb{E}[M_{-0} \mid \mathcal{F}_{-\infty}]$ both in L^1 and a.e.*

We have the following analog of Theorem 6.5 and Corollary 6.8.

Theorem 11.2. *Fix a computable probability space $(\mathbb{X}, \mu)$.*

1. *If (M_{-k}) is an L^1-computable backwards martingale, and the limit $M_{-\infty}$ is L^1-computable, then $M_{-k} \to M_{-\infty}$ converges effectively in the L^1-norm and effectively a.e. Hence, $\widetilde{M}_{-k}(z) \to \widetilde{M}_{-\infty}(z)$ on Schnorr randoms z.*
2. *If (M_{-k}) is an L^2-computable backwards martingale, and $\|M_{-\infty}\|_{L^2} = \inf_k \|M_{-k}\|_{L^2}$ is computable, then $M_{-k} \to M_{-\infty}$ converges effectively in the L^2-norm and effectively a.e. Hence, $M_{-\infty}$ is L^2-computable, and $\widetilde{M}_{-k}(z) \to \widetilde{M}_{-\infty}(z)$ on Schnorr randoms z.*

Proof. In the L^1 case, the proof is basically the same as that of Lemma 6.5. Since the limit is known, there is an effectively convergent subsequence. Further, since the inequalities (6.1) and (6.2) only apply to finite intervals of indices, they remain true for backwards martingales.

In the L^2-case the argument resembles Corollary 6.8. For any $k \in \mathbb{N}$, $M_{-\infty} = \mathbb{E}[M_{-k} \mid \mathcal{F}_{-\infty}]$. By the Pythagorean theorem,

$$\|M_{-k} - M_{-\infty}\|_{L^2} = \|M_{-k}\|_{L^2} - \|M_{-\infty}\|_{L^2}.$$

So $M_{-k} \to M_{-\infty}$ effectively in L^2. The rest follows from the L^1-case. □

Remark 11.3. Theorem 11.2 is analogous to both the effective ergodic theorem (Theorem 10.2) and the effective Lévy 0-1 law (Theorem 6.3 and Corollary 6.8), as seen in Table 3. All three theorems concern the convergence of averages of some function f. The effective versions are analogous for each theorem.

Moreover, each theorem takes place on a *structured* probability space. A (COMPUTABLE) MEASURE PRESERVING SYSTEM $(\mathbb{X}, \mu, T)$ is a (computable) probability space $(\mathbb{X}, \mu)$ with a(n) (effectively measurable) measure-preserving action T. A (COMPUTABLE) FILTERED PROBABILITY SPACE $(\mathbb{X}, \mu, (\mathcal{F}_n))$ or $(\mathbb{X}, \mu, (\mathcal{F}_{-n}))$ is a (computable) probability space $(\mathbb{X}, \mu)$ with a (computable) filtration $(\mathcal{F}_n)$ or $(\mathcal{F}_{-n})$ (see Remark 7.6). Each such space has a σ-algebra which determines the limit. A sufficient condition for effective convergence is that both the structured probability space and the σ-algebra are computable. (In this case say a σ-algebra $\mathcal{G}$ is COMPUTABLE if $f \mapsto \mathbb{E}[f \mid \mathcal{G}]$ is a computable operator $L^1 \to L^1$).

	Ergodic averages	Backwards martingales	Lévy 0-1 law
Space	$(\mathbb{X}, \mu, T)$	$(\mathbb{X}, \mu, (\mathcal{F}_{-n}))$	$(\mathbb{X}, \mu, (\mathcal{F}_n))$
Averages	$\frac{1}{n}\sum_{k<n} f \circ T^k$	$\mathbb{E}[f \mid \mathcal{F}_{-n}]$	$\mathbb{E}[f \mid \mathcal{F}_n]$
Limit σ-algebra	$\mathcal{I}nv(T)$	$\mathcal{F}_{-\infty}$	$\mathcal{F}_\infty$
Limit	$\mathbb{E}[f \mid \mathcal{I}nv(T)]$	$\mathbb{E}[f \mid \mathcal{F}_{-\infty}]$	$\mathbb{E}[f \mid \mathcal{F}_\infty]$
"Nicest" system	ergodic system	$\mathcal{F}_{-\infty}$ is trivial	$\mathcal{F}_\infty$ is Borel σ-alg.

TABLE 3. Comparison of three convergence theorems.

Backwards martingales are quite useful. I will give three applications. The first application is a variation of Kučera's theorem for Schnorr randomness (Corollary 10.4). However, this version does not follow directly from the ergodic theoretic Corollary 10.3.

Corollary 11.4. *On $(2^{\mathbb{N}}, \lambda)$, assume A is effectively measurable and $\lambda(A) > 0$. Then for all Schnorr random $x \in 2^{\mathbb{N}}$, there is some Schnorr random $y \in \widetilde{A}$ such that y is a permutation of* finitely-many *bits of x. In particular, if A is an effectively closed set of computable positive measure, then $y \in A$.*

Proof. Let $\mathcal{F}_{-n}$ be the sigma-algebra of sets invariant under permutations of the first n bits. Notice that $\mathcal{F}_{-0} \supseteq \mathcal{F}_{-1} \supseteq \ldots$ and that $\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_{-n}] = \frac{1}{n!}\sum_T \mathbf{1}_A \circ T$ where $T\colon 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ ranges over all permutations of bits which permute only the first n bits. Hence $\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_{-n}]$ is an L^1-computable backwards martingale. Further, $\mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_{-n}$ is trivial (i.e. all sets in $\mathcal{F}_{-n}$ are measure one or measure zero). Let x be Schnorr random. By Theorem 11.2,

$$\frac{1}{n!}\sum_T \mathbf{1}_{\widetilde{A}}(T(x)) = \widetilde{\mathbb{E}}[\mathbf{1}_A \mid \mathcal{F}_{-n}](x) \xrightarrow[n\to\infty]{} \widetilde{\mathbb{E}}[\mathbf{1}_A \mid \mathcal{F}_{-\infty}](x) = \lambda(A) > 0.$$

Therefore, $\frac{1}{n!}\sum_T \mathbf{1}_{\widetilde{A}}(T(x)) > 0$ for some n, and, moreover, $T(x) \in \widetilde{A}$ for some T which permutes the first n bits. □

Corollary 11.4 is an effective version of the Hewitt-Savage 0-1 law, any set invariant under exchangeability of bits (or finite permutations of bits) is measure 0 or 1. Indeed, let S be the smallest set containing $\widetilde{A}$ closed under exchangeability.

Either S is measure 0 and contains no Schnorr randoms, or S is measure 1 and contains every Schnorr random.

Corollary 11.4 can be expressed as a special case of the following generalization of Corollary 10.3 to infinitely many maps $T_0, T_1, \ldots$

Corollary 11.5. *Assume $T_0, T_1, \ldots$ is a computable sequence of effectively measurable, commuting, measure preserving maps on a computable probability space $(\mathbb{X}, \mu)$. Assume $\mathcal{I}nv(T_0, \ldots, T_{d-1})$ is computable for all d (as in Remark 11.3). Assume the system is ergodic ($\mathcal{I}nv(T_0, T_1, \ldots)$ is trivial), and A is an effectively measurable set. Then for all Schnorr randoms x, there exists infinitely-many d and infinitely-many tuples $(k_0, \ldots, k_{d-1})$ of length d such that $\widetilde{T}_0^{k_0} \cdots \widetilde{T}_{d-1}^{k_{d-1}}(x) \in \widetilde{A}$.*

Proof. Notice that

$$\mathcal{I}nv(T_0, T_1 \ldots) = \bigcap_d \mathcal{I}nv(T_0, \ldots, T_{d-1}).$$

Fix a Schnorr random x. Since $\mathcal{I}nv(T_0, T_1 \ldots)$ is trivial, we have by Theorem 11.2 that

$$\widetilde{\mathbb{E}}[\mathbf{1}_A \mid \mathcal{I}nv(T_0, \ldots, T_{d-1})](x) \xrightarrow[d\to\infty]{} \mu(A) > 0.$$

There must exist infinitely many d such that

$$\widetilde{\mathbb{E}}[\mathbf{1}_A \mid \mathcal{I}nv(T_0, \ldots, T_{d-1})](x) \geq 0.$$

Fix such a d. By Theorem 10.2,

$$\frac{1}{n^d} \sum_{k_0<n} \cdots \sum_{k_{d-1}<n} \widetilde{f} \circ \widetilde{T}_0^{k_0} \circ \cdots \circ \widetilde{T}_{d-1}^{k_{d-1}} \xrightarrow[n\to\infty]{} \widetilde{\mathbb{E}}[\mathbf{1}_A \mid \mathcal{I}nv(T_0, \ldots, T_{d-1})](x) \geq 0.$$

Hence, there are infinitely many tuples $(k_0, \ldots, k_{d-1})$ such that $\widetilde{T}_0^{k_0} \cdots \widetilde{T}_{d-1}^{k_{d-1}}(x) \in \widetilde{A}$. □

Before giving the next two examples, recall the probabilistic mindset. For the remainder of this section we fix a computable probability space $(\Omega, \mathbb{P})$ as our sample space. We are not concerned with what this space is. A measurable function $X\colon (\Omega, \mathbb{P}) \to \mathbb{R}$ is called a RANDOM VARIABLE. Recall its DISTRIBUTION (or PUSH-FORWARD PROBABILITY MEASURE) $\mathbb{P}_X$ is a probability measure on $\mathbb{R}$ defined by

$$\int \varphi \, d\mathbb{P}_X = \mathbb{E}[\varphi(X)] \tag{11.1}$$

for any bounded continuous $\varphi\colon \mathbb{R} \to \mathbb{R}$. (This equation then extends to all $\varphi \in L^1(\mathbb{R}, \mathbb{P}_X)$.) Given a sequence of random variables $X = (X_i)_{i\in\mathbb{N}}$, the joint distribution of X is the probability measure $\mathbb{P}_X$ on $\mathbb{R}^{\mathbb{N}}$ given by the equation

$$\int \varphi \, d\mathbb{P}_X = \mathbb{E}[\varphi(X_0, \ldots, X_{d-1})]$$

for any bounded continuous $\varphi\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ depending only on the first d coordinates.. *In other words, one may just think of $(\Omega, \mathbb{P})$ as $(\mathbb{R}^{\mathbb{N}}, \mathbb{P}_X)$. Then X_i just becomes the ith coordinate of $\mathbb{R}^{\mathbb{N}}$.*[17] A sequence $X = (X_i)$ is INDEPENDENT AND IDENTICALLY DISTRIBUTED (I.I.D.) if the joint distribution $\mathbb{P}_X$ is the product measure $\mu^{\mathbb{N}} :=$

[17] This intuition also holds in computable probability. A probability measure μ on $\mathbb{R}^{\mathbb{N}}$ is computable if and only if there is a sequence $X = (X_i)$ of uniformly effectively measurable (even a.e. computable) random variables on $(2^{\mathbb{N}}, \lambda)$ such that $\mu = \mathbb{P}_X$ [44, 25].

$\mu \otimes \mu \otimes \cdots$ where $\mu = \mathbb{P}_{X_0}$. Equivalently, for all bounded continuous functions $\varphi\colon \mathbb{R}^d \to \mathbb{R}$,

$$\mathbb{E}[\varphi(X_0, \dots, X_{d-1})] = \int \varphi \, d\mu^d. \tag{11.2}$$

Fact 11.6 (Strong law of large numbers, see [13])**.** *Let (X_i) be a sequence of i.i.d. integrable random variables with partial sums $S_k = \sum_{i=0}^{k-1} X_i$. Then $S_k/k \to \mathbb{E}[X_0]$ a.e. (and in the L^1-norm).*

Corollary 11.7 (Effective strong law of large numbers)**.** *Let (X_i) be sequence of i.i.d. L^1-computable random variables with partial sums $S_k = \sum_{i=0}^{k-1} X_i$. Then $S_k/k \to \mathbb{E}[X_0]$ effectively a.e. and effectively in the L^1-norm. Hence, $\widetilde{S}_k(\omega)/k \to \mathbb{E}[X_0]$ on Schnorr randoms ω.*

Proof. It is known that $M_{-k} := S_k/k$ is a backwards martingale adapted to the filtration $\mathcal{F}_{-k} = \sigma(S_k, S_{k+1}, \dots) = \sigma(S_k, X_{k+1}, X_{k+2}, \dots)$ [13, Example 5.61]. ($\mathcal{F}_{-k}$ is the σ-algebra of sets invariant under permuting $X_0, \dots, X_{k-1}$.) Clearly (M_{-k}) is L^1-computable. By the strong law of large numbers, we know (M_{-k}) converges to $\mathbb{E}[X_0]$, and the expectation is a computable real number. Hence by Theorem 11.2 $S_k/k \to \mathbb{E}[X_0]$ effectively in the L^1-norm and effectively a.e. Hence, by Lemma 3.19, $\widetilde{S}_k/k \to \mathbb{E}[X_0]$ on Schnorr randoms. □

Remark 11.8. Taking $(\Omega, P) = (2^{\mathbb{N}}, \lambda)$ and $X_i(x) = x(i)$, the previous corollary implies that all Schnorr randoms z have an equal density of 1s and 0s—a fact which is well known. In Section 12, I use an extension of this fact to show that the strong law of large numbers characterizes Schnorr randomness. Corollary 11.7 could also be proved using the effective ergodic theorem (Theorem 10.2). Indeed, this is another similarity between backwards martingales and ergodic averages.

Now, I consider de Finetti's theorem. A sequence of random variables $X = (X_i)$ is EXCHANGEABLE if the joint distribution of $(X_0, \dots, X_{d-1})$ is the same as that of $(X_{\sigma(0)}, \dots, X_{\sigma(d-1)})$ for any permutation σ. In other words, the joint distribution $\mathbb{P}_X$ is unchanged by permuting coordinates. De Finetti's theorem says that every exchangeable sequence is a convex combination of i.i.d. sequences.

Fact 11.9 (de Finetti's theorem, see [13, 18])**.** *Every exchangeable sequence of random variables $X = (X_i)$ is i.i.d. conditioned on some random measure μ. That is there is a $(\Omega, \mathbb{P})$-measurable random map $\mu\colon \omega \mapsto \mu_\omega$ where μ_ω is a probability measure on $\mathbb{R}$, such that for any bounded continuous function $\varphi\colon \mathbb{R}^d \to \mathbb{R}$,*

$$\mathbb{E}\left[\varphi(X_0, \dots, X_{d-1}) \mid \mu\right](\omega) = \int \varphi \, d\mu_\omega^d \quad (\mathbb{P}\text{-}a.e.\ \omega) \tag{11.3}$$

where $\mathbb{E}[\,\cdot \mid \mu]$ is conditioning on the least σ-algebra for which the map $\omega \mapsto \mu_\omega$ is measurable. This random measure $\omega \mapsto \mu_\omega$, called the DIRECTING MEASURE, *is $\mathbb{P}$-a.s. unique.*

Moreover, the following a.e. convergence theorems hold. For every $f \in L^1(\mathbb{R}^d, \mathbb{P}_{X_0})$,

$$\frac{1}{k} \sum_{i=0}^{k-1} f(X_i(\omega)) \to \mathbb{E}\left[f(X_0) \mid \mu\right](\omega) \quad (\mathbb{P}\text{-}a.e.\, \omega) \tag{11.4}$$

This can be extended to all $f \in L^1(\mathbb{R}^d, \mathbb{P}_{X_0,\dots,X_{d-1}})$ as follows.
(11.5)

$$A_k(f) = \frac{1}{k!/(k-d)!}\sum_\sigma f(X_{\sigma(0)},\dots,X_{\sigma(d-1)}) \to \mathbb{E}\left[f(X_0,\dots,X_{d-1}) \mid \mu\right](\omega) \quad (\mathbb{P}\text{-}a.e.\,\omega)$$

where the average is over all $\frac{k!}{(k-\ell)!}$ many injections $\sigma\colon \{0,\dots,d-1\} \to \{0,\dots,k-1\}$.

First, note the connection with the strong law of large numbers. If $X = (X_i)$ is i.i.d., then $\omega \mapsto \mu_\omega$ is constant. Therefore, the strong law of large numbers follows from equation (11.4) using $f(x) = x$. Second, note the similarity between equations (11.5) and the ergodic theorem.

Theorem 11.10 (Computable de Finetti's theorem (Freer, Roy [18]))**.** *If $X = (X_i)$ is a sequence of exchangeable random variables with computable distribution $\mathbb{P}_X$, then the distribution $\mathbb{P}_\mu$ of the directing measure μ is computable from $\mathbb{P}_X$ and vice versa.*

We now can show this effective a.e. convergence theorem.

Corollary 11.11. *Let $X = (X_i)$ be a uniformly computable sequence of effectively measurable, exchangeable random variables with directing measure μ. Then for all $f \in L^1_{comp}(\mathbb{R}^d, \mathbb{P}_{X_0,\dots,X_{d-1}})$,*

$$A_k(f) \to \mathbb{E}\left[f(X_0,\dots,X_{d-1}) \mid \mu\right]$$

both effectively a.e. and effectively in L^1. Hence, for all Schnorr random ω,

$$\widetilde{A_k}(\widetilde{f})(\omega) \to \widetilde{\mathbb{E}}\left[f(X_0,\dots,X_{d-1}) \mid \mu\right](\omega)$$

where

$$\widetilde{A_k}(f) := \frac{1}{k!/(k-d)!}\sum_\sigma f(\widetilde{X}_{\sigma(0)},\dots,\widetilde{X}_{\sigma(d-1)}).$$

Proof. Since $X = (X_i)$ is uniformly effectively measurable, the distribution $\mathbb{P}_X$ is computable (Propositions 3.26 and 3.28). Then by Theorem 11.10, the distribution $\mathbb{P}_\mu$ is also computable. Also, for any $f \in L^1(\mathbb{R}^d, \mathbb{P}_{X_0,\dots,X_{d-1}})$, we have $M_{-k} = A_k(f)$ is a backwards martingale [13, Chapter 5].

Let $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be a bounded computable function. Then $M_{-k} = A_k(\varphi)$ is an L^2-computable backwards martingale. By Theorem 11.2, it is enough to compute the (square of the) L^2-norm of the limit

$$\begin{aligned}\|\mathbb{E}\left[\varphi(X_0,\dots,X_{d-1}) \mid \mu\right]\|^2_{L^2(\Omega,P)} &\overset{\text{eq. (11.3)}}{=} \int \left(\int \varphi\, d\mu_\omega\right)^2 dP(\omega)\\ &\overset{\text{eq. (11.1)}}{=} \int \left(\int \varphi\, d\nu\right)^2 dP_\mu(\nu).\end{aligned}$$

This last integral is computable since $\nu \mapsto \left(\int \varphi\, d\nu\right)^2$ is a computable map.

Hence we have proved the result for bounded computable $\varphi\colon \mathbb{R}^d \to \mathbb{R}$. For $f \in L^1_{comp}(\mathbb{R}^d, P_{X_0,\dots,X_{d-1}})$, take some φ which approximates f in the L^1-norm, then

$$\begin{aligned}&\|\mathbb{E}\left[\varphi(X_0,\ldots,X_{d-1}) \mid \mu\right] - \mathbb{E}\left[f(X_0,\ldots,X_{d-1}) \mid \mu\right]\|_{L^1(\Omega,P)}\\ &\quad= \|\mathbb{E}\left[(\varphi - f)(X_0,\ldots,X_{d-1}) \mid \mu\right]\|_{L^1(\Omega,P)}\\ &\quad\le \|(\varphi - f)(X_0,\ldots,X_{d-1})\|_{L^1(\Omega,P)}\\ &\quad= \|\varphi - f\|_{L^1(\mathbb{R}^d,P_{X_0,\ldots,X_{d-1}})}.\end{aligned}$$

Since the last term is uniformly computable, we can compute the limit $\mathbb{E}\left[f(X_0,\ldots,X_{d-1}) \mid \mu\right]$ in the L^1-norm. By Theorem 11.2, this completes the proof. □

Example 11.12 (Pólya's urn)**.** Consider an urn with one black ball and one red ball. At each stage k we take a ball from the urn, then return that ball to the urn along with another ball of the same color. Let X_k be the color of the kth ball drawn (0 for red, 1 for black). It turns out the sequence of random variables (X_k) is exchangeable. Let $S_k = \sum_{i=0}^{k-1} X_i$. By de Finetti's theorem the average S_k/k converges a.s., meaning that the ratio of red balls to black balls approaches a limit a.s. Now suppose, Pólya's Urn is modeled on a computer such that the random variables (X_k) are a.e. computable with respect to a uniformly distributed random real $x \in [0,1]$. Then if x is Schnorr random, the simulation of Pólya's urn is guaranteed to converge to a fixed ratio of red and black balls.

Remark 11.13. There are other computable aspects of the ergodic theorem that could be explored for de Finetti's theorem. For one, the map $\omega \mapsto \mu_\omega$ is a form of ergodic decomposition. Hoyrup [22] has a number of results about the computability of the ergodic decomposition. In particular, I suspect that the the map $x \mapsto \mu_x$ is effectively measurable. I also suspect Schnorr random points ω satisfy the following "typicalness" property (similar to [20]) for de Finetti's theorem: for all bounded continuous (not necessarily computable) functions $\varphi\colon \mathbb{R}^d \to [0,1]$, we have

$$\lim_{k\to\infty} \widetilde{A_k}(\varphi)(\omega) = \int \varphi \, d\widetilde{\mu}^d_\omega.$$

Pursuing this, however, would take me too far afield.

12. Characterizing Schnorr randomness

In this section, I show that most of the effective a.e. convergence theorems in this paper are optimal in that Schnorr randomness cannot be strengthened to another form of randomness. In other words, combined with the effective a.e. convergence theorems in this paper, these examples characterize Schnorr randomness. See Table 1 in the introduction for how to match these examples to the corresponding a.e. convergence theorem(s).

12.1. **Monotone convergence, the Lebesgue differentiation theorem, absolutely continuous functions and measures, and uniformly integrable martingales.**

Example 12.1. Fix $(\mathbb{X},\mu)$ and let (U_n) be a Schnorr test. Consider the following function f. By Remark 2.10, we may assume (U_n) is decreasing, and also assume $\mu(U_n) \le 2^{-2n}$ by taking a subsequence. Let $f = \sum_n \mathbf{1}_{U_n}$. The following calculation shows that $f \in L^2_{comp}$.

$$\left\| f - \sum_{n=0}^{m-1} \mathbf{1}_{U_n} \right\|_{L^2} = \left\| \sum_{n=m}^{\infty} \mathbf{1}_{U_n} \right\|_{L^2} \leq \sum_{n=m}^{\infty} \|\mathbf{1}_{U_n}\|_{L^2} = \sum_{n=m}^{\infty} \mu(U_n)^{1/2}$$

$$\leq \sum_{n=m}^{\infty} 2^{-n} = 2^{-m+1}$$

Clearly, $f(x) = \infty$ if x is covered by (U_n).

This example is similar to the Schnorr integral tests of Miyabe [33]. This example will allow me to characterize Schnorr randomness using the monotone convergence theorem, the Lebesgue differentiation theorem, differentiation of absolutely continuous functions, differentiation of absolutely continuous measures, and convergence of uniformly integrable martingales.

Theorem 12.2 (Example of monotone convergence)**.** *Let (U_n) be a Schnorr test on $(\mathbb{X}, \mu)$. There is an increasing sequence of bounded computable functions (f_n) such that $\sup_n \|f_n\|_{L^2} = \infty$ and $\sup_n f_n(x) = \infty$ for all x covered by (U_n).*

Proof. Let $f = \sum_n \mathbf{1}_{U_n}$ be as in Example 12.1. Define $g_n = \sum_{k<n} \mathbf{1}_{U_n}$. We can find a computable $f_n \leq g_n$ such that $\|g_n - f_n\|_{L^2} \leq 2^{-n}$ and $\sup_n f_n = \sup_n g_n = f$. Namely, by effective inner regularity (Proposition 3.22) find a closed set $C_n \subseteq U_n$ of computable measure such that $\mu(U_k - C_k) \leq 2^{-(k+1)}$. Then, using the effective Tietze extension theorem [52] we can find a computable function $h_k \leq \mathbf{1}_{U_k}$ such that $h_k = 0$ on U_k^c and $h_k = 1$ on C_k. Then $f_n = \sum_{k<n} h_n$ is as desired. □

Theorem 12.3 (Example of Lebesgue differentiation theorem)**.** *For any Schnorr test (U_i) on $([0,1]^d, \lambda)$, there is an $f \in L^2_{comp}([0,1]^d, \lambda)$ such that $\frac{1}{\lambda(B(x,r))} \int_{B(x,r)} f \, d\lambda \to \infty$ for all x covered by (U_i). (This holds as well for $\mathbb{T}^d$ and for the dyadic version on $2^{\mathbb{N}}$.)*

Proof. Take the L^2-computable f from Example 12.1. Let x be covered by (U_i). Then for each k, there is some r_k such that $B(x, r_k) \subseteq U_k$. Since (U_k) is decreasing, $f(y) \geq k$ for all $y \in B(x, r_k)$. Hence, $\frac{1}{\lambda(B(x,r))} \int_{B(x,r)} f \, d\lambda \geq k$. Hence $\limsup_{r \to 0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} f \, d\lambda = \infty$. □

Theorem 12.4 (Example of absolutely continuous measure)**.** *Let (U_n) be a Schnorr test on $([0,1]^d, \lambda)$. There is an absolutely continuous, positive measure μ with L^2-computable derivative $\frac{d\mu}{d\lambda}$ such that $\frac{\mu(B(z,r))}{\lambda(B(z,r))} \xrightarrow[r \to 0]{} \infty$ for all z covered by (U_n).*

Proof. Take the L^2-computable f from Example 12.1. let μ be defined by $\mu(A) = \int_A f \, d\lambda$. The rest of the proof is the same as the previous one. □

Theorem 12.5 (Example of absolutely continuous function)**.** *Let (U_n) be a Schnorr test on $([0,1], \lambda)$. There is an increasing, absolutely continuous, computable function F with L^2-computable derivative such that $\frac{d}{dx} F|_{x=z} = \infty$ for all z covered by (U_n).*

Proof. Take the L^2-computable f from Example 12.1. Let $F(x) = \int_0^x f(t) \, dt$. Then F is computable, increasing, and absolutely continuous. By the same argument as in Theorem 12.3, $\frac{d}{dx} F|_{x=z} = \infty$ for all z covered by (U_n). □

Theorem 12.6 (Example of a dyadic uniformly integrable martingale)**.** *Let (U_n) be a Schnorr test on $(2^{\mathbb{N}}, \lambda)$. There is a nonnegative, computable, dyadic, uniformly integrable, martingale (M_k) with limit $M_\infty \in L^2_{comp}$ (and hence in L^1_{comp}) such that $M_k(x) \to \infty$ on all x covered by (U_n).*

Proof. Take the L^2-computable f from Example 12.1. Then let $M_k = f^{(k)} = \mathbb{E}[f \mid \mathcal{B}_k]$ as in Example 6.1. This is a computable, dyadic martingale with limit $M_\infty = f$. If x is covered by (U_n) then $M_k(x) \to \infty$ by the same argument as Theorem 12.3. □

In this next theorem, $x \in (\mathbb{X}, \mu)$ is Kurtz random if it is not in any Σ^0_2 null set. Every Kurtz random is Schnorr random. All a.e. computable functions f are defined on Kurtz randoms, since the the domain of f is a measure one Π^0_2 set. Further, no Kurtz randoms are on the boundary of a ball in *Basis*$(\mathbb{X}, \mu)$, since the set of boundaries is a Σ^0_2 null set. Therefore for each decomposition of $\mathbb{X}$ into finitely many cells, a Kurtz random x is in the interior of one of the cells. See Rute [41], for more discussion.

Theorem 12.7 (Example of a uniformly integrable martingale)**.** *Fix $(\mathbb{X}, \mu)$. Let (U_n) be a Schnorr test. From (U_n) we can construct an a.e. computable, uniformly integrable, L^2-computable (and hence L^1-computable) martingale (M_k) with limit $M_\infty \in L^2_{comp}$ (and hence in L^1_{comp}) such that $M_k(x)$ diverges for all Kurtz random x covered by U_n. (Since M_k is a.e. computable, it is well-defined on Kurtz randoms.)*

Proof. The idea is the same as the previous proof, except that one needs a "canonical filtration" for the space $(\mathbb{X}, \mu)$. Recall the collection *Basis*$(\mathbb{X}, \mu)$ from Lemma 3.5 which has an enumeration $\{B_i\}$. Let $\mathcal{P}_k$ be the partition generated by $\{B_0, \dots, B_k\}$. This generates a filtration $\sigma(\mathcal{P}_k)$ of $\mathbb{X}$ such that $\sigma(\mathcal{P}_k) \uparrow \mathcal{B}(\mathbb{X})$ (the Borel σ-algebra).

Now let f be as in Example 12.1. Let $M_k = \mathbb{E}[f \mid \mathcal{P}_k]$ although we will define it in an a.e. computable manner as follows. To compute $M_k(x)$, just find the atom $Q \in \mathcal{P}_k$ that x is in, and then computing $\frac{1}{\mu(Q)} \int_Q f \, d\mu$. This can be done for almost every x, namely all x in the interior of some $Q \in \mathcal{P}_k$ with positive measure (all Kurtz randoms x have this property).

If x is Kurtz random and covered by (U_n), then take the intersection of the first N many sets U_n that contain x. There is a ball $B(x, r)$ in the intersection (since we are assuming the U_n are decreasing). Since *Basis*$(\mathbb{X}, \mu)$ is an effective basis, there is a computable sequence of sets $\{Q_i\}$ from $\bigcup_k \mathcal{P}_k$ such that $B(x, r) = \bigcup_i Q_i$ μ-a.e. If x is Kurtz random, then $x \in Q$ for some $Q = Q_i \in \mathcal{P}_k$ for some k. Then we have for all $\ell \geq k$,

$$M_\ell(x) = \mathbb{E}[f \mid \mathcal{P}_\ell](x) \geq \mathbb{E}[N \cdot \mathbf{1}_Q \mid \mathcal{P}_\ell](x) = N \cdot \mathbf{1}_Q(x) = N.$$

Hence $M_k(x) \xrightarrow[k \to \infty]{} \infty$. □

12.2. **Singular martingales, functions of bounded variation, and measures.** Consider these two examples of nonnegative, dyadic, singular martingales (the limit is zero) corresponding to a Schnorr test (U_n). The main idea is to bet when it looks like x is in another U_n, and then to "bet away" the money back down to zero. One puts all its mass (bets all its money) on a countable set of points. The other puts its mass on a measure-zero set, without atoms.

Example 12.8 (Singular "atomic" martingale)**.** Let (U_n) be a Schnorr test on $(2^{\mathbb{N}}, \lambda)$. Assume (U_n) is decreasing, and assume $\mu(U_n) \leq 2^{-n}$. Effectively partition

$U_n = \bigcup_m [\sigma^n_m]$ (that is, a prefix-free representation of U_n). Let $(\sigma_i)_i$ be a reordering of $(\sigma^n_m)_{n,m}$. If x is covered by (U_n) then $x \in [\sigma_i]$ for infinitely-many i.

For each i, create a martingale as follows. For each i, let a_i be the "midpoint" of $[\sigma_i]$ (that is $a_i = \sigma_i 100...$). Let $b_i = \lambda(\sigma_i)/\sqrt{\lambda(U_n)}$ for the n such that $\sigma_i \subseteq U_n$. Then define a computable dyadic martingale $M^{(i)}$ as the one that puts all its money on the point a_i and has starting capital b_i. That is, for each $\tau \in 2^{<\omega}$, define

$$M^{(i)}(\tau) = \begin{cases} b_i/\lambda(\tau) & \text{if } a_i \in [\tau] \\ 0 & \text{otherwise} \end{cases}.$$

It is easy to verify each $M^{(i)}$ is a computable, nonnegative, singular, dyadic martingale. Define $M = \sum_i M^{(i)}$. This is also a nonnegative, singular, dyadic martingale, and M is finite and computable since $\sum_i b_i = \sum_n \lambda(U_n)/\sqrt{\lambda(U_n)} \leq \sum_n 2^{-n/2}$ and $\sum_i b_i$ is computable. If x is covered by (U_n) then for every n we have some i such that $x \in \sigma_i \subseteq U_n$ and $M(\sigma_i) \geq M^{(i)}(\sigma_i) = b_i/\lambda(\sigma_i) = 1/\sqrt{\lambda(U_n)} \geq 2^{n/2}$. Hence $\limsup_k M_k(x) = \infty$.

This first example allows us to characterize Schnorr randomness by singular martingales, atomic measures, and bounded variation functions consisting only of jumps.

Theorem 12.9 (Example of singular martingale)**.** *Let (U_n) be a Schnorr test. There is a nonnegative, computable, singular, dyadic martingale (M_k) such that $\limsup_k M_k(x) = \infty$ for all x covered by U_n.*

Proof. Use the martingale in Example 12.8 (or in Example 12.13 below). □

On $([0,1]^d, \lambda)$ redefine $I_k(x)$ to be the OPEN DYADIC SET containing x (in the absolutely continuous case, it did not much matter if $I_k(x)$ was open or half-open). Define $\overline{I_k}(x)$ as the corresponding closed set.

Lemma 12.10. *Let μ be a computable positive measure on $2^{\mathbb{N}}$. There is a corresponding computable positive measure ν on $[0,1]^d$ such that $\nu(I_k(x)) \leq \mu(x \upharpoonright dk) \leq \nu(\overline{I_k}(x))$ for all vectors x with no dyadic rational coordinates.*

Proof. Let $T\colon 2^{\mathbb{N}} \to [0,1]^d$ be the (usual) computable map $T(x) = (y_0, \ldots, y_{d-1})$ where $y_i = 0.x(i)x(d+i)x(2d+1)\ldots$. in particular, $T^{-1}(I_k(x)) \subsetneq [x \upharpoonright dk] \subsetneq T^{-1}(\overline{I_k}(x))$. That is, the first d bits of x correspond to the first bit of each coordinate in $(x_1, \ldots, x_d)$. Define ν as the push-forward measure of μ along T, hence $\nu(\overline{I_k}(x)) \leq \mu(x \upharpoonright dk) \leq \nu(I_k(x))$. By Proposition 3.26, ν is computable. □

Theorem 12.11 (Example of atomic, singular measure)**.** *Let (U_n) be a Schnorr test on $([0,1]^d, \lambda)$. There is an atomic, singular positive measure ν such that $\limsup_r \frac{\nu(B(z,r))}{\lambda(B(z,r))} = \infty$ for all z covered by (U_n).*

Proof. Let (V_n) be a test on $(2^{\mathbb{N}}, \lambda)$ which covers the points in $2^{\mathbb{N}}$ corresponding to the points that (U_n) covers in $2^{\mathbb{N}}$. (Partition each U_n into closed dyadic sets and replace each with the corresponding basic open set $[\sigma]$.)

Let μ be the computable positive measure on $2^{\mathbb{N}}$ associated with the martingale M in Example 12.8. That is, $\mu(\sigma) = M(\sigma)\lambda(\sigma)$. Notice that μ is atomic. Let ν be the computable positive measure on $[0,1]^d$ as in Lemma 12.10; ν is still atomic. Without loss of generality, we assume in Example 12.8 that $\limsup_k M_{dk}(x) = \infty$ for the x covered by U_n. (Just require the σ_i to be of length dk for some k.) Then by Lemma 12.10, for all x covered by (U_n). We have $\limsup_k \nu(\overline{I_k}(x))/\lambda(\overline{I_k}(x)) \geq \limsup_k M_k(x) = \infty$. By an geometric argument similar to the proof of Lemma 4.9, we have $\limsup_{r\to 0} \frac{\nu(B(z,r))}{\lambda(B(z,r))} = \infty$. □

Theorem 12.12 (Example of bounded variation function with jumps)**.** *Let (U_n) be a Schnorr test on $([0,1],\lambda)$. There is a nondecreasing function F and a computable sequence of pairs of reals (a_i, b_i) such that $F(x) = \sum_{a_i \leq x} b_i$ (F only consists of jumps), $V(F) = \sum_i b_i$ is computable, and $\frac{d}{dx}F|_{x=z}$ does not exist for all z covered by (U_n).*

Proof. Let a_i, b_i be from Example 12.8 (except a_i is now the corresponding real in $[0,1]$). Let ν be the measure from the previous example. Notice that each a_i is an atom of ν with weight b_i. Hence, $F(x) = \sum_{a_i\leq x} b_i = \nu([0,x])$. By the previous proof, the derivative of F does not exist at z covered by (U_n). □

Now for the second example martingale.

Example 12.13 (Singular "continuous" martingale)**.** Define σ_i and b_i the same as in Example 12.8. But now, we want to put the mass on a set of points in $[\sigma]$. Define $N^{(i)}$ as follows. If $|\tau| \leq |\sigma|$ then bet all the money on σ.

$$N^{(i)}(\tau) = \begin{cases} b_i/\lambda(\tau) & \text{if } \tau \preceq \sigma \\ 0 & \text{otherwise} \end{cases}.$$

If τ is incomparable with σ then $N^{(i)}(\tau) = 0$. If $\tau \succeq \sigma$, then bet that the even bits are all $1s$, ignoring the odd bits. That is,

$$N^{(i)}(\tau 0) = \begin{cases} 0 & |\tau| \text{ is even} \\ N^{(i)}(\tau) & |\tau| \text{ is odd} \end{cases}$$
$$N^{(i)}(\tau 1) = \begin{cases} 2\cdot N^{(i)}(\tau) & |\tau| \text{ is even} \\ N^{(i)}(\tau) & |\tau| \text{ is odd} \end{cases}.$$

This nonnegative dyadic martingale will almost surely converge to 0 and is therefore singular. Define $M = \sum_i M^{(i)}$. As before, M is computable. If x is covered by (U_n), by the same argument as in Example 12.8, we have $\limsup_k M_k(x) = \infty$.

Theorem 12.14. *Let (U_n) be a Schnorr test on $([0,1]^d,\lambda)$. There is a continuous, singular, positive measure μ such that $\limsup_{r\to 0} \frac{\mu(B(z,r))}{\lambda(B(z,r))} = \infty$ for all z covered by (U_n).*

Proof. Follow the proof of Theorem 12.11, except use Example 12.13 to get a continuous measure. □

Theorem 12.15. *Let (U_n) be a Schnorr test on $([0,1],\lambda)$. There is a continuous, nondecreasing function F with zero derivative almost surely such that $\frac{d}{dx}F|_{x=z}$ does not exist for all z covered by (U_n).*

Proof. Let $F(x) = \nu([0,x])$ where ν is the measure in Theorem 12.14. Therefore, the derivative of F does not exist for all z covered by (U_n). □

12.3. **Backwards martingales, the strong law of large numbers, de Finetti's theorem, and the ergodic theorem.** Consider the following fact found in Schnorr's book [43, Theorem 12.1] which says that each non-Schnorr random can fail to satisfy the law of large numbers. (The proof can also be found in Gács, Hoyrup, Rojas [20].)

Proposition 12.16 ((Schnorr))**.** *If (U_n) is a Schnorr test, then there is an a.e. computable measure preserving transformation $\varphi\colon 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ such that for all x not covered by (U_n), if $y = \varphi(x)$, then $\limsup_k \frac{1}{k}\sum_{i<k} y(i) \geq \frac{2}{3}$.*

This fact will allow us to use the strong law of large numbers, de Finetti's theorem, backwards martingale convergence, and the ergodic theorem to characterize Schnorr randomness on $(2^{\mathbb{N}}, \lambda)$.

Corollary 12.17. *If (U_n) is a Schnorr test on $(2^{\mathbb{N}}, \lambda)$, then the following hold.*

(1) *There is a computable i.i.d. sequence of i.i.d. $0,1$-valued random variables (X_i) such that $\frac{1}{k}\sum_{i<k} X_i(x)$ diverges for all x covered by (U_n).*
(2) *There is a computable exchangeable sequence of a.e. computable random variables (X_i) and a bounded computable $\psi\colon \mathbb{R} \to \mathbb{R}$, such that $\frac{1}{k}\sum_{i<k} \psi(X_i(x))$ diverges for all x covered by (U_n).*
(3) *There is a bounded a.e. computable backwards martingale (M_{-k}) with a constant, computable limit $M_{-\infty}$ such that $M_{-k}(x)$ diverges for all x covered by (U_n).*
(4) *(Gács, Hoyrup, Rojas [20]) There is a bounded a.e. computable function $f\colon 2^{\mathbb{N}} \to \mathbb{R}$ and an a.e. computable, ergodic, measure preserving $T\colon 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ such that $\frac{1}{k}\sum_{i<k} f(T^n(x))$ diverges for all x covered by (U_n).*

Proof. Take φ as in Proposition 12.16. Slightly modify φ so that $\frac{1}{k}\sum_{i<k}\varphi(n)(i)$ diverges on all x covered by (U_n). (Just swap the 0s and 1s whenever $\frac{1}{k}\sum_{i<k}\varphi(n)(i) > \frac{6}{5}$ or $< \frac{4}{5}$.)

For (1) and (2), let $X_n(x)$ be the nth bit of $\varphi(x)$. This sequence is i.i.d. (and therefore exchangeable). For (2), also let ψ be the identity map.

For (3), let $M_{-k} = \frac{1}{k}\sum_{j<k} X_k$ from (1). Recall from Corollary 11.7, this is a backwards martingale with limit $\frac{1}{2}$.

For (4), Gács, Hoyrup, and Rojas [20] showed that φ can be constructed to have an a.e. computable inverse. Set $T = \varphi \circ \sigma \circ \varphi^{-1}$ where σ is the left shift map, and set f to be the first bit of $\varphi(x)$. Then $\frac{1}{n}\sum_{k<n} f(T^n(x))$ is equal to $\frac{1}{k}\sum_{j<k} X_k$ which diverges. □

12.4. **Convergence of test functions to 0.**

Theorem 12.18. *Let (U_n) be a Schnorr test in $(2^{\mathbb{N}}, \lambda)$. There is a computable sequence (φ_n) of dyadic test functions, such that $\|\varphi_n\|_{L^2} < 2^{-n}$ but $\limsup_n \varphi_n(x) = \infty$ on all x covered by (U_n).*

Proof. By Remark 2.10, we may assume that $\lambda(U_n) \leq 2^{-(2n+2)}$. We may also computably break up (U_n) into a disjoint union of dyadic intervals $U_n = \bigcup_m [\sigma^n_m]$. (For each n, the set $\{\sigma^n_m\}$ be infinite or finite—it is enough to know it is computably enumerable uniformly in n.) Then

$$\sum_{n,m} \left(1/\sqrt{\lambda(U_n)}\right) \cdot \lambda(\sigma^n_m) = \sum_n \left(1/\sqrt{\lambda(U_n)}\right) \cdot \lambda(U_n) \leq 1,$$

and the sum is computable. Renumber $\{\sigma_i\}_i = \{\sigma_{n,m}\}_{n,m}$ using a computable pairing function. Effectively partition the double sequence $(\sigma_i)_i$ into finite sequences $(\sigma_{i(k)}, \sigma_{i(k)+1}, \dots, \sigma_{i(k+1)-1})$ such that $\sum_{j=\sigma(k)}^{i(k)-1} \left(1/\sqrt{\lambda(U_n)}\right) \cdot \lambda(\sigma_i) \le 2^{-k}$ where i codes the pair (n, m) (break up the $[\sigma_i]$ into smaller intervals if needed).

Let $\varphi_k = \sum_{j=\sigma(k)}^{i(k)-1} (1/\sqrt{\lambda(U_n)}) \cdot \mathbf{1}_{[\sigma_i]}$. By the pigeonhole principle, if x is covered by (U_n) then for each n, $\varphi_k(x) > n$ for infinitely many k. □

Theorem 12.19. *Let (U_n) be a Schnorr test on $(\mathbb{X}, \mu)$. There is a computable sequence (φ_n) of test functions, such that $\|\varphi_n\|_{L^2} < 2^{-n}$ but $\limsup_n \varphi_n(x) = \infty$ on all Kurtz random x covered by (U_n).*

Proof. The proof is the same as the previous one. Just replace dyadic intervals $[\sigma]$ with finite Boolean combinations of *Basis*$(\mathbb{X}, \mu)$ from Lemma 3.5. (Also, make the sets slightly larger to cover their measure-zero boundaries.) □

Theorem 12.20. *Let (U_n) be a Schnorr test on $(\mathbb{X}, \mu)$. There is a computable sequence (f_k) of computable functions such that $\|f_k\|_{L^2} < 2^{-k}$ but $\limsup_k f_k(x) = \infty$ on all x covered by (U_n).*

Proof. Take the test functions (φ_k) from the previous two theorems. Approximate them with computable functions f_k as follows. For each $[\sigma]$ in φ_k (or the corresponding finite Boolean combination B of basis elements), find a computable function h_σ such that on $h_\sigma = 1$ on $[\sigma]$ and $\|h_\sigma - \mathbf{1}_\sigma\|_{L^2}$ is sufficiently small. This can be done by defining $h_\sigma = 1$ on $[\sigma]$ (or in the other case, on the closure $\overline{B}$ which has the same measure), using effective outer regularity (Proposition 3.22) to find an open set $V \supseteq [\sigma]$ of similar measure, defining $h_\sigma = 0$ on V^c and then using the effective Tietze extension theorem [52] to extend this to a computable function. □

Theorem 12.21. *Let (U_n) be a Schnorr test on $([0,1]^d, \lambda)$. There is a computable sequence (p_k) of rational polynomials such that $\|p_k\|_{L^2} < 2^{-k}$ but $\limsup_k p_k(x) = \infty$ on all x covered by (U_n).*

Proof. Take the computable functions (f_n) in the last theorem. Effectively approximate (f_n) by polynomials using the effective Weierstrass approximation theorem [40]. Since they are close in the uniform norm, they are close in the L^2-norm. □

Appendix A. Proofs from Section 3.

A.1. **Useful facts.** The following set of calculations are straightforward, but useful.

Fact A.1. *If $f \le g$ (a.e.), then*

$$\mu\{f > \varepsilon\} \le \mu\{g > \varepsilon\}.$$

Also

$$\mu\{f_1 + f_2 > \varepsilon_1 + \varepsilon_2\} \le \mu\{f_1 > \varepsilon_1\} + \mu\{f_2 > \varepsilon_2\}.$$

and

$$\mu\left\{\sum_i f_i > \sum_i \varepsilon_i\right\} \le \sum_i \mu\{f_i > \varepsilon_i\}.$$

Also, recall Markov's inequality and a useful variation for the metric d_{meas}.

Fact A.2 (Markov's inequality, see [47]). *Assume f is an integrable function and $\varepsilon > 0$. Then*

$$\mu\{x \mid |f| \geq \varepsilon\} \leq \frac{\|f\|_{L^1}}{\varepsilon}.$$

Also given $\mathbb{Y}$-valued measurable functions f and g and $0 < \varepsilon \leq 1$,

$$\mu\{x \mid d_{\mathbb{Y}}(f,g) \geq \varepsilon\} = \mu\{x \mid \min\{d_{\mathbb{Y}}(f,g),1\} \geq \varepsilon\} \leq \frac{d_{meas}(f,g)}{\varepsilon}.$$

A.2. **Integrable functions, measurable functions, and measurable sets.**

Restatement of Proposition 3.7. *The measure of each cell of Basis$(\mathbb{X},\mu)$ is computable from its code σ.*

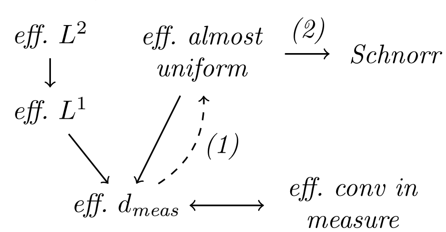
Proof. Given a cell $C = A_1 \cap \ldots \cap A_\ell \cap B_1^c \cap \ldots \cap B_k^c$ (where $A_1, \ldots A_\ell, B_1, \ldots, B_k \in$ *Basis*$(\mathbb{X},\mu)$, that is balls of null boundary), then C is in between the effectively open and effectively closed sets $A_1 \cap \ldots \cap A_\ell \cap \overline{B_1}^c \cap \ldots \cap \overline{B_k}^c$ and $\overline{A}_1 \cap \ldots \cap \overline{A}_\ell \cap B_1^c \cap \ldots \cap B_k^c$ which have the same measure. Since the measure of effectively open sets is lower semicomputable and closed sets is upper semicomputable (Proposition 2.5), the measure of C is computable (uniformly from its code σ). □

Proposition A.3. *Let A be a set formed by combining elements of Basis$(\mathbb{X},\mu)$ using finitely-many connectives Boolean connectives $\cup, \cap, ^c$ as well as the closure operator. Then $\mu(A)$ is computable from (the code for) A.*

Proof. A finite Boolean combination can be decomposed into a finite union of pairwise disjoint cells (basically disjunctive normal form). Since the boundaries of the cells have measure zero, the closure operator does not effect the measure. □

A.3. **Effective modes of convergence.**

Restatement of Proposition 3.15 (Modes of effective convergence). *On a computable probability space $(\mathbb{X},\mu)$, the following implications are effective—in that a rate of convergence for the latter is computable from the former. (L^1 and L^2 only apply to real-valued functions.)*



(1) *The dotted arrow represents that if $f_i \to f$ with a geometric rate of convergence in the metric d_{meas}, e.g. $\forall j \geq i\ d_{meas}(f_j, f) \leq 2^{-i}$, then $f_i \to f$ effectively almost uniformly.*
(2) *For the arrow going to "Schnorr", see Lemma 3.19.*

Proof. ($L^2 \to L^1 \to d_{meas}$): Use that $d_{meas}(f_i, f) \leq \|f_i - f\|_{L^1} \leq \|f_i - f\|_{L^2}$.

($d_{meas} \to$ measure): Assume $n(\varepsilon)$ is a rate of convergence in the metric d_{meas}. I claim $m(\varepsilon_1, \varepsilon_2) = n(\varepsilon_1\varepsilon_2)$ is a rate of convergence in measure (assuming $0 < \varepsilon < 1$). Indeed, for $i \geq n(\varepsilon_1\varepsilon_2)$, $d_{meas}(f_i, f) \leq \varepsilon_1\varepsilon_2$ and by Markov's inequality (Fact A.2),

$$\mu\left\{d_{\mathbb{Y}}(f_i, f) > \varepsilon_1\right\} \leq \frac{d_{meas}(f_i, f)}{\varepsilon_1} \leq \frac{\varepsilon_1\varepsilon_2}{\varepsilon_1} = \varepsilon_2.$$

(Measure $\to d_{meas}$): Let $m(\varepsilon_1, \varepsilon_2)$ be a rate of convergence in measure. I claim that $n(\varepsilon) = m(\varepsilon/2, \varepsilon/2)$ is a rate of convergence in the metric d_{meas}. Indeed, for $i \geq m(\varepsilon/2, \varepsilon/2)$ we have that

$$d_{meas}(f_i, f) = \int \min\{d_{\mathbb{Y}}(f_i, f), 1\}\, d\mu \leq \mu\{d_{\mathbb{Y}}(f_i, f) > \varepsilon/2\} + \varepsilon/2 \leq \varepsilon/2 + \varepsilon/2.$$

(Almost uniform $\to$ measure): A rate of effective almost uniform convergence $n(\varepsilon_1, \varepsilon_2)$ is also a rate of convergence in measure since if $i \geq n(\varepsilon_1, \varepsilon_2)$,

$$\mu\{d_{\mathbb{Y}}(f_i, f) > \varepsilon_1\} \leq \mu\left\{\sup_{i \geq n(\varepsilon_1, \varepsilon_2)} d_{\mathbb{Y}}(f_i, f) > \varepsilon_1\right\}.$$

(1): Assume $\forall j \geq i\ d_{meas}(f_j, f) \leq 2^{-i}$. Based on this rate of convergence in the metric d_{meas}, the rate of convergence in measure is $n(\varepsilon_1, \varepsilon_2) \geq -\log_2(\varepsilon_1\varepsilon_2)$. I claim $n(\varepsilon_1, \varepsilon_2) = -\log_2\left(\frac{\varepsilon_1\varepsilon_2}{(2+\sqrt{2})^2}\right)$ is a rate of almost uniform converge. Indeed, if $n = -\log_2\left(\frac{\varepsilon_1\varepsilon_2}{(2+\sqrt{2})^2}\right)$, by Facts A.1 and A.2 and the fact that $\sum_{i\geq 0} 2^{-i/2} = 2 + \sqrt{2}$ we have

$$\begin{aligned}
\mu\left\{\sup_{i\geq n} d_{\mathbb{Y}}(f_i, f) > \varepsilon_1\right\} &\leq \mu\left\{\sum_{i\geq n} d_{\mathbb{Y}}(f_i, f) > \varepsilon_1\right\} \\
&\leq \sum_{i\geq 0} \mu\left\{d_{\mathbb{Y}}(f_{i+n}, f) > \frac{2^{-i/2}}{2+\sqrt{2}} \cdot \varepsilon_1\right\} \\
&\leq \sum_{i\geq 0} \frac{d_{meas}(f_{i+n}, f)}{\frac{2^{-i/2}}{2+\sqrt{2}} \cdot \varepsilon_1} \leq \sum_{i\geq 0} \frac{2^{-(i+n)}}{\frac{2^{-i/2}}{2+\sqrt{2}} \cdot \varepsilon_1} \\
&= \frac{(2+\sqrt{2})}{2^n \varepsilon_1} \sum_{i\geq 0} 2^{-i/2} = \frac{(2+\sqrt{2})^2}{2^n\varepsilon_1} = \varepsilon_2. \quad \square
\end{aligned}$$

Restatement of Proposition 3.16. *Let (f_n) and f be uniformly effectively measurable real-valued functions.*

1. *If $f_n \to f$ effectively a.e.. and $g_n \to g$ effectively a.e.., then $f_n + g_n \to f + g$ effectively a.e..*
2. *If $f_n^j \to f^j$ effectively a.e.. ($j \in \{0, \ldots, k-1\}$), and g is computable with a uniform modulus of continuity, then $g(f_n^0, \ldots, f_n^{k-1}) \to g(f^0, \ldots, f^{k-1})$ effectively a.e..*
3. *(Squeeze theorem) Assume $f_n \leq g_n \leq h_n$ a.e. and that $f_n \to g$ effectively a.e.. and $h_n \to g$ effectively a.e.., then $g_n \to g$ effectively a.e.*

Further, in all cases the rates of convergence for the latter are computable from the former (in (2) use the modulus of continuity for g). Indeed, we do not need to assume the functions are effectively measurable, just that the rates of convergence are computable. The same results hold for continuous convergence, e.g. $f_r \to f$ as $r \to 0$.

Proof. (1): Assume $f_i \to f$ and $g_i \to g$ with rates $n(\varepsilon_1, \varepsilon_2)$ and $n'(\varepsilon_1, \varepsilon_2)$, respectively, of a.e.. convergence. I claim $m(\varepsilon_1, \varepsilon_2) = \max\left\{n(\frac{\varepsilon_1}{2}, \frac{\varepsilon_2}{2}), n'(\frac{\varepsilon_1}{2}, \frac{\varepsilon_2}{2})\right\}$ is a rate of almost uniform convergence for $f_i + g_i \to f + g$. Indeed, if $m = m(\varepsilon_1, \varepsilon_2)$ then

$$\mu\left\{\sup_{i\geq m} |(f_i+g_i)-(f+g)| > \varepsilon_1\right\} \leq \mu\left\{\left(\sup_{i\geq m}|f_i - f|\right) + \left(\sup_{i\geq m}|g_i-g|\right) > \varepsilon_1\right\}$$
$$\leq \mu\left\{\sup_{i\geq m}|f_i-f| > \frac{\varepsilon_1}{2}\right\} + \mu\left\{\sup_{i\geq m}|g_i - g| > \frac{\varepsilon_1}{2}\right\} \leq \frac{\varepsilon_2}{2} + \frac{\varepsilon_2}{2} = \varepsilon_2.$$

(2): Assume $f_i^j \to f^j$ with a rate of a.e. convergence $n_j(\varepsilon_1, \varepsilon_2)$. Also assume $g\colon \mathbb{R}^k \to \mathbb{R}$ is a continuous function with a computable modulus of continuity $\delta(\varepsilon)$, that is for all $x_0, \ldots, x_{k-1}, y_0, \ldots, y_{k-1} \in \mathbb{R}$.

$$\sum_{j=0}^{k-1} |x_j - y_j| \leq \delta(\varepsilon) \quad \to \quad |g(x_1,\ldots,x_j) - g(y_1,\ldots,y_k)| \leq \varepsilon.$$

Fix $\varepsilon_1, \varepsilon_2 > 0$. Let $m = \max_{j<k} n_j(\frac{\delta(\varepsilon_1)}{k}, \frac{\varepsilon_2}{k})$. Then

$$\begin{aligned}
&\mu\left\{\sup_{i\geq m}\left|g(f_i^0,\ldots,f_i^{k-1}) - g(f^0,\ldots,f^{k-1})\right| > \varepsilon_1\right\}\\
&\qquad \leq \mu\left\{\sum_{j<k}\sup_{i\geq m}|f_i^j - f^j| > \delta(\varepsilon_1)\right\}\\
&\qquad \leq \sum_{j<k}\mu\left\{\sup_{i\geq m}|f_n^j - f| > \frac{\delta(\varepsilon_1)}{k}\right\}\\
&\qquad \leq \sum_{j<k}\frac{\varepsilon_2}{k} = \varepsilon_2.
\end{aligned}$$

(3): Assume $f_n \leq g_n \leq h_n$ a.e. and $f_i \to g$ and $h_i \to g$ with computable rates of a.e.. convergence. Let a rate of a.e.. convergence for $f_i \to g$ be $n(\varepsilon_1, \varepsilon_2)$. By part (2), a rate of a.e.. convergence $n'(\varepsilon_1, \varepsilon_2)$ for $(h_i - f_i) \to 0$ is computable. We claim that $g_i \to g$ with a rate of $m(\varepsilon_1, \varepsilon_2) = \max\left\{n(\frac{\varepsilon_1}{2}, \frac{\varepsilon_2}{2}), n'(\frac{\varepsilon_1}{2}, \frac{\varepsilon_2}{2})\right\}$. Indeed, for $n = n(\varepsilon_1, \varepsilon_2)$, we have

$$\mu\left\{\sup_{i\geq m}|g_i - g| > \varepsilon_1\right\} \leq \mu\left\{\sup_{i\geq m}\left(|f_i - g| + (h_i - f_i)\right) > \varepsilon_1\right\}$$
$$\leq \mu\left\{\sup_{i\geq m}|f_i - g| > \frac{\varepsilon_1}{2}\right\} + \mu\left\{\sup_{i\geq m}(h_i - f_i) > \frac{\varepsilon_1}{2}\right\} \leq \frac{\varepsilon_2}{2} + \frac{\varepsilon_2}{2} = \varepsilon_2.$$

As for continuous convergence, the proofs are the same. □

A.4. **Convergence on Schnorr randomness.**

Remark A.4. A Solovay test for Schnorr randomness (U_n) is a computable sequence of effectively open sets U_n such that the sum $\sum_n \mu(U_n)$ is finite and computable. (This follows when $\mu(U_n)$ is computable uniformly from n and $\mu(U_n) \leq 2^{-n}$ or any other sequence with a finite sum.) If $x \in U_n$ for infinitely-many n, then say n is Solovay covered by (U_n). Then x is Schnorr random if and only if it is not Solovay-covered by any Solovay test for Schnorr randomness [10, 20]. (This is an effective version of the Borel-Cantelli lemma.)

Lemma A.5. *Suppose (φ_n) is a computable sequence of test functions which converge effectively a.e. to $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$.*

(1) *(Existence) The limit* $\lim_{n\to\infty} \varphi_n(x)$ *exists on all Schnorr randoms* x.
(2) *(Uniqueness) Given another sequence of test functions* (ψ_n) *converging effectively a.e. to* f,
$$\lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} \psi_n(x) \quad (\textit{on Schnorr random } x).$$

Proof. First we show existence by showing that $\varphi_n(x)$ is Cauchy for Schnorr randoms x. The main idea is to break up the indices into finite intervals. Since the rate of effectively a.e. convergence is computable, effectively choose n_k so that
$$\mu\left\{\sup_{n\geq n_k} d_{\mathbb{Y}}(\varphi_n, \varphi_{n_k}) > 2^{-(k+1)}\right\} \leq 2^{-(k+1)}.$$
Our Solovay test for Schnorr randomness is
$$U_k = \left\{x \;\middle|\; \max_{n\in[n_k,n_{k+1}]} d_{\mathbb{Y}}(\varphi_n(x), \varphi_{n_k}(x)) > 2^{-(k+1)}\right\}.$$
Each set is effectively open uniformly in k. (As a technicality, let x only range over the interiors of the cells in φ_n. This guarantees that U_k is effectively open. It is also sufficient for our purposes since the boundary of each cell is a measure zero effectively closed set and therefore cannot contain Schnorr randoms.) This is a Solovay test since $\mu(U_k)$ is computable and by our choice of n_k,
$$\mu(U_k) \leq \mu\left\{\sup_{n\geq n_k} d_{\mathbb{Y}}(\varphi_n, \varphi_{n_k}) > 2^{-(k+1)}\right\} \leq 2^{-(k+1)}.$$

Now, let x be Schnorr random (and hence is not on the boundary of any cell). We have that x is in at most finitely many U_k. Hence for some k_0 large enough, for all $k \geq k_0$ and all $n \in [n_k, n_{k+1}]$ we have $d_{\mathbb{Y}}(\varphi_n(x), \varphi_{n_k}(x)) \leq 2^{-(k+1)}$. It follows that for all $k \geq k_0$ and for all $n \geq n_k$ that
$$d_{\mathbb{Y}}(\varphi_n(x), \varphi_{n_k}(x)) \leq \sum_{j\geq k} 2^{-(j+1)} \leq 2^{-k}.$$
Hence $\varphi_n(x)$ is Cauchy.

For uniqueness, take (φ_n) and (ψ_n) and interleave them, $\varphi_0, \psi_0, \varphi_1, \psi_1, \ldots$. It is easy to see this sequence still has an effectively rate of a.e. convergence. Hence it converges on Schnorr randoms and each subsequence must converge to the same value. □

Restatement of Proposition 3.18. *Suppose* $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ *is effectively measurable with Cauchy-name* (φ_n) *(in the metric* d_{meas}, L^1*-norm, or* L^2*-norm).*

(1) *(Existence) The limit* $\lim_{n\to\infty} \varphi_n(x)$ *exists on all Schnorr randoms* x.
(2) *(Uniqueness) Given another Cauchy-name* (ψ_n) *for* f,
$$\lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} \psi_n(x) \quad (\textit{on Schnorr random } x).$$

Proof. A Cauchy-name has an effective rate of a.e. convergence by Proposition 3.15 and the rest follows from Lemma A.5. □

Restatement of Lemma 3.19 (Convergence Lemma). *Suppose that* (f_k) *and* f *are uniformly effectively measurable. If*
$$f_k \to f \quad (\textit{effectively a.e.})$$
then
$$\widetilde{f}_k(x) \to \widetilde{f}(x) \quad (\textit{for all Schnorr random } x).$$

Proof. First, I will approximate (f_k) with a sequence of test functions (ψ_k) which converges effectively a.e. to f, and then show that $(\widetilde{f}_k)$ is close to (ψ_k) on Schnorr randoms.

For each k, let $(\varphi^k_n)_{n\in\mathbb{N}}$ be a Cauchy-name for f_k. Since a rate of a.e. convergence of $(\varphi^k_n)_{n\in\mathbb{N}}$ is computable from k, effectively choose $(n_{k,i})_{k,i\in\mathbb{N}}$ so that

$$\mu\left\{\sup_{n\geq n_{k,i}} d_{\mathbb{Y}}(\varphi^k_n, \varphi^k_{n_{k,i}}) > 2^{-(k+i+1)}\right\} \leq 2^{-(k+i+1)}.$$

Consider the sequence $\psi_k = \varphi^k_{n_{k,0}}$. I will show that $\psi_k \xrightarrow[k\to\infty]{} f$ effectively a.e. as follows . Choose ε and δ. Since $f_k \to f$ effectively a.e., we can effectively choose k' such that

$$\mu\left\{\sup_{k\geq k'} d_{\mathbb{Y}}(f_k, f) > \frac{\varepsilon}{2}\right\} \leq \frac{\delta}{2}.$$

Let $k(\varepsilon,\delta) = \max\{-2\log_2\varepsilon, -2\log_2\delta, k'\}$. Then $\sum_{k\geq k(\varepsilon,\delta)} 2^{-(k+1)} = 2^{-k(\varepsilon,\delta)} \leq \min\{\varepsilon/2, \delta/2\}$

$$\begin{aligned}\mu\left\{\sup_{k\geq k(\varepsilon,\delta)} d_{\mathbb{Y}}(\varphi^k_{n_{k,0}}, f) > \varepsilon\right\} &\leq \sum_{k\geq k(\varepsilon,\delta)} \mu\left\{d_{\mathbb{Y}}(\varphi^k_{n_{k,0}}, f_k) > 2^{-(k+1)}\right\} + \mu\left\{\sup_{k\geq k'} d_{\mathbb{Y}}(f_k, f) > \frac{\varepsilon}{2}\right\} \\ &\leq \sum_{k\geq k(\varepsilon,\delta)} 2^{-(k+1)} + \frac{\delta}{2} \leq \delta.\end{aligned}$$

Hence, $\psi_k(x) \to \widetilde{f}(x)$ on Schnorr randoms.

To show convergence of $\widetilde{f}_k$, consider the Solovay test

$$U_{k,i} = \left\{x \;\middle|\; \max_{n\in[n_{k,i}, n_{k,i+1}]} d_{\mathbb{Y}}(\varphi_n(x), \varphi_{n_{k,i}}(x)) > 2^{-(k+i+1)}\right\}.$$

(Again, as in Lemma A.5, use the convention that x only ranges over the interiors of the cells.) This is a Solovay test since each $\mu(U_{k,i})$ is computable from k, i and since

$$\sum_k\sum_i \mu(U_{k,i}) \leq \sum_k\sum_i 2^{-(k+i+1)} = 2.$$

Now, let x be Schnorr random (and hence not on the boundary of any cell). We have that x is in at most finitely many $U_{k,i}$. Hence for some k_0 large enough, for all $k \geq k_0$, for all $i \geq 0$, and for all $n \in [n_{k,i}, n_{k,i+1}]$ we have $d_{\mathbb{Y}}(\varphi^k_n(x), \varphi^k_{n_{k,i}}(x)) \leq 2^{-(k+i+1)}$. It follows that for all $k \geq k_0$ and for all $n \geq n_{k,0}$ that

$$d_{\mathbb{Y}}(\varphi^k_n(x), \psi_k(x)) = d_{\mathbb{Y}}(\varphi^k_n(x), \varphi^k_{n_{k,0}}(x)) \leq \sum_{i\geq 0} 2^{-(k+i+1)} \leq 2^{-k}.$$

Hence $d_{\mathbb{Y}}(\widetilde{f}_k(x), \psi_k(x)) \leq 2^{-k}$. Therefore, $\lim_k \widetilde{f}_k(x) = \lim_k \psi_k(x) = \widetilde{f}(x)$. □

A.5. **Properties of effectively measurable functions.**

Restatement of Proposition 3.20. *The following implications hold for real-valued functions (and all the computations are uniform).*

(1) $f \in L^2_{comp} \Rightarrow f \in L^1_{comp} \Rightarrow f \in L^0_{comp}$. *(The converses do not hold in general.)*
(2) *If* $0 \leq f \leq 1$*, then* $f \in L^2_{comp} \Leftrightarrow f \in L^1_{comp} \Leftrightarrow f \in L^0_{comp}$.

(3) *$f \in L^1_{comp} \Leftrightarrow$ ($f \in L^0_{comp}$ and $\|f\|_{L^1}$ is computable).*
(4) *$f \in L^2_{comp} \Leftrightarrow$ ($f \in L^0_{comp}$ and $\|f\|_{L^2}$ is computable).*
(5) *If $f \in L^1_{comp}$ then $\int f\,d\mu$ is computable.*
(6) *If B is effectively measurable, then $\mu(B)$ is computable.*
(7) *If $0 \leq g \leq 1$, $g \in L^1_{comp}$, and $f \in L^1_{comp}$, then $g \cdot f \in L^1_{comp}$.*

Proof. (1): Use that $\|f-\varphi\|_{L^2} \geq \|f-\varphi\|_{L^1} \geq d_{meas}(f,\varphi)$.

(2): In this case, $\|f-\varphi\|^2_{L^2} \leq \|f-\varphi\|_{L^1} = d_{meas}(f,\varphi) \leq \|f-\varphi\|_{L^2}$.

(3): Given f effectively measurable, break up $\max\{f,0\} = \sum_{n\in\mathbb{N}} f_n$ where $f_n = \min\{\max\{f,n\},n{+}1\}-n$ and similarly for $\min\{-f,0\}$. By (2), f_n is L^1-computable from n. Use $\|f\|_{L^1}$ to approximate f in L^1 with finite sums of (f_n).

(4): Same as (4).

(5): Use $\int f\,d\mu = \|\max\{f,0\}\|_{L^1} + \|\min\{f,0\}\|_{L^1}$ and that L^1 is a computable lattice.

(6): Use $\mu(B) = \mu(B\triangle\varnothing) = d(B,\varnothing)$ and that $\varnothing$ is effectively measurable.

(7): Use that $g \in L^1_{comp}$ by (2) and

$$\begin{aligned}\|g\cdot f - \psi\cdot\varphi\|_{L^1} &\leq \|g\cdot(f-\varphi)\|_{L^1} + \|(g-\psi)\cdot\varphi\|_{L^1}\\ &\leq \|f-\varphi\|_{L^1} + \|g-\psi\|_{L^1}\cdot\|\varphi\|_\infty.\end{aligned}$$

Approximate f with a test function φ and then approximate g with ψ. □

Restatement of Proposition 3.21 (Effective Lusin's theorem). *Given an effectively measurable $f\colon(\mathbb{X},\mu)\to\mathbb{Y}$, and some rational $\varepsilon \geq 0$, there is an effectively closed set K of computable measure $\mu(K) \geq 1-\varepsilon$ and a computable function $g\colon K\to\mathbb{Y}$ such that $g = \widetilde{f}\restriction K$ on Schnorr randoms. (Further, g and K are computable uniformly from ε and any name for f.) Moreover, if $\mathbb{Y}=\mathbb{R}$, then $g\colon K\to\mathbb{Y}$ can be extended (uniformly from its name) to a total computable function $g\colon\mathbb{X}\to\mathbb{Y}$ such that $g = \widetilde{f}\restriction K$ on Schnorr randoms.*

Proof. Let (φ_n) be the Cauchy-name for f in the metric d_{meas}. Let (U_k) be the Solovay test for Schnorr randomness from Lemma A.5, that is

$$U_k = \{x \mid \max_{n\in[n_k,n_{k+1}]} d_{\mathbb{Y}}(\varphi_n(x),\varphi_{n_k}(x)) > 2^{-k}\}$$

for some computable sequence (n_k). Again, we ignore the boundaries of the cells corresponding to φ_n for $n\in[n_k,n_{k+1}]$. Recall, $\mu(U_k)$ is computable from k and $\mu(U_k)\leq 2^{-k}$. To handle the boundaries, we can find an effectively open set V_k of computable measure $\mu(V_k)\leq 2^{-k}$ such that V_k covers the boundaries of the cells corresponding to φ_n for $n\in[n_k,n_{k+1}]$.

Let

$$K = \left(\bigcup_{k\geq 2-\log_2\varepsilon} U_k\cup V_k\right)^c.$$

Then

$$1-\mu(K) \leq \sum_{k\geq 2-\log_2\varepsilon} \mu(U_k\cup V_k) \leq \sum_{k\geq 2-\log_2\varepsilon} 2\cdot 2^{-k} \leq \varepsilon$$

and $\mu(K)$ is computable (the measure of every finite union is computable, and the measure of the remaining tail can be made arbitrarily small).

As in the proof of Lemma A.5, it follows that for all $x \in K$ and all $k \geq 2 - \log_2 \varepsilon$, that $x \in U_k$ and is not on the boundaries of the relevant cells. Therefore

$$d_{\mathbb{Y}}(\varphi_n(x), \varphi_{n_k}(x)) \leq \sum_{j \geq k} 2^{-(j+1)} \leq 2^{-k}.$$

Use this to compute the value of $g(x) := \lim_n \varphi_n(x)$ for $x \in K$. If x is Schnorr random this is equal to $\widetilde{f}(x)$.

If $\mathbb{Y} = \mathbb{R}$, then by the effective Tietze extension theorem [52], we can extend g to a total computable function. □

Restatement of Proposition 3.22 (Effective inner/outer regularity). *Given $A \subseteq (\mathbb{X}, \mu)$ effectively measurable, and some rational $\varepsilon > 0$, there is an effectively open set U and an effectively closed set C both of computable measure such that $C \subseteq \widetilde{A} \subseteq U$ for Schnorr randoms such that $\mu(U) - \mu(C) \leq \varepsilon$. (The sets U, C and their measures $\mu(U), \mu(C)$ are uniformly computable from ε and any name for A.)*

Proof. From the effective Lusin's theorem (Proposition 3.21), we can choose an effectively closed K of computable measure $\mu(K) \geq 1 - \varepsilon$ and a computable function $g \colon K \to \{0,1\}$ such that $\widetilde{\mathbf{1}_A} \upharpoonright K = g$ on Schnorr randoms. Then let $C = \{x \in K \mid g(x) = 1\}$ and $U = \mathbb{X} \smallsetminus \{x \in K \mid g(x) = 0\}$. These are effectively closed and open. Then $C \subseteq \widetilde{A} \subseteq U$ for Schnorr randoms since , and $\mu(U) - \mu(C) = 1 - \mu(K) \leq \varepsilon$.

The measures $\mu(C)$ and $\mu(U)$ are computable as follows. From a name for g, we can enumerate a sequence of balls $\{B^0_i\}$ and $\{B^1_j\}$ from $\mathit{Basis}(\mathbb{X}, \mu)$ such that if $x \in B^0_i$ and $x \in K$ then $f(x) = 0$ and similarly for B^1_i. Notice $\bigcup_i B^0_i \cup \bigcup_j B^1_j$ covers K.

Let $V = K^c$ and enumerate a sequence of balls $\{A_i\}$ from $\mathit{Basis}(\mathbb{X}, \mu)$ such that $V = \bigcup_i A_i$ and hence $\mathbb{X} = \bigcup_i B^0_i \cup \bigcup_j B^1_j \cup \bigcup_k A_k$. Find a finite subsequence of these balls such that

$$\mu(B^0_1 \cup \ldots \cup B^0_\ell \cup B^1_1 \cup \ldots \cup B^1_n \cup A_1 \cup \ldots \cup A_m) \approx 1.$$

Then $\mu(C) \approx \mu((B^1_1 \cup \ldots \cup B^1_n) \smallsetminus (A_1 \cup \ldots \cup A_m))$ and $\mu(U) \approx 1 - \mu((B^0_1 \cup \ldots \cup B^0_\ell) \smallsetminus (A_1 \cup \ldots \cup A_m))$. □

Restatement of Proposition 3.24 (Schnorr layerwise computability). *Consider a (pointwise-defined) measurable function $f \colon \mathbb{X} \to \mathbb{Y}$ that is* Schnorr layerwise computable*, that is, there is a computable sequence (C_n) of effectively closed sets of computable measure $\mu(C_n) \leq 2^{-n}$, such that $f \upharpoonright C_n$ is computable on C_n uniformly in n. Then there is an effectively measurable $g \colon (\mathbb{X}, \mu) \to \mathbb{Y}$ such that $\widetilde{g} = f$ on Schnorr randoms.*

Proof. Fix $\varepsilon > 0$. Choose C_n such that $\mu(C_n) \geq 1 - \varepsilon$. From a name for $f \upharpoonright C_n$, we can enumerate a sequence of balls $\{B_i\}_i$ from $\mathit{Basis}(\mathbb{X}, \mu)$ and values c_i for which if $x \in B_i$ and $x \in C_n$ then $d_{\mathbb{Y}}(f(x), c_i) \leq \varepsilon$. Note that $\{B_i\}_i$ covers C_n, so we can compute a subsequence $B_0, \ldots, B_{k-1}$ such that $\mu(B_0, \ldots, B_{k-1}) \geq 1 - 2\varepsilon$.

Let φ be the test function made from all cells of $B_0, \ldots, B_{k-1}$ (except the cell $B^c_0 \cup \ldots \cup B^c_{k-1}$). Use the approximations c_i to determine the value of φ on each cell. Then $d_{\mathbb{Y}}(\varphi(x), f(x)) \leq \varepsilon$ unless $x \notin C_n$ or $x \notin B_0 \cup \ldots \cup B_{k-1}$. Therefore,

$$d_{meas}(\varphi, f) \leq \varepsilon + (1 - \mu(C_n)) + (1 - \mu(B_0 \cup \ldots \cup B_{k-1})) \leq 4\varepsilon.$$

Hence, f is almost-everywhere equal to an effectively measurable function g with Cauchy name φ. Moreover, $\varphi_n(x) \to f(x)$ for all x in all but finitely-many

C_n. This is true of all Schnorr randoms x, since (C_n^c) forms a Solovay test for Schnorr randomness. ☐

Restatement of Proposition 3.25 (Examples of effectively measurable functions and sets). *All of these functions $f\colon \mathbb{X} \to \mathbb{Y}$ and sets $A \subseteq \mathbb{X}$ are effectively measurable, and $\widetilde{f} = f$ and $\widetilde{A} = A$ on Schnorr randoms.*

1. *Test functions and test sets as in Propositions 3.1 and 3.3 and in Definition 3.8.*
2. *Computable functions and decidable sets (i.e., computable 0,1-valued functions).*
3. *Almost-everywhere computable functions $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ and almost-everywhere decidable sets (i.e., almost everywhere computable 0,1-valued functions).*
4. *Nonnegative lower semicomputable functions $f\colon \mathbb{X} \to \mathbb{R}$ with a computable integral, effectively open sets $U \subseteq \mathbb{X}$ of computable measure, and effectively closed sets $C \subseteq \mathbb{X}$ of computable measure.*

Proof. (1): This is obvious from the definition of effectively measurable and of $\widetilde{f}$.

(2): See (3).

(3): We will show that almost-everywhere computable functions are Schnorr layerwise computable. From a name for f from n, we can enumerate a sequence of balls $\{B_i^n\}_i$ from $Basis(\mathbb{X}, \mu)$ and values c_i^n for which if $x \in B_i^n$ then $d_{\mathbb{Y}}(f(x), c_i) \leq 2^{-n}$. Moreover, $\mu(\bigcup_i B_i^n) = 1$. Choose ε. For each n, find a subsequence $(B_0^n, \ldots, B_{k(n)-1}^n)$ such that $\mu(B_0^n \cup \ldots \cup B_{k(n)-1}^n) \geq 1 - \varepsilon/2^n$. Then let $C_\varepsilon = \bigcap_n (\overline{B}_0^n \cup \ldots \cup \overline{B}_{k(n)-1}^n)$.

I will show that C_ε is an effectively closed set of computable measure $\mu(C_\varepsilon) \geq 1 - 2\varepsilon$ such that f is computable on C_ε. It is clearly effectively closed. It has computable measure since $\mu(\bigcap_{n \leq m} (\overline{B}_0^n \cup \ldots \cup \overline{B}_{k(n)-1}^n))$ is computable and

$$\mu(C_\varepsilon) - \mu\left(\bigcap_{m>n} (\overline{B}_0^n \cup \ldots \cup \overline{B}_{k(n)-1}^n)\right) \leq \sum_{m>n} \left(1 - \mu(\overline{B}_0^n \cup \ldots \cup \overline{B}_{k(n)-1}^n)\right) \leq \sum_{m>n} \varepsilon/2^n = \varepsilon/2^m.$$

Similarly, $1 - \mu(C_\varepsilon) \leq \sum_n (1 - \mu(\overline{B}_0^n \cup \ldots \cup \overline{B}_{k(n)-1}^n)) \leq \sum_n \varepsilon/2^n \leq 2\varepsilon$. Finally, f is computable on C_ε since for any n and x in C_ε we can wait until $x \in B_i^n$ for some i, and we know that $f(x)$ is within 2^{-n} of c_i^n.

(4): Let $f = \sup g_n$ where (g_n) is a computable sequence of computable functions. Then $\|f - g_n\|_{L^1} = \int f - g_n \, d\mu$ and from monotonicity we can compute an effective rate of a.e. convergence of g_n to f. Therefore f is effectively measurable and $\widetilde{f} = \lim_n \widetilde{g_n} = \lim_n g_n = f$. For effectively open U of computable measure, just use $f = \mathbf{1}_U$ which is lower semi computable. The same for effectively closed C of computable measure. ☐

Restatement of Proposition 3.26 (Push-forward measures). *Iff $\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ is effectively measurable, then the push-forward measure $(\mathbb{Y}, \mu_* f)$ is a computable probability space (uniformly from $(\mathbb{X}, \mu)$, $\mathbb{Y}$, and f).*

Proof. It is enough to compute $\int \varphi \, d\mu_* f = \int \varphi \circ f \, d\mu$ uniformly from a computable function $\varphi\colon \mathbb{Y} \to [0, 1]$. By the effective Lusin's theorem (Proposition 3.21) $\widetilde{f}$ is Schnorr layerwise computable. Since φ is a computable function, we have that $\varphi \circ \widetilde{f}$ is Schnorr layerwise (since from the definition of Schnorr layerwise computable, the composition of a computable function with a Schnorr layerwise computable function is still Schnorr layerwise computable). By Proposition 3.24, $\varphi \circ f$ is effectively measurable. Since $\varphi \circ f$ is effectively measurable and bounded, the integral $\int \varphi \circ f \, d\mu$ is computable (Proposition 3.20). $\square$

Restatement of Proposition 3.27 (Preservation of Schnorr randomness). *If $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ is effectively measurable and x is Schnorr random, then $\widetilde{f}(x)$ is Schnorr random on $(\mathbb{Y}, \mu_* f)$.*

Proof. Assume not. Let (U_n) be a $(\mathbb{Y}, \mu_* f)$-Schnorr test which covers $\widetilde{f}(x)$. Let $g = \sum_n \mathbf{1}_{U_n}$. Then g is a lower semicomputable function and hence $g = \sup_n \varphi_n$ for a computable sequence of computable functions. We can also assume that $0 \le \varphi_n \le n$. By the same argument as in the previous proof, $\varphi_n \circ f$ is effectively measurable uniformly in n and $\widetilde{\varphi_n \circ f} = \varphi_n \circ \widetilde{f}$ on $(\mathbb{X}, \mu)$-Schnorr randoms x.

Moreover, we can show that $\varphi_n \circ f \to g \circ f$ effectively in measure since $d_{meas}(\varphi_n \circ f, g \circ f) = d_{meas}(\varphi_n, g) \le \|g - \varphi_n\|_{L^1}$ which is computable since $\int g \, d\mu_* f = \sum_n \mu_* f(U_n)$ is computable and $\int \varphi_n \, d\mu_* f$ is computable since φ_n is computable and bounded. Restricting to a subsequence (n_k) we have that $\varphi_{n_k} \circ f \to g \circ f$ converges effectively a.e. By Lemma 3.19, $(\varphi_{n_k} \circ \widetilde{f})(x)$ must converge (to something in $\mathbb{R}$) since x is Schnorr random. However, $\lim_k (\varphi_{n_k} \circ \widetilde{f})(x) = \infty \notin \mathbb{R}$. $\square$

Restatement of Proposition 3.28 (Composition and tuples).

(1) *(Composition) Given $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ and $g\colon (\mathbb{Y}, \mu_* f) \to \mathbb{Z}$ effectively measurable, the composition $g \circ f$ is effectively measurable (uniformly from f and g) and*

$$\widetilde{f \circ g} = \widetilde{f} \circ \widetilde{g} \quad \textit{(on Schnorr randoms).}$$

(2) *(Tuples) Given $f_n\colon (\mathbb{X}, \mu) \to \mathbb{Y}_n$ effectively measurable (uniformly in n), the tuples*

$$(f_0, \ldots, f_{k-1})\colon (\mathbb{X}, \mu) \to \mathbb{Y}_0 \times \cdots \times \mathbb{Y}_{k-1}$$

and

$$(f_n)_{n \in \mathbb{N}}\colon (\mathbb{X}, \mu) \to \prod_{n \in \mathbb{N}} \mathbb{Y}_n$$

are effectively measurable (uniformly from (f_n)) and

$$\widetilde{(f_0, \ldots, f_{k-1})} = (\widetilde{f}_0, \ldots, \widetilde{f}_{k-1}) \quad \textit{and} \quad \widetilde{(f_i)_{i \in \mathbb{N}}} = (\widetilde{f}_i)_{i \in \mathbb{N}} \quad \textit{(on Schnorr randoms).}$$

Proof. (1): Consider f and g with Cauchy-names, (φ_n) and (ψ_n). First we show that $\psi_n \circ f$ is effectively measurable uniformly in f and ψ_n. Fix $\varepsilon > 0$. We can effectively choose some small $\varepsilon' > 0$ such that all but a small μ-measure of x are more than ε' from the boundary of the cells which make up ψ_n. Then choose φ_m such that $\mu\{d_{\mathbb{Y}}(\varphi_m, f) > \varepsilon'\} < \varepsilon$. Outside of this bad set, we have the $\varphi_m(x)$ and $f(x)$ are in the same cell of $\psi_n(x)$, and hence $\psi_n \circ \varphi_m - \psi_n \circ f = 0$. Hence $d_{meas}(\varphi_m, f) \le \varepsilon$. Therefore, $\psi_n \circ \varphi_m \xrightarrow[m \to \infty]{} \psi_n \circ f$ effectively a.e. and therefore $\psi_n \circ f$ is effectively measurable uniformly from f and ψ_n. (This required that $\psi_n \circ \varphi_m$ is a test function. To ensure this, one may need to slightly modify φ_m to avoid hitting the boundary of the cells in ψ_n.) By Lemma 3.19, $\widetilde{\psi_n \circ f} = \lim_m \psi_n \circ \varphi_m = \psi_n \circ \widetilde{f}$ on Schnorr randoms.

Next, we show that $g \circ f$ is effectively measurable uniformly in f and g. This is straightforward since $d_{meas}(g \circ f, \psi_n \circ f) = d_{meas}(g, \psi_n)$. Moreover, by Lemma 3.19,

$\widetilde{g \circ f} = \lim_n \widetilde{\psi_n \circ f} = \lim_n \psi_n \circ \widetilde{f} = \widetilde{g} \circ \widetilde{f}$ on Schnorr randoms x (since $\widetilde{f}(x)$ is Schnorr random).

(2): I just do the infinite case. Let $\mathbb{Y} = \prod_{n\in\mathbb{N}} \mathbb{Y}_n$ with metric $d_{\mathbb{Y}} = \sum_{n\in\mathbb{N}} 2^{-(n+1)} \min\{d_{\mathbb{Y}_n}, 1\}$. For each n, let $(\varphi_n^j)_{j\in\mathbb{N}}$ be the Cauchy-name for f_n in the metric d_{meas}. Then approximate $f = (f_n)_{n\in\mathbb{N}}$ with $\psi_k = (\varphi_n^{2^{k+n+2}})_{n\in\mathbb{N}}$. Then

$$d_{meas}(f, \psi_k) = \int d_{\mathbb{Y}}(f, \psi_k)\, d\mu = \sum_{n\in\mathbb{N}} 2^{-(n+1)} \int \min\left\{ d_{\mathbb{Y}_n}(f_n, \varphi_n^{2^{k+n+2}}), 1\right\} d\mu \le 2^{-k}.$$

Therefore f is effectively measurable and $\widetilde{(f_i)_{i\in\mathbb{N}}} = \lim_{j\to\infty} (\varphi_i^j)_{i\in\mathbb{N}} = (\widetilde{f_i})_{i\in\mathbb{N}}$. □

Restatement of Proposition 3.29 (Combinations of measurable functions).

(1) *(Computable pointwise operations). All computable pointwise operations, including vector, lattice, and Boolean algebra operations preserve effective measurability. Moreover, given* $f, g\colon (\mathbb{X}, \mu) \to \mathbb{R}$ *and* $A, B \subseteq (\mathbb{X}, \mu)$ *effectively measurable, we have*

$$\widetilde{f+g} = \widetilde{f} + \widetilde{g}, \quad \widetilde{af} = a\widetilde{f}, \quad \widetilde{f\cdot g} = \widetilde{f} \cdot \widetilde{g}$$
$$\widetilde{\min(f,g)} = \min(\widetilde{f}, \widetilde{g}), \quad \widetilde{\max(f,g)} = \max(\widetilde{f}, \widetilde{g}), \quad \widetilde{|f|} = \left|\widetilde{f}\right|$$
$$\widetilde{A \cup B} = \widetilde{A} \cup \widetilde{B}, \quad \widetilde{A \cap B} = \widetilde{A} \cap \widetilde{B}, \quad \widetilde{A^c} = \widetilde{A}^c, \quad \widetilde{\mathbb{X}} = \mathbb{X}, \quad \widetilde{\varnothing} = \varnothing$$

on Schnorr randoms, and

$$f \le g \text{ a.e.} \quad \Leftrightarrow \quad \widetilde{f} \le \widetilde{g} \text{ (on Schnorr randoms)}$$
$$A \subseteq B \text{ a.e.} \quad \Leftrightarrow \quad \widetilde{A} \subseteq \widetilde{B} \text{ (on Schnorr randoms).}$$

(2) *(Inverse image) Given* $f\colon (\mathbb{X}, \mu) \to \mathbb{Y}$ *and* $B \subseteq (\mathbb{Y}, \mu_* f)$ *effectively measurable then* $f^{-1}(B)$ *is effectively measurable and* $\widetilde{f^{-1}(B)} = \widetilde{f}^{-1}(\widetilde{B})$ *on Schnorr randoms.*

(3) *(Rotations) Given* $f\colon (\mathbb{T}^d, \lambda) \to \mathbb{R}$ *effectively measurable, and a computable vector* $t \in \mathbb{T}^d$*, then* $h(x) := f(x - t)$ *is effectively measurable and* $\widetilde{h}(x) = \widetilde{f}(x - t)$ *on Schnorr randoms.*

(4) *(Indicator functions) Given* $A \subseteq (\mathbb{X}, \mu)$*,* A *is effectively measurable if and only if* $\mathbf{1}_A\colon (\mathbb{X}, \mu) \to \mathbb{R}$ *is effectively measurable (equivalently,* L^1*-computable by Proposition 3.20 (2)) and* $x \in \widetilde{A}$ *if and only if* $\widetilde{\mathbf{1}_A}(x) = 1$ *on Schnorr randoms. (Notice the codomain of* $\mathbf{1}_A$ *is* $\mathbb{R}$ *here rather than* $\{0, 1\}$ *as in Definition 3.17.)*

Proof. (1): This is a direct application of Propositions 3.25 and 3.28. Also if $f \le g$ a.e., then $g - f = \max\{g - f, 0\}$ a.e., and $\widetilde{g} - \widetilde{f} = \widetilde{g - f} = \widetilde{\max\{g - f, 0\}} = \max\{\widetilde{g} - \widetilde{f}, 0\} \ge 0$ on Schnorr randoms. Similarly for $A \subseteq B$.

(2): Use that $\mathbf{1}_{f^{-1}(B)} = \mathbf{1}_B \circ f$. The rest follows from Proposition 3.28.

(3): Let $g(x) := x - t$. Then g is computable and measure preserving, that is $\lambda_* g = \lambda$. Hence, $h = f \circ g$, and $\widetilde{h} = \widetilde{f} \circ g$ by Propositions 3.25 and 3.28.

(4): Consider the computable inclusion map $i\colon \{0, 1\} \to \mathbb{R}$. We have $(\mathbf{1}_A\colon (\mathbb{X}, \mu) \to \mathbb{R}) = i \circ (\mathbf{1}_A\colon (\mathbb{X}, \mu) \to \{0, 1\})$. By Proposition 3.28, if A is effectively measurable then $\mathbf{1}_A\colon (\mathbb{X}, \mu) \to \mathbb{R}$ is. For the other direction, if $\mathbf{1}_A\colon (\mathbb{X}, \mu) \to \mathbb{R}$ is effectively measurable, then consider the partial computable map $g\colon \mathbb{R} \to \{0, 1\}$ which sends

$0 \mapsto 0$ and $1 \mapsto 1$. This maps is almost-everywhere computable on $(\mathbb{R}, \mu_* \mathbf{1}_A)$ (which only has mass on 0 and 1). Now, $(\mathbf{1}_A : (\mathbb{X}, \mu) \to \{0,1\}) = g \circ (\mathbf{1}_A : (\mathbb{X}, \mu) \to \mathbb{R})$. The rest follows from Proposition 3.28. □

Restatement of Proposition 3.30. *The following implications hold for real-valued functions (and all the computations are uniform).*

1. *If $f \in L^1_{comp}$ and A is effectively measurable, then $\int_A f \, d\mu$ is computable.*
2. *If $\mathbb{X}$ is effectively compact (see* [35]*)—as is $[0,1]^d$, $\mathbb{T}^d$, and $2^{\mathbb{N}}$—and $g \colon \mathbb{X} \to \mathbb{R}$ is computable, then g is L^1-computable (since it has computable bounds).*
3. *If $f \colon (\mathbb{X}, \mu) \to \mathbb{Y}$ is effectively measurable and $g \in L^1_{comp}(\mathbb{Y}, \mu_* f)$ (resp. $L^2_{comp}(\mathbb{Y}, \mu_* f)$), then $g \circ f \in L^1_{comp}(\mathbb{X}, \mu)$ (resp. $L^2_{comp}(\mathbb{X}, \mu)$).*

Proof. (1): By Proposition 3.29 (4), $\mathbf{1}_A \in L^1_{comp}$. Then use 3.20 (7).

(2): By Proposition 3.25, g is effectively measurable. Since $\mathbb{X}$ is effectively compact, $\max_{x \in \mathbb{X}} g(x)$ and $\min_{x \in \mathbb{X}} g(x)$ are computable from g [35]. Now apply Proposition 3.20 (2).

(3): By Proposition 3.28 $g \circ f \in L^0_{comp}$, and moreover $\|g \circ f\|_{L^1(\mathbb{Y}, \mu_* f)} = \|g\|_{L^1(\mathbb{X}, \mu)}$ (similarly for L^2). Apply Proposition 3.20 (3). □

Restatement of Proposition 3.31. *Given a measurable map $f \colon (\mathbb{X}, \mu) \to \mathbb{Y}$, the following are equivalent.*

1. *f is effectively measurable.*
2. *The push-forward measure $(\mathbb{Y}, \mu_* f)$ is computable and one (or all) of the following "pull-back" maps are computable:*
   (a) *(L^1 functions) $g \in L^1(\mathbb{Y}, \mu_* f) \;\mapsto\; g \circ f \in L^1(\mathbb{Y}, \mu_* f)$.*
   (b) *(L^2 functions) $g \in L^2(\mathbb{Y}, \mu_* f) \;\mapsto\; g \circ f \in L^2(\mathbb{Y}, \mu_* f)$.*
   (c) *(Measurable sets) $B \subseteq (\mathbb{Y}, \mu_* f) \;\mapsto\; f^{-1}(B) \subseteq (\mathbb{X}, \mu)$.*

Proof. (1) ⇒ (2) follows from Propositions 3.26, 3.28, 3.29, and 3.30.

For the other direction, assume (2)(a) or (2)(b). Then $B \mapsto \mathbf{1}_B \mapsto \mathbf{1}_B \circ f \mapsto f^{-1}(B)$ is a chain of computable operators (using Proposition 3.29 4), and therefore (2)(c) holds.

For (2)(c) ⇒ (1), fix $\varepsilon > 0$. Since $(\mathbb{Y}, \mu_* f)$ is computable, effectively choose finitely many balls $B_0, \ldots, B_{k-1}$ of radius at most $\varepsilon/2$ from $\mathit{Basis}(\mathbb{Y}, \mu_* f)$ such that $\mu_* f(B_0 \cup \cdots \cup B_{k-1}) \geq 1 - \varepsilon/2$. Let $C_0, \ldots, C_{2^k - 1}$ be the cells formed by combining the elements of $B_0, \ldots, B_{k-1}$. Let C_0 denote the cell $B_0^c \cap \cdots \cap B_{k-1}^c$ which is the only cell without a diameter bounded by $\varepsilon/2$. For $i \geq 1$, effectively choose a point y_i inside the cell C_i (by choosing the center of the lowest indexed ball B_j for which $C_i \subseteq B_j$). Let $A_i = f^{-1}(C_i)$. By assumption, these are effectively measurable. Define $\varphi \colon (\mathbb{X}, \mu) \to \mathbb{Y}$ as the effectively measurable function which has value y_i on A_i for $i \geq 1$ and 0 otherwise. Notice on A_i ($1 \leq i \leq 2^k - 1$), that φ and f both take values in C_i.

$$\begin{aligned} d_{meas}(f, \varphi) &= \int \max\{d_{\mathbb{Y}}(f, \varphi), 1\} \, d\mu \\ &\leq 1 \cdot \mu(A_0) + \sum_{i=1}^{2^k - 1} \int_{A_i} d_{\mathbb{Y}}(f, \varphi) \, d\mu \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \cdot \sum_{i=1}^{k-1} \mu(A_i) \leq \varepsilon. \end{aligned}$$

Hence f is effectively measurable. □

References

[1] Jeremy Avigad, Edward T. Dean, and Jason Rute. Algorithmic randomness, reverse mathematics, and the dominated convergence theorem. *Ann. Pure Appl. Logic*, 163(12):1854–1864, 2012.

[2] Jeremy Avigad, Philipp Gerhardy, and Henry Towsner. Local stability of ergodic averages. *Trans. Amer. Math. Soc.*, 362(1):261–288, 2010.

[3] Laurent Bienvenu, Adam R. Day, Mathieu Hoyrup, Ilya Mezhirov, and Alexander Shen. A constructive version of Birkhoff's ergodic theorem for Martin-Löf random points. *Inform. and Comput.*, 210:21–30, 2012.

[4] Errett Bishop. *Foundations of constructive analysis.* McGraw-Hill Book Co., New York, 1967.

[5] Volker Bosserhoff. Notions of probabilistic computability on represented spaces. *J. Universal Computer Science*, 14(6):956–995, 2008.

[6] Vasco Brattka, Peter Hertling, and Klaus Weihrauch. A tutorial on computable analysis. In *New computational paradigms*, pages 425–491. Springer, New York, 2008.

[7] Vasco Brattka, Joseph S. Miller, and André Nies. Randomness and differentiability. Submitted.

[8] Claude Dellacherie and Paul-André Meyer. *Probabilities and potential B: Theory of martingales*, volume 72 of *North-Holland Mathematics Studies.* North-Holland Publishing Co., Amsterdam, 1982. Translated from the French by J. P. Wilson.

[9] O. Demut. The differentiability of constructive functions of weakly bounded variation on pseudo numbers. *Comment. Math. Univ. Carolinae*, 16(3):583–599, 1975.

[10] Rodney G. Downey and Evan J. Griffiths. Schnorr randomness. *J. Symbolic Logic*, 69(2):533–554, 2004.

[11] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity.* Theory and Applications of Computability. Springer, New York, 2010.

[12] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[13] Rick Durrett. *Probability: theory and examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[14] Abbas Edalat. A computable approach to measure and integration theory. *Inform. and Comput.*, 207(5):642–659, 2009.

[15] Johanna N. Y. Franklin, Noam Greenberg, Joseph S. Miller, and Keng Meng Ng. Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. *Proc. Amer. Math. Soc.*, 140(10):3623–3628, 2012.

[16] Johanna N.Y. Franklin and Henry Towsner. Randomness and non-ergodic systems. Submitted.

[17] Cameron E. Freer, Bjørn Kjos-Hanssen, André Nies, and Frank Stephan. Algorithmic aspects of Lipschitz functions. Submitted.

[18] Cameron E. Freer and Daniel M. Roy. Computable de Finetti measures. *Annals of Pure and Applied Logic*, 2011.

[19] Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoret. Comput. Sci.*, 341(1-3):91–137, 2005.

[20] Peter Gács, Mathieu Hoyrup, and Cristóbal Rojas. Randomness on computable probability spaces—a dynamical point of view. *Theory Comput. Syst.*, 48(3):465–485, 2011.

[21] Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. Computing the speed of convergence of ergodic averages and pseudorandom points in computable dynamical systems. In Xizhong Zheng and Ning Zhong, editors, Proceedings Seventh International Conference on *Computability and Complexity in Analysis,* Zhenjiang, China, 21-25th June 2010, volume 24 of *Electronic Proceedings in Theoretical Computer Science*, pages 7–18. Open Publishing Association, 2010.

[22] Mathieu Hoyrup. Computability of the ergodic decomposition. *Annals of Pure and Applied Logic*, 164(5):542–549, 2013. Computability in Europe 2011.

[23] Mathieu Hoyrup and Cristóbal Rojas. An application of Martin-Löf randomness to effective probability theory. In *Mathematical theory and computational practice*, volume 5635 of *Lecture Notes in Comput. Sci.*, pages 260–269. Springer, Berlin, 2009.

[24] Mathieu Hoyrup and Cristóbal Rojas. Applications of effective probability theory to Martin-Löf randomness. In *Automata, languages and programming. Part I*, volume 5555 of *Lecture Notes in Comput. Sci.*, pages 549–561. Springer, Berlin, 2009.

[25] Mathieu Hoyrup and Cristóbal Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Inform. and Comput.*, 207(7):830–847, 2009.

[26] Mathieu Hoyrup and Cristóbal Rojas. Absolute continuity of measures and preservation of randomness. Preprint.

[27] Mathieu Hoyrup, Cristóbal Rojas, and Klaus Weihrauch. Computability of the Radon-Nikodym derivative. In *Models of computation in context*, volume 6735 of *Lecture Notes in Comput. Sci.*, pages 132–141. Springer, Heidelberg, 2011.

[28] Tahereh Jafarikhah, Massoud Pourmahdian, and Klaus Weihrauch. Computable Jordan decomposition of linear continuous functionals on $C[0;1]$. In Mathieu Hoyrup, Ker-I Ko, Robert Rettinger, and Ning Zhong, editors, *CCA 2013 Tenth International Conference on Computability and Complexity in Analysis*, volume 367-7/2013, pages 48–59. Informatik Berichte FernUniversität in Hagen, 2013.

[29] Yitzhak Katznelson. *An introduction to harmonic analysis*. John Wiley & Sons Inc., New York, 1968.

[30] Ker-I Ko and Harvey Friedman. Computational complexity of real functions. *Theoret. Comput. Sci.*, 20(3):323–352, 1982.

[31] Hong Lu and Klaus Weihrauch. Computable Riesz representation for the dual of $C[0;1]$. *MLQ Math. Log. Q.*, 53(4-5):415–430, 2007.

[32] Wolfgang Merkle, Nenad Mihailović, and Theodore A. Slaman. Some results on effective randomness. *Theory Comput. Syst.*, 39(5):707–721, 2006.

[33] Kenshi Miyabe. L^1-computability, layerwise computability and Solovay reducibility. *Computability*. To appear.

[34] Michał Morayne and Sławomir Solecki. Martingale proof of the existence of Lebesgue points. *Real Anal. Exchange*, 15(1):401–406, 1989/90.

[35] Takakazu Mori, Yoshiki Tsujii, and Mariko Yasugi. Computability structures on metric spaces. In *Combinatorics, complexity, & logic (Auckland, 1996)*, Springer Ser. Discrete Math. Theor. Comput. Sci., pages 351–362. Springer, Singapore, 1997.

[36] Satyadev Nandakumar. An effective ergodic theorem and some applications. In *STOC'08*, pages 39–44. ACM, New York, 2008.

[37] André Nies. *Computability and randomness*, volume 51 of *Oxford Logic Guides*. Oxford University Press, Oxford, 2009.

[38] Noopur Pathak. A computational aspect of the Lebesgue differentiation theorem. *J. Log. Anal.*, 1(9):15, 2009.

[39] Noopur Pathak, Cristóbal Rojas, and Stephen G. Simpson. Schnorr randomness and the Lebesgue differentiation theorem. *Proceedings of the American Mathematical Society*. To appear.

[40] Marian B. Pour-El and J. Ian Richards. *Computability in analysis and physics*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1989.

[41] Jason Rute. Computable randomness and betting for computable probability spaces. *Submitted*.

[42] N. A. Šanin. *Constructive real numbers and constructive function spaces*. Translated from the Russian by E. Mendelson. Translations of Mathematical Monographs, Vol. 21. American Mathematical Society, Providence, R.I., 1968.

[43] Claus-Peter Schnorr. *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*. Lecture Notes in Mathematics, Vol. 218. Springer-Verlag, Berlin, 1971.

[44] Matthias Schröder. Admissible representations for probability measures. *MLQ Math. Log. Q.*, 53(4-5):431–445, 2007.

[45] Robert I. Soare. *Recursively enumerable sets and degrees*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1987. A study of computable functions and computably generated sets.

[46] Hayato Takahashi. Bayesian approach to a definition of random sequences with respect to parametric models. In M.J. Dinneen, editor, *Proc. of IEEE ISOC ITW2005 on Coding and Complexity*, pages 217–220, 2005.
[47] Terence Tao. *An introduction to measure theory*, volume 126 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2011.
[48] Michael E. Taylor. *Measure theory and integration*, volume 76 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006.
[49] V. V. V'yugin. Ergodic theorems for individual random sequences. *Theoret. Comput. Sci.*, 207(2):343–361, 1998.
[50] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982.
[51] Klaus Weihrauch. *Computable analysis: An introduction*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2000.
[52] Klaus Weihrauch. On computable metric spaces Tietze-Urysohn extension is computable. In *Computability and complexity in analysis (Swansea, 2000)*, volume 2064 of *Lecture Notes in Comput. Sci.*, pages 357–368. Springer, Berlin, 2001.
[53] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.
[54] Yongcheng Wu and Decheng Ding. Computability of measurable sets via effective topologies. *Arch. Math. Logic*, 45(3):365–379, 2006.
[55] Yongcheng Wu and Klaus Weihrauch. A computable version of the Daniell-Stone theorem on integration and linear functionals. *Theoret. Comput. Sci.*, 359(1-3):28–42, 2006.
[56] Xiaokang Yu. *Measure theory in weak subsystems of second-order arithmetic*. ProQuest LLC, Ann Arbor, MI, 1987. Thesis (Ph.D.)–The Pennsylvania State University.
[57] Xiaokang Yu. Lebesgue convergence theorems and reverse mathematics. *Math. Logic Quart.*, 40(1):1–13, 1994.
[58] Xiaokang Yu and Stephen G. Simpson. Measure theory and weak König's lemma. *Arch. Math. Logic*, 30(3):171–180, 1990.