RANDOMNESS AND INITIAL SEGMENT COMPLEXITY FOR MEASURES

ANDRÉ NIES AND FRANK STEPHAN

Abstract. We study algorithmic randomness properties for probability measures on Cantor space. We say that a measure μ on the space of infinite bit sequences is Martin-Löf absolutely continuous if the non-Martin-Löf random bit sequences form a null set with respect to μ. We think of this as a weak randomness notion for measures.

We begin with examples, and provide a robustness property related to Solovay tests. The initial segment complexity of a measure μ at a length n is defined as the μ-average over the descriptive complexity of strings of length n, in the sense of either C or K. We relate this weak randomness notion for a measure to the growth of its initial segment complexity. We show that a maximal growth implies the weak randomness property, but also that both implications of the Levin-Schnorr theorem fail. We discuss C-triviality and K-triviality for measures and relate these two notions with each other. Here, triviality means that the initial segment complexity grows as slowly as possible.

We show that every measure that is Martin-Löf random in the sense of Hoyrup and Rojas is Martin-Löf absolutely continuous; the converse fails because only the latter property is compatible with having atoms. In a final section we consider weak randomness relative to a general ergodic computable measure. We seek appropriate effective versions of the Shannon-McMillan-Breiman theorem and the Brudno theorem where the bit sequences are replaced by measures. We conclude with several open questions.

Contents

1991 *Mathematics Subject Classification.* 03D32,68Q30.

André Nies is partially supported by the Marsden Fund of the Royal Society of New Zealand, 13-UOA-184 and 19-UOA-346. Address: Department of Computer Science, University of Auckland, Private Bag 92019, Auckland, New Zealand. Email: andre@cs.auckland.ac.nz.

Frank Stephan is partially supported by supported by the Singapore Ministry of Education Academic Research Fund Tier 2 grants MOE2016-T2-1-019 / R146-000-234-112 (2017–2020) and MOE2019-T2-2-121 / R146-000-304-112 (2020–2023). Address: Department of Mathematics and Department of Computer Science, National University of Singapore, 10 Lower Kent Ridge Road, Block S17, Singapore 117690, Republic of Singapore. Email: fstephan@comp.nus.edu.sg.

1. Introduction

The theory of algorithmic randomness is usually developed for bit sequences, or equivalently, reals in the unit interval. A central randomness notion based on algorithmic tests is the one due to Martin-Löf [17].

Let $\{0,1\}^{\mathbb{N}}$ denote the topological space of infinite bit sequences. A probability measure μ on $\{0,1\}^{\mathbb{N}}$ can be seen as a statistical superposition of bit sequences. A single bit sequence Z forms an extreme case: the corresponding measure μ is the Dirac measure δ_Z, i.e., μ is concentrated on $\{Z\}$. At the opposite extreme is the uniform measure λ which independently gives each bit value the probability $1/2$. The uniform measure represents maximum disorder as no bit sequence is preferred over any other.

Recall that a measure μ on $\{0,1\}^{\mathbb{N}}$ is called absolutely continuous if each λ-null set is a μ-null set. We introduce Martin-Löf absolute continuity, an algorithmic randomness notion for probability measures that is a *weakening* of absolute continuity: we require that the λ-null set in the hypothesis be effective in the sense of Martin-Löf. Given that there is a universal Martin-Löf test, and hence a largest effective null set, all we have to require is that $\mu(\mathcal{C}) = 0$ where $\mathcal{C}$ is the class of bit sequences that are not Martin-Löf random. (By a "measure" we mean a probability measure on Cantor space unless otherwise stated.)

Our research is motivated in part by a recent definition of Martin-Löf randomness for quantum states corresponding to infinitely many qubits, due to the first author and Scholz [26]. Let M_n be the algebra of complex $2^n \times 2^n$ matrices. Using the terminology of [26], measures correspond to the quantum states ρ where the matrix $\rho \upharpoonright_{M_n}$ is diagonal for each n. For such diagonal states, Tejas Bhojraj [3, Section 2.3], starting from an early announcement of the conference version of the present work [28], has shown that for measures, the randomness notion defined there is equivalent to the one proposed here. So, measures form a useful testing ground for conjectures in the subtler setting of quantum states. This applies, for instance, to the Shannon-McMillan-Breiman (SMB) theorem discussed below, which has been studied in the setting of quantum states by the first author and Tomamichel (see [25, Section 6]). In Proposition 7.2 we will show that a certain effective form of the SMB theorem holds for measures; the more general case of quantum states remains open. In Proposition 7.3 we give an example based on measures which shows that [25, Conjecture 6.3] cannot hold as stated, because an additional boundedness hypothesis is missing.

Growth of the initial segment complexity. Given a binary string x, by $C(x)$ we denote its plain descriptive complexity, and by $K(x)$ its prefix-free descriptive complexity. Some of our motivation is derived from the classical theory. Randomness of infinite bit sequences is linked to the growth of the descriptive complexity of their initial segments. For instance, the Levin-Schnorr theorem intuitively says that randomness of Z means incompressibility (up to the same constant b) of all the initial segments of Z. We want to study how much of this is retained in the setting of measures μ. The formal growth condition in the Levin-Schnorr theorem says that $K(x) \geq |x| - b$ for each initial segment x of Z. The "n-th initial segment" of a measure μ is given by its values $\mu[x]$ for all strings x of length n, where $[x]$ denotes the set of infinite sequences extending x. We define the complexities $C(\mu\upharpoonright_n)$ and $K(\mu\upharpoonright_n)$ of such an initial segment as the μ-average of the individual complexities of the strings of length n. With this definition, in Section 4 we show that both implications of the analog of the Levin-Schnorr theorem fail. However by Proposition 7.4 discussed below, for measures that are random in our weak sense, $C(\mu\upharpoonright_n)/n$, and hence also $K(\mu\upharpoonright_n)/n$, converge to 1. Thus, such measures have effective dimension 1. (See Downey and Hirschfeldt [9, Section 12.3] for background on effective dimension.)

It turns that the initial complexity of measures can exhibit interesting growth behaviour that is not possible for bit sequences. We will verify in Fact 3.3 that the initial segment complexity of the uniform measure λ has maximal growth in the sense of K, namely $K(\lambda\upharpoonright_n) \geq^+ n + K(n)$. This implies that the initial complexity of λ in the sense of C also has maximal growth. For bit sequences, such a growth rate is ruled out by a result of Katseff [13]. To prove this maximal growth rate for K, we use that, in a sense to be made precise, most strings of a given length are incompressible. We also provide an example of such a measure other than λ. On the other hand, we show that a maximal growth rate in the sense of C implies ML absolute continuity.

K-trivial and C-trivial measures. Opposite to random bit sequences lie the K-trivial sequences A, where the initial segment complexity grows no faster than that of a computable set, in the sense that $K(A\upharpoonright_n) \leq^+ K(n)$. C-trivial sequences are defined analogously. Chaitin [7] proved that each C-trivial sequence is computable, while Solovay [36] showed that a K-trivial sequence can be noncomputable. For background see e.g. [24, Section 5.3]. In Section 5 we extend these notions to statistical superpositions of bit sequences: we introduce C- and K-trivial measures. We first show that such measures have a countable support. This means that they are countable convex combinations of Dirac measures (based on bit sequences that necessarily satisfy the same notion of triviality). Thereafter, again guided by the case of bit sequences, we provide closure properties of the two notions: downward closure under truth-table reducibility, and closure under products.

We leave the question open whether each C-trivial measure is K-trivial. In Prop. 5.8 we study three conditions relating C-complexity and K-complexity, and show that they are all equivalent. If these conditions hold then each C-trivial measure is K-trivial. (Whether they indeed hold was discussed at length in a working group at the 2020 American Institute of Mathematics workshop on randomness and applications. A weaker version of such a condition was considered in Bruno Bauwen's thesis [1].)

We will also consider slightly stronger forms of the notions of C- and K-triviality, for which this implication C-trivial $\rightarrow$ K-trivial holds, and is proper.

Martin-Löf random measures. Measures can be viewed as points in a canonical computable probability space, in the sense of Hoyrup and Rojas [12]. This yields a notion of Martin-Löf randomness for measures. Bienvenu and Culver (see [8, Theorem 2.7.1]) have shown that no measure μ that is Martin-Löf random is absolutely continuous; in fact μ is orthogonal to λ in the sense that some λ-null set is co-null with respect to μ. In contrast, in Section 6 we show that this notion implies our weak notion of randomness, Martin-Löf absolute continuity. The stronger randomness notion forces the measure to be atomless, so the converse implication fails. The different randomness notions for measures we have discussed beg further questions. For instance, does the randomness notion studied by Bienvenu and Culver imply that the initial segment complexity for the measure (in the sense of C or of K) is maximal infinitely often?

Shannon-McMillan-Breiman Theorem. This theorem from the 1950s says informally that for an ergodic measure ρ on $\{0,1\}^{\mathbb{N}}$, outside of an appropriate ρ-null set, every bit sequence Z knows the entropy of the measure ρ: it is the limit of the normalized information content w.r.t. ρ of initial segments of Z, namely, $-\frac{1}{n}\log_2 \rho[Z\upharpoonright_n]$. See e.g. [34], where the result is called the Entropy Theorem. In the short Section 7, we replace Z by a measure μ that is Martin-Löf a.c. with respect to ρ, and take the μ-average of the information contents of strings of the same length. We only obtain a partial analog of the theorem. We use as a hypothesis that the normalized information content of strings w.r.t. ρ is bounded, and also provide an example showing that this hypothesis is necessary. However, in a similar vein, in Proposition 7.4 we establish an analog for measures of the effective Brudno's theorem [10, 11], which states that the entropy of ρ is given as the limit of $C(Z\upharpoonright_n)/n$, for any Z that is ρ-ML-random. We show that one can take the limit of $C(\mu\upharpoonright_n)/n$ for any μ that is ML absolutely continuous in ρ.

We end the paper with conclusions and open questions.

For general background on recursion theory and algorithmic randomness we refer the readers to the textbooks of Calude [5], Downey and Hirschfeldt [9], Li and Vitányi [16], Nies [24], Odifreddi [29, 30] and Soare [35]. Lecture notes on recursion theory are also available online, e.g. [37].

A conference version of the paper has appeared in the proceedings of STACS 2020 [28]. The present paper is a much improved and expanded version. This applies in particular to Sections 3 and 5.

2. Measures and Randomness

In this section we will formally define one of our main notions, Martin-Löf absolute continuity (Definition 2.3), and collect some basic facts concerning it. In particular, we verify that the well-known equivalence of Martin-Löf test and Solovay tests extends to measures. We begin by briefly discussing algorithmic randomness for bit sequences [9, 24]. We use standard notation: letters $Z, X, \ldots$ denote elements of the space of infinite bit sequences $\{0,1\}^{\mathbb{N}}$, σ, τ denote finite bit strings, and $[\sigma] = \{Z\colon\ Z \succ \sigma\}$ is the set of infinite bit sequences extending σ. $Z\upharpoonright_n$ denotes the string consisting of the first n bits of Z. For quantities r, s depending on the same parameters, one writes $r \leq^+ s$ for $r \leq s + O(1)$. A subset G of $\{0,1\}^{\mathbb{N}}$ is called *effectively open* if $G = \bigcup_i [\sigma_i]$ for a computable sequence $\langle \sigma_i \rangle_{i\in\mathbb{N}}$ of strings. A measure ρ on $\{0,1\}^{\mathbb{N}}$ is computable if the map $\{0,1\}^{<\omega} \to \mathbb{R}$ given by $\sigma \mapsto \rho[\sigma]$ is computable. That is, each real $\rho[\sigma]$ is computable, uniformly in σ.

Definition 2.1. Let ρ be a computable measure on $\{0,1\}^{\mathbb{N}}$. A ρ-Martin-Löf test (ρ-ML-test, for short) is a sequence $\langle G_m \rangle_{m\in\mathbb{N}}$ of uniformly effectively open sets such that $\rho G_m \leq 2^{-m}$ for each m. A bit sequence Z *fails* the test if $Z \in \bigcap_m G_m$; otherwise it *passes* the test. A bit sequence Z is ρ-Martin-Löf random (ρ-ML-random for short) if Z passes each ρ-Martin-Löf test.

By λ one denotes the uniform measure on $\{0,1\}^{\mathbb{N}}$. So $\lambda[\sigma] = 2^{-|\sigma|}$ for each string σ. If no measure ρ is provided it will be tacitly assumed that $\rho = \lambda$, and we will use the term "Martin-Löf random" instead of "λ-Martin-Löf random" etc. Let $K(x)$ denote the prefix free version of descriptive (i.e., Kolmogorov) complexity of a bit string x.

Theorem 2.2 (Levin [15], Schnorr [32])**.**
Z is Martin-Löf random $\Leftrightarrow \exists b \forall n\, [K(Z\upharpoonright_n) \geq n - b]$.

Using the notation of [24, Chapter 3], let $\mathcal{R}_b$ denote the set of bit sequences Z such that $K(Z\upharpoonright_n) < n - b$ for some n. It is easy to see that $\langle \mathcal{R}_b \rangle_{b\in\mathbb{N}}$ is a Martin-Löf test. The Levin-Schnorr theorem says that this test is universal: Z is Martin-Löf random iff it passes the test.

Recall that unless otherwise stated, all measures are probability measures on Cantor space. We use the letters μ, ν, ρ for measures (and recall that λ denotes the uniform measure). We now provide the formal definition of our weak randomness notion for measures.

Definition 2.3. Let ρ be a measure. A measure μ is called *Martin-Löf absolutely continuous in* ρ if $\inf_m \mu(G_m) = 0$ for each ρ-Martin-Löf test $\langle G_m \rangle_{m\in\mathbb{N}}$. We denote this by $\mu \ll_{ML} \rho$.

If $\inf_m \mu(G_m) = 0$ we say that μ *passes* the test. If $\inf_m \mu(G_m) \geq \delta$ where $\delta > 0$ we say μ *fails the test at level* δ.

In the case that $\rho = \lambda$, we say that μ is ML absolutely continuous, for short.

Martin-Löf absolute continuity is a *weakening* of the usual notion that μ is absolutely continuous in ρ, usually written $\mu \ll \rho$. In fact, $\mu \ll \rho$ holds iff μ is ρ-ML^X absolutely continuous for each oracle X.

In the definition it suffices to consider ρ-ML tests $\langle G_m \rangle_{m\in\mathbb{N}}$ such that $G_m \supseteq G_{m+1}$ for each m, because we can replace $\langle G_m \rangle$ by the ρ-ML test given by $\widehat{G}_m = \bigcup_{k>m} G_k$, and of course $\inf_m \mu(\widehat{G}_m) = 0$ implies that $\inf_m \mu(G_m) = 0$. So we can change the definition above, replacing the condition $\inf_m G_m = 0$ by the only apparently stronger condition that $\lim_m G_m = 0$.

It is well known that there is a universal ρ-ML test. The intersection of all the components of such a test consists of the non-ρ-ML-random sequences. This implies:

Fact 2.4. *$\mu \ll_{ML} \rho$ iff the sequences that are not ρ-ML-random form a μ-null set.*

We have already mentioned the two diametrically opposite types of examples of measure that are ML absolutely continuous in a measure ρ:

Example 2.5. (a) $\rho \ll_{ML} \rho$ for each measure ρ.
(b) For a Dirac measure δ_Z, we have $\delta_Z \ll_{ML} \rho$ iff Z is ρ-ML-random.

Next we provide a source of non-examples where $\rho = \lambda$.

Example 2.6. A Bernoulli measure on $\{0,1\}^{\mathbb{N}}$ assigns a fixed probability p to a 0, independently of the position. Such a measure is not ML absolutely continuous in the case that $p \neq 1/2$. To see this, note that each Martin-Löf random sequence Z satisfies the law of large numbers

$$\lim_n \frac{1}{n} |\{i < n \colon Z(i) = 1\}| = 1/2;$$

see e.g. [24, Proposition 3.2.13]. So if μ is ML absolutely continuous, then μ-almost surely, Z satisfies the law of large numbers. This is not the case when μ is a Bernoulli measures for $p \neq 1/2$, because μ-almost surely, Z satisfies the law of large numbers for p instead of $1/2$.

Definition 2.7. For a measure ν and string σ with $\nu[\sigma] > 0$, let ν_σ denote the localisation to $[\sigma]$:

$$\nu_\sigma(A) = \nu(A \cap [\sigma])/\nu[\sigma].$$

Clearly if ν is ML absolutely continuous then so is ν_σ.

A set S of probability measures is called *convex* if $\mu_i \in S$ for $1 \leq i \leq k$ implies that the convex combination $\mu = \sum_{i=1}^k \alpha_i \mu_i$ is in S, where the α_i are reals in $[0,1]$ and $\sum_i \alpha_i = 1$. The *extreme points* of S are the ones that can only be written as convex combinations of elements of S where $k = 1$.

Proposition 2.8. *The ML absolutely continuous probability measures form a convex set. Its extreme points are the ML absolutely continuous Dirac measures, i.e. the measures δ_Z where Z is a ML-random bit sequence.*

Proof. Let $\mu = \sum_{i=1}^k \alpha_i \mu_i$ as above where the μ_i are ML absolutely continuous measures. Suppose $\langle G_m \rangle$ is a Martin-Löf test. Then $\lim_m \mu_i(G_m) = 0$ for each i, and hence $\lim_m \mu(G_m) = 0$.

Suppose that μ is Martin-Löf a.c. If μ is a Dirac measure then it is an extreme point of the Martin-Löf a.c. measures. Conversely, if μ is not Dirac, there is a least number t such that the decomposition

$$\mu = \sum_{|\sigma|=t, \mu[\sigma]>0} \mu[\sigma] \cdot \mu_\sigma$$

is nontrivial. Hence μ is not an extreme point. $\square$

Recall that a *ρ-Solovay-test*, for a computable probability measure ρ, is a sequence $\langle S_k \rangle_{k \in \mathbb{N}}$ of uniformly Σ^0_1 sets such that $\sum_k \rho(S_k) < \infty$. If $\rho = \lambda$ we will simply use the term *Solovay test.* A bit sequence Z passes such a test if $Z \notin S_k$ for almost every k. (Each Martin-Löf test is also a Solovay test, but Solovay tests have a stronger passing condition). A basic fact from the theory of algorithmic randomness (e.g. [24, Prop. 3.2.19]) states that Z is Martin-Löf random iff Z passes each Solovay test. This is usually asserted only relative to the uniform measure λ, but in fact carries over to any computable measure ρ.

The following characterises the ML absolutely continuous measures with countable support.

Fact 2.9. *Let $N \in \mathbb{N} \cup \{\infty\}$. Let $\mu = \sum_{k<N} c_k \delta_{Z_k}$ where $0 < c_k \leq 1$ for each k, and $\sum_{k<N} c_k = 1$. Then μ is ML absolutely continuous iff all the Z_k are Martin-Löf random.*

Proof. The implication from left to right is immediate. For the converse implication, given a Martin-Löf test $\langle G_m \rangle$, note that the Z_k pass this test as a Solovay test. Hence for each r, there is M such that for each $k \leq r$ and each $m \geq M$ we have $Z_k \notin G_m$. This implies that $\mu(G_m) \leq \sum_{k>r} c_k$ for each $m \geq M$. So $\lim_m \mu(G_m) = 0$. □

We say that a measure μ *passes* a Solovay test $\langle S_k \rangle_{k \in \mathbb{N}}$ if $\lim_k \mu(S_k) = 0$. The fact that passing all Martin-Löf tests is equivalent to passing all Solovay tests generalises from bit sequences to measures. We note that Tejas Bhojraj [3, Thm. 2.11] has proved such a result in even greater generality in the setting of quantum states, where the proof is considerably more involved.

Proposition 2.10. *For probability measures μ, ρ such that ρ is computable, we have*

$\mu \ll_{ML} \rho$ iff μ passes each ρ-Solovay-test.

Proof. Relative to ρ, each Martin-Löf test is a Solovay test, and the passing condition $\lim_m \mu(G_m) = 0$ works for both types of tests by the remark after Definition 2.3. This yields the implication from right to left.

For the implication from left to right, suppose that $\mu \ll_{ML} \rho$, and let $\langle S_k \rangle_{k \in \mathbb{N}}$ be a ρ-Solovay-test. By $\limsup_k S_k$ one denotes the set of bit sequences Z such that $\exists^\infty k\, [Z \in S_k]$, that is, the sequences that fail the test. By the basic fact (e.g. [24, Prop. 3.2.19]) mentioned above, the set of ρ-Martin-Löf random sequences is disjoint from $\limsup_k S_k$. Hence, by hypothesis on μ, we have $\mu(\limsup_k S_k) = 0$. By Fatou's Lemma, $\limsup_k \mu(S_k) \leq \mu(\limsup_k S_k)$. So μ passes the Solovay test. □

3. A fast growing initial segment complexity implies being ML-a.c.

In this section we formally define our two variants of initial segment complexity for measures, and first their connections with ML absolute continuity.

Recall that in this paper we use $C(x)$ to denote the plain Kolmogorov complexity of a binary string x and $K(x)$ to denote the prefix-free Kolmogorov complexity. Note that other letters are also used in the literature; for example, the textbook by Calude [5] uses K and H, a notation going back to Chaitin.

Recall the upper bounds $C(x) \leq^+ n$ and $K(x) \leq^+ n + K(n)$ for all strings x of length n. We assume for notational convenience the strict lower bounds that $C(n) \leq C(x)$ and $K(n) \leq K(x)$. This convention reduces the number of constants needed in the statements and proofs of our results. To achieve this, we replace some given universal machine V by a new universal machine U such that whenever $V(p) = x$ then $U(0p) = 0^{|x|}$ and $U(1p) = x$ with 0^n standing for n, and U is undefined on all other inputs. It is easy to see that if V is universal by adjunction, so is U; if V is prefix-free so is U. The descriptions for U are only by one bit longer than those for V and universality requires only that the length of the shortest descriptions for some x goes up at most by a constant.

Definition 3.1. Let $K(\mu\upharpoonright_n) := \sum_{|x|=n} K(x)\mu[x]$ be the μ-average of all the $K(x)$ over all strings x of length n. In a similar way we define $C(\mu\upharpoonright_n)$. We also define the descriptive complexity conditioned on n, namely $C(\mu\upharpoonright_n|\, n)$, as expected.

It may help to think of $K(\mu\upharpoonright_n)$ as $\int K_n d\mu_n$ where K_n and μ_n are the respective functions restricted to the set of strings of length n. Note that for each bit sequence Z, we have $K(\delta_Z\upharpoonright_n) = K(Z\upharpoonright_n)$.

In this section we write $K(n)$ for $K(0^n)$. We tacitly rely on standard inequalities such as $C(x) \leq^+ K(x)$, $K(x) \leq^+ |x| + 2\log|x|$ and $K(n) \leq^+ 2\log n$. We also use that for each $r \in \mathbb{N}$, there are at most $2^r - 1$ strings such that $C(x) < r$. See e.g. [24, Chapter 2].

Note that for each string x of length n, the deviations from the upper bounds are related by the inequality $n - C(x) \leq^+ 2(n + K(n) - K(x))$. For this, see e.g. [24, 2.5.5] which yields the inequality after some elementary algebraic manipulation. Taking the μ-average, this implies the corresponding inequality for the initial segment complexity of measures on Cantor space:

Fact 3.2. *Let μ be a measure. For each n one has*

$$n - C(\mu\upharpoonright_n) \leq^+ 2(n + K(n) - K(\mu\upharpoonright_n)).$$

The following says that the uniform measure λ on $\{0,1\}^{\mathbb{N}}$ has the fastest growing initial segment complexity that is possible in the sense of K, and therefore also of C.

Fact 3.3. *$K(\lambda\upharpoonright_n) \geq^+ n + K(n)$, and therefore also $C(\lambda\upharpoonright_n) \geq^+ n$.*

Proof. Chaitin [6] showed that there is a constant c (dependent only on the universal prefix-free machine) such that, for all d, among the strings of length n, there are at most 2^{n+c-d} strings with $K(x) \leq n + K(n) - d$; also see [24, Thm. 2.2.26].

Fix n. For $d \in \mathbb{N} - \{0\}$ let A_d be the set of strings of length n such that $K(x) \leq n + K(n) - d$. Identifying sets of strings of the same length and the clopen sets they generate, by Chaitin's result we have $\lambda A_d \leq 2^{c-d}$.

Clearly $K(x) = n + K(n) - \sum_{d>0} 1_{A_d}(x)$ whenever $|x| = n$. Taking the λ-average over all strings of length n we obtain

$$\begin{aligned} K(\lambda\upharpoonright_n) = 2^{-n}\sum_{|x|=n} K(x) \quad &\geq \quad n + K(n) - \sum_{d>0} \lambda(A_d) \\ &\geq \quad n + K(n) - \sum_{d>0} 2^{c-d} \\ &= \quad n + K(n) - 2^{c+1}. \end{aligned}$$

□

We provide an example of a computable probability measure $\mu \neq \lambda$ that satisfies the hypothesis of Fact 3.3. Informally, along a string, μ puts 1/3 on 0 until it hits the first 1; from then on μ behaves like the uniform measure λ.

Recall from 2.7 on the localization ν_σ of a measure ν to a string σ. For $r \in \mathbb{N} - \{0\}$ let $\sigma(r)$ denote the string $0^{r-1}1$. Note that for a string y of length $n - r$ one has $K(\sigma(r)\hat{\ }y) \geq^+ K(y)$ (where the additive constant is independent of both n and r). So by Fact 3.3, for $n \geq r$ one has $K(\lambda_{\sigma(r)} \upharpoonright n) \geq^+ n - r + K(n-r)$. Trivially, for $n < r$ one has $K(\lambda_{\sigma(r)} \upharpoonright n) = K(n)$.

Fact 3.4. *The computable measure $\mu = 2\sum_{r>0} 3^{-r}\lambda_{\sigma(r)}$ satisfies $K(\mu\upharpoonright_n) \geq^+ n + K(n)$. Moreover, μ is absolutely continuous.*

Proof. Keeping in mind that $2\sum_{r=1}^{n} 3^{-r} = 1 - 3^{-n}$, we calculate

$$\begin{aligned} K(\mu\upharpoonright_n) &= 2\sum_{r>0} 3^{-r} K(\lambda_{\sigma(r)} \upharpoonright n) \\ &\geq^+ 2\sum_{r=1}^{n} 3^{-r}(n - r + K(n-r)) + 3^{-n}K(n) \\ &\geq^+ n - 2\sum_{r=1}^{n} r \cdot 3^{-r} + 2\sum_{r=1}^{n} 3^{-r}K(n-r) + 3^{-n}K(n) \\ &\geq^+ n + K(n) \end{aligned}$$

The last inequality holds because $n(1 - 3^{-n}) \geq n - 1$, and $K(n-r) \geq^+ K(n) - 2\log r$, and $2\sum_{r=1}^{\infty} r \cdot 3^{-r}$ is finite.

Clearly μ is computable; $\mu \ll \lambda$ because when viewed as a measure on the unit interval via the binary expansion, we have $d\mu = g d\lambda$ where the function $g\colon [0,1] \to \mathbb{R}$ is given by $g(0) = 0$, $g(x) = 2 \cdot (\frac{2}{3})^r$ for x in the interval $(2^{-r}, 2^{-r+1}]$ where $r > 0$ (this interval corresponds to $\sigma(r)$). We also have $\lambda \ll \mu$, albeit via an unbounded μ-integrable function. □

Li and Vitanyi [16] called a bit sequence $Z \in \{0,1\}^{\mathbb{N}}$ Kolmogorov random if there is r such that $C(Z\upharpoonright_n) \geq n - r$ for infinitely many n. One says that Z is strongly Chaitin random if there is r such that $K(Z\upharpoonright_n) \geq n + K(n) - r$ for infinitely many n. For bit sequences these notions are equivalent to 2-randomness (i.e., ML-randomness relative to the halting problem) by [27] and [19], respectively; also see [24, Theorem 8.1.14] or [9].

One can extend these notions to measures. We call μ strongly Chaitin random if $K(\mu\upharpoonright_n) \geq^+ n + K(n)$ for infinitely many n. Similarly we define Kolmogorov randomness of a measure. By Fact 3.2 the following is immediate:

Fact 3.5. *If a measure μ is strongly Chaitin random, then μ is Kolmogorov random.*

We leave open the question whether the converse implication holds as well. The potentially weaker notion already implies being ML absolutely continuous:

Theorem 3.6. *Suppose that μ is a Kolmogorov random measure. Then μ is ML absolutely continuous.*

Proof. By hypothesis there is $r \in \mathbb{N}$ such that $C(\mu\restriction_n) \geq n - r$ for infinitely many n. Assume for a contradiction that μ is not ML absolutely continuous. So there is a Martin-Löf test $\langle G_d \rangle_{d\in\mathbb{N}}$ and $\delta > 0$ such that $\mu(G_d) \geq \delta$ for each d. We view G_d as given by an enumeration of strings, uniformly in d; thus $G_d = \bigcup_i [\sigma_i]$ for a sequence $\langle \sigma_i \rangle_{i\in\mathbb{N}}$ that is computable uniformly in d. Let $G_d^{\leq n}$ denote the clopen set generated by the strings in this enumeration of length at most n. (Note that this set is not effectively given as a clopen set, but we effectively have a description of it as a Σ^0_1 set). Let c be a constant such that, for each x of length n, one has $C(x) \leq n + c$.

Lemma 3.7. *If x is a string of length n such that $[x] \subseteq G_d^{\leq n}$ then $C(x \mid d) \leq^+ n - d$.*

To verify this, let N be a fixed plain machine that on input y and auxiliary input d prints out the y-th string x of length $n = |y| + d$ such that the enumeration of $G_d^{\leq n}$ asserts that $[x] \subseteq G_d^{\leq n}$. (Here we view y as the binary representation of a number, with leading zeros allowed.) Since $\lambda G_d \leq 2^{-d}$, sufficiently many strings are available to print all such x. This machine shows that $C(x \mid d) \leq^+ n - d$ for any x such that $[x] \subseteq G_d^{\leq n}$, as required. This verifies the lemma.

Note that (after possibly increasing the constant c) any x as above satisfies $C(x) \leq n-d+2\log d+c$. For each d, n, letting x range over strings of length n, we have

$$C(\mu\restriction_n) = \sum_{|x|=n} C(x)\mu[x] = \sum_{[x]\subseteq G_d^{\leq n}} C(x)\mu[x] + \sum_{[x]\not\subseteq G_d^{\leq n}} C(x)\mu[x].$$

The first summand is bounded above by $\mu(G_d^{\leq n})(n - d + 2\log d + c)$ via the lemma, the second by $(1 - \mu(G_d^{\leq n}))(n + c)$. We obtain

$$C(\mu\restriction_n) \leq n + c - \mu(G_d^{\leq n})d/2.$$

Now for each d, for sufficiently large n we have $\mu(G_d^{\leq n}) \geq \delta$. So given r letting, $d = 2r/\delta$, for large enough n we have $C(\mu\restriction_n) \leq n + c - r$. This contradicts the hypothesis on μ. ☐

It would be interesting to find out whether the above-mentioned coincidences of randomness notions for bit sequences lift to measures; for instance, do the conditions in the theorem above actually imply that the measure is ML absolutely continuous relative to the halting problem $\emptyset'$?

4. Both implications of the Levin-Schnorr Theorem fail for measures

We will show that both implications of the analog of the Levin-Schnorr Theorem 2.2 fail for measures. The implication from left to right would say that a ML absolutely continuous measure cannot have an initial segment complexity in the sense of K growing slower than $n - O(1)$. This can be disproved by a simple example of a measure with countable support.

Note that, by Proposition 7.4 below, we have $\lim_n K(\mu\restriction_n)/n = 1$ for each ML absolutely continuous measure μ, which provides a lower bound on the growth.

Example 4.1. There is a ML absolutely continuous measure μ such that for each $\theta \in (0,1)$, one has $K(\mu \upharpoonright_n) \le^+ n - n^\theta$.

Proof. We let $\mu = \sum_k c_k \delta_{Z_k}$ where Z_k is Martin-Löf random and $0^{n_k} \prec Z_k$ for a sequence $\langle c_k \rangle_{k\in\mathbb{N}}$ of reals in $[0,1]$ that add up to 1, and a sufficiently fast growing computable sequence $\langle n_k \rangle_{k\in\mathbb{N}}$ to be determined below. Then μ is ML absolutely continuous by Fact 2.9.

For n such that $n_k \le n < n_{k+1}$ we have

$$\begin{aligned} K(\mu \upharpoonright_n) &\le^+ (\sum_{l=0}^{k} c_l) \cdot (n + 2\log n) + (\sum_{l=k+1}^{\infty} c_l) \cdot 2\log n \\ &\le^+ (1 - c_{k+1})n + 2\log n. \end{aligned}$$

Hence, to achieve $K(\mu \upharpoonright_n) \le^+ n - n^\theta$ it suffices to ensure that $c_{k+1} n_k \ge n_{k+1}^\theta + 2\log n_{k+1}$ for almost all k. For instance, we can let $c_k = \frac{1}{(k+1)(k+2)}$ and $n_k = 2^{k+4}$. □

To disprove the implication from right to left in a potential generalisation to measures of the Levin-Schnorr Theorem, we need to provide a measure μ such that μ is not Martin-Löf a.c. and $K(\mu \upharpoonright_n) \ge^+ n$. This is immediate from the following fact on the growth of the initial segment complexity for certain bit sequences X, Y, letting $\mu = (\delta_X + \delta_Y)/2$.

Theorem 4.2. *There are a Martin-Löf random bit sequence X and a non-Martin-Löf random bit sequence Y such that, for all n,*

$$K(X \upharpoonright_n) + K(Y \upharpoonright_n) \ge^+ 2n.$$

Proof. Let X be a low Martin-Löf random set (i.e., $X' \equiv_T \emptyset'$). We claim that there is a strictly increasing function f such that the complement of the range of f is a recursively enumerable set E, and $K(X \upharpoonright_m) \ge m + 3n$ for all $m \ge f(n)$. To see this, recall that $\lim_n K(X \upharpoonright_n) - n = \infty$. Since X is low there is a computable function p such that for all n, $\lim_s p(n,s)$ is the maximal m such that $K(X \upharpoonright_m) \le m + 3n$.

Define $f(n,s)$ for $n \le s$ as follows. $f(n,0) = n$; for $s > 0$ let n be least such that $p(n,s) \ge f(n,s-1)$ or $n = s$. If $m \ge n$ and $n < s$ then let $f(m,s) = s + m - n$ else let $f(m,s) = f(m,s-1)$.

Note that for each n there are only finitely many $s > 0$ with $f(n,s) \ne f(n,s-1)$ and that almost all s satisfy $f(n,s) > p(n,s)$, as otherwise $f(n,s)$ would be modified either at n or some smaller value. Furthermore, $f(n,s) \ne f(n,s-1)$ can only happen if there is an $m \le n$ with $f(m,s-1) \le p(m,s)$ and that happens only finitely often, as all the $p(m,s)$ converge to a fixed value and every change of an $f(m,s)$ at some time s leads to a value above s. Furthermore, once an element is outside the range of f, it will never return, and so the complement of the range of f is recursively enumerable. So $f(n) = \lim_s f(n,s)$ is a function as required, which verifies the claim. (The complement E of the range of f is called a Dekker deficiency set in the literature [29].)

Now let $g(n) = \max\{m : f(m) \le n\}$ (with the convention that $\max(\emptyset) = 0$). Since g is unbounded, by a result of Miller and Yu [22, Corollary 3.2] there is a Martin-Löf random Z such that there exist infinitely many n with

$K(Z\upharpoonright_n) \leq n + g(n)/2$; note that the hypothesis of this result of Miller and Yu does not make any effectivity requirements on g. Let

$$Y = \{n + g(n) : n \in Z\}.$$

Note that $K(Z\upharpoonright_n) \leq K(Y\upharpoonright_n) + g(n) + K(g(n))$, as one can enumerate the set E until there are, up to n, only $g(n)$ many places not enumerated and then one can reconstruct $Z\upharpoonright_n$ from $Y\upharpoonright_n$ and $g(n)$ and the last $g(n)$ bits of $Z\upharpoonright_n$. As Z is Martin-Löf random, $K(Z\upharpoonright_n) \geq^+ n$, so

$$K(Y\upharpoonright_n) \geq^+ n - g(n) - K(g(n)) \geq^+ n - 2g(n).$$

The definitions of X, f, g give $K(X\upharpoonright_n) \geq n + 3g(n)$. This shows that, for almost all n, $K(X\upharpoonright_n) + K(Y\upharpoonright_n) \geq 2n$.

However, the set Y is not Martin-Löf random, because there are infinitely many n such that $K(Z\upharpoonright_n) \leq^+ n + g(n)/2$. Now $Y\upharpoonright_{n+g(n)}$ can be computed from $Z\upharpoonright_n$ and $g(n)$, as one needs only to enumerate E until the $g(n)$ nonelements of E below n are found, and they allow to see where the zeroes have to be inserted into the string $Z\upharpoonright_n$ in order to obtain $Y\upharpoonright_{n+g(n)}$. We have $K(g(n)) \leq g(n)/4$ for almost all n and thus $K(Y\upharpoonright_{n+g(n)}) \leq^+ n+3/4\cdot g(n)$ for infinitely many n. So the set Y is not Martin-Löf random, as required. $\square$

Failing a strong Solovay test implies non-complex initial segments. In contrast to the negative results above, we show that failing a stronger type of tests leads to considerable dips in the intitial segment complexity of a measure in the sense of C (and hence, also in the sense of K by Fact 3.2).

We say that a Solovay test $\langle S_r\rangle_{r\in\mathbb{N}}$ is *strong* if each S_r is clopen and there is a recursive function g such that $g(r)$ is a strong index for a finite set of strings X_r such that $[X_r]^\prec = S_r$. For bit sequences, this means no restriction of the corresponding randomness notion: in that case any Solovay test $\langle G_m\rangle$ can be replaced by a strong Solovay test, listing strings that make up the uniformly Σ^0_1 sets G_m one-by-one. We conjecture that this equivalence of test notions no longer holds for measures.

The following is a weak version for measures of one implication of the Miller-Yu theorem [21, Theorem 7.1]. We show that if a measure μ is far from ML absolutely continuous, then its initial segment complexity $C(\mu\upharpoonright_n |n)$ defined in Definition 3.1 has infinitely many dips of size $\delta f(n)$ for some positive constant δ and a computable function f that grows somewhat fast in that $\sum_n 2^{-f(n)} < \infty$. We use elements of its proof in Bienvenu, Merkle and Shen [4]. We note that the result is a variation on (the contrapositive of) Theorem 3.6, proving a stronger conclusion from a stronger hypothesis. A version of the result for quantum states is [26, Thm. 4.4].

Proposition 4.3. *Suppose that μ fails a strong Solovay test $\langle S_r\rangle_{r\in\mathbb{N}}$ at level δ, namely $\exists^\infty r\,[\mu(S_r) \geq \delta]$. Then there is a computable function f such that $\sum_n 2^{-f(n)} < \infty$ and*

$$\exists^\infty n\,[C(\mu\upharpoonright_n|\, n) \leq^+ n - \delta f(n)].$$

Proof. Let $\langle X_r\rangle$ be as in the definition of a strong Solovay test. We may assume that $\sum \lambda S_r \leq 1/2$, and all strings in X_r have the same length n_r, where $n_r < n_{r+1}$ for each r. We write $\mu(X_r)$ for $\mu[X_r]^\prec$. Let f be the computable function such that

$$2^{-f(n_r)} \geq \mu(X_r) > 2^{-f(n_r)-1}$$

and $f(m) = m$ for each m that is not of the form n_r. There is a constant d such that each bit string x in the set X_r satisfies (where $n = n_r$)

$$C(x \mid n) \leq n - f(n) + d =: g(r). \tag{1}$$

For, r can be computed from $n = n_r$, and each string $x \in X_r$ is determined by r and its position $i < 2^n\lambda(X_r)$ in the lexicographical listing of X_r. We can determine i by $\log(2^n\lambda(X_r)) \leq n - f(n) + d$ bits for some fixed d. In fact we may assume the description has exactly that many bits. Thus, there is a Turing machine L with two inputs such that for each $\sigma \in X_r$, we have $L(v_\sigma; n) = \sigma$ for some bit string v_σ of length $g(r)$.

Let c be a constant such that $C(x) \leq |x| + c$ for each string x. Now suppose that $\mu(X_r) \geq \delta$. Then for $n = n_r$, where σ ranges over strings of length n,

$$\begin{aligned} C(\mu\restriction_n \mid n) &\leq^+ \sum_{\sigma \in X_r} (n - f(n))\mu[\sigma] + \sum_{\sigma \notin X_r} (n + c)\mu[\sigma] \\ &\leq^+ n - f(n) \sum_{\sigma \in X_r} \mu[\sigma] \\ &\leq^+ n - \delta f(n). \end{aligned}$$

There are infinitely many such r by hypothesis. This completes the proof. □

5. K- and C-triviality for measures

Recall that a set A is called K-trivial if $K(A\restriction_n) \leq K(n) + b$ for each n; one says that b is a K-triviality constant for A. In a similar way one defines C-trivial sets. This section generalises these definitions to measures.

Definition 5.1. A measure μ is called K-trivial if $K(\mu\restriction_n) \leq^+ K(n)$ for each n; μ is called C-trivial if $C(\mu\restriction_n) \leq^+ C(n)$ for all n.

Suppose μ is a Dirac measures δ_A. Then μ is K-trivial iff A is K-trivial in the usual sense, and μ is C-trivial iff A is C trivial in the usual sense, which is equivalent to being recursive by a well-known result of Chaitin [7] (for a more recent proof see e.g. [24, 5.2.20]). More generally, any finite convex combination of K-trivial Dirac measures is K-trivial, and similarly for C. Now if α, β are any positive real numbers such that $\alpha + \beta = 1$, then $\alpha \cdot \delta_\emptyset + \beta \cdot \delta_\mathbb{N}$ is a C-trivial and K-trivial measure such that any measure representation in the sense of [24, 1.9.2] computes α. Thus one cannot bound the computational complexity of K-trivial or of C-trivial measures in any way.

For a further example, let $\langle R_i \rangle$ be a sequence of uniformly recursive sets, and let $\langle \alpha_i \rangle$ be a sequence of uniformly recursive positive real numbers such that $\sum_i \alpha_i = 1$ and $S = \sum_i \alpha_i \log i$ is finite. Let $\mu = \sum_i \alpha_i \delta_{R_i}$. Since $K(R_i\restriction_n) \leq^+ K(n) + K(i)$ and $K(i) \leq^+ 2\log i$, for each n we have

$$K(\mu\restriction_n) = \textstyle\sum_i \alpha_i K(R_i\restriction_n) \leq^+ K(n) + 2S.$$

Hence μ is K-trivial.

In the following, recall from the third paragraph of Section 3 that our choice of universal machines implies that $C(n) \leq C(x)$ and $K(n) \leq K(x)$ for all strings x of length n.

Atoms and Triviality. We show that both K-trivial and C-trivial measures are discrete in the sense of [33]; namely, they are concentrated on the set of their atoms. Such measures were further studied by Kautz [14], where they were called "trivial measures", and more recently by Porter [31].

Proposition 5.2. *If a measure μ is K-trivial, then μ is supported by its set of atoms.*
In fact the weaker hypothesis $\exists p\, \exists^\infty n\, [K(\mu\restriction_n) \leq K(n) + p]$ suffices for this. A similar statement holds for C.

Proof. For a set $R \subseteq \{0,1\}^*$, by $[R]^\prec$ one denotes the open set $\{Z\colon\ \exists n\, Z\restriction_n \in R]\}$.

Assume for a contradiction that μ gives a measure greater than $\varepsilon > 0$ to the set of its non-atoms. Fix c arbitrary, with the goal of showing that $K(\mu\restriction_n) \geq K(n) + \varepsilon c/2$ for large enough n.

There is a constant d (in fact $d = O(2^c)$) such that for each n there are at most d strings x of length n with $K(x) \leq K(n) + c$ (see e.g. [24, Theorem 2.2.26(ii)]).

Let $S_n = \{x\colon\ |x| = n \wedge \mu[x] \leq \varepsilon/2d\}$. By hypothesis we have $\mu[S_n]^\prec \geq \varepsilon$ for large enough n. Therefore by choice of d we have

$$\mu[S_n \cap \{x\colon\ K(x) > K(|x|) + c\}]^\prec \geq \varepsilon/2.$$

Now we can give a lower bound for the μ-average of $K(x)$ over all strings x of length n:

$$\sum_{|x|=n} K(x)\mu[x] \geq (1 - \frac{\varepsilon}{2})K(n) + \frac{\varepsilon}{2}(K(n) + c) \geq K(n) + \varepsilon c/2,$$

as required. Notice that we have only used the weaker hypothesis. ☐

Thus, if μ is K-trivial for constant p, then μ has the form $\sum_{r<N} \alpha_r \delta_{A_r}$ where $N \leq \infty$ and each α_r is positive and $\sum_{r<N} \alpha_r = 1$. The following observation will show that in addition all the A_r are K-trivial. In particular, K-trivial measures are not Martin-Löf absolutely continuous.

Fact 5.3. *Suppose $\mu = \sum_r \alpha_r \delta_{A_r}$ is K-trivial for the constant p. Then each A_r is K-trivial for the constant p/α_r. Similarly, if μ is C-trivial for the constant p, then each A_r is C-trivial for the constant p/α_r.*

Proof. We have

$$K(n) + p \geq K(\mu\restriction_n) = \sum_s \alpha_s K(A_s\restriction_n) \geq K(n) + \alpha_r(K(A_r\restriction_n) - K(n)).$$

Therefore, $K(A_r\restriction_n) - K(n) \leq p/\alpha_r$, as required.

The proof for C is obtained via replacing K by C everywhere. ☐

Above we built a computable K-trivial measure with infinitely many atoms. On the other hand, the following example shows that not every infinite convex combination $\mu = \sum_k \alpha_k \delta_{A_k}$ of K-trivial Dirac measures for constants

b_k yields a K-trivial measure, even if there is a finite bound on the values $\alpha_k b_k$. Identifying Chaitin's Ω with its binary expansion, let

$$A_k = \{\ell : \ell \in \Omega \wedge \ell < k\} \text{ and } \alpha_k = (k+1)^{-1/2} - (k+2)^{-1/2}.$$

All sets A_k are finite and thus K-trivial for constant $2k+O(1)$. Furthermore, the sum of all α_k is 1 and $\alpha_k = O(1/k)$. We have

$$K(\mu\restriction_n) = \sum_{|x|=n} K(x)\mu(x) \geq (\sum_{m\geq n} \alpha_m)\cdot K(\Omega \restriction n) \geq (n+2)^{-1/2}\cdot(n+2) = \sqrt{n+2}$$

for almost all n, and thus the average grows faster than $K(n) + c$. So the measure is not K-trivial.

In a sense, an atomless measure can be arbitrarily close to being C-trivial and K-trivial.

Proposition 5.4. *For each nondecreasing unbounded function f which is recursively approximable from above there is a non-atomic measure μ such that $C(\mu\restriction_n) \leq^+ C(n) + f(n)$ and $K(\mu\restriction_n) \leq^+ K(n) + f(n)$.*

Proof. The proof for C and K is essentially the same; we present the one for K first.

There is a recursively enumerable set A such that, for all n, $A \cap \{0, \ldots, n\}$ has up to a constant $f(n)/2$ non-elements. One lets μ be the measure such that $\mu(x) = 2^{-m}$ in the case that all ones in x are not in A and $\mu(x) = 0$ otherwise, where m is the number of non-elements of A below $|x|$. One can see that when $\mu(x) = 2^{-m}$ then x can be computed from $|x|$ and the string $b_0 b_1 \ldots b_{m-1}$ which describes the bits at the non-elements of A. Thus

$$K(x) \leq^+ K(|x|) + K(b_0 b_1 \ldots b_{m-1}) \leq^+ K(|x|) + 2m.$$

It follows that $K(\mu\restriction_n) \leq^+ K(n) + f(n)$, as the μ-average of strings $x \in \{0,1\}^n$ with $K(x) \leq^+ K(n) + f(n)$ is at most $K(n) + f(n)$ plus a constant.

Now the bound for C follows along the same lines; note that for C one gives a prefix-free coding of the string $b_0 b_1 \ldots b_{m-1}$ above followed by a C-description for $|x|$ plus a constant-length coding prefix – assuming that the universal machine is universal by adjunction. Again the length of the code is $C(n) + f(n)$ plus a constant. $\square$

Closure properties of C- and of K-triviality. In the following suppose Γ is a total Turing reduction, that is, a truth-table reduction, mapping $\{0,1\}^{\mathbb{N}}$ to $\{0,1\}^{\mathbb{N}}$. Given a measure ν, let $\mu = \Gamma(\nu)$ denote the image measure, defined as usual by $\mu(\mathcal{A}) = \nu(\Gamma^{-1}(\mathcal{A}))$, where $\mathcal{A}$ is a subset of $\{0,1\}^{\mathbb{N}}$. As Γ is a truth-table reduction, some strictly increasing computable function f bounds the use of Γ.

Proposition 5.5. *Let $\mu = \Gamma(\nu)$ as above. We have $K(\mu\restriction_n) \leq^+ K(\nu\restriction_{f(n)})$ and similarly for C. In particular, if ν is K-trivial then so is μ, and if ν is C-trivial so is μ.*

Proof. Let x range over strings of length n and y over strings of length $f(n)$. Since $K(\Gamma^y) \le K(y)$ we have

$$\begin{aligned} K(\mu\restriction_n) &= \sum_x K(x) \sum_{\Gamma^y \succeq x} \nu[y] = \sum_x \sum_{\Gamma^y \succeq x} K(x)\nu[y] \\ &\le^+ \sum_x \sum_{\Gamma^y \succeq x} K(y)\nu[y] = K(\nu\restriction_{f(n)}). \end{aligned}$$

Using that n and $f(n)$ have, up to a constant, the same descriptive complexity in the sense of K, this completes the proof. The proof for C is similar to the proof for K with only the obvious changes. □

Note that $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ is effectively isomorphic to $\{0,1\}^{\mathbb{N}}$ via the map $A, B \to A \oplus B$. So a product of two measures on $\{0,1\}^{\mathbb{N}}$ can itself be viewed as a measure on $\{0,1\}^{\mathbb{N}}$. We show that K- and C-triviality of measures is closed under taking products, generalising well-known facts for bit sequences. The following fact will be needed for the case of C.

Fact 5.6. *$K(x \mid n, C(n)) \le^+ 2(C(x) - C(n))$ for each n and each $x \in \{0,1\}^n$.*

Proof. One suffices to show that there is a constant d such that for each b,

$$C(x) \le C(n) + b \to K(x|n, C(n)) \le 2b + d.$$

Following an argument of Chaitin [7], Nies [24, Proof of Lemma 5.2.21] showed that for every plain machine M there are at most $O(b^2 \cdot 2^b)$ descriptions p with $M(p) = n$ and $|p| \le C(n) + b$. Consider the machine M such that $M(p) = 0^{|U(p)|}$ where U is the universal plain machine. There are at most $O(b^2 \cdot 2^b)$ many $x \in \{0,1\}^n$ with $C(x) \le C(n) + b$, because any such x has the form $U(p)$ for some p with $M(p) = n$ and $|p| \le C(n) + b$. Uniformly in b, n and $C(n)$, the set of such x is recursively enumerable. Hence one can provide a prefix-free code of length up to $2b$ plus constant for any pair $\langle b, a\rangle$ with $a \in O(b^2 \cdot 2^b)$ selecting the a-th such x. This shows that a constant d as above exists. □

Theorem 5.7. *Let μ and ν be measures.*
(a) If μ and ν are K-trivial then $\mu \times \nu$ is K-trivial.
(b) If μ and ν are C-trivial then $\mu \times \nu$ is C-trivial.

Proof. We assume in both cases (a) and (b) that μ and ν are trivial with respect to the corresponding descriptive string complexity and then show that $\mu \times \nu$ is trivial for the same complexity notion.

For (a), by Prop 5.2 one may assume that $\mu = \sum_i \alpha_i A_i$ and $\nu = \sum_j \beta_j B_j$ with $1 = \sum_i \alpha_i$ and $1 = \sum_j \beta_j$, where the A_i and B_j are K-trivial sets. For a string x of length n and a shortest description p of (a code for) n, $K(x) =^+ K(x|p) + |p|$; this follows easily from [16, Theorem 3.9.1]. Note that

$$K(A_i \oplus B_j\restriction_{2n}) \le^+ K(n) + K(A_i\restriction_n |p) + K(B_j\restriction_n |p).$$

Since μ and ν are K-trivial we have $K(A_i\restriction_n) =^+ K(n) + K(A_i\restriction_n |p)$ and $K(B_j\restriction_n) =^+ K(n) + K(B_j\restriction_n |p)$, so $\sum_i \alpha_i K(A_i\restriction_n |p)$ and $\sum_j \beta_j K(B_j\restriction_n |p)$

are both bounded by some constant independent of n. Let c be a common upper bound of such constants. Now, for some constant d,

$$\begin{aligned} K(\mu \times \nu \restriction_{2n}) &= \sum_{i,j} \alpha_i \cdot \beta_j \cdot K(A_i \oplus B_j \restriction_{2n}) \\ &\leq \sum_{i,j} \alpha_i \cdot \beta_j \cdot (K(n) + K(A_i \restriction_n |p) + K(B_j \restriction_n |p) + d) \\ &= K(n) + d + (\sum_i \alpha_i \cdot K(A_i \restriction_n |p) \cdot (\sum_j \beta_j)) \\ &\qquad + (\sum_i \alpha_i) \cdot (\sum_j \beta_j \cdot K(B_j \restriction_n |p)) \\ &\leq K(n) + 2c + d. \end{aligned}$$

Since $K(n) =^+ K(2n)$, one can conclude that $\forall n\,[K(\mu \times \nu \restriction_{2n}) \leq^+ K(2n)]$. The same can be proven for $2n+1$ with only a slight more notational complexity. Thus $\mu \times \nu$ is K-trivial.

For (b), as in (a), assume that $\mu = \sum_i \alpha_i A_i$ and $\nu = \sum_j \beta_j B_j$ with $1 = \sum_i \alpha_i$ and $1 = \sum_j \beta_j$, where now A_i, B_j are C-trivial sets. Let c be the supremum over n of all $\sum_i \alpha_i \cdot (C(A_i \restriction_n) - C(n))$ and $\sum_j \beta_j \cdot (C(B_j \restriction_n) - C(n))$. Note that $c < \infty$ because μ and ν are C-trivial.

One can produce descriptions for $A_i \restriction_n$ and $B_j \restriction_n$ from prefix-free conditional descriptions for $A_i \restriction_n$ and $B_j \restriction_n$, given n and $C(n)$ together with a coding of these two parameters n and $C(n)$. Note that $n, C(n)$ can be given by a shortest C-description of n in one go, which needs $C(n)$ bits. With the other two prefix-free conditional descriptions appended, one obtains for a suitable constant d' that

$$C(A_i \oplus B_j \restriction_{2n}) \leq C(n) + K(A_i \restriction_n |n, C(n)) + K(B_j \restriction_n |n, C(n)) + d'.$$

Now one uses this to show the following as in (a):

$$\begin{aligned} C(\mu \times \nu \restriction_{2n}) &= \sum_{i,j} \alpha_i \cdot \beta_j \cdot C(A_i \oplus B_j \restriction_{2n}) \\ &\leq \sum_{i,j} \alpha_i \cdot \beta_j \cdot (C(n) + K(A_i \restriction_n |n, C(n)) + K(B_j \restriction_n |n, C(n)) + d') \\ &\leq \sum_{i,j} \alpha_i \cdot \beta_j \cdot (C(n) + 2 \cdot ((C(A_i \restriction_n) - C(n)) + (C(B_j \restriction_n) - C(n))) + 2d + d') \\ &= C(n) + 2d + d' + 2 \cdot (\sum_i \alpha_i \cdot (C(A_i \restriction_n) - C(n))) \cdot (\sum_j \beta_j) \\ &\qquad + 2 \cdot (\sum_i \alpha_i) \cdot (\sum_j \beta_j \cdot (C(B_j \restriction_n) - C(n))) \\ &\leq C(n) + 4c + 2d + d'. \end{aligned}$$

The inequality from the first to the second line follows from the definition of d'; the inequality from the second to the third line follows from Fact 5.6.

Thereafter, one uses the distributivity of absolutely converging sums, and the definition of c. The same can be proven for $2n+1$ with only a slight increase in notational complexity. $\square$

The next proposition shows that three conditions involving the descriptive complexities of strings are equivalent. It is unknown whether these equivalent conditions are true. These conditions relate fundamental properties of plain and prefix-free Kolmogorov complexity; properties of similar type have been investigated quite a lot in algorithmic information theory; fairly recent sample references are [1, 2, 20, 16]. The last condition states informally that there are finitely many potential ways to compute $C(n)$ from n and $K(n)$, and for each n one of them succeeds. (In fact at present it is unknown whether there is a single way.)

Proposition 5.8. *The following conditions are equivalent:*

(a) *There is a constant c such that for all n and all $x \in \{0,1\}^n$, $K(x) - K(n) \leq 2 \cdot (C(x) - C(n)) + c$;*
(b) *For each c' there is d' such that for all n and all $x \in \{0,1\}^n$, if $C(x) - C(n) \leq c'$ then $K(x) - K(n) \leq d'$;*
(c) *There is a constant c'' with $K(C(n)|n, K(n)) \leq c''$ for all n.*

Proof. The implication (a) to (b) is straightforward: if c' bounds $C(x)-C(n)$ then $d' = 2c' + c$ bounds $K(x) - K(n)$.

The implication (b) to (c) can be seen as follows: Note that for almost all n it holds that $C(n) < n$. So given n sufficiently large, one considers the string $x_n = (\{C(n)\} \upharpoonright_n)$, the n-bit string that has a 1 in position $C(n)$ and 0s elsewhere. This string has plain Kolmogorov complexity satisfying $C(x_n) =^+ C(n)$, so let c' be such that $C(x_n) \leq C(n) + c'$. Thus $K(x_n) \leq K(n) + d'$ for all n, where d' is defined by condition (b). By [24, Theorem 2.2.26] there are at most $O(2^{d'})$ strings $y \in \{0,1\}^n$ with $K(y) \leq K(n) + d'$ and one can enumerate the first, the second, ..., the k-th until x_n comes up; let k be the corresponding sequence-number. Now $k \leq 2^{d''}$ for some constant d'' independent of n and therefore $K(x_n|n, K(n)) \leq^+ d''$. As one can compute $C(n)$ from x_n, there is a further constant c'' with $K(C(n)|n, K(n)) \leq c''$ for all n.

The implication (c) to (a) is based on Fact 5.6 which says that $K(x|n, C(n)) \leq^+ 2(C(x) - C(n))$ and that the description p_x witnessing it is prefix-free. Now by condition (c) one has that $K(C(n)|n, K(n))$ is bounded by a constant. From a shortest K-description r_n of n one can compute both n and $K(n)$ and therefore one can compute all of $C(n), K(n), n$ from a suitable prefix-free description q_n with $|q_n| \leq^+ K(n)$. Now one can define a prefix-free machine M such that $M(q_n \hat{\ } p_x) = x$ for each string x of length n where $|q_n| =^+ K(n)$ and $|p_x| \leq^+ 2(C(x) - C(n))$. It follows that $K(x) - K(n) \leq^+ |p_x| \leq^+ 2(C(x) - C(n))$. This yields condition (a). □

If these conditions are true then the question whether all C-trivial measures are K-trivial has an affirmative answer. To see this note that if $\mu = \sum_r \alpha_r \cdot \delta_{A_r}$ is C-trivial then there is a constant b such that $\sum_r \alpha_r \cdot (C(A_r \upharpoonright_n) -$

$C(n)) \leq b$ for all n. It follows from Condition (a) that

$$\begin{aligned} K(\mu\restriction_n) &= K(n) + \sum_r \alpha_r \cdot (K(A_r\restriction_n) - K(n)) \\ &\leq K(n) + \sum_r \alpha_r \cdot (2 \cdot (C(A_r\restriction_n) - C(n)) + c) \\ &\leq K(n) + c + 2 \cdot \sum_r \alpha_r \cdot (C(A_r\restriction_n) - C(n)) \leq K(n) + 2b + c. \end{aligned}$$

Strong C- and strong K-triviality for measures.
Given the examples above we consider a strengthening of our notions.

Definition 5.9. Call a measure μ *strongly C-trivial* if it can be written as a convex combination $\sum_s \alpha_s \delta_{A_s}$ such that the A_s are C-trivial via constants b_s such that $\sum_s \alpha_s \cdot b_s$ is finite. Similarly, one defines *strongly K-trivial* measures.

Fact 5.10. *Every strongly C-trivial measure is C-trivial. Every strongly K-trivial measure is K-trivial.*

Proof. By hypothesis $\mu = \sum_s \alpha_s \cdot \delta_{A_s}$ where each A_s is C-trivial with constant b_s, and $\sum \alpha_s \cdot b_s < c$. For all n we have $C(\mu\restriction_n) = \sum_s \alpha_s \cdot C(A_s\restriction_n) \leq \sum_s \alpha_s \cdot (C(n) + b_s) \leq C(n) + \sum_s \alpha_s \cdot b_s \leq C(n) + c$. The proof is analogous for K. □

While we ignore whether every C-trivial measure is K-trivial, we show that this implication holds for the strong versions of the two notions. For this we prove a lemma about sets of interest on its own. Its proof owes to the proof that each C-trivial is recursive (Chaitin [7], also see [24, Proof of Thm 5.2.20(i)]).

Lemma 5.11. *There is an absolute constant d such that every set that is C-trivial via a constant b is K-trivial via the constant $2b + d$.*

Proof. As mentioned in the proof of Fact 5.6, there are only $O(b^2 \cdot 2^b)$ binary strings x of length n with $C(x) \leq C(n) + b$. Now one considers the following recursively enumerable tree T_b of binary strings: a string x of length n is in T_b iff there is an y extending x of length 2^{n+1} such that for all $m \leq 2^{n+1}$ it holds that $C(y\restriction_m) \leq \log(m) + b + 1$.

Note that there is an m with $2^n \leq m \leq 2^{n+1}$ such that $C(m) > n$. Thus for that length m, there are only $O(b^2 \cdot 2^b)$ strings which qualify and therefore, for length n, there are at most $O(b^2 \cdot 2^b)$ many strings in T_b. Furthermore, note that the membership in T_b is closed under prefixes and therefore T_b is a recursively enumerable tree.

For each length n, one can enumerate from given n, b the $O(b^2 \cdot 2^b)$ strings of that length in T_b. So one can describe these strings by any description of n plus prefix-free codes of size $2b + d$ which compute the strings given n, b. Note that these codes just say "the first string of length n enumerated into T_b", "the second string of length n enumerated into T_b" and so on. As there are $O(2^b \cdot b^2)$ of these strings, one can represent them and b in a prefix-free way with $2b$ plus constant bits. Thus they can be described in a prefix-free way with a description of size $K(n) + 2b + d$ where d is a constant independent of b, n. Hence $K(A\restriction_n) \leq K(n) + 2b + d$. □

Proposition 5.12. *Every strongly C-trivial measure is strongly K-trivial.*

Proof. Suppose that a measure μ is given as a convex combination $\sum_s \alpha_s \cdot \delta_{A_s}$ where each A_s has the C-triviality constant b_s, and $\sum_s \alpha_s \cdot b_s$ is finite. By the foregoing lemma every A_s is K-trivial via the constant $d + 2b_s$. Clearly $\sum_s \alpha_s \cdot (d + 2b_s) = d + 2 \cdot \sum_s \alpha_s b_s$ is finite. So μ is strongly K-trivial. □

One might ask whether every K-trivial measure is strongly K-trivial, and whether every C-trivial measure is strongly C-trivial. The answer to this question is "no" as the following result shows.

Proposition 5.13. *There is a measure μ which is C-trivial and K-trivial but not strongly K-trivial (and hence not strongly C-trivial).*

Proof. For each n, let p_n be the description of length up to 2^n for the prefix-free universal machine which produces the longest output string of form 0^{m_n}; here recall that m_n and 0^{m_n} are identified with each other. Note that $K(p_n)$ and $|p_n|$ are both approximately 2^n. Now one defines the sets $A_n = \{m_n + k : 1p_{n+1} \text{ has a 1 at position } k\}$. The sets A_n are uniformly limit recursive and the approximations converge in time m_{n+1} to A_n; furthermore, $\max(A_n) \leq m_n + 2^{n+1}$. Every set A_n is finite and thus K-trivial. Let $\mu = \sum_n 2^{-n-1} \cdot \delta_{A_n}$.

We show that μ is K-trivial. One sees that for $\mu \restriction_\ell$ with $m_n \leq \ell < m_{n+1}$, 2^{-n-1} of the mass is concentrated on $\emptyset \restriction_\ell$ and for $h = 0, 1, \ldots, n$, 2^{-h-1} of the mass is concentrated on $A_h \restriction_\ell$. For $h < n$, the string $A_h \restriction_\ell$ can be computed from h and ℓ. Thus it has prefix-free complexity $K(\ell) + 2\log(h)$ which contributes to the average measure $K(\mu \restriction_\ell)$ a term bounded by $2^{-h-1} \cdot K(\ell) + 2^{-h-1} \cdot 2\log(h)$. As the sum $2^{-h-1} \cdot 2 \cdot \log(h)$ converges to a constant c, one can estimate the overall sum as

$$K(\ell) + c + 2^{-n-1} \cdot (K(A_n \restriction_\ell) - K(\ell)),$$

and the latter is bounded by $K(\ell) + c + 2$, as $K(A_n \restriction_\ell) \leq K(\ell) + 2^{n+2}$. Thus μ is K-trivial. By a similar argument, μ is C-trivial.

The K-triviality constant b_n for A_n is bounded from below by $K(A_n \restriction_{m_n + 2^{n+1}}) - K(m_n + 2^{n+1})$ and that term is at least 2^n, as $K(p_{n+1}) =^+ 2^{n+1}$ and $K(m_n) =^+ 2^n$. It follows that $\sum_n 2^{-n-1} b_n \geq \sum_n 1/2 = \infty$. Thus μ is not strongly K-trivial. □

By the facts and propositions above, in Proposition 5.5 the condition that Γ be a truth-table reduction is necessary; downward closure of the K-trivial measures does not hold for Turing reductions in general.

Proposition 5.14. *There are measures μ, ν and a Turing reduction Γ such that ν is strongly K-trivial, Γ is defined on all the atoms of ν and $\mu = \Gamma(\nu)$ is not K-trivial.*

Proof. Consider the set $\Omega \oplus \overline{\Omega}$, that is, the join of Ω with its complment. This set is left-r.e. and initial segments of length $2m$ have approximately complexity m. Now let

$$A_m = \{x \in \Omega \oplus \overline{\Omega} : x < 2n_m\}$$

and $\gamma_m = \frac{1}{(m+1)\cdot(m+2)} = \frac{1}{m+1} - \frac{1}{m+2}$. Furthermore, let n_m be the least integer greater than $(m+2)^{3/2}$ and let $k_m = \langle m, c_\Omega(n_m) \rangle$ where $c_\Omega(n_m)$ is the time to follow the recursive left-enumeration of Ω until the first n_m bits are correct (and therefore remain correct from then on). Note that k_m can be computed from $A_m \upharpoonright_{2n_m+2}$ and $A_m \upharpoonright_{2n_m+2}$ can be computed from k_m. Let $B_m = \{k_m\}$ and $\nu = \sum_m \gamma_m \cdot \delta_{B_m}$ and $\mu = \sum_m \gamma_m \cdot \delta_{A_m}$.

Note that for $\ell \geq k_m$, k_m can be computed from ℓ and m, as ℓ is an upper bound of $c_\Omega(n_m)$; for $\ell < k_m$, $B_m \upharpoonright_\ell = 0^\ell$. Thus the coding constant of B_m is $2\log(m)$ plus some constant. As $\sum_m \gamma_m \cdot 2 \cdot \log(m)$ is a convergent sum – almost all terms are bounded by $(m+2)^{-3/2}$ which is a convergent sum – the measure ν is K-trivial by Fact 5.10. Actually it is by definition even strongly K-trivial.

The measure μ is not K-trivial, as all A_k with $k \geq m$ have the common prefix $u = \Omega \oplus \overline{\Omega} \upharpoonright_{2n_m}$ of length $2n_m$ and that prefix has, for almost all m, the Kolmogorov complexity $(m+2)^{3/2}$ or more. Note that $\mu(\{u\} \times \{0,1\}^{\mathbb{N}})$ is at least $\frac{1}{m+2}$, so this string contributes to the average $\mu \upharpoonright_{2n_m}$ at least $(m+2)^{1/2}$. Thus μ is not K-trivial, as the function $m \mapsto K(2n_m)$ cannot be bounded from below by any increasing unbounded recursive function.

Now one defines a Turing functional Γ that translates each B_m to A_m: the functional searches in the oracle B_m for the first element x of B_m; as long as the oracle is not empty, such an element is found. Then it determines the m, s with $x = \langle m, s\rangle$ and computes the corresponding entry in the set $\{x \in \Omega_s \oplus \overline{\Omega}_s : x < 2n_m\}$. In the case that the oracle has the set B_m, the search gives k_m and this Turing reduction provides the set A_m. Thus Γ has the desired properties. Note that Γ is undefined only on $\emptyset$, and the measure of $\{\emptyset\}$ with respect to ν is 0. $\square$

While we know that every strongly C-trivial measure is strongly K-trivial, we ignore whether every C-trivial measure is K-trivial. The converse implication fails even for Dirac measures because each C-trivial sequence is computable by the aforementioned result of Chaitin. We next show that this failure of the converse implication can be witnessed by a strongly K-trivial measure where all the atoms are recursive, and thus C-trivial (as sequences).

Proposition 5.15. *There is a strongly K-trivial measure which is not C-trivial and which has only recursive atoms.*

Proof. Let A be an r.e. K-trivial set which is not C-trivial, that is, which is not recursive. Let a_k be the least number m such that $C(A \upharpoonright_m) \geq C(m) + 4^k$; this is defined because as A is not C-trivial. Furthermore, as A and the function C are both limit-recursive, the numbers a_k can be computed in the limit and have uniformly recursive approximations $a_{k,s}$ which converge to a_k, though not monotonously.

One next defines a sequence $b_{k,s}$ such that $b_{k,0} = 0$ and $b_{k,s+1} \in \{b_{k,s}, s+1\}$ where the value $s+1$ is taken if

$$a_{k,s+1} > b_{k,s} \text{ or } a_{k,s} \neq a_{k,s+1} \text{ or there is } h < k \text{ with } b_{h,s+1} \geq b_{k,s}.$$

Note that the definition of these values implies that

$$a_k \leq a_h \text{ and } a_k \leq b_k < b_h \text{ whenever } k \leq h.$$

Now let $A_k = \{a \in A : a < b_k\}$. The next claim will show that the A_k are all K-trivial with a small K-triviality constant and that for this, it is not needed to compute b_k uniformly from k; instead one computes the b_k from s and k only for $s \geq b_k$.

Claim 5.16. *The set A_k has the K-triviality constant $2\log(k) + c$ for some constant c.*

To see this, it will be shown that $A_k \restriction_s$ can be computed nonuniformly from $A \restriction_s$ and k. Suppose that one knows k and $A \restriction_s$ and suppose that $s \geq b_k$. Then one can compute s from $A \restriction_s$ as s is the length of $A \restriction_s$. Furthermore, b_k equals $b_{k,s}$ and $A_k \restriction_s$ equals $A_k = \{a \in A : a < b_k\} \restriction_s$, thus both can be computed from $A \restriction_s$ and k. Below b_s, $A_k \restriction_s = A \restriction_s$ and no computations are needed. As A is K-trivial, it follows that

$$K(A_k \restriction_s) \leq K(A \restriction_s) + K(k) + c' \leq K(s) + 2\log(k) + c.$$

Thus each A_k is K-trivial with the K-trivially constant bounded by $2\log(k)+$ c for some constant c independent of k. This shows the claim.

The measure $\mu = \sum_k 2^{-k-1} \cdot \delta_{A_k}$ is strongly K-trivial by the result of the claim and the fact that $\sum_k 2^{-k-1} \cdot (2\log(k) + c)$ is finite. However, μ is not C-trivial, as $\mu \restriction_{a_k}$ satisfies for $h \geq k$ that $A_h \restriction_{a_k} = A \restriction_{a_k}$ and

$$C(\mu \restriction_{a_k}) \geq C(a_k) + 2^{-k} \cdot (C(A \restriction_{a_k}) - C(a_k)) \geq C(a_k) + 2^k;$$

note that the values 2^k are obviously not bounded by a constant. Furthermore, all atoms of μ are the sets A_k which are finite and thus recursive. □

6. Martin-Löf random measures

Let $\mathcal{M}(\{0,1\}^{\mathbb{N}})$ be the space of probability measures on Cantor space (which is canonically a compact topological space). Probability measures on this space have been defined by Mauldin and Monticino [18]. A special case is the uniform measure $\mathbb{P}$ on $\mathcal{M}(\{0,1\}^{\mathbb{N}})$.

To define $\mathbb{P}$, first let $\mathcal{R}$ be the set of representations of probability measures; namely, $\mathcal{R}$ consists of the functions $X\colon \{0,1\}^* \to [0,1]$ such that $X_\emptyset = 1$ and $X_\sigma = X_{\sigma 0} + X_{\sigma 1}$ for each string σ. We note that $\mathcal{R}$ is closed in the product topology. $\mathbb{P}$ is the unique measure on $\mathcal{R}$ such that for each string σ and $r, s \in [0,1]$, we have $\mathbb{P}(X_{\sigma 0} \leq r \mid X_\sigma = s) = \min(1, r/s)$. Intuitively, we choose $X_{\sigma 0}$ at random w.r.t. the uniformly distribution on the interval $[0, X_\sigma]$, and the choices made at different strings are independent. In the language of [18], the transition kernel τ takes the value λ for each dyadic rational (corresponding to a string σ).

We remark that for $n > 0$ and a string σ of length n, the function v_n on $[0,1]$ given by $x \to \mathbb{P}(\{\mu\colon \mu[\sigma] \leq x\})$ only depends on $n = |\sigma|$. In fact, the v_n are given by the recursion $v_1(x) = x$ and $v_{n+1}(x) = x + \int_x^1 v_n(x/t)dt$. E.g., $v_2(x) = x(1 - \ln x)$.

Culver's thesis [8] shows that the measure $\mathbb{P}$ is computable in the sense of Hoyrup and Rojas [12]. So the framework provided in [12] yields a notion of Martin-Löf randomness for points in the space $\mathcal{M}(\{0,1\}^{\mathbb{N}})$.

Proposition 6.1. *Every probability measure μ that is Martin-Löf random w.r.t. $\mathbb{P}$ is Martin-Löf absolutely continuous.*

For the duration of this proof let μ range over $\mathcal{M}(\{0,1\}^{\mathbb{N}})$. For an open set $G \subseteq \{0,1\}^{\mathbb{N}}$, let

$$r_G = \int \mu(G) d\mathbb{P}(\mu).$$

Our proof of Proposition 6.1 is based on two facts.

Fact 6.2. *$r_G = \lambda(G)$.*

Proof. Clearly, for each n we have

$$\sum_{|\sigma|=n} r_{[\sigma]} = \int \sum_{|\sigma|=n} \mu([\sigma]) d\mathbb{P}(\mu) = 1.$$

Furthermore, $r_\sigma = r_\eta$ whenever $|\sigma| = |\eta| = n$ because by the remark above there is a $\mathbb{P}$-preserving transformation T of $\mathcal{M}(\{0,1\}^{\mathbb{N}})$ such that $\mu([\sigma]) = T(\mu)([\eta])$. Therefore $r_{[\sigma]} = 2^{-|\sigma|}$.

If σ, η are incompatible then $r_{[\sigma]\cup[\eta]} = r_{[\sigma]} + r_{[\eta]}$. Now it suffices to write $G = \bigcup_i [\sigma_i]$ where the strings σ_i are incompatible, so that $\lambda G = \sum_i 2^{-|\sigma_i|}$. □

Fact 6.3. *Let $\mu \in \mathcal{M}(\{0,1\}^{\mathbb{N}})$ and let $\langle G_m \rangle_{m\in\mathbb{N}}$ be a Martin-Löf test such that there is $\delta \in \mathbb{Q}^+$ with $\forall m\, [\mu(G_m) > \delta]$. Then μ is not Martin-Löf random w.r.t. $\mathbb{P}$.*

Proof. Observe that by the foregoing fact

$$\delta \cdot \mathbb{P}(\{\mu \colon \mu(G_m) \geq \delta\}) \leq \int \mu(G_m) d\mathbb{P}(\mu) = \lambda(G_m) \leq 2^{-m}.$$

Let $\mathcal{G}_m = \{\mu \colon \mu(G_m) > \delta\}$ which is uniformly effectively open in the space of measures $\mathcal{M}(\{0,1\}^{\mathbb{N}})$. Fix k such that $2^{-k} \leq \delta$. By the inequality above, we have $\mathbb{P}(\mathcal{G}_m) \leq 2^{-m}/\delta \leq 2^{-m+k}$. Hence $\langle \mathcal{G}_{m+k} \rangle_{m\in\mathbb{N}}$ is a Martin-Löf test w.r.t. $\mathbb{P}$ that succeeds on μ. □

This argument also works for randomness notions stronger than Martin-Löf's. For instance, if there is a weak-2 test $\langle G_m \rangle_{m\in\mathbb{N}}$ such that $\mu G_m > \delta$ for each m, then μ is not weakly 2-random with respect to $\mathbb{P}$. The converse of Proposition 6.1 fails. Culver [8] shows that each measure μ that is Martin-Löf random w.r.t. $\mathbb{P}$ is non-atomic. So a measure δ_Z for a Martin-Löf random bit sequences Z is ML absolutely continuous but not Martin-Löf random with respect to $\mathbb{P}$.

7. Being ML-a.c. relative to computable ergodic measures

We review some notions from the field of symbolic dynamics, a mathematical area closely related to Shannon information theory. See e.g. [34] for more detail. Thereafter we will consider effective "almost-everywhere theorems" related to that area in the framework of randomness for measures.

In symbolic dynamics it can be useful to admit alphabets other than the binary one. Let $\mathbb{A}^\infty$ denote the topological space of one-sided infinite sequences of symbols in an alphabet $\mathbb{A}$. Randomness notions etc. carry over from the case of $\mathbb{A} = \{0,1\}$. A dynamics on $\mathbb{A}^\infty$ is given by the shift operator T, which erases the first symbol of a sequence. A measure ρ on $\mathbb{A}^\infty$ is called *shift invariant* if $\rho(G) = \rho(T^{-1}(G))$ for each open (and hence each measurable) set G. The *empirical entropy* of a measure ρ along $Z \in \mathbb{A}^\infty$ is given by the sequence of random variables

$$h_n^\rho(Z) = -\frac{1}{n}\log_{|\mathbb{A}|}\rho[Z\restriction_n].$$

A shift invariant measure ρ on $\mathbb{A}^\infty$ is called *ergodic* if every ρ integrable function f with $f \circ T = f$ is constant ρ-almost surely. The following equivalent condition can be easier to check: for any strings $u, v \in \mathbb{A}^*$,

$$\lim_N \frac{1}{N}\sum_{k=0}^{n-1}\rho([u] \cap T^{-k}[v]) = \rho[u]\rho[v].$$

For ergodic ρ, the entropy $H(\rho)$ is defined as $\lim_n H_n(\rho)$, where

$$H_n(\rho) = -\frac{1}{n}\sum_{|w|=n}\rho[w]\log\rho[w].$$

Thus, $H_n(\rho) = \mathbb{E}_\rho h_n^\rho$ is the expected value with respect to ρ. By the concavity of the logarithm function the limit exists and equals the infimum of the sequence. This limit is denoted $H(\rho)$, the entropy of ρ.

A well-known result from the 1950s due to Shannon, McMillan and Breiman states that for an ergodic measure ρ, for ρ-a.e. Z the empirical entropy along Z converges to the entropy of the measure. See e.g. Shields [34, Chapter 1], but note that the result is called the Entropy Theorem there.

Theorem 7.1 (SMB theorem)**.** *Let ρ be an ergodic measure on the space $\mathbb{A}^\infty$. For ρ-almost every Z we have $\lim_n h_n^\rho(Z) = H(\rho)$.*

Recall from Fact 2.4 that a measure μ is *ML absolutely continuous with respect to ρ* iff $\mu(\mathcal{C}) = 0$ where $\mathcal{C}$ is the class of sequences in $\mathbb{A}^\infty$ that are not Martin-Löf random with respect to ρ.

If a computable measure ρ is shift invariant, then $\lim_n h_n^\rho(Z)$ exists for each ρ-Martin-Löf random Z by a result of Hochman [10]. Hoyrup [11, Theorem 4.1] gave an alternative proof for ergodic ρ, and also showed that in that case we have $\lim_n h_n^\rho(Z) = H(\rho)$ for each ρ-ML-random Z. We extend this result to measures μ that are ML absolutely continuous with respect to ρ, under the additional hypothesis that the h_n^ρ are uniformly bounded. This holds e.g. for Bernoulli measures and the measures given by a Markov process.

Proposition 7.2. *Let ρ be a computable ergodic measure on the space $\mathbb{A}^\infty$ such that for some constant D, each h_n^ρ is bounded above by D. Suppose a measure μ is ML absolutely continuous with respect to ρ. Write $s = H(\rho)$. Then $\lim_n E_\mu|h_n^\rho - s| = 0$.*

Proof. By Hoyrup's result, $\lim_n h_n^\rho(Z) = s$ for each ρ-ML-random Z. Since the sequences that are not Martin-Löf random w.r.t. ρ form a null set w.r.t. μ, we infer that $\lim_n |h_n^\rho(Z) - s| = 0$ for μ-a.e. Z. Since the exception set is measurable and the $h_n\rho$ are bounded, the Dominated Convergence Theorem now shows that $\lim_n \mathbb{E}_\mu|h_n^\rho - s| = 0$, as required. $\square$

We now given an example showing that the boundedness hypothesis on the h_n^ρ is necessary. We provide a computable ergodic measure ρ such that some finite measure $\mu \ll \rho$ makes the sequence $E_\mu h_n^\rho$ converge to ∞.

Proposition 7.3. *There are an ergodic computable measure ρ (associated to a binary renewal process) and a computable measure $\mu \ll \rho$ such that $\lim_n \mathbb{E}_\mu h_n^\rho = \infty$. (We can then normalise μ to become a probability measure, while maintaining the same conclusion.)*

Proof. Let k range over positive natural numbers. The real $c = \sum_k 2^{-k^4}$ is computable. Let $p_k = 2^{-k^4}/c$ so that $\sum p_k = 1$. Let $b = \sum_k k \cdot p_k$, and note that b is also computable.

Let ρ be the shift invariant measure associated with the binary renewal process corresponding to the sequence $\langle p_k \rangle_{k \in \mathbb{N}}$ and $1/b$, which is given by the conditions

$$\rho[Z_0 = 1] = 1/b \text{ and } \rho(10^k 1 \prec Z \mid Z_0 = 1) = p_k.$$

Informally, the process has initial value 1 with probability $1/b$, and after each 1, with probability p_k it takes k many 0s until it reaches the next 1. See again e.g. [34, Section 1.2c] where it is shown that a measure ρ of this kind is ergodic. Write $v_k = 10^k 1$. Note that $\rho[v_k] = p_k/b$.

Define a function f in $L_1(\rho)$ by $f(v_k\hat{\ }Z) = k^{-2}/p_k$ and $f(X) = 0$ for any X not extending any v_k. It is clear that f is $L_1(\rho)$-computable, in the usual sense that there is an effective sequence of basic functions $\langle f_n \rangle$ converging effectively to f in the L_1-norm: let $f_n(X) = f(X)$ in case $v_k \prec X$, $k \le n$, and $f_n(X) = 0$ otherwise. Define the measure μ by $d\mu = f d\rho$, i.e. $\mu(A) = \int_A f d\rho$. Thus $\mu[v_k] = k^{-2}/b$. Since ρ is computable and f is $L_1(\rho)$-computable, μ is computable. Also note that $\mu(\{0,1\}^{\mathbb{N}}) = \int f d\rho$ is finite.

For any $n > 2$, letting $k = n - 2$, we have

$$E_\mu h_n^\rho \ge -\frac{1}{n}\mu[v_k]\log\rho[v_k] = -\frac{1}{nk^2 b}(-k^4 - \log(bc)) \ge \frac{k^2}{nb} - O(1).$$

This completes the proof. □

The next observation shows that the asymptotic initial segment complexity of a ML absolutely continuous measure relative to ρ obeys some lower bound. Note that $H(\lambda) = 1$. So for $\rho = \lambda$, this shows that in Example 4.1 we cannot subtract, say, $n/4$ instead of n^θ.

Proposition 7.4. *Let ρ be a computable ergodic measure, and suppose μ is a ML absolutely continuous measure with respect to ρ. Then*

$$\lim_n \frac{1}{n} K(\mu\restriction_n) = \lim_n \frac{1}{n} C(\mu\restriction_n) = H(\rho).$$

Proof. We can use K and C interchangeably because $C(x) \le^+ K(x) \le^+ C(x) + K(C(x))$ [24, Proposition 2.4.1]. We settle on K.

Let $k_n(Z) = K(Z\restriction_n)/n$. The argument is very similar to the one in Proposition 7.2 above, replacing the functions h_n by the k_n. Note that k_n is bounded above by a constant because $K(x) \le^+ |x| + 2\log|x|$. Hoyrup's result [11, Theorem 4.1] states that $\lim_n k_n(Z) = H(\rho)$ for each ρ-ML-random Z. Now we can apply the Dominated Convergence Theorem as in the proof of Proposition 7.2. □

8. Conclusions and open questions

In the present paper, we studied probability measures on Cantor space with respect to algorithmic randomness and triviality properties.

A main property we studied was that a measure μ is ML absolutely continuous: for every Martin-Löf test the μ-measure of its components converges to 0. We provided several examples and showed a robustness property related to Solovay tests. In Section 4 we introduced strong Solovay tests for measures. We leave the following question open.

Question 8.1. *If a measure passes every* strong *Solovay test, does it pass every Solovay test?*

Thereafter, we studied measures μ with respect to the growth behaviour of the μ-averages of the descriptive string complexities $C(x)$ and $K(x)$, taken over all strings x of length n. We considered measures where the growth is as fast as possible, and others where it is as slow as possible. In the second direction, we looked at C- and at K-trivial measures, where the averages are bounded up to a constant by $C(n)$ or $K(n)$, respectively. Such measures are atomic, that is, they are weighted, possibly infinite sums of Dirac measures. The atoms of such measure satisfy the corresponding triviality property; furthermore, the weighted sum of the atoms has to converge fast, though we could not find a precise criterion for this. In this context, we mention two open questions.

Question 8.2. *Is every C-trivial measure also K-trivial?*

An affirmative answer to the next question would lead to an affirmative answer to Question 8.2; Proposition 5.8 provided more detail.

Question 8.3. *Is there a constant bound on the function $n \mapsto K(C(n)|n, K(n))$?*

Measures μ on the upper end of the growth spectrum are ML absolutely continuous. However, we showed that also here there is no exact correspondence between the growth-rate of the μ averages of the string complexity and the property of being ML absolutely continuous; instead, in the border area, there are measures that are smaller with respect to the growth-rate of the averages which are ML absolutely continuous, and measures that are larger but which are not ML absolutely continuous. In particular, an analogue of the Levin-Schnorr Theorem does not hold for measures. We noted that near the upper end of the growth spectrum for C are the Kolmogorov random measures μ, the ones for which there is a constant d with $C(\mu\restriction_n) \geq n - d$ for infinitely many n. Besides the usual uniform measure, there are other examples, such as the Dirac measure of a set that is 2-random. While the classes of C-trivial and K-trivial measures are closed under finite convex combinations, this condition is not known to hold for classes of measures at the upper end of the growth spectrum.

Question 8.4. *Assume that measures μ, ν are Kolmogorov random. Is every convex combination of μ and ν Kolmogorov random as well?*

As mentioned in Section 3, a sequence X is 2-random iff it is Kolmogorov random as defined in [16]. So an affirmative answer to the question above implies an affirmative answer to the following:

Question 8.5. *If $X, Y \in \{0,1\}^{\mathbb{N}}$ are both 2-random, does there exist a constant d such that infinitely many n satisfy both $C(X\restriction_n) \geq n - d$ and $C(Y\restriction_n) \geq n - d$?*

The following diagram summarizes the implications obtained between (weak) randomness notions for measures. Non-labelled implication arrows are trivial. We ignore at present whether the implications given by the arrows labelled 3.5 and 3.2 are proper. All the other implication arrows are proper. In most cases this can be seen already for Dirac measures. For instance, Martin-Löf randomness of a measure implies that the measure has no atoms, so the converse of the implication labeled Prop. 6.1 fails. We can rule out two further implications. Each ML-random μ is orthogonal to λ by Bienvenu and Culver [8, Section 2.6], and hence not absolutely continuous. We ignore whether ML-randomness implies one of the condition stronger than ML-a.c. displayed in the lower part of the diagram.

$$\begin{tikzcd}
\mu \text{ is abs. continuous} \arrow[drr] & & \\
\mu \text{ is Martin-Löf random} \arrow[rr, "\text{Prop. 6.1}"] & & \mu \text{ is ML a.c.} \\
\exists^\infty n\, K(\mu\restriction_n) \geq^+ n + K(n) \arrow[r, "3.5"'] & \exists^\infty n\, C(\mu\restriction_n) \geq^+ n \arrow[ur, "\text{Thm. 3.6}"'] & \\
\forall n\, K(\mu\restriction_n) \geq^+ n + K(n) \arrow[u] \arrow[r, "3.2"'] & \forall n\, C(\mu\restriction_n) \geq^+ n \arrow[u] &
\end{tikzcd}$$

In Section 7 we considered ML absolute continuity relative to a general ergodic computable measure. We proved appropriate effective versions of the Shannon-McMillan-Breiman theorem and the Brudno theorem where the bit sequences are replaced by measures; in the former case we needed an additional boundedness hypothesis.

One possible effective version of the Birkhoff ergodic theorem states that for an ergodic computable ρ, if $f\{0,1\}^{\mathbb{N}} \to \mathbb{R}$ is ρ-integrable and lower semicomputable and Z is ρ-Martin-Löf random, then the limit of the usual ergodic averages $A_n f(Z) = \frac{1}{n}\sum_{k<n}(f \circ T^k)(Z)$ equals $\int f d\rho$. For background see e.g. [23] which contains references to original work. If in addition the $A_n f$ are bounded then an argument similar to the one in the proof of Proposition 7.2 shows that $\lim_n \int (A_n f) d\mu = \int f d\rho$ for any measure $\mu \ll_{ML} \rho$. However, it is unknown whether this additional hypothesis is necessary.

Acknowledgements. The authors would like to thank the anonymous referees of STACS 2020 for useful suggestions. They also thank the journal referee for very helpful comments. They would like to thank Benjamin Weiss and Bruno Bauwens for helpful correspondence.

Some problems from this paper were presented at the 2020 AIM workshop on randomness and applications, for instance Question 8.3. The authors thank the participants, in particular Joseph S. Miller, Alexander Shen and Elvira Mayordomo, for their interest and their useful comments.

References

[1] Bruno Bauwens. Computability in statistical hypotheses testing, and characterizations of independence and directed influences in time series using Kolmogorov complexity. PhD Thesis, Universiteit Gent, Faculteit Ingenieurswetenschappen, February 2010. https://biblio.ugent.be/publication/1107852/file/4335294

[2] Bruno Bauwens. Relating and contrasting plain and prefix Kolmogorov complexity. *Theory of Computing Systems*, 58(3):482–501, 2016.

[3] Tejas Bhojraj. *Algorithmic Randomness and Kolmogorov Complexity for Qubits.* Ph.D. Dissertation, University of Wisconsin at Madison, 2021. Available at arXiv:2106.14280.

[4] Laurent Bienvenu, Wolfgang Merkle and Alexander Shen. A simple proof of the Miller-Yu Theorem. *Fundamenta Informaticae*, 83(1-2):21–24, 2008.

[5] Cristian S. Calude. *Information and Randomness - An Algorithmic Perspective.* Second Edition, Springer, Heidelberg, 2002.

[6] Gregory J. Chaitin. A theory of program size formally identical to information theory. *Journal of the Association for Computing Machinery*, 22:329-340, 1975.

[7] Gregory J. Chaitin. Information-theoretical characterizations of recursive infinite strings. *Theoretical Computer Science*, 2:45–48, 1976.

[8] Quinn Culver. *Topics in Algorithmic Randomness and Effective Probability.* PhD thesis, University of Notre Dame, 2015.

[9] Rodney G. Downey and Denis Hirschfeldt. *Algorithmic Randomness and Complexity.* Springer-Verlag, Berlin, 2010. 855 pages.

[10] Michael Hochman. Upcrossing inequalities for stationary sequences and applications. *Annals of Probability*, 37(6):2135–2149, 2009.

[11] Mathieu Hoyrup. The dimension of ergodic random sequences. In Christoph Dürr and Thomas Wilke, editors, *STACS 2012*, pages 567–576, 2012.

[12] Mathieu Hoyrup and Cristóbal Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.

[13] Howard P. Katseff. Complexity dips in random infinite binary sequences. *Information and Control*, 38(3):258–263, 1978.

[14] Steven M. Kautz, *Degrees of random sets*, PhD thesis, Cornell University, August 1991.

[15] Leonid A. Levin. The concept of a random sequence. *Doklady Akademii Nauk SSSR*, 212:548–550, 1973.

[16] Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and its Applications.* Texts in Computer Science. Springer, New York, third edition, 2008.

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

[18] Daniel Mauldin and Michael Monticino. Randomly generated distributions. *Israel Journal of Mathematics*, 91(1-3):215–237, 1995.

[19] Joeseph S. Miller. The K-degrees, low for K-degrees, and weakly low for K sets. *Notre Dame Journal of Formal Logic*, 50(4):381–391, 2009.

[20] Joseph S. Miller. Contrasting plain and prefix-free Kolmogorov complexity. *Research Note on the author's homepage.*

[21] Joseph S. Miller and Liang Yu. On initial segment complexity and degrees of randomness. *Transactions of the American Mathematical Society*, 360:3193–3210, 2008.

[22] Joseph S. Miller and Liang Yu. Oscillation in the initial segment complexity of random reals. *Advances in Mathematics*, 226(6):4816–4840, 2011.

[23] Kenshi Miyabe, André Nies and Jing Zhang. Using almost-everywhere theorems from analysis to study randomness. *Bull. Symb. Logic*, 22:305–331, 2016.

[24] André Nies. *Computability and Randomness*, volume 51 of *Oxford Logic Guides.* Oxford University Press, Oxford, 2009. 444 pages. Paperback version 2011.

[25] André Nies (editor). Logic Blog 2017. Available at `http://arxiv.org/abs/1804.05331`, 2017.

[26] André Nies and Volkher Scholz. Martin-Löf random quantum states. *Journal of Mathematical Physics*, 60(9):092201, 2019. available at doi.org/10.1063/1.5094660.

[27] André Nies, Frank Stephan and Sebastiaan Terwijn. Randomness, relativization and Turing degrees. *The Journal of Symbolic Logic*, 70(2):515–535, 2005.

[28] André Nies and Frank Stephan. Randomness and Initial Segment Complexity for Probability Measures. In Christophe Paul and Markus Bläser, editors, *37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020)*, volume 154 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 55:1–55:14, Dagstuhl, Germany, 2020.

[29] Piergiorgio Odifreddi. *Classical Recursion Theory.* North-Holland, Amsterdam, 1989.

[30] Piergiorgio Odifreddi. *Classical Recursion Theory, Volume II.* Elsevier, Amsterdam, 1999.

[31] Christopher P. Porter. Trivial measures are not so trivial. *Theory of Computing Systems*, 56(3):487–512, 2015.

[32] Claus-Peter Schnorr. Process complexity and effective random tests. *Journal on Computer and System Sciences*, 7:376–388, 1973.

[33] Claus-Peter Schnorr. A survey of the theory of random sequences. *Basic Problems in Methodology and Linguistics*, (*Proceedings of the Fifth International Congress on Logic, Methodology and Philosophy of Science*, Part III, University Western Ontario, London, Ontario, 1975), University Western Ontario Series Philosophy of Science, volume 11, pages 193–211. Reidel, Dordrecht, 1977.

[34] Paul C. Shields. *The Ergodic Theory of Discrete Sample Paths*. Graduate Studies in Mathematics 13. American Mathematical Society, 1996.

[35] Robert I. Soare. *Recursively Enumerable Sets and Degrees. A Study of Computable Functions and Computably Generated Sets*. Springer, Heidelberg, 1987.

[36] Robert Solovay. Handwritten manuscript related to Chaitin's work. IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 215 pages, 1975.

[37] Frank Stephan. *Recursion Theory.* Lecture notes, School of Computing, National University of Singapore, Technical Report TR10/12, 2012.