

# Solovay functions and their applications in algorithmic randomness

Laurent Bienvenu<sup>a</sup> Rod Downey<sup>b</sup> André Nies<sup>c</sup>  
Wolfgang Merkle<sup>d</sup>

<sup>a</sup>*Laboratoire CNRS J.-V. Poncelet, Moscow, Russia*

<sup>b</sup>*School of Mathematics, Statistics and Operations Research, Victoria University of Wellington, New Zealand*

<sup>c</sup>*Department of Computer Science, University of Auckland, New Zealand*

<sup>d</sup>*Institut für Informatik, Heidelberg University, Germany*

---

## Abstract

Classical versions of Kolmogorov complexity are incomputable. Nevertheless, in 1975 Solovay showed that there are computable functions  $f \geq K + O(1)$  such that for infinitely many strings  $\sigma$ ,  $f(\sigma) = K(\sigma) + O(1)$ , where  $K$  denotes prefix-free Kolmogorov complexity (while  $C$  denotes plain Kolmogorov complexity). Such an  $f$  is now called a Solovay function. We prove that many classical results about  $K$  can be obtained by replacing  $K$  by a Solovay function. For example, the three following properties of a function  $g$  all hold for the function  $K$ .

- (i) The sum of the terms  $\sum_n 2^{-g(n)}$  is a Martin-Löf random real.
- (ii) A sequence  $A$  is Martin-Löf random if and only if  $C(A \upharpoonright n) > n - g(n) - O(1)$ .
- (iii) A sequence  $A$  is  $K$ -trivial if and only if  $K(A \upharpoonright n) < g(n) + O(1)$ .

We show that when fixing any of these three properties, then among all computable functions exactly the Solovay functions possess this property. Furthermore, this characterization extends accordingly to the larger class of right-c.e. functions.

*Key words:* Kolmogorov complexity, Algorithmic randomness, Solovay functions.

---

## 1 Introduction and overview

### 1.1 Introduction

A fundamental aspect of Kolmogorov complexity is its inherent noncomputability. That is, standard complexities measuring compressibility, such as

plain (Kolmogorov) complexity  $C(\sigma)$  or prefix-free (Kolmogorov) complexity  $K(\sigma)$ , are as complicated to calculate as the halting problem, and hence undecidable. This nonalgorithmic aspect of Kolmogorov complexity is useful in the sense that it can enable new proofs of various undecidability results such as Gödel's Incompleteness Theorems (see [24, Chapter 2] for the First Incompleteness Theorem, and [19] for the Second Incompleteness Theorem). On the other hand, this aspect also hinders the use of Kolmogorov complexity as a tool for measuring common information in real data. This hindrance necessitates the use of replacements for the classical complexities by computable text compressions (see for example, Cilibrasi and Vitányi [11]).

Nevertheless, it is remarkable how much of the classical theory of Kolmogorov complexities can be carried out by using *good effective upper bounds*. For example, for plain complexity we know that the identity machine describing  $\sigma$  by  $\sigma$  itself will give us  $f(\sigma) = |\sigma|$  as a simple computable upper bound, and Kolmogorov's basic counting argument shows that this bound is achieved for each  $n$  by some  $\sigma$  of length  $n$ . Thus the function  $f(\sigma) = |\sigma|$  is a good computable upper bound for  $C(\sigma)$  in the sense that

- (i)  $C(\sigma) \leq f(\sigma) + O(1)$ , and
- (ii)  $\exists^\infty \sigma f(\sigma) \leq C(\sigma) + O(1)$ .

This upper bound has a number of uses. One is Chaitin's result [9] that if  $A$  is *C-trivial*, which means that  $C(A \upharpoonright n) \leq C(n) + O(1)$  for all  $n$ , then  $A$  is computable.

The situation for prefix-free complexity is not so clear. Similar to the simple upper bound  $|\sigma|$  for  $C(\sigma)$ , the function  $|\sigma| + K(|\sigma|)$  is an upper bound for  $K(\sigma)$  that is infinitely often tight up to an additive constant. However, this upper bound is not computable but only right-c.e., i.e., the binary relation " $g(n) < k$ " is c.e. Nevertheless, by a theorem of Solovay there are such infinitely often tight upper bounds for  $K$  that are computable and such functions are now called Solovay functions.

**Definition 1.1** *A function  $g$  is a Solovay function if  $g$  is computable and it holds that*

- (i)  $K(\sigma) \leq g(\sigma) + O(1)$ , and
- (ii)  $\exists^\infty \sigma g(\sigma) \leq K(\sigma) + O(1)$ .

*A function  $g$  is a weak Solovay function if  $g$  is right-c.e. and satisfies (i) and (ii).*

In what follows we derive a number of fundamental results about Solovay functions and weak Solovay functions. Recall that the three following properties

of a function  $g$  all hold for the function  $K$ , by well-known results or, in case of the third one, by definition.

- (i) The sum  $\sum_n 2^{-g(n)}$  is a Martin-Löf random real.
- (ii) A sequence  $A$  is Martin-Löf random if and only if  $C(A \upharpoonright n) \geq n - g(n) - O(1)$ .
- (iii) A sequence  $A$  is  $K$ -trivial if and only if  $K(A \upharpoonright n) \leq g(n) + O(1)$ .

Our main results are that when fixing any of these three properties, then the property is true for all weak Solovay functions but indeed is false for all right-c.e. functions that are not weak Solovay functions. That is, among all right-c.e. functions exactly the weak Solovay functions possess this property. This characterization extends trivially to the special case of Solovay functions in the sense that among all computable functions exactly the Solovay functions possess the property under consideration. Note that in case of the first property the latter characterization was shown in a conference article by the first two authors of this paper [5], whereas the corresponding characterization of weak Solovay functions was subsequently demonstrated by Hölzl, Kräling and Merkle [17], see Section 2.2 for further details.

Quite aside from their intrinsic interest, Solovay functions have found many uses in the study of algorithmic randomness, many using the results proven in this paper some of which had been reported in earlier conference papers [5,6]. Recall that  $A$  is called  $K$ -trivial if and only if there is a constant  $b \in \mathbb{N}$  such that  $K(A \upharpoonright n) \leq K(n) + b$  for all  $n$ . Barmpalias and Sterkenburg [2] used Solovay functions to show that calculating the number of  $K$ -trivials with constant  $d$  is sharply  $\Delta_3^0$ . More recently, Solovay functions were used for the characterization of  $K$ -trivial points in computable metric spaces by Melnikov and Nies [25]. A research announcement by Bienvenu, Day, et al. [4] contains a new, golden-run-free proof that  $K$ -triviality implies lowness for  $K$ . One ingredient is our new, golden-run-free proof, based on Solovay functions, that every  $K$ -trivial is Turing below a c.e.  $K$ -trivial (Section 4.4). For background on the golden run method see Nies [32, Section 5.4].

## 1.2 Overview

The properties (i), (ii), and (iii) of  $K$  stated in the last section, which characterize weak Solovay functions, are discussed in Section 2, 3, and 4, respectively. In Section 2.1 we review Solovay's construction of a Solovay function before we discuss in Section 2.2 the characterization of weak Solovay functions by property (i), i.e., by the fact that the sum of the terms  $2^{-g(n)}$  is a Martin-Löf random real.

In Section 3, for a start we investigate into conditions formulated in terms of complexity gaps that are sufficient or necessary for being Martin-Löf ran-

dom. We derive several negative results in this direction, for example, there is no function  $h$  that tends to infinity such that  $K(A \upharpoonright n) \geq n - h(n) - O(1)$  implies that  $A$  is Martin-Löf random. These results will be used in subsequent proofs but have some interest in their own. In particular, they contrast a positive result by Csisma and Montalbán [12], who gave a sufficient condition for K-triviality in terms of a complexity gap. Section 3.2 is then about characterizations of Martin-Löf randomness via weak Solovay functions and vice versa. Recall that the Gács-Miller-Yu Theorem [15,28] gives a characterization of Martin-Löf randomness in terms of plain Kolmogorov complexity  $C$ : a sequence  $A$  is Martin-Löf random if and only if  $C(A \upharpoonright n) \geq n - K(n) - O(1)$ ; moreover, there exists a computable upper bound  $f$  for  $K$  such that  $A$  is Martin-Löf random if and only if  $C(A \upharpoonright n) \geq n - f(n) - O(1)$ . We show that any right-c.e. function  $f$  characterizes Martin-Löf randomness by the latter condition if and only if  $f$  is a weak Solovay function.

In Section 4, we explore the relationship between K-triviality and Solovay functions, showing that for any weak Solovay function  $g$ , a sequence  $A$  is K-trivial if and only if  $K(A \upharpoonright n) \leq g(n) + O(1)$ . Again the characterization works both ways. Namely, if  $f$  is right-c.e. and has the property that  $K(A \upharpoonright n) \leq f(n) + O(1)$  is equivalent to  $A$  being K-trivial, then  $f$  must be a weak Solovay function, and a similar characterization holds for Solovay functions. Nies [31] used the golden run method to prove that every K-trivial is truth-table reducible to a c.e. K-trivial. Using the connection between Solovay functions and K-trivials, we get a new and easy proof of the slightly weaker result that every K-trivial is Turing reducible to a c.e. K-trivial.

Finally, in Section 5, we look at the *c-hitting set* of a Solovay function  $g$ , which is the set of numbers  $n$  such that  $g(n) \leq K(n) + c$ . Any such set, unless finite, is shown to be Turing complete, and sparse, in the sense that it is hyperimmune.

### 1.3 Notation and preliminaries

Here we gather some notation that will be used throughout the paper. String refers to a binary string, and  $2^{<\omega}$  denotes the set of all strings. The length of a string  $x$  is denoted by  $|x|$ . Both, plain and prefix-free Kolmogorov complexity are defined on the set of finite strings, but as usual are also applied to other objects such as integers, rational numbers, pairs of strings, etc., as long as they can be effectively encoded by finite strings. Unless explicitly stated otherwise, sequence refers to an infinite binary sequence. A subset  $A$  of the natural numbers may be identified with its characteristic sequence  $A(0)A(1)\dots$ . Occasionally, sequences are referred to as *reals*, see the discussion on page 7. The set of all sequences is denoted by  $2^\omega$ . For  $i = 0, 1, \dots$ , the  $i$ -th bit of a sequence  $A$  is denoted by  $A(i)$ , while the prefix of  $A$  of length  $i$  is written  $A \upharpoonright i$ ,

that is,  $A = A(0)A(1)\dots$ , and  $A \upharpoonright i = A(0)\dots A(i-1)$  for  $i > 0$ , while  $A \upharpoonright 0$  is equal to the empty string. For a string  $\sigma$ , the *cylinder*  $[\sigma]$  is the set of sequences  $A$  such that  $\sigma$  is a prefix of  $A$ . If  $S$  is a *set* of strings, we write  $[S]$  for the set of sequences having some prefix in  $S$ , i.e.,  $S = \bigcup_{\sigma \in S} [\sigma]$ . We denote by  $\lambda$  the uniform measure on  $2^\omega$ , which is the probability measure one gets when each bit of a sequence is chosen at random with probabilities  $(1/2, 1/2)$  and independently of all the other bits.

An *order* is a nondecreasing, unbounded function from  $\mathbb{N}$  to  $\mathbb{N}$ . For a given order  $h$ , let  $h^{-1}(k)$  be the largest integer  $n$  such that  $h(n) \leq k$ . Note that for any order  $h$ , the function  $h^{-1}$  is itself an order, and is computable if and only if  $h$  is.

A function  $f: D \rightarrow \mathbb{R}$  is *left-c.e.* (also known as *approximable* or *semi-computable from below*) if there is a computable function  $(x, t) \mapsto f_t(x)$  such that for all  $x \in D$ ,  $f_t(x)$  is nondecreasing in  $t$  and converges to  $f(x)$ . The value  $f_t(x)$  is called the approximation of  $f(x)$  at stage  $t$ . The notion of a *right-c.e.* function is defined accordingly, where now the approximations  $f_t(x)$  are required to be nonincreasing in  $t$ . Note that for any left-c.e. or right-c.e. function one can assume that the approximations  $f_t(x)$  are given as a pair of natural numbers  $p$  and  $q$ , which represent the dyadic rational  $p2^{-q}$ . Furthermore, in the important case of right-c.e. functions with values in the natural numbers such as plain and prefix-free Kolmogorov complexity, by rounding approximations down to the next natural number below, one can assume that the approximations  $f_0(x), f_1(x), \dots$  are natural numbers, and that hence for each  $x$  all but finitely many of the approximations  $f_t(x)$  are equal to  $f(x)$ .

We will assume that the reader is familiar with the rudiments of algorithmic randomness, the definitions of plain and prefix-free complexity, and the like. We only remind the reader of some of the most salient points. A *bounded request set*, also known as Kraft-Chaitin or KC-set, is a computably enumerable set  $W$  of pairs  $(\sigma, n)$  where the first coordinate is a string and the second an integer and such that  $\sum_{(\sigma, n) \in W} 2^{-n}$  is finite. Enumerating a pair  $(\sigma, n)$  into a request set is often said to incur a *cost* of  $2^{-n}$ ; the request set being bounded refers to the fact that the total cost is finite. Given such a bounded request set  $W$ , the *Kraft-Chaitin theorem*, due to Levin [22] and Chaitin [8], asserts that  $K(\sigma) \leq n + O(1)$  for all  $(\sigma, n) \in W$ .

The classical application of the Kraft-Chaitin Theorem is the Levin-Schnorr characterization of Martin-Löf randomness.

**Theorem 1.2 (Levin-Schnorr [23,33])** *A sequence  $X$  is Martin-Löf random if and only if  $K(X \upharpoonright n) \geq n - O(1)$ .*

Another consequence of the Kraft-Chaitin Theorem is that every right-c.e. function of finite weight is an upper bound for Kolmogorov complexity.

**Lemma 1.3 (Levin [22], Chaitin [8])** *Let  $f: \mathbb{N} \rightarrow \mathbb{N}$  be a right-c.e. function. Then the two following assertions are equivalent.*

- (i)  $K \leq f + O(1)$ .
- (ii)  $\sum_n 2^{-f(n)}$  is finite.

**Proof** The implication (i) $\Rightarrow$ (ii) is trivial as  $\sum_n 2^{-K(n)} \leq 1$ . For the implication (ii) $\Rightarrow$ (i), fix some computable function  $(n, t) \mapsto f_t(n)$  with values in the natural numbers such that for all  $n$ , the sequence  $f_0(n), f_1(n), \dots$  converges nonincreasingly to  $f(n)$ , and let the set  $D_n$  be equal to  $\{f_0(n), \dots, f(n)\}$ , hence  $\sum_{\ell \in D_n} 2^{-\ell} \leq 2^{-f(n)+1}$ . Then the set of all pairs of the form  $(\ell, n)$  such that  $\ell \in D_n$  is a bounded request set and we are done by applying the Kraft-Chaitin theorem.  $\square$

## 2 Basics on Solovay functions

### 2.1 Existence

**Theorem 2.1 (Solovay [34])** *There exists a Solovay function.*

**Proof** Let us start by an observation. For some constant  $c$ , given strings  $p$  and  $x$  such that the standard universal prefix-free machine  $\mathbb{U}$  on input  $p$  outputs  $x$  after exactly  $t$  steps of computation, we have

$$K(\langle x, p, t \rangle) \leq |p| + c,$$

where  $\langle \cdot, \cdot, \cdot \rangle$  is the usual effective bijection from  $2^{<\omega} \times 2^{<\omega} \times 2^{<\omega}$  to  $2^{<\omega}$ . This holds because there is a prefix-free machine that on input  $p$  simulates  $\mathbb{U}$  on input  $p$  and outputs  $\langle x, p, t \rangle$  in case the simulated computation terminates in exactly  $t$  steps with output  $x$ . Suppose now that  $p$  is in addition a *shortest*  $\mathbb{U}$ -description for  $x$ , i.e.,  $\mathbb{U}(p) = x$  and  $K(x) = |p|$ . We then have

$$|p| = K(x) \leq K(\langle x, p, t \rangle) + O(1) \leq |p| + O(1).$$

Now let the function  $g_S$  be defined by

$$g_S(\langle x, p, t \rangle) = \begin{cases} |p| & \text{if } \mathbb{U} \text{ on input } p \text{ outputs } x \text{ in} \\ & \text{exactly } t \text{ steps of computations,} \\ 2|\langle x, p, t \rangle| & \text{otherwise.} \end{cases}$$

By construction, we have  $K \leq g_S + O(1)$ . Furthermore,  $g_S(\langle x, p, t \rangle)$  is at most  $K(\langle x, p, t \rangle) + O(1)$  for all triples  $\langle x, p, t \rangle$  such that  $p$  is a shortest  $\mathbb{U}$ -description for  $x$  and  $\mathbb{U}$  on input  $p$  outputs  $x$  in exactly  $t$  steps of computation. Thus,  $g_S$  is as desired.  $\square$

## 2.2 A criterion for being a Solovay function

The next theorem provides the fundamental characterization of Solovay functions in terms of the sum  $\sum_n 2^{-f(n)}$ . It was first proven in an earlier conference paper by Bienvenu and Downey [5] for computable functions, and then extended by Hölzl, Kräling and Merkle [17] to right-c.e. functions. The survey paper [7] by Bienvenu and Shen gives a further extension, namely to *real-valued* right-c.e. functions which only change finitely often at each value.

**Theorem 2.2** *Let  $f: \mathbb{N} \rightarrow \mathbb{N}$  be a right-c.e. function. Then  $f$  is a weak Solovay function if and only if  $\sum_n 2^{-f(n)}$  is finite and is a Martin-Löf random real.*

There are two different ways to define the notion of a Martin-Löf random real, however, both options lead to the same randomness notion. One option is to define Martin-Löf randomness for infinite binary sequences first, and define a real number to be Martin-Löf random if its binary expansion is. One other option is to directly adapt the definition of Martin-Löf randomness to real numbers, by saying that an open set  $\mathcal{U}$  of  $\mathbb{R}$  is effectively open, or  $\Sigma_1^0$ , if it can be written as  $\bigcup_i (a_i, b_i)$ , for a c.e. set of pairs  $\{(a_i, b_i)\}_{i \in \omega}$ , and considering the Lebesgue measure on  $\mathbb{R}$  instead of the uniform measure on  $2^\omega$ , the rest of the definition remaining the same. The second approach has been applied elegantly by Kučera and Slaman [21] for proving that left-computable Martin-Löf random reals are Solovay complete. Hölzl, Kräling and Merkle [17] use a similar argument in their proof of Theorem 2.2. We include this proof for ease of reference.

**Proof of Theorem 2.2** We first show the backwards direction of the equivalence asserted in the theorem. We assume that  $f$  is not a Solovay function and construct a sequence  $U_0, U_1, \dots$  of sets that is a Martin-Löf test and covers  $\Omega_f = \sum_n 2^{-f(n)}$ . First, we fix an appropriate computable approximation from above to  $f$ , i.e., a computable function  $(n, s) \mapsto f_s(n)$  such that for all  $n$  the sequence  $f_0(n), f_1(n), \dots$  is a nonascending sequence of natural numbers that converges to  $f(n)$ , where we assume in addition that  $f_s(n) - f_{s+1}(n)$  is always equal to either 0 or 1. Let  $a_0 = 0$  and given  $a_i$ , let  $a_{i+1} = a_i + d_i$ , where  $d_i$  is obtained as follows. With some appropriate ordering of pairs understood, search for the next pair that either is of the form  $(n, 0)$  or is of the

form  $(n, s + 1)$  where  $f_s(n) - f_{s+1}(n) = 1$ , and then let

$$d_i = 2^{-f_0(n)} \quad \text{or} \quad d_i = 2^{-f_{s+1}(n)} - 2^{-f_s(n)} = 2^{-f_s(n)},$$

respectively. In this situation, say that the increase from  $a_i$  to  $a_{i+1}$  of size  $d_i$  occurs due to  $n$ . Furthermore, let  $b_i$  be the sum of all increases  $d_j$  such that  $j \leq i$  and  $d_j$  and  $d_i$  occur due to the same  $n$ . By construction, all  $d_i$  and  $a_i$  are dyadic rationals and the  $a_i$  converge nondecreasingly to  $\Omega_f$ .

For given  $c$ , the component  $U_c$  is obtained as follows. Say that the index  $i$  is  $c$ -matched if  $d_i$  occurs due to  $n$  and it holds that

$$2^{c+2}b_i \leq 2^{-K(n)}. \tag{1}$$

For every index  $i$  for which it could be verified that  $i$  is  $c$ -matched by approximating  $K$  from above, add an interval of size  $2d_i$  to  $U_c$ , where this interval either starts at  $a_i$  or at the supremum of all reals that are already in  $U_c$ , whichever is larger.

By construction,  $2^{c+2}$  times the sum of all  $d_i$  such that  $i$  is matched is at most  $\Omega_K < 1$ , hence the measure of  $U_c$  is at most  $2^{-(c+1)}$ . Moreover, the sets  $U_0, U_1, \dots$  are uniformly c.e., hence they form a Martin-Löf test. This test covers  $\Omega_f$  because by the assumption that  $f$  is not a Solovay function, and by the fact stated above that information content measures are upper bounds for  $K$  up to an additive constant, it holds that

$$\lim_{n \rightarrow \infty} (f(n) - K(n)) = \infty.$$

Hence for every  $c$  there is an index  $i(c)$  such that all  $i \geq i(c)$  will eventually become  $c$ -matched, where then  $U_c$  will accordingly be increased by an interval of length  $2d_i$  to the right of  $a_{i(c)}$ . The lengths of these intervals add up to  $\beta = 2(\Omega_f - a_{i(c)})$ , hence  $U_c$  will cover, except for some gaps, the open interval between  $a_{i(c)}$  and  $a_{i(c)} + \beta$ , which contains  $\Omega_f$ .

It remains to show that  $\Omega_f$  cannot be a member of such a gap. Such a gap can only occur in case an interval of size  $2d_i$  is added to  $U_c$  in a situation where  $a_i$  is strictly larger than the supremum of all reals that are already in  $U_c$ , where then the supremum of the gap is  $a_i \leq \Omega_f$ . Consequently, the real  $\Omega_f$  cannot be a member of the gap unless it is equal to the rational number  $a_i$  and is hence not Martin-Löf random.

Next we show the forward direction of the equivalence asserted in the theorem. If  $f$  is a weak Solovay function, we already know by definition that  $\alpha = \sum_n 2^{-f(n)}$  is finite. Let us now prove that  $\alpha$  is a Martin-Löf random real. Suppose it is not. Then for arbitrarily large  $c$  there exists  $k$  such that  $K(\alpha|k) \leq$

$k - c$  (this because of the Levin-Schnorr theorem, Theorem 1.2). Given  $\alpha \upharpoonright k$ , one can effectively find some  $s$  such that

$$\sum_{n>s} 2^{-f(n)} \leq 2^{-k}.$$

Thus, by a standard bounded request argument, one has  $K(n|\alpha \upharpoonright k) \leq f(n) - k + O(1)$  for all  $n > s$ . Thus, for all  $n > s$ :

$$K(n) \leq f(n) + K(\alpha \upharpoonright k) - k + O(1) \leq f(n) + (k - c) - k - O(1) \leq f(n) - c - O(1).$$

And since  $c$  can be taken arbitrarily large, this shows that  $\lim_{n \rightarrow +\infty} f(n) - K(n) = +\infty$ , i.e.,  $f$  is not a Solovay function.  $\square$

The characterization of weak Solovay functions in Theorem 2.2 holds also when relativized to any oracle by virtually the same proof. In fact, Hölzl, Kräling and Merkle [17] demonstrated the relativized version, as a joint generalization of, first, the already mentioned corresponding characterization of Solovay functions by Bienvenu and Downey [5], and, second, a remarkable characterization of the sequences that are weakly low for  $K$  by Miller [29]. For the sake of completeness, we shortly review the latter result, which now becomes a special case of Theorem 2.2. Recall that the  $K$ -trivial sequences are exactly those that are low for  $K$ , i.e., a sequence  $A$  is  $K$ -trivial if and only if  $K^A \geq K - O(1)$ . The notion of lowness for  $K$  can be weakened as follows. We say that  $A$  is *weakly low for  $K$*  if  $K^A(n) \geq K(n) - O(1)$  for infinitely many  $n$ .

**Theorem 2.3 (Miller)** *The two following assertions are equivalent for all sequences  $A$ .*

- (i)  $A$  is weakly low for  $K$ .
- (ii)  $\Omega_K = \sum_n 2^{-K(n)}$  is Martin-Löf random relative to  $A$ .

**Proof** The function  $K$  is right-c.e., hence  $A$ -right-c.e., and is an upper bound for  $K^A$  up to an additive constant. So by definition of weakly low the first statement holds if and only if  $K$  is a weak Solovay function relative to  $A$ . The theorem is then immediate from the relativization of Theorem 2.2 to  $A$ .  $\square$

Weak Solovay functions are by definition upper bounds for  $K$ , hence always tend to infinity. We demonstrate next that indeed some Solovay functions do so nondecreasingly.

**Corollary 2.4** *There exists a Solovay function that is an order.*

**Proof** Take a left-c.e. Martin-Löf random real  $\Omega$  and write  $\Omega = \sum_n 2^{-k_n}$  where  $(k_n)$  is a computable sequence of integers. Now, transform this sum as follows: rewrite each term  $2^{-k_n}$  as a sum of finitely many terms  $2^{-k'_n}$  where  $k'_n = \max\{k_i \mid i \leq n\}$ . Let  $l_0, l_1, \dots$  be the sequence of values  $k'_n$  with repetitions obtained this way and consider the function  $f: i \mapsto l_i$ . Then the function  $f$  is a computable order, and the sum  $\sum_i 2^{-f(i)}$  is equal to the Martin-Löf random number  $\Omega$ , hence  $f$  is a Solovay function by Theorem 2.2.  $\square$

### 3 Connections to Martin-Löf randomness

#### 3.1 A “no-gap” theorem for randomness

In this section we will see how Solovay functions relate to Martin-Löf randomness, via the Gács-Miller-Yu theorem, which is stated below as Theorem 3.4. To do so, we first investigate into gap phenomena for Martin-Löf randomness, which are interesting in their own right.

By definition, a sequence is K-trivial if there is a constant  $c$  such that for all  $n$  we have that  $K(A \upharpoonright n) \leq K(n) + c$ . Csima and Montalbán [12] proved that there is an order  $h$  such that  $K(A \upharpoonright n) \leq K(n) + h(n)$  implies that  $A$  is K-trivial; in fact their function  $h$  is  $\Delta_4^0$ . Since there is a prefix-free machine that on input  $p$  outputs  $|\mathbb{U}(p)|$  whenever  $\mathbb{U}(p)$  is defined, there is a constant  $c$  such that for all strings  $\sigma$  it holds that  $K(|\sigma|) - c \leq K(\sigma)$ . Hence their result can be restated as follows: a sufficient condition for a sequence to be K-trivial is that for all  $n$  the prefix-free Kolmogorov complexity of the length  $n$  prefix of the sequence lies in the gap formed by  $K(n) - c$  and  $K(n) + h(n)$ . In what follows, we will show “no-gap theorems” for the concept of Martin-Löf randomness in the sense that for the latter concept it is not possible to obtain similarly sufficient or necessary conditions in terms of gap functions.

Chaitin [10] proved that when a sequence  $A$  is Martin-Löf random, then one does not just have  $K(A \upharpoonright n) \geq n - O(1)$ , but in fact  $K(A \upharpoonright n) - n$  tends to infinity. Together with the Levin-Schnorr characterization, this shows a dichotomy: given a sequence  $A \in 2^\omega$ , either  $A$  is not Martin-Löf random, in which case  $K(A \upharpoonright n) - n$  takes arbitrarily large negative values, or  $A$  is Martin-Löf random, in which case  $K(A \upharpoonright n) - n$  tends to  $+\infty$ . This means for example that there is no sequence  $A \in 2^\omega$  such that  $K(A \upharpoonright n) = n + O(1)$ . One may ask whether this dichotomy is due to a gap phenomenon, that is: is there a function  $h$  that tends to infinity, such that for every Martin-Löf random sequence  $A$ ,  $K(A \upharpoonright n) \geq n + h(n) - O(1)$ ? Similarly, is there a function  $h$  that tends to infinity such that for every sequence  $A$ ,  $K(A \upharpoonright n) \geq n - h(n) - O(1)$  implies that  $A$  is Martin-Löf random? We answer both questions, as well as their plain

complexity counterparts, in the negative.

**Theorem 3.1** *There exists no function  $h: \mathbb{N} \rightarrow \mathbb{N}$  that tends to infinity and such that*

$$K(A \upharpoonright n) \geq n - h(n) - O(1)$$

*is a sufficient condition for Martin-Löf randomness of  $A$ .*

*Similarly, there is no function  $h: \mathbb{N} \rightarrow \mathbb{N}$  that tends to infinity and such that*

$$C(A \upharpoonright n) \geq n - K(n) - h(n) - O(1)$$

*is a sufficient condition for Martin-Löf randomness of  $A$ .*

In fact, for both statements we will build counterexamples that are not even Church stochastic (see Downey and Hirschfeldt [13] for the definition of Church stochasticity).

**Proof** First, notice that since we want to prove this for *any* function that tends to infinity, we can restrict our attention to the nondecreasing ones. Indeed, if  $h$  is a function that tends to infinity, the function

$$\tilde{h}(n) = \min\{h(i) \mid i \geq n\}$$

also tends to infinity and  $\tilde{h} \leq h$ .

Now, assume we are in the simple case where the function  $h$  is nondecreasing and computable. A standard technique to get a nonrandom binary sequence  $B$  such that  $K(B \upharpoonright n) \geq n - h(n) - O(1)$  is the following: take a Martin-Löf random sequence  $A$ , and insert zeroes into  $A$  in positions  $h^{-1}(0), h^{-1}(1), h^{-1}(2), \dots$ . It is easy to see that the resulting sequence  $B$  is not Martin-Löf random (indeed, not even Church stochastic), and that the Kolmogorov complexity of its initial segments is as desired. This approach was refined by Merkle, Miller et al. [27] where the authors used an insertion of zeroes on a co-c.e. set of positions in order to construct a left-c.e. sequence  $B$  that is not Mises-Wald-Church stochastic, but has initial segments of very high complexity.

Of course, the problem here is that the function  $h$  in the hypothesis may be noncomputable, and in particular may grow slower than any computable nondecreasing function. In that case, the direct construction we just described does not necessarily work: indeed, inserting zeroes at a noncomputable set of positions may not affect the complexity of  $A$ . To overcome this problem, we invoke the Kučera-Gács theorem, see Kučera [20], Gács [16], or Merkle and Mihailović [26]. This theorem states that any sequence, and thus any function

from  $\mathbb{N}$  to  $\mathbb{N}$ , is Turing-reducible to some Martin-Löf random sequence. Hence, instead of choosing *any* Martin-Löf sequence  $A$ , we pick one that computes the function  $h^{-1}$  and then insert zeroes into  $A$  at positions  $h^{-1}(0), h^{-1}(1), \dots$  (by this we mean that we construct a sequence  $B$  by adding  $h^{-1}(0)$  bits of  $A$ , then a zero, then  $h^{-1}(1)$  bits of  $A$ , etc.). Intuitively, the resulting sequence  $B$  should not be random, as the bits of  $A$  can be used to compute the places where the zeroes have been inserted. This intuition however is not quite correct, as inserting the zeroes may destroy the Turing reduction  $\Phi$  from  $A$  to  $h^{-1}$ . In other words, looking at  $B$ , we may not be able to distinguish the bits of  $A$  from the inserted zeroes.

The trick to solve this last problem is to delay the insertion of the zeroes to “give enough time” to the reduction  $\Phi$  to compute the positions of the inserted zeroes. More precisely, we insert the  $k$ -th zero in position  $n_k = h^{-1}(k) + t(k)$  where  $t(k)$  is the time needed by  $\Phi$  to compute  $h^{-1}(k)$  from  $A$ . This way,  $n_k$  is computable from  $A \upharpoonright n_k$  in time at most  $n_k$ . From this, it is not too hard to construct a computable selection rule that selects precisely the inserted zeroes, witnessing that  $B$  is not Church stochastic, hence is not Martin-Löf random. Moreover, since the “insertion delay” only makes the inserted zeroes more sparse, we have  $K(B \upharpoonright n) \geq n - h(n) - O(1)$ . And similarly, since  $A$  is Martin-Löf random, we have by the Gács-Miller-Yu theorem (Theorem 3.4 below):  $C(B \upharpoonright n) \geq n - K(n) - h(n) - O(1)$ .

The formal details are as follows. Let  $h$  be a nondecreasing function. By the Kučera-Gács theorem, let  $A$  be a Martin-Löf random sequence and  $\Phi$  be a Turing functional such that  $\Phi^A(n) = h^{-1}(n)$  for all  $n$ . Let  $t(n)$  be the computation time of  $\Phi^A(n)$ , where we can assume that  $t$  is an increasing function. Let  $B \in 2^\omega$  be the sequence obtained by inserting zeroes into  $A$  in positions  $h^{-1}(n) + t(n)$ . To show that  $B$  is not Church stochastic, we construct a (total) computable selection rule that filters the inserted zeroes from  $B$ . Let  $S$  be the selection rule that works as follows on a given sequence  $X \in 2^\omega$ . We proceed by induction; we call  $k_n$  the number of bits selected by  $S$  from  $X \upharpoonright n$  and  $x_n$  the prefix  $X \upharpoonright n$  of  $X$  from which these  $k_n$  bits are deleted ( $x_0$  is thus the empty string, and  $k_0 = 0$ ).

At stage  $n+1$ , having already read  $X \upharpoonright n$ ,  $S$  computes  $\Phi_n^{x_n}(k_n)$ . If the computation halts after  $s$  steps,  $S$  checks whether  $\Phi_n^{x_n}(k_n) + s$  returns  $n$  (the subscript of  $\Phi$  refers to the number of computation steps allowed). If so,  $S$  selects the  $n$ -th bit of  $X(n)$  of  $X$  and then sets  $x_{n+1} = x_n$  and  $k_{n+1} = k_n + 1$ . Otherwise,  $S$  just reads the bit  $X(n)$ , extends  $x_n$  by this bit, i.e.,  $x_{n+1} = x_n X(n)$ , and lets  $k_{n+1} = k_n$ .

It is clear that  $S$  is a total computable selection rule. Now suppose that we run it on  $B$ . We argue that  $S$  selects exactly the zeroes that have been inserted into  $A$  to get  $B$ . We prove this by induction. If  $S$  has already selected from  $B$  the first  $i$  inserted zeroes, then the next selected bit is the bit in position  $n = \Phi^{x_n}(k_n) + s$  where  $\Phi^{x_n}(k_n)$  is computed in  $s$  steps. But since the selected bits are exactly the zeroes that were inserted in  $A$ , we have  $k_n = i$  and  $x_n = A \upharpoonright n - i$ , and thus  $s$  is the computation time of  $\Phi^{x_n}(k_n) = \Phi^{A \upharpoonright n - i}(i)$ , which we called  $t(i)$ . And by definition of  $\Phi$ ,  $\Phi^{A \upharpoonright n - i}(i) = h^{-1}(i)$ . Therefore,  $n = h^{-1}(i) + t(i)$ , i.e., the selected bit was an inserted zero. This proves that  $S$  only selects bits that belong to the zeroes that were inserted into  $A$ . Conversely, we need to prove that all such bits are indeed selected by  $S$ . Let  $i \in \mathbb{N}$ . The  $i+1$ -th inserted zero is in position  $n = h^{-1}(i) + t(i)$ . At stage  $n$ , we have by the induction hypothesis  $x_n = A \upharpoonright n - i$  and  $k_n = i$ . Thus,  $\Phi_n^{x_n}(k_n) = \Phi_{h^{-1}(i)+t(i)}^{A \upharpoonright t(i)+h^{-1}(i)-i}(i)$ , which has to halt because both quantities  $t(i) + h^{-1}(i) - i$  and  $h^{-1}(i) + t(i)$  are greater than  $t(i)$ , which is the computation time of  $\Phi^A(i)$ . Thus the bit in position  $n$  is indeed selected. Therefore,  $S$  satisfies the desired properties, and witnesses the fact that  $B$  is not Church stochastic.

Finally, for all  $n$ , calling  $i$  the number of inserted zeroes in  $B \upharpoonright n$ , we easily see that  $B \upharpoonright n$  and  $A \upharpoonright n - i$  can each be computed from the other one, by successive insertion or deletion of zeroes. Thus:  $K(B \upharpoonright n) = K(A \upharpoonright n - i) + O(1) \geq n - i - O(1)$  since  $A$  is Martin-Löf random. And by definition of the positions where the zeroes are inserted, we have  $n \geq h^{-1}(i - 1) + t(i - 1)$ , hence  $i \leq h(n) + O(1)$ . Therefore:

$$K(B \upharpoonright n) \geq n - i - O(1) \geq n - h(n) - O(1)$$

for all  $n$ , which completes the proof of the first assertion. We omit the almost identical argument for the second assertion on C-complexity.  $\square$

Although we do not discuss them in this paper, the above construction can also be applied to two other variants of Kolmogorov complexity, namely monotone complexity  $K_m$  and a-priori complexity  $K_M$  (see [13] for a definition of these complexities). A sequence  $A$  is Martin-Löf random if and only if  $K_M(A \upharpoonright n) = K_m(A \upharpoonright n) + O(1) = n + O(1)$ , but there is no way to get a weaker sufficient condition.

**Proposition 3.2** *There exists no function  $h: \mathbb{N} \rightarrow \mathbb{N}$  that tends to infinity and such that*

$$K_M(A \upharpoonright n) \geq n - h(n) - O(1)$$

*is a sufficient condition for  $A$  to be Martin-Löf random. Since  $K_m \geq K_M$ , this remains true with  $K_m$  in place of  $K_M$ .*

The proof of Proposition 3.2 is virtually identical to the proof of Theorem 3.1. Note that the proposition is in fact stronger than the theorem, as  $K \geq K_M$ .

Another consequence of the construction performed in this proof is the dual version of Theorem 3.1 stated in Proposition 3.3. The first part of the proposition has been obtained earlier on and in different ways by Miller and Yu [30, Corollary 3.2], and in fact with the weaker hypothesis that  $h$  is unbounded.

**Proposition 3.3** *There exists no function  $h: \mathbb{N} \rightarrow \mathbb{N}$  that tends to infinity and such that*

$$K(A \upharpoonright n) \geq n + h(n) - O(1)$$

*is a necessary condition for  $A$  to be Martin-Löf random.*

*Similarly, there is no function  $h: \mathbb{N} \rightarrow \mathbb{N}$  that tends to infinity and such that*

$$C(A \upharpoonright n) \geq n - K(n) + h(n) - O(1)$$

*is a necessary condition for  $A$  to be Martin-Löf random.*

**Proof** Suppose for the sake of contradiction that there exists a function  $h'$  which tends to infinity and such that  $K(A \upharpoonright n) \geq n + h'(n) - O(1)$  is a necessary condition for  $A$  to be Martin-Löf random. Once again, we can assume that  $h'$  is nondecreasing. Then, we perform the exact same construction as in the proof of Theorem 3.1 for a given function  $h$ . At the end of proof, when evaluating the complexity of  $B$ , we have  $K(B \upharpoonright n) = K(A \upharpoonright n - i) + O(1)$ , with  $i \leq h(n) + O(1)$ , and since  $A$  is Martin-Löf random,  $K(A \upharpoonright n - i) \geq (n - i) + h'(n - i) - O(1)$ . It follows that

$$K(B \upharpoonright n) \geq n - h(n) + h'(n - h(n)) - O(1).$$

By assumption on  $h'$ , we have  $h'(n) \leq n/3$  for almost all  $n$ , hence if we let  $h(n) = h'(n/2)$ , we have  $K(B \upharpoonright n) \geq n - O(1)$ . This is a contradiction since by the Levin-Schnorr theorem, this would imply that the sequence  $B$  is Martin-Löf random, which it is not by construction. The proof of the second part of the proposition is almost identical.  $\square$

### 3.2 The Gács-Miller-Yu theorem

We now turn to the link between Solovay functions and the Gács-Miller-Yu theorem. This theorem gives a characterization of Martin-Löf random sequences in terms of the C-complexity of their initial segments, even though the condition still involves  $K$ .

**Theorem 3.4 (Gács-Miller-Yu)**  *$A \in 2^\omega$  is Martin-Löf random if and only if*

$$C(A \upharpoonright n) \geq n - K(n) - O(1).$$

Moreover there exists a computable upper bound  $f$  of  $K$  such that  $A$  is Martin-Löf random if and only if  $C(A\upharpoonright n) \geq n - f(n) - O(1)$ .

Gács [15] actually gives a variant of the first part, with conditional complexity  $C(A\upharpoonright n|n)$  instead of  $C(A\upharpoonright n)$ . Miller and Yu [28] proved the first part as stated above, as well as the second part about the existence of a computable  $f$  with the given properties.

The second part of the theorem indicates the existence of “tight enough” computable upper bounds for  $K$ . These turn out to be *exactly* the Solovay functions.

Using the “no-gap” theorems of the previous section, we first show that any such function must be a weak Solovay function, even if we only assume the function to be merely right-c.e.

**Theorem 3.5** *Let  $f$  be a right-c.e. function such that*

$$C(A\upharpoonright n) \geq n - f(n) - O(1) \Leftrightarrow A \text{ is Martin-Löf random.}$$

*Then  $f$  is a weak Solovay function. In particular,  $f$  is a Solovay function in case  $f$  is computable.*

**Proof** For a start suppose that  $f$  is an upper bound for  $K$  up to an additive constant. Then  $f$  must be a Solovay function because otherwise  $h = f - K$  tends to infinity and by assumption on  $f$ , one has for all sequences  $A$  that

$$C(A\upharpoonright n) \geq n - K(n) - h(n) - O(1)$$

implies that  $A$  is Martin-Löf random, which contradicts the no-gap result stated in Theorem 3.1. In particular, in this case the function  $f$  is a Solovay function in case it is computable. (We note in passing that a similar argument shows for any, not necessarily right-c.e. function  $f$  that satisfies the forward implication of the equivalence in the theorem that in case the function  $f$  is an upper bound for  $K$ , then this upper bound must be infinitely often tight.)

By the preceding discussion, it suffices to show that  $f$  is an upper bound for  $K$  up to an additive constant. For a right-c.e. function, this is equivalent to  $\sum_n 2^{-f(n)} < \infty$ . For the sake of contradiction, suppose that  $\sum_n 2^{-f(n)} = \infty$ . We already know by the Gács-Miller-Yu theorem that every Martin-Löf random real  $A$  satisfies  $C(A\upharpoonright n) \geq n - K(n) - O(1) \geq n - 2 \log(n) - O(1)$ . Thus, after replacing  $f(n)$  by  $\min(2 \log n, f(n))$  (a transformation which preserves the right-c.e.-ness and the property  $\sum_n 2^{-f(n)} = \infty$ ), we can assume that  $f(n) \leq 2 \log n$ .

Let us now build an auxiliary right-c.e. function  $g$  as follows. Since  $\sum_n 2^{-f(n)} = \infty$  and  $f$  is right-c.e., one can effectively find a partition of  $\mathbb{N}$  into consecutive intervals  $I_0, I_1, I_2, \dots$  together with stages  $t_0 < t_1 < t_2 \dots$  such that  $\sum_{n \in I_k} 2^{-f_{t_k}(n)} \geq 2^{2^k}$  for all  $k$ . One can further assume that  $\min I_k \geq 2^k - 1$  (indeed one can always add more elements to  $I_{k-1}$  in order to increase  $\min I_k$  if necessary). Now for all  $k$ , the function  $g$  is defined on  $I_k$  by  $g(n) = f_{t_k}(n) + k$ . This implies that  $g$  is computable, that  $g - f \rightarrow \infty$  and

$$\sum_n 2^{-g(n)} = \sum_k \sum_{n \in I_k} 2^{-f_{t_k}(n) - k} \geq \sum_k 2^k = \infty.$$

Finally, observe that on every interval  $I_k$ , we have  $g(n) = f_{t_k}(n) + k$ , and by our assumption that  $\min I_k \geq 2^k - 1$ , this means that  $g(n) = f_{t_k}(n) + O(\log n) = O(\log n)$  (for the last inequality we use the fact that  $f = O(\log n)$ , and assume without loss of generality that the enumeration of  $f$  from above is  $O(\log n)$  at all stages). This last property ensures that  $C(n|n - g(n)) = O(1)$  because the function  $n \mapsto n - g(n)$  is computable and  $O(1)$ -to-one. We can then use a well-known result due to Martin-Löf (see [13, Theorem 3.11.2]), which states that when  $g$  is a computable function such that  $C(n|n - g(n)) = O(1)$ , then for any  $A \in 2^\omega$ , there are infinitely many  $n$  such that  $C(A \upharpoonright n) \leq n - g(n) - O(1)$ . But when  $A$  is Martin-Löf random, this contradicts the hypothesis that  $C(A \upharpoonright n) \geq n - f(n) - O(1)$  (because  $g - f \rightarrow \infty$ ). This gives us the desired contradiction, and thus proves that  $f$  is an upper bound for  $K$  up to an additive constant.  $\square$

We now show the converse of Theorem 3.5.

**Theorem 3.6** *Let  $g$  be a weak Solovay function. Then  $A \in 2^\omega$  is Martin-Löf random if and only if*

$$C(A \upharpoonright n) \geq n - g(n) - O(1).$$

We begin our proof with a combinatorial lemma.

**Lemma 3.7** *Let  $\sigma$  be a string. Let  $I = [s, t]$  be a finite interval of integers with  $s \geq |\sigma|$ . Let  $(a_i)_{i \in I}$  be a finite set of integers such that*

$$\sum_{i \in I} a_i 2^{-i} \geq 2^{-|\sigma|+1}.$$

*Then, there exists a subset  $J$  of  $I$  and a finite set of strings  $S$  such that*

- (i)  $[S] = [\sigma]$ ,
- (ii) for all  $\tau \in S$ ,  $|\tau| \in J$ ,
- (iii) for all  $j \in J$ ,  $|S \cap \{0, 1\}^{\leq j}| \leq a_j$ .

*Moreover,  $J$  and  $S$  can be constructed effectively given  $\sigma$ ,  $I$  and  $(a_i)_{i \in I}$ .*

**Proof** We construct  $J$  and  $S$  via the following procedure. We initialize  $J$  and  $S$  to  $\emptyset$ . Now the procedure is as follows:

For all  $i$  from  $s$  to  $t$  do

If  $|S| \geq a_i$  do nothing. Otherwise:

- (1) Put  $i$  into  $J$ .
- (2) Split  $[\sigma] \setminus [S]$  into cylinders of measure  $2^{-i}$ . Let  $T$  be the set of strings of length  $i$  generating those cylinders.
- (3) Let  $T'$  be the set containing the  $c_i = a_i - |S|$  first strings of  $T$  in the lexicographic order (if  $c_i > |T|$  then let  $T' = T$ ).
- (4) Enumerate all strings of  $T'$  into  $S$ .

We now verify that this procedure works, i.e., that the algorithm is well-defined and that the set  $S$  we obtain after the  $t$ -loop is as wanted. First, notice that at the beginning of the  $i$ -loop,  $S$  contains only strings of length smaller than  $i$ , therefore  $[S]$  can be split into cylinders of measure  $2^{-i}$ . Since  $|\sigma| \leq s \leq i$ , this is also the case for  $[\sigma]$ , hence for  $[\sigma] \setminus [S]$ , so step (2) is well-defined. We also immediately see that the conditions (ii) and (iii) of the lemma are satisfied: indeed, we only enumerate strings of a given length  $i$  after enumerating  $i$  into  $J$ , and if we do so, we ensure that at the end of the  $i$ -loop, the cardinality of  $S \cap \{0, 1\}^{\leq i}$  is at most  $a_i$ . It remains to verify condition (i). First it is clear that  $S \subseteq [\sigma]$  as we only enumerate cylinders that are contained in  $[\sigma]$ . Suppose that this inclusion is strict. Then, when running the above procedure, at step 3, we are never in the case where  $c_i > |T|$ , hence for all  $i$ , at the end of the  $i$ -loop, we have  $|S \cap \{0, 1\}^{\leq i}| \geq a_i$ , whether  $i$  is in  $J$  or not. Therefore, at the end of the procedure, we have

$$\begin{aligned} \sum_{i=s}^t a_i 2^{-i} &\leq \sum_{i=s}^t |S \cap \{0, 1\}^{\leq i}| 2^{-i} \leq \sum_{i=s}^t \sum_{k=s}^i |S \cap \{0, 1\}^k| 2^{-i} \\ &\leq \sum_{k=s}^t |S \cap \{0, 1\}^k| \sum_{i=k}^t 2^{-i} < \sum_{k=s}^t |S \cap \{0, 1\}^k| 2^{-k+1} \\ &< 2\lambda([S]) < 2\lambda([\sigma]) < 2^{-|\sigma|+1}, \end{aligned}$$

and this contradicts the hypothesis of the lemma.  $\square$

**Proof of Theorem 3.6** Let  $g$  be a weak Solovay function. In the equivalence to be proved, the implication from left to right follows directly from the Gács-Miller-Yu theorem. In order to prove the reverse implication, let  $A$  be a sequence that is not Martin-Löf random. We shall prove that  $C(A \upharpoonright n) \leq n - g(n) - k$  holds for infinitely many  $n$  and arbitrarily large  $k$ .

By Corollary 4.6 below, for every weak Solovay function  $h$  there is a Solovay function  $\tilde{h} \leq h$ . The proof of the corollary does not depend on Theorem 3.6

or any result demonstrated by using this theorem, hence we can apply the lemma already now and can assume that  $g$  is computable. We further assume, for technical reasons which will become clear at the end of the proof, that for all  $i$ , either  $g(i) \leq 2 \log(i)$  or  $g(i) = +\infty$ . If it is not the case, replace  $g$  by the bigger function  $\tilde{g}$  defined by  $\tilde{g}(i) = g(i)$  if  $g(i) \leq 2 \log(i)$ , and  $\tilde{g}(i) = +\infty$  otherwise. Then we have

$$\sum_i 2^{-\tilde{g}(i)} = \sum_i 2^{-g(i)} - \sum_{\substack{i \\ g(i) \geq 2 \log i}} 2^{-g(i)},$$

where the third sum is a computable real number as the  $i$ -th term is bounded by  $1/i^2$ . Thus  $\sum_i 2^{-\tilde{g}(i)}$  is equal to a Martin-Löf random real minus a computable real, hence is a Martin-Löf random real and thus  $\tilde{g}$  is still a Solovay function.

Now, let  $(\mathcal{U}_k)_{k \in \mathbb{N}}$  be a Martin-Löf test covering  $A$  and such that  $\lambda(\mathcal{U}_k) \leq 2^{-2k-1}$  for all  $k$ . We design a procedure  $(P_k)$  which for all  $k$  tries to enumerate a set of strings  $S_k$  such that  $[S_k] = \mathcal{U}_k$ , with additional properties on the length of the strings it contains. We ensure that this procedure succeeds for almost all  $k$  by building an auxiliary test  $\mathcal{V}_k$  which tests the randomness of  $\sum_i 2^{-g(i)}$ . The procedure  $(P_k)$  works as follows.

- (1) Wait for a new cylinder  $[\sigma]$  to be enumerated into  $\mathcal{U}_k$ .
- (2) Choose a large integer  $s$ , say larger than  $2^N$  with  $N$  larger than any integer mentioned so far in the construction (including  $k$ ).
- (3) Enumerate into  $\mathcal{V}_k$  the dyadic real interval

$$\left[ \sum_{i < s} 2^{-g(i)}, 2^{-|\sigma|+1+k} + \sum_{i < s} 2^{-g(i)} \right].$$

- (4) Wait for a stage  $t$  such that  $\sum_{i \leq t} 2^{-g(i)} > 2^{-|\sigma|+1+k} + \sum_{i < s} 2^{-g(i)}$ .
- (5) When this happens, we have  $\sum_{i=s}^t 2^{-g(i)} > 2^{-|\sigma|+1+k}$ . We then apply Lemma 3.7 with  $a_i = 2^{i-g(i)-k}$  to get a finite set of strings  $S_k^\sigma$  and a finite set of integers  $J_k^\sigma$  such that  $[S_k^\sigma] = [\sigma]$ , for all  $\tau \in S_k^\sigma$ ,  $|\tau| \in J_k^\sigma$  and for all  $j \in J_k^\sigma$ ,  $|S^\sigma \cap \{0, 1\}^{\leq j}| \leq a_j$ . We then put all strings of  $S_k^\sigma$  into  $S_k$  and go back to step 1.

It is possible that for some  $k$ ,  $(P_k)$  will at some point reach step 4 and wait there forever. We claim that this can only happen for finitely many  $k$ . Indeed, for a given  $k$ , we have  $\lambda(\mathcal{V}_k) \leq 2^{-k}$ , because whenever a cylinder  $[\sigma]$  enters  $\mathcal{U}_k$  at step 1, an interval of length  $2^{-|\sigma|+1+k}$  enters  $\mathcal{V}_k$ , hence  $\lambda(\mathcal{V}_k) \leq 2^{k+1} \lambda(\mathcal{U}_k) \leq 2^{-k}$ . Thus,  $(\mathcal{V}_k)_{k \in \mathbb{N}}$  is a Martin-Löf test. Furthermore, if the procedure for  $S_k$  waits forever at some step 4, this precisely means that  $\sum_i 2^{-g(i)}$  belongs to the dyadic interval which was put into  $\mathcal{V}_k$  at step 3, and thus in that case  $\sum_i 2^{-g(i)} \in \mathcal{V}_k$ . Since  $\sum_i 2^{-g(i)}$  is random, it can only belong to finitely many  $\mathcal{V}_k$ , hence for

almost all  $k$  the procedure  $(P_k)$  never waits forever at step 4. In that case, the c.e. set  $S_k$  it builds does satisfy  $[S_k] = \mathcal{U}_k$  by construction.

To finish the proof, let  $k$  be such that  $(P_k)$  succeeds. Since  $A$  is not Martin-Löf random,  $A$  belongs to  $\mathcal{U}_k$ , hence to  $[S_k]$ . This means that for some  $n$ ,  $A \upharpoonright n$  belongs to  $S_k$ . To describe  $A \upharpoonright n$ , it suffices to describe  $k$  (this can be done with  $2 \log k + O(1)$  bits), and its position inside  $S_k$ . For its position inside  $S_k$ , we simply describe the position of  $A \upharpoonright n$  inside the  $S_k^\sigma$  it belongs to, when the latter is sorted in the length-lexicographic order. By construction of  $S_k^\sigma$ ,  $n$  must be in  $J_k^\sigma$  (otherwise  $S_k^\sigma$  would be empty), and there are at most  $a_n = 2^{n-g(n)-k}$  strings of length less than or equal to  $n$  in  $S_k^\sigma$ , and therefore we can specify the position of  $A \upharpoonright n$  inside  $S_k^\sigma$  with  $n - g(n) - k$  bits. Thus, our description of  $A \upharpoonright n$  has total length  $n - g(n) - k + 2 \log k + O(1)$ . Since  $k$  can be taken as large as wanted, this will be enough to prove the theorem, but one last thing we need to check is that this description is enough to retrieve  $A \upharpoonright n$ . Indeed, while we give the index of  $A \upharpoonright n$  inside the  $S_k^\sigma$  it belongs to, we do not describe  $\sigma$  explicitly. However,  $\sigma$  can be found as follows. The description of  $A \upharpoonright n$  we give has length  $n - g(n) - k + 2 \log k + O(1)$ . By assumption,  $g(n) \leq 2 \log n$  and by construction of  $S_k^\sigma$ ,  $k \leq \log s \leq \log n$ . Hence our description has length between  $n - 3 \log n + O(1)$  and  $n + O(1)$ . Hence the length of our description gives us  $n$  with logarithmic precision. This is enough to find the string  $\sigma$  such that  $A \upharpoonright n$  belongs  $S_k^\sigma$  because by construction of  $S_k$ , if  $l$  is the length of some string in  $S_k^{\sigma'}$  with  $\sigma' \neq \sigma$ , then either  $2^l < n$  or  $2^n < l$ , and hence either  $l < n - 3 \log n$  or  $n < l - 3 \log l$ .  $\square$

## 4 Connections to K-triviality

### 4.1 $K$ -trivial sequences

From their incompressibility characterization, it can be seen that the Martin-Löf random sequences are those which have initial segments of roughly maximal Kolmogorov complexity. It is natural to ask which sequences  $A$  have initial segments of *minimal* prefix free Kolmogorov complexity  $K(A \upharpoonright n) \leq K(n) + O(1)$ . Chaitin [8] proved that any such sequence is computable from the halting problem, and Solovay [34] was able to construct such a sequence that is noncomputable and computably enumerable. The class of such sequences was further studied by Downey, Hirschfeldt, Nies and Stephan [14,31], who called them *K-trivial*.

The  $K$ -trivial sequences turned out to have remarkable properties. Perhaps the most striking fact is that they can be characterized as the sequences that are low for Martin-Löf randomness, or, alternatively, as the sequences that

are low for prefix-free Kolmogorov complexity. In other words, a sequence  $A$  is  $K$ -trivial if and only if Martin-Löf randomness relativized to  $A$  coincides with Martin-Löf randomness, if and only if the prefix-free Kolmogorov complexity relativized to  $A$  is within an additive constant of the unrelativized one.

In this section we will show that in the definition of the notion of  $K$ -trivial, the upper bound  $K(n) + O(1)$  can be equivalently replaced by any weak Solovay function, and that in fact the ability to do so characterizes the Solovay functions and the weak Solovay functions. Using this characterization, we give an easy, golden-run-free proof for the fact that every  $K$ -trivial is Turing below a c.e.  $K$ -trivial. Some of the results of this section were announced in an earlier conference paper [6].

#### 4.2 Solovay functions characterize $K$ -triviality

In what follows, we show that weak Solovay functions can be used in place of prefix-free Kolmogorov complexity to characterize  $K$ -triviality. This means that  $K$ -triviality is equivalent to  $g$ -triviality in the sense of the following definition, for any weak Solovay function  $g$ .

**Definition 4.1** *Given a function  $g: \mathbb{N} \rightarrow \mathbb{N}$  and an integer  $c$ , a sequence  $A$  is  $g$ -trivial with constant  $c$  if  $K(A \upharpoonright n) \leq g(n) + c$  holds for all  $n$ . A sequence is  $g$ -trivial if it is  $g$ -trivial for some  $c$ .*

The notion of  $K$ -triviality in the sense of Definition 4.1 coincides with the usual notion of  $K$ -triviality. However, in the usual concept of  $K$ -triviality the reference to  $K$  is with respect to the upper bound  $K(n)$  but surely also to the fact that we bound the  $K$ -complexity of the initial segments of the sequence under consideration. This problem could be resolved by taking  $g$ -trivial as an abbreviation for  $K$ - $g$ -trivial, where a sequence  $A$  is  $f$ - $g$ -trivial if  $f(A \upharpoonright n) \leq g(n) + c$  holds for some constant  $c$  and all  $n$ .

We start by proving the equivalence of  $K$ -triviality and  $g_S$ -triviality, where  $g_S$  is Solovay's original Solovay function as constructed in the proof of Theorem 2.1. We will in Theorem 4.4 below see how to extend this equivalence to any weak Solovay function.

**Theorem 4.2** *Let  $g_S$  be the Solovay function constructed by Solovay. Then a sequence is  $K$ -trivial if and only if it is  $g_S$ -trivial.*

**Proof** One direction is easy: if  $A$  is  $K$ -trivial, then  $K(A \upharpoonright n) \leq K(n) + O(1)$ , and by definition a Solovay function is an upper bound of  $K$  up to an additive constant, hence  $A$  is  $g_S$ -trivial.

For the other direction, let  $A$  be  $g_S$ -trivial for some constant  $c$ . Fix  $n$ . Let  $p$  be a shortest prefix description for  $n$  and let  $t$  be the running time of  $p$  on  $\mathbb{U}$ , i.e.,  $|p| = K(n)$  and  $\mathbb{U}(p) = n$  in exactly  $t$  steps. Let  $m = \langle n, p, t \rangle$ . By definition of  $g_S$ , we have  $g_S(m) = |p|$ , hence it holds that

$$K(A \upharpoonright m) \leq g_S(m) + c = |p| + c = K(n) + c.$$

The result then follows by observing that  $n$  can be retrieved from  $m$ , and thus  $K(A \upharpoonright n) \leq K(A \upharpoonright m) + O(1) \leq K(n) + c + O(1)$ .  $\square$

The proof of Theorem 4.2 actually shows a bit more than asserted in the theorem.

**Remark 4.3** The equivalence of  $K$ -triviality and  $g_S$ -triviality stated in Theorem 4.2 holds in the strong form that triviality constants are preserved up to an additive constant. More precisely, there is a constant  $c_S$  such that if a sequence is  $K$ -trivial with constant  $c$ , then it is  $g_S$ -trivial with constant  $c + c_S$ , and a similar remark holds for the reverse implication.

**Theorem 4.4** *Let  $g$  be a weak Solovay function. Then any sequence is  $K$ -trivial if and only if it is  $g$ -trivial.*

Again, the implication from  $K$ -trivial to  $g$ -trivial is immediate. The difficulty resides in the converse. The core of the corresponding proof is the following technical lemma, which guarantees that building a bounded request set to ensure that a sequence is  $g$ -trivial does not “cost more” (in a specific sense to be explained below) than building a bounded request set to ensure that it is  $h$ -trivial for a weak Solovay function  $h$ . A first consequence of the lemma will be Corollary 4.6 below, which asserts that for every weak Solovay function  $h$  there is a Solovay function  $\tilde{h} \geq h$ . The lemma and the corollary will then be applied in order to demonstrate Theorem 4.4.

**Lemma 4.5** *Let  $g$  be a Solovay function, and  $h$  a weak Solovay function. There exists a positive constant  $c$  and a computable partition of  $\mathbb{N}$  into subsequent nonempty intervals  $(I_n)_{n \in \mathbb{N}}$  such that for all  $n$  we have  $n < \min I_n$  and*

$$2^{-g(n)} \leq 2^c \sum_{i \in I_n} 2^{-h(i)}.$$

**Proof** We will actually use the weaker hypothesis that  $g$  is computable and  $\sum_n 2^{-g(n)}$  is finite. We design a procedure which uniformly in  $p$  tries to construct a partition  $(I_n^p)_{n \in \mathbb{N}}$  such that  $2^{-g(n)} \leq 2^p \sum_{i \in I_n^p} 2^{-h(i)}$ . The procedure works as follows:

For  $n$  from 0 to  $\infty$  do

- (1) Let  $s(p, n) \in \mathbb{N}$  be the least integer  $> n$  which does not belong to one of the previously constructed intervals  $I_j^p$  for  $j < n$ .
- (2) Search for some  $t > s(p, n)$  large enough to have

$$\sum_{i=s(p,n)}^t 2^{-h_t(i)} \geq 2^{-p}2^{-g(n)}.$$

- (3) When  $t$  is found, define  $I_n^p$  to be  $[s(p, n), t]$ .

It is possible that the procedure of parameter  $p$ , for some  $n$ , never finds  $t$  at Step 2. When this happens, we have by construction:

$$\sum_{i \geq s(p,n)} 2^{-h(i)} \leq 2^{-p}2^{-g(n)}.$$

Hence by the Kraft-Chaitin theorem, for all  $i \geq s(p, n)$ :

$$K(i) \leq h(i) - p + K(p, n, s(p, n)) - g(n) + O(1). \quad (2)$$

Using the construction,  $s(p, n)$  can be described via the pair  $(p, n)$  alone, hence

$$K(p, n, s(p, n)) \leq K(p, n) + O(1) \leq K(n) + 2 \log p + O(1). \quad (3)$$

Since  $\sum_n 2^{-g(n)}$  is finite we have  $K(n) \leq g(n) + O(1)$ . Then (2) and (3) yield for all  $i \geq s(p, n)$ :

$$K(i) \leq h(i) - p + 2 \log p + O(1).$$

Now, recall that  $h$  is a weak Solovay function so  $K(i) \geq h(i) + O(1)$  for infinitely many  $i$ . Therefore the above situation can only happen for a finite number of  $p$ . In other words, for all  $p$  large enough, the procedure never waits forever at step 2 and hence produces effectively a partition  $(I_n^p)_{n \in \mathbb{N}}$  of  $\mathbb{N}$  into intervals such that for all  $n$ , and each  $I_n^p = [s, t]$  we obtain as wanted

$$2^{-p}2^{-g(n)} \leq \sum_{i=s}^t 2^{-h_t(i)} \leq \sum_{i=s}^t 2^{-h(i)}. \quad \square$$

Our first application of Lemma 4.5 is the following.

**Corollary 4.6** *Let  $h$  be a weak Solovay function. There exists a Solovay function  $\tilde{h}$  such that  $h \leq \tilde{h}$ .*

**Proof** Let  $g$  be any Solovay function (for example, the one constructed in the proof of Theorem 2.1). By Lemma 4.5, there exists a constant  $c$  and a computable partition  $(I_n)_{n \in \mathbb{N}}$  of  $\mathbb{N}$  into intervals such that for all  $n$

$$2^{-g(n)} \leq 2^c \sum_{i \in I_n} 2^{-h(i)}.$$

Let  $\tilde{h}: \mathbb{N} \rightarrow \mathbb{N}$  be the function defined as follows. For a given  $i$ , let  $I_n$  be the interval to which  $i$  belongs, and set

$$\tilde{h}(i) = h_t(i) \text{ where } t \text{ is the least integer s.t. } 2^{-g(n)} \leq 2^c \sum_{i \in I_n} 2^{-h_t(i)}.$$

It is clear that  $\tilde{h}$  is computable and  $h \leq \tilde{h}$ . Next,

$$\sum_i 2^{-\tilde{h}(i)} = \sum_n \sum_{i \in I_n} 2^{-\tilde{h}(i)}$$

is random. Indeed, by construction for all  $n$ ,  $2^{-g(n)} = O(\sum_{i \in I_n} 2^{-\tilde{h}(i)})$ . Hence  $\sum_n 2^{-g(n)}$  is Solovay reducible to  $\sum_i 2^{-\tilde{h}(i)}$  (see [13, p. 405] for the definition of Solovay reducibility). Since the former is random, the latter must be random as well by the Kučera-Slaman theorem [21]. Therefore  $\tilde{h}$  is a Solovay function.  $\square$

**Proof of Theorem 4.4** Recall that  $\mathbb{U}$  is the universal prefix-free machine defining  $K$ ; a  $\mathbb{U}$  “description” of  $\tau$  is a string  $p$  such that  $\mathbb{U}(p) = \tau$ .

Let  $h$  be a weak Solovay function,  $d$  a constant and  $A$  a sequence such that  $K(A \upharpoonright n) \leq h(n) + d$  for all  $n$ . We want to prove that  $A$  is  $K$ -trivial. By Corollary 4.6, we may assume that  $h$  is computable. We apply Lemma 4.5 to obtain a constant  $c$  and a computable partition of  $\mathbb{N}$  into intervals  $(I_n)_{n \in \mathbb{N}}$  such that  $n < \min I_n$  and  $2^{-g_S(n)} \leq 2^c \sum_{i \in I_n} 2^{-h(i)}$  for all  $n$ .

We show that  $A$  is  $K$ -trivial by building a bounded request set. For all  $n$  and all strings  $\sigma$  of length  $n$ , we wait until we find an extension  $\tau$  of  $\sigma$  whose length is  $\max I_n$  and such that for all  $i \in I_n$ , some description of  $\tau \upharpoonright i$  of length at most  $h(i) + d$  is in the domain of  $\mathbb{U}$ . Since  $h$  is computable we can recognize when this happens. In this case, we enumerate a pair  $(\sigma, g_S(n) + c + d)$  into our request set. The cost of this for us is  $2^{-g_S(n) - c - d}$ , which we can account against the cost for  $\mathbb{U}$  to enumerate descriptions of  $\tau \upharpoonright i$  as above. That cost is at least  $\sum_{i \in I_n} 2^{-h(i) - d}$ , which in turn is at least  $2^{-g_S(n) - c - d}$  by construction of the intervals  $I_n$ . Hence, we never spend more than  $\mathbb{U}$  does, which ensures that our request set is bounded. Now, by assumption on  $A$ , for every  $n$ , for every  $i \in I_n$ , the universal machine must issue a description of  $A \upharpoonright i$  of length at most  $h(i) + d$ , hence some pair  $(A \upharpoonright n, g_S(n) + c + d)$  enters our bounded request set at some point. Therefore, for all  $n$ ,  $K(A \upharpoonright n) \leq g_S(n) + c + d + O(1)$ . By Theorem 4.2, we can conclude that  $A$  is  $K$ -trivial.  $\square$

**Remark 4.7** As in the corresponding Theorem 4.2, the equivalence of  $K$ -triviality and  $g$ -triviality stated in Theorem 4.4 holds in the strong form that triviality constants are preserved up to an additive constant.

More precisely, every weak Solovay function  $g$  is an upper bound for  $K$  up to some additive constant  $c_g$ , hence any sequence that is  $K$ -trivial with constant  $c$  is  $g$ -trivial with constant  $c + c_g$ . Conversely, in the proof of Theorem 4.4 it is shown that for every weak Solovay function  $h$  there is a constant  $c$  such that  $K(A \upharpoonright n) \leq h(n) + d$  implies  $K(A \upharpoonright n) \leq g_S(n) + c + d + O(1)$ , and applying Remark 4.3, we get that  $A$  is  $K$ -trivial via constant  $d + c_h$  for  $c_h = c + c_S + O(1)$ .

Another interesting corollary can be derived from the proof of Theorem 4.4

**Remark 4.8** It is known that given a computable strictly increasing function  $l: \mathbb{N} \rightarrow \mathbb{N}$ , if a sequence  $X$  satisfies  $K(A \upharpoonright l(n)) \leq K(n) + O(1)$ , then  $X$  is  $K$ -trivial (see [13, Proposition 11.1.4] or [32, Exercise 5.2.9 and Solution]). This fact can be extended to weak Solovay functions: For such a function  $l$ , if  $K(A \upharpoonright l(n)) \leq g(n) + O(1)$  for some weak Solovay function  $g$ , then  $A$  is  $K$ -trivial. The proof works in the exact same way: first prove it for the particular case of the function  $g$  defined on the range of  $l$  by  $g(l(\langle n, p, t \rangle)) = |p|$  if  $\mathbb{U}(p)$  outputs  $l(n)$  in exactly  $t$  steps of computation, and outputs  $2|\langle n, p, t \rangle|$  otherwise (the values outside of the range of  $l$  do not matter). The rest of the argument for the function  $g$  is the same as in Theorem 4.2). Then, extend it to all weak Solovay functions, with the same proof as Theorem 4.4, only restricted to the  $n$ 's that are in the range of  $l$ . Details are left to the reader.

### 4.3 $K$ -triviality characterizes Solovay functions

Next we prove that any right-c.e. function  $g$  that makes the equivalence

$$A \text{ is } K\text{-trivial} \quad \text{if and only if} \quad K(A \upharpoonright n) \leq g(n) + O(1) \quad (4)$$

true is a weak Solovay function, and hence a Solovay function in case  $g$  is computable. In the proof of our result, we need only to consider the case where  $g$  is an upper bound for  $K$  up to an additive constant because otherwise the class of sequences  $A$  that satisfy the right-hand side of equivalence (4) is empty. We then prove the stronger fact that in the case  $g$  is such a right-c.e. upper bound for  $K$  but is not a weak Solovay function, there are uncountably many sequences  $A$  such that  $K(A \upharpoonright n) \leq g(n) + O(1)$ . This is enough for our purposes, since there are only countably many  $K$ -trivial sequences (indeed, as we mentioned earlier, they are all computable in the halting problem).

**Theorem 4.9** *Suppose  $g$  is a right-c.e. function such that  $K(n) \leq g(n) + O(1)$  but  $g$  is not a weak Solovay function. Then the set  $\{A \mid K(A \upharpoonright n) \leq g(n) + O(1)\}$  is uncountable.*

**Proof** We will build an increasing sequence  $a_1 < a_2 < a_3 < \dots$  of integers such that any subset  $A$  of  $\{a_1, a_2, a_3, \dots\}$  satisfies  $K(A \upharpoonright n) \leq g(n) + O(1)$ .

The sequence is defined by induction (but not effectively), where we set  $a_1 = 0$  and where we ensure by induction that for all  $k$ , for any subset  $B$  of the finite set  $\{a_1, \dots, a_k\}$  and for all  $n \geq a_k$ , for some constant  $d$  that depends neither on  $B$  nor on  $k$  we have that

$$K(B \upharpoonright n) \leq g(n) + d. \quad (5)$$

This suffices to prove the desired result: let  $A$  be any subset of  $\{a_1, a_2, a_3, \dots\}$ , and let  $n$  be some natural number. Let  $k$  be such that  $a_k \leq n < a_{k+1}$  and let  $B = A \cap \{a_1, \dots, a_k\}$ . Then  $B \upharpoonright n = A \upharpoonright n$ , hence  $K(A \upharpoonright n) \leq g(n) + d$  by (5).

We now explain the inductive definition of the sequence  $a_k$ . Suppose we have already defined  $a_1, \dots, a_k$  with the property (5). Let us choose  $c$  to be a very large integer, say  $c > 2a_k + k + 1$ . Consider the sum  $\Omega_g = \sum_n 2^{-g(n)}$ . By Theorem 2.2, this is not a random real as  $g$  is not a weak Solovay function. Hence, there exists a prefix  $\sigma$  of  $\Omega_g$  such that  $K(\sigma) \leq |\sigma| - c$ . Let  $p$  be a shortest description for  $\sigma$ . Knowing  $p$ , one can effectively perform the following operations: first, retrieve  $\sigma = \mathbb{U}(p)$ ; then, enumerate  $\Omega_g$  from below and wait until it becomes larger than the real value  $0.\sigma$  (treated as a real number written in binary) using the approximation of the values  $g(n)$  from above; when this happens, let  $a_{k+1}$  be the least number  $m$  such that for all  $i \geq m$ , so far there has been no contribution to  $\Omega_g$  by the value  $g(i)$  (more precisely, via the approximation of these values from above). Since  $\sigma$  is a prefix of  $\Omega_g$ , this means in particular that  $\sum_{n \geq a_{k+1}} 2^{-g(n)}$  does not exceed  $2^{-|\sigma|}$ , so by the Kraft-Chaitin theorem, any integer  $n \geq a_{k+1}$  can be described by  $p$  and some prefix-free code of length  $g(n) - |\sigma|$ . Therefore, if  $n \geq a_{k+1}$  and  $B$  is a subset of  $\{a_1, \dots, a_{k+1}\}$ , then  $B \upharpoonright n$  can be described in a prefix-free way by

- $B \upharpoonright a_k$ ,
- $p$  (from which  $a_{k+1}$  can be retrieved),
- the single bit  $B(a_{k+1})$ ,
- some additional  $g(n) - |\sigma|$  bits.

Thus  $K(B \upharpoonright n) \leq 2a_k + |p| + 1 + g(n) - |\sigma| + O(1) \leq g(n) + O(1)$ , using the fact that  $c > 2a_k + 1$  and  $|p| \leq |\sigma| - c$ . This concludes the inductive step.  $\square$

A corollary of Theorem 4.9 is that there is no  $\Delta_2^0$  “gap” for K-triviality. In the proof of this corollary we use a folklore fact about approximable orders that is stated in the following lemma.

**Lemma 4.10** *For every  $\Delta_2^0$  order  $h$  there is a right-c.e. order  $g$  where  $g \leq h$ .*

**Proof** Given an order  $h$  and a natural number  $k$ , let the  $k$ -block of  $h$  be the finite and possibly empty set of all  $n$  such that  $h(n) = k$ . Observe that we have  $h_0 \leq h_1$  for two given orders  $h_0$  and  $h_1$  in case for all natural numbers  $k$  the  $k$ -block of  $h_0$  is at least as large as the  $k$ -block of  $h_1$ . Similarly, by increasing the size of any block of any order, the order is transformed into a strictly smaller order.

Given a  $\Delta_2^0$  order  $h$ , write  $h$  as the pointwise limit of a uniformly computable sequence of functions  $(h_s^0)$  and let

$$h_s(n) = \max_{i \in \{0, \dots, n\}} h_{n+s}^0(i).$$

The sequence  $(h_s)$  is uniformly computable and converges pointwise to  $h$  because for each  $n$  and for almost all  $s$ , each of the values  $h_{n+s}^0(i)$  where  $i \leq n$  agrees with  $h(i)$ , and since  $h$  is an order, their maximum  $h_s(n)$  agrees with  $h(n)$ . Furthermore, since the  $h_s^0$  converge to the order  $h$ , for every natural number  $k$  there are  $n_0$  and  $s_0$  such that  $k < h_s^0(n_0)$  for all  $s \geq s_0$ , hence  $k < h_s(n)$  holds for all  $n \geq \max\{n_0, s_0\}$  and all  $s$ .

Let  $z_{\langle s, n \rangle} = h_s(n)$  where  $\langle \cdot, \cdot \rangle$  is the usual effective and effectively invertible pairing function. Then the sequence  $z_0, z_1, \dots$  is computable and each natural number  $k$  occurs in the sequence only finitely often but at least the size of the  $k$ -block of  $h$  many times. Let  $g_0: n \mapsto n$  and for each  $i \geq 0$  let  $g_{i+1}$  be the order obtained from the order  $g_i$  by extending the size of the  $z_i$ -block of  $g_i$  by 1. Then the  $g_i$  form a uniformly computable sequence. By the preceding discussion, for all  $i$  the order  $g_{i+1}$  is strictly smaller than the order  $g_i$  and, in particular,  $g_{i+1}(n) \leq g_i(n)$  for all  $n$ . Moreover, the  $g_i$  converge to a right-c.e. order  $g$ . Furthermore, we have  $g \leq h$  because by construction for all  $k$  the  $k$ -block of  $g$  is at least as large as the  $k$ -block of  $h$ .  $\square$

**Corollary 4.11** *There is no  $\Delta_2^0$  order  $h$  such that for all sequences  $A$ , the sequence  $A$  is  $K$ -trivial if and only if  $K(A \upharpoonright n) \leq K(n) + h(n) + O(1)$ .*

**Proof** If  $h$  were right-c.e. this would follow directly from Theorem 4.9 because in this case  $K + h$  would be a right-c.e. function that is an upper bound for  $K$  but is not a weak Solovay function since the order  $h$  tends to infinity. In case  $h$  is merely  $\Delta_2^0$ , fix a right-c.e. order  $g \leq h$  according to Lemma 4.10. By the previous discussion there is a sequence  $A$  which is not  $K$ -trivial and satisfies  $K(A \upharpoonright n) \leq K(n) + g(n) + O(1)$  and a fortiori  $K(A \upharpoonright n) \leq K(n) + h(n) + O(1)$ .  $\square$

Corollary 4.11 was independently proven by Barmpalias and Vlek [3]. Furthermore, Baartse and Barmpalias [1] showed that there *does exist* a  $\Delta_3^0$  gap.

#### 4.4 Covering K-trivials by c.e. K-trivials

An important property of the class of K-trivial sequences is that its c.e. members form a Turing cover of the whole class. That is, for every K-trivial sequence  $A$ , there is a c.e. K-trivial sequence  $B$  such that  $A \leq_T B$ . The original proof [31] is difficult and uses advanced techniques such as “cost functions”, but yields the stronger result  $A \leq_{tt} B$  (also see [32, Corollary 5.5.3]). Using Solovay functions, we obtain an elementary proof for the case of Turing reducibility, where the core of the argument relies on the following proposition.

**Proposition 4.12** *Let  $A$  be a K-trivial sequence. Then  $A$  is a path of some K-trivial c.e. tree  $T$  which only has finitely many paths.*

**Proof** By Corollary 2.4, let  $g$  be a computable order which is also a Solovay function. For each  $k$ , let  $n_k = g^{-1}(k)$ . Note that the sequence  $(n_k)$  is computable and nondecreasing (but not necessarily increasing). Let  $c$  be a constant such that  $K(A \upharpoonright n) \leq g(n) + c$  for all  $n$ , and  $g(n) \leq K(n) + c$  for infinitely many  $n$ . Consider the set of strings

$$S = \{\tau \mid (\exists k) |\tau| = n_k \wedge (\forall \sigma \preceq \tau) K(\sigma) \leq g(|\sigma|) + c\}.$$

The set  $S$  is c.e. and contains all the initial segments of  $A$  of type  $A \upharpoonright n_k$  for some  $k$ . Let  $T$  be the closure under prefixes of  $S$ ; this makes  $T$  a c.e. tree such that  $S$  is the restriction of  $T$  to levels of type  $n_k$  for some  $k$ .

We claim that  $T$  is as wanted. First of all,  $A$  is a path of  $T$  by construction.  $T$  has only finitely many paths because  $g$  is a Solovay function: any path  $B$  of  $T$  satisfies  $K(B \upharpoonright n) \leq g(n) + c$  for all  $n$ . Hence, by Remark 4.7,  $B$  is K-trivial via a constant  $c + O(1)$ . There are at most  $2^{c+O(1)}$  such sequences [35].

It remains to show that  $T$  is K-trivial. By Remark 4.8, we only need to prove that  $K(T_k) \leq g(n_k) + O(1)$  for all  $k$ , where  $T_k$  is the restriction of  $T$  to strings of length at most  $n_k$ . Fix a  $k$ .  $S$  being c.e. let  $\tau$  be the last string of length at most  $n_k$  enumerated into  $S$ . By definition of  $S$ , we have  $K(\tau) \leq g(|\tau|) + c$ . Let  $p$  be a description for  $\tau$  of length at most  $g(|\tau|) + c$  which in turn is at most  $g(n_k)$  since  $g$  is nondecreasing. Up to padding  $p$  with meaningless bits, we can assume that  $p$  has length  $g(n_k) + c + O(1)$ . Now, given  $p$ , one can retrieve  $n_k$ ,  $\tau$ , the enumeration stage  $s$  of  $\tau$  in  $S$ , all strings of  $S$  of length at most  $n_k$  (enumerating  $S$  during  $s$  steps) and finally, closing under prefixes, all strings  $T$  of length at most  $n_k$ . Thus

$$K(T_k) \leq |p| + O(1) \leq g(n_k) + O(1),$$

which is what we wanted to prove.  $\square$

**Corollary 4.13** *Every K-trivial sequence is computable in some c.e. K-trivial sequence.*

**Proof** Let  $A$  be a K-trivial. By the previous proposition, let  $A$  be a path of some K-trivial c.e. tree  $T$  with only finitely many paths. Since a tree with finitely many paths computes all its paths,  $A$  is computable from  $T$ .  $\square$

## 5 The $c$ -hitting set of a Solovay function

**Definition 5.1** *Let  $f: \mathbb{N} \rightarrow \mathbb{N}$  be a Solovay function and let  $c$  be an integer. The  $c$ -hitting set of  $f$  is the set*

$$\mathcal{H}(f, c) = \{n \mid f(n) \leq K(n) + c\}.$$

Note that sets of the form  $\mathcal{H}(f, c)$  might be empty or finite but for a fixed Solovay function  $f$ , the set  $\mathcal{H}(f, c)$  is infinite for all sufficiently large  $c$ .

**Proposition 5.2** *Let  $f$  be a Solovay function and  $c$  be an integer such that  $\mathcal{H}(f, c)$  is infinite. Then the set  $\mathcal{H}(f, c)$  is hyperimmune and Turing-complete.*

**Proof** Suppose  $\mathcal{H}(f, c) = \{a_0 < a_1 < a_2 < \dots\}$  is not hyperimmune, i.e., there is a computable function  $F$  such that  $a_n < F(n)$  for all  $n$ . Under this assumption, we shall get a contradiction by proving that all K-trivial sequences are computable (which is not the case!). The argument is the same as Chaitin's [9] proof that a sequence  $A$  is computable if and only if  $C(A \upharpoonright n) \leq \log n + O(1)$ .

Let  $G$  be the function defined inductively by  $G(0) = 0$  and  $G(n+1) = F(G(n))$ . Consider the computable partition of  $\mathbb{N}$  made of the intervals  $I_n = [G(n), G(n+1) - 1]$ . An easy induction shows that  $a_{G(n)} \in I_n$  for all  $n$ , hence  $\mathcal{H}(f, c) \cap I_n \neq \emptyset$  for all  $n$ .

Let  $A$  be a noncomputable K-trivial set and, by Theorem 4.4, let  $d$  be a constant such that  $K(A \upharpoonright n) \leq f(n) + d$ . Consider the c.e. tree

$$T = \{\tau \mid (\forall \sigma \preceq \tau) K(\sigma) \leq f(|\sigma|) + d\}.$$

As in the proof of Proposition 4.12,  $T$  has finitely many paths, among which is  $A$ . For each  $n$ , let

$$m_n = \min_{i \in I_n} |T \cap \{0, 1\}^i|.$$

We first note that the sequence  $(m_n)$  is left-c.e. (as  $T$  is c.e.). Furthermore, it is bounded: indeed, by construction of the  $I_n$ , there is for each  $n$  some  $i \in I_n \cap \mathcal{H}(f, c)$ , meaning that  $f(i) \leq K(i) + c$ . Thus any string  $\tau \in T \cap \{0, 1\}^i$  is such that  $K(\tau) \leq f(|\tau|) + d \leq K(|\tau|) + c + d$ . By the coding theorem the number of strings of the latter type of any given length is bounded from above by some constant that depends only on  $c$  and  $d$ . Let then  $l = \limsup_n m_n$  and  $N$  such that  $m_n \leq l$  for all  $n \geq N$ . Since the sequence  $(m_n)$  is left-c.e. one can computably find two sequences  $N < n_0 < n_1 < n_2 < \dots$  and  $s_0 < s_1 < s_2 < \dots$  such that for all  $k$

$$\min_{i \in I_{n_k}} |T_{s_k} \cap \{0, 1\}^i| = l$$

(where  $T_s$  is the c.e. approximation of  $T$  at stage  $s$ , where we assume that  $T_s$  is a tree for all  $s$ ). Consider the tree

$$T^* = \{\tau \mid (\forall i < |\tau|)(\forall k) i \in I_{n_k} \Rightarrow \tau \upharpoonright i \in T_{s_k}\}.$$

The tree  $T^*$  is computable, and all its paths are paths of  $T$ . In fact  $T^*$  has exactly the same paths as  $T$ , because  $T^*$  coincides with  $T$  on infinitely many levels (namely on each  $a_k \in I_{n_k}$  such that  $|T_{s_k} \cap \{0, 1\}^i| = l$ ). Thus  $A$  is a path of  $T^*$ , a computable tree with finitely many paths, and therefore  $A$  is computable, a contradiction.

It remain to show that  $\mathcal{H}(f, c)$  is Turing complete. Observe that for any Solovay function and constant  $c$  the set  $\mathcal{H}(f, c)$  is co-c.e. From  $\mathcal{H}(f, c)$ , one can compute a sequence  $(n_k)_{k \in \mathbb{N}}$  of integers such that  $K(n_k) \geq k$  for all  $k$ . Indeed, given  $k$ , since any Solovay function tends to  $+\infty$ , it suffices to find  $n \in \mathcal{H}(f, c)$  such that  $f(n) \geq k + c$ . Then one has  $k + c \leq f(n) \leq K(n) + c$ , thus  $K(n) \geq k$ . By a result of Kjos-Hanssen, Merkle and Stephan [13,18, Theorem 2.7], the ability to compute such a sequence is equivalent to being of diagonally noncomputable degree. By Arslanov's completeness criterion, a (co)-c.e. set of diagonally noncomputable degree is Turing complete.  $\square$

**Acknowledgement.** Downey and Nies were supported by the Marsden fund of New Zealand. We would like to thank Nan Fang for helpful discussion. Furthermore, we are grateful to the anonymous referees of *Theory of Computing Systems* for their insightful remarks.

## References

- [1] Martijn Baartse and George Barmpalias. On the gap between trivial and nontrivial initial segment prefix-free complexity. *Theory of Computing Systems*,

In press.

- [2] George Barmpalias and Thomas Sterkenburg. On the number of infinite sequences with trivial initial segment complexity. *Theoretical Computer Science*, 412(52):7133–7146, 2011.
- [3] George Barmpalias and Charlotte Vlek. Kolmogorov complexity of initial segments of sequences and arithmetical definability. *Theoretical Computer Science*, 412(41):5656–5667, 2011.
- [4] Laurent Bienvenu, Alan Day, Noam Greenberg, Antonin Kučera, Joseph S. Miller, André Nies, and Dan Turetsky. Computing  $K$ -trivial sets by incomplete random sets. *Bull. Symb. Logic*, 20:80–90, 2014.
- [5] Laurent Bienvenu and Rodney Downey. Kolmogorov complexity and Solovay functions. In *STACS*, volume 3 of *LIPICs*, pages 147–158. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2009.
- [6] Laurent Bienvenu, Wolfgang Merkle, and André Nies. Solovay functions and  $K$ -triviality. In *STACS*, volume 9 of *LIPICs*, pages 452–463. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.
- [7] Laurent Bienvenu and Alexander Shen. Random semicomputable reals revisited. In Michael J. Dinneen, Bakhadyr Khossainov, and André Nies, editors, *Computation, Physics and Beyond*, volume 7160 of *Lecture Notes in Computer Science*, pages 31–45. Springer, 2012.
- [8] Gregory Chaitin. A theory of program size formally identical to information theory. *Journal of the Association for Computing Machinery*, 22:329–340, 1975.
- [9] Gregory Chaitin. Information-theoretical characterizations of recursive infinite strings. *Theoretical Computer Science*, 2:45–48, 1976.
- [10] Gregory Chaitin. Incompleteness theorems for random reals. *Advances in Applied Mathematics*, 8:119–146, 1987.
- [11] Rudi Cilibrasi and Paul Vitanyi. Clustering by compression. *IEEE Transactions on Information Theory*, 51(4):1523–1545, 2005.
- [12] Barbara Csima and Antonio Montalbán. A minimal pair of  $K$ -degrees. *Proceedings of the American Mathematical Society*, 134:1499–1502, 2005.
- [13] Rodney Downey and Denis Hirschfeldt. *Algorithmic randomness and complexity*. Theory and Applications of Computability. Springer, 2010.
- [14] Rodney Downey, Denis Hirschfeldt, André Nies, and Frank Stephan. Trivial reals. In *Proceedings of the 7th and 8th Asian Logic Conferences*, pages 103–131. Singapore University Press, 2003.
- [15] Peter Gács. Exact expressions for some randomness tests. *Z. Math. Log. Grdl. M.*, 26:385–394, 1980.
- [16] Peter Gács. Every set is reducible to a random one. *Information and Control*, 70:186–192, 1986.

- [17] Rupert Hölzl, Thorsten Kräling, and Wolfgang Merkle. Time-bounded Kolmogorov complexity and Solovay functions. *Theory of Computing Systems*, 52(1):80–94, 2013.
- [18] Bjørn Kjos-Hanssen, Wolfgang Merkle, and Frank Stephan. Kolmogorov complexity and the recursion theorem. *Transactions of the American Mathematical Society*, 363(10):5465–5480, 2011.
- [19] Shira Kritchman and Ran Raz. The surprise examination paradox and the second incompleteness theorem. *Notices of the AMS*, 57:1454–1458, 2010.
- [20] Antonin Kučera. Measure,  $\Pi_1^0$  classes, and complete extensions of PA. *Lecture Notes in Mathematics*, 1141:245–259, 1985.
- [21] Antonin Kučera and Ted Slaman. Randomness and recursive enumerability. *SIAM Journal on Computing*, 31:199–211, 2001.
- [22] Leonid Levin. *Some theorems on the algorithmic approach to probability theory and information theory*. Dissertation in mathematics, Moscow, 1971.
- [23] Leonid Levin. The concept of random sequence. *Doklady Akademii Nauk SSSR*, 212:548–550, 1973.
- [24] Ming Li and Paul Vitányi. *An introduction to Kolmogorov complexity and its applications*. Texts in Computer Science. Springer-Verlag, New York, 3rd edition, 2008.
- [25] Alexander Melnikov and André Nies. K-triviality in computable metric spaces. *Proceedings of the American Mathematical Society*, 141(8):2885–2899, 2013.
- [26] Wolfgang Merkle and Nenad Mihailović. On the construction of effective random sets. *Journal of Symbolic Logic*, 69:862–878, 2004.
- [27] Wolfgang Merkle, Joseph S. Miller, André Nies, Jan Reimann, and Frank Stephan. Kolmogorov-Loveland randomness and stochasticity. *Annals of Pure and Applied Logic*, 138(1-3):183–210, 2006.
- [28] Joseph Miller and Liang Yu. On initial segment complexity and degrees of randomness. *Transactions of the American Mathematical Society*, 360(6):3193–3210, 2008.
- [29] Joseph 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, 2010.
- [30] Joseph S. Miller and Liang Yu. Oscillation in the initial segment complexity of random reals. *Advances in Mathematics*, 226(6):4816–4840, 2011.
- [31] André Nies. Lowness properties and randomness. *Advances in Mathematics*, 197(1):274–305, 2005.
- [32] André Nies. *Computability and randomness*. Oxford Logic Guides. Oxford University Press, 2009.

- [33] Claus Schnorr. Process complexity and effective random tests. *Journal of Computer and System Sciences*, 7:376–388, 1973.
- [34] Robert Solovay. Draft of a paper (or series of papers) on Chaitin’s work. Unpublished notes, 215 pages, 1975.
- [35] Domenico Zambella. Sequences with simple initial segments. Technical report ML-1990-05, ILLC Publications, Mathematical Logic and Foundations (ML) Series, 1990.