KOLMOGOROV COMPLEXITY AND COMPUTABLY ENUMERABLE SETS

GEORGE BARMPALIAS AND ANGSHENG LI

Abstract. We study the computably enumerable sets in terms of the:
(a) Kolmogorov complexity of their initial segments;
(b) Kolmogorov complexity of finite programs when they are used as oracles.
We present an extended discussion of the existing research on this topic, along with recent developments and open problems. Besides this survey, our main original result is the following characterization of the computably enumerable sets with trivial initial segment prefix-free complexity. A computably enumerable set A is K-trivial if and only if the family of sets with complexity bounded by the complexity of A is uniformly computable from the halting problem.

1. Introduction

The study of computably enumerable sets is a major part of classical computability theory. The main focus in Kolmogorov complexity on the other hand is arguably strings and sequences of high complexity, hence not (segments of) the characteristic sequences of computably enumerable sets. Despite this, the study of the initial segment Kolmogorov complexity of computably enumerable (c.e. for short) sets dates back to the work of Barzdins [Bar68] and hence is nearly as old as the theory of Kolmogorov complexity itself. Moreover, as we argue in the following, it is motivated by natural questions about c.e. sets and has interesting nontrivial interactions with the traditional study of c.e. sets from computability theory.

This paper is concerned with the study of the computably enumerable sets in terms of:

(a) the complexity of their initial segments;
(b) the complexity of finite programs when they are used as oracles.

Kolmogorov complexity is a well known measure of complexity of strings that is based on the intuitive idea that complicated strings do not have short descriptions. We will focus on two variants, the plain and the prefix-free complexity, which are formally defined in Section 2. However our discussions, as well as some of the arguments we present (in particular the arguments of Section 6), are often applicable to different variations like monotone or process complexity (see [DH10,

---

Last revision: 20-07-2012.

Key words and phrases. Computably enumerable sets, Kolmogorov complexity, relativization.

This research was partially done during the programme 'Semantics and Syntax' in the Isaac Newton Institute for the Mathematical Sciences, Cambridge U.K. Barmpalias was supported by the *Research fund for international young scientists* number 611501-10168 from the National Natural Science Foundation of China, and an *International Young Scientist Fellowship* number 2010-Y2GB03 from the Chinese Academy of Sciences. Li was supported by the *hundred talent program* of the Chinese Academy of Sciences. Both authors received partial support by the *Grand project: Network Algorithms and Digital Information* (number ISCAS2010-01) of the Institute of Software, Chinese Academy of Sciences.

Section 3.15] for an introduction). We will not be talking about the complexity of left or right c.e. reals, namely the binary expansions of reals in $(0, 1)$ whose left or right Dedekind cut is computably enumerable. This is already a well developed area (see [DH10, Chapter 5] for an elaborate presentation) and is not directly relevant to our topic. Moreover we do not discuss the interesting topic of the resource bounded versions of clauses (a), (b) (see [LV08, Chapter 7] for a general introduction on resource bounded Kolmogorov complexity and [LV08, Theorem 7.1.3] which refers to the resource bounded initial segment complexity of a c.e. set).

We start with a brief overview of the basics of Kolmogorov complexity in Section 2. In Section 3 we motivate the topic of this paper with various intuitive questions and a survey of the relevant work in the literature. More specifically, in Section 3.1 we study the c.e. sets according to clause (a) and we discuss to what extend the motivating questions have been addressed in the literature. Furthermore, we present the main original result in this paper (whose proof is given in Section 5), as an answer to one of these questions. Section 4 consists of an analogous discussion of the study of c.e. sets according to clause (b). Finally Section 6 is devoted to the special topic of c.e. splittings with respect to (a) and (b). Throughout the paper we point to several open questions in the context of each discussion.

2. Background on Kolmogorov complexity and randomness

A standard measure of the complexity of a finite string was introduced by Kolmogorov in [Kol65]. The basic idea behind this approach is that simple strings have short descriptions relative to their length while complex or random strings are hard to describe concisely. Kolmogorov formalized this idea using the theory of computation. In this context, Turing machines play the role of our idealized computing devices, and we assume that there are Turing machines capable of simulating any mechanical process which proceeds in a precisely defined and algorithmic manner. Programs can be identified with binary strings.

2.1. **The definition of complexity and randomness.**

A string τ is said to be a description of a string σ with respect to a Turing machine M if this machine halts when given program τ and outputs σ. Then the Kolmogorov complexity of σ with respect to M (denoted by $C_M(\sigma)$) is the length of its shortest description with respect to M. It can be shown that there exists an *optimal* machine V, i.e. a machine which gives optimal complexity for all strings, up to a certain constant number of bits. This means that for each Turing machine M there exists a constant c such that $C_V(\sigma) < C_M(\sigma) + c$ for all finite strings σ. Hence the choice of the underlying optimal machine does not change the complexity distribution significantly and the theory of Kolmogorov complexity can be developed without loss of generality, based on a fixed underlying optimal machine U. We let C denote the Kolmogorov complexity with respect to a fixed optimal machine.

When we come to consider randomness for infinite strings, it becomes important to consider machines whose domain satisfies a certain condition; the machine M is called *prefix-free* if it has prefix-free domain (which means that no program for which the machine halts and gives output is an initial segment of another). Similarly to the case of ordinary Turing machines, there exists an *optimal* prefix-free machine U so that for each prefix-free machine M the complexity of any string with respect to U is up to a constant number of bits larger than the complexity of it with respect to M. We let K denote the prefix-free complexity with respect to a fixed optimal prefix-free machine.

In order to define randomness for infinite sequences, we consider the complexity of all finite initial segments. A finite string σ is said to be *c-incompressible* if $K(\sigma) \geq |\sigma| - c$, where $K = K_U$. Levin [Lev73] and Chaitin [Cha75] defined an infinite binary sequence X to be random (also called 1-random) if there exists some constant c such that all of its initial segments are c-incompressible. By identifying subsets of $\mathbb{N}$ with their characteristic sequence we can also talk about randomness of sets of numbers. Moreover the above definitions and facts relativize to an arbitrary oracle X when the machines that we use have access to this external source of information. For example, in this case we write K^X for the corresponding function of prefix-free complexity.

This definition of randomness of infinite sequences is independent of the choice of underlying optimal prefix-free machine, and coincides with other definitions of randomness like the definition given by Martin-Löf in [ML66]. The coincidence of the randomness notions resulting from various different approaches may be seen as evidence of a robust and natural theory.

2.2. **Sets of descriptions and construction of machines.** The *weight* of a prefix-free set S of strings, denoted $\texttt{wgt}(S)$, is defined to be the sum $\sum_{\sigma \in S} 2^{-|\sigma|}$. The *weight* of a (oracle) prefix-free machine M^X is defined to be the weight of its domain and is denoted $\texttt{wgt}(M^X)$.

Prefix-free machines are most often built in terms of *request sets*. A request set L is a set of pairs $\langle \rho, \ell \rangle$ where ρ is a string and ℓ is a positive integer. A 'request' $\langle \rho, \ell \rangle$ represents the intention of describing ρ with a string of length ℓ. We define the *weight of the request* $\langle \rho, \ell \rangle$ to be $2^{-\ell}$. We say that L is a *bounded request set* if the sum of the weights of the requests in L is less than 1. This sum is the *weight of the request set* L and is denoted by $\texttt{wgt}(L)$.

The Kraft-Chaitin theorem (see e.g. [DH10, Section 2.6]) says that for every bounded request set L which is c.e., there exists a prefix-free machine M such that for each $\langle \rho, \ell \rangle \in L$ there exists a string τ of length ℓ such that $M(\tau) = \rho$. The same holds when L is c.e. relative to an oracle X, giving a machine M^X. In Section 5 and the proof of Proposition 3.1 we freely use this method of construction without explicit reference to the Kraft-Chaitin theorem.

3. Measuring the complexity of a computably enumerable set

3.1. **Initial segment complexity of computably enumerable sets.** Computable sets have trivial Kolmogorov complexity. In order to produce the first n bits of a computable set it suffices to have a description of n, since all other information can be coded in a fixed program.[1] Computably enumerable sets may not be computable, but it is not hard to see that the information they may absorb in their initial segments is quite limited.

(3.1) How complex can the segments of the characteristic sequence of a c.e. set be?

Barzdins [Bar68] observed that $2 \log n$ is an upper bound (up to an additive constant) of the plain Kolmogorov complexity of the first n bits of any computably enumerable set. Moreover he constructed a c.e. set A such that $C(A \restriction_n) \geq \log n$

[1]Note that we may consider $C(n)$ either by identifying it with $C(0^n)$ or by assuming that the underlying optimal machine prints numbers, as well as strings. Similarly for $K(n)$.

for all n while Chaitin [Cha76a] (utilizing a result of Meyer, see [LV08, Exercise 2.3.4]) showed that if $\forall n\ (C(X\restriction_n) \leq \log n + c)$ for some constant c then X is computable. On the other hand Solovay [Sol75] observed that there is no c.e. set such that $2\log n$ is a lower bound of the initial segment complexity of it (even up to an additive constant). Note that $\log n$ is an upper bound of $C(n)$. Hölzl, Kräling and Merkle [HKM09] observed that, in fact, for every c.e. set A there are infinitely many n such that $C(A\restriction_n)$ is bounded by $C(n)$ (plus an additive constant). Finally Kummer [Kum96] showed that there are c.e. sets A such that $C(A\restriction_n) \geq 2\log n - c$ for some constant c and infinitely many n. Moreover he showed that the Turing degrees that contain such complex c.e. sets are exactly the array non-computable c.e. degrees (a well studied class of degrees from computability theory). For a more detailed overview of these results and their proofs we refer to [DH10, Section 16.1].

We are not aware of a thorough study of question (3.1) in terms of the prefix-free complexity, but some facts may be obtained from the above results and the known relationships between C and K (see [DH10, Section 4.2] for a presentation of the equations of Solovay from [Sol75] that relate C and K). Moreover the observation from [HKM09] holds for K invariably.

Our next question concerns the relation between the overall information that is coded into a c.e. set and the way that this information affects the Kolmogorov complexity of its initial segments. Note that $C(n)$ and $K(n)$ are lower bounds for the plain and prefix-free initial segment complexity of any sequence. Hence we may say that the plain or prefix-free initial segment complexity of a set is *trivial* if it is bounded by one of these $C(n)$ or $K(n)$ respectively (up to an additive constant). Such sets are also known as C-trivial or K-trivial respectively.

(3.2) How much information can be coded into a c.e. set with trivial or 'low' initial segment complexity?

As we already discussed, Chaitin [Cha76a] showed that sequences with trivial plain initial segment complexity are computable. The case of prefix-free complexity turned out to be more interesting. In [DHNS03] it was shown that sequences with trivial prefix-free initial segment complexity cannot compute the halting problem. Hirschfeldt and Nies extended this result in [Nie05] and showed that the amount of information that can be coded into K-trivial sequences is in fact quite limited. On the other hand, a number of results say that there are Turing complete sets of 'very low' initial segment complexity. For example, given a nondecreasing unbounded Δ^0_2 function g there exists a complete c.e. set A and a constant c such that $K(A\restriction_n) \leq K(n) + g(n) + c$ for all n. This was demonstrated by Frank Stephan, see [BV11, Section 5]. In [BB10, BV11] it was shown that this result is optimal, in the sense that it is no longer true if one of the conditions on g is removed. A stronger result was obtained in [Bar11b]. It was shown that there are complete c.e. sets of arbitrarily low complexity, with respect to the nontrivial complexities of the c.e. sets. A more precise statement of this result is given in Section 3.2 where a formal way for comparing the initial segment complexities of two sets is discussed.

Another topic of interest concerns the study of the ways in which sufficiently random oracles are no better than computable oracles for performing certain computational tasks. An early observation from [dLMSS55] is that if a set is c.e. relative to a sufficiently random oracle then it is c.e. without the use of an oracle. Moreover it is well known (e.g. see [BLS08a, Section 3]) that if a sequence X is random, then it is random relative to every sufficiently random sequence. We give an example of such a result in the context of this paper. Recall that $K(n)$ is the trivial complexity and a set is K-trivial if its complexity is trivial (modulo an additive constant). The following observation says that if the initial segment complexity of a c.e. set relative to a sufficiently random oracle is trivial, then the set is already K-trivial. The level of randomness that is required for this result is weak 2-randomness. An oracle is weakly 2-random if it is not a member of any null Π^0_2 class.

Proposition 3.1. *Suppose that X is weakly 2-random and A is a c.e. set. If $\exists c \forall n\ K^X(A\restriction_n) \le K(n)+c$ then A is K-trivial.*

Proof. Given a c.e. set A and a constant c, the class of oracles X such that $\forall n\ K^X(A\restriction_n) \le K(n)+c$ is a Π^0_2 class. Hence it suffices to show that if this class is not null then A is K-trivial. On this assumption, by Kolmogorov's 0-1 law there exists a constant d such that the measure of the oracles X such that $\forall n\ K^X(A\restriction_n) \le K(n)+d$ is larger than $1/2$. Let U be the underlying optimal oracle machine such that $\mu(U^X) < 1/2$ for all oracles X. We also assume that any computations of U at stage s use less than s bits of the oracle. Without loss of generality we may assume that for all X, if there is a U^X description of length n that describes some string τ then for each $i > n$ there exists a U^X description of τ of length i. Given a computable enumeration $A[s]$ of A we construct a prefix-free machine M as follows. At stage $s+1$ let n be the least number $\le s$ such that

- $K_M(A\restriction_n)[s] > K(n)[s]+d$
- $K^X(A\restriction_n)[s] \le K(n)[s]+d$ for a set of oracles X of measure $> 1/2$

(if there is no such n, do nothing). Then enumerate an M-description of $A\restriction_n[s]$ of length $K(n)[s]+d$.

It remains to prove that the request set for M is bounded. Let σ_i be the ith description enumerated in M. It suffices to show that $\sum_{i\le n} 2^{-|\sigma_i|} < 1$ for all n. For a contradiction, suppose that $\sum_{i\le n_0} 2^{-|\sigma_i|} \ge 1$ for some n_0, and let s_0 be the stage where σ_{n_0} was enumerated into our machine M. Each string σ_i, $i \le n_0$ contributes at least $2^{-|\sigma_i|-1}$ to the expected weight of the machine $U^{X\restriction_{s_0}}$, where $X\restriction_{s_0}$ is any oracle of length s_0. Since $\mu(U^X) < 1/2$ for all oracles X, this expected value is also less than $1/2$. This gives the required contradiction. □

Note that Proposition 3.1 is no longer true if we replace 'weakly 2-random' with '1-random' since there are 1-random sequences which compute all c.e. sets. Moreover the same proof shows the result in the more general case when A is Δ^0_2.

We would like to make a note of another way to study the complexity of c.e. sets which was introduced by Chaitin [Cha76b] and will not be studied in this paper. The (algorithmic) entropy of a c.e. set A was defined as the probability that a universal c.e. operator enumerates A. Most of the research around this concept has to do with the relationship to another measure of complexity, namely the length of the shortest prefix-free description of a c.e. index of A. Solovay [Sol75, Sol77] obtained an upper bound of the latter in terms of the entropy function and Vereshchagin [Ver07] improved it for the special case of finite sets. For a more elaborate overview of this research on the algorithmic entropy of c.e. sets we refer to [DH10, Section 16.2].

3.2. **Measures of relative complexity on computably enumerable sets.** A fruitful way to study the complexity of a sequence is to compare it with the complexities of other sequences. Measures of relative complexity provide a formal way to do this. In the case of initial segment complexity, a number of such measures were introduced by Downey, Hirschfeld and LaForte in [DHL04]. One such measure is the $\leq_K$ reducibility defined as

$$X \leq_K Y \iff \exists c \forall n \; (K(X \upharpoonright_n) \leq K(Y \upharpoonright_n) + c)$$

as well as its plain complexity version $\leq_C$ which is defined similarly. We may express the fact that $X \leq_K Y$ simply by saying that the initial segment complexity of X is less than (or equal to) the complexity of Y. The result from [Bar11b] that we discussed in relation to question (3.2) may be formally stated as follows. Given any c.e. set B which is not K-trivial, there exists a Turing complete c.e. set A such that $A <_K B$.

A central topic in the study of c.e. sets in computability theory are the 'c.e. splittings', see [DS93]. We say that a pair B, C of c.e. sets is a c.e. splitting of A if $B \cap C = \emptyset$ and $B \cup C = A$. One of the simplest questions that we can ask about the initial segment complexity of c.e. splittings is the following.

(3.3) Given a c.e. set can we split it into two c.e. sets of strictly less initial segment complexity?

A positive answer was given in [Bar11a, Section 5] for both the plain and the prefix-free complexity, provided that the given set has non-trivial complexity. In Section 6 we give an extension of this splitting theorem, showing that the splitting may avoid bounding the complexity of any given nontrivial Δ^0_2 set. Note that such results can be viewed as analogues of the classic Sacks splitting theorem that was proved in [Sac63] for the Turing degrees.

(3.4) Given a c.e. set can we split it into two c.e. sets of the same initial segment complexity?

This question is formally addressed in Section 6, but remains open. We note that the sets that satisfy the analogue of (3.4) in the Turing degrees are called mitotic and have been studied extensively in computability theory, see [Lad73b, Lad73a, DS89]. The existence of non-mitotic sets was shown by Lachlan in [Lac67].

(3.5) Is there a c.e. set whose initial segment complexity is maximal amongst the c.e. sets?

This question has been answered negatively in [Bar05] with respect to a stronger measure $\leq_{\mathrm{cl}}$, where $X \leq_{\mathrm{cl}} Y$ if $X \upharpoonright_n$ can be computed from $Y \upharpoonright_{n+c}$ for some constant c and all n. In [ASDFM11] an easier proof of this result was given. By [DHL04] the partial order $\leq_{\mathrm{cl}}$ coincides with the Solovay reducibility on the c.e. sets, which is a standard measure of relative randomness for the larger class of c.e. reals. In particular, it does express in some (crude) sense the relative complexity of sequences. Other aspects of $\leq_{\mathrm{cl}}$ on the c.e. sets were studied in [BL06, Day10]. A similar study would be interesting for $\leq_K$ and $\leq_C$. For example, we do not know if the corresponding structures of c.e. degrees are dense. Another question is whether for every pair of c.e. sets A, B there exists a c.e set C such that $A \leq_K C$ and $B \leq_K C$.

(3.6) How 'large' is the class of sets with initial segment complexity bounded by the complexity of a c.e. set?

There are many ways to measure the largeness of a class of reals, including determining the cardinality of the class. In the case of $\leq_C, \leq_K$ it turns out that the lower cones below a c.e. sets are always subsets of Δ^0_2, hence countable (e.g. see [BV11, Section 2]). It is interesting to examine if these are uniform subclasses of Δ^0_2, in the sense that they can be indexed by a single $\emptyset'$-computable predicate. In Section 5 we prove the following characterization.

Theorem 3.2. *The following are equivalent for a computably enumerable set A.*

(a) *A is K-trivial;*
(b) *every set $X \leq_K A$ is truth-table reducible to $\emptyset'$.*

Moreover the above clauses are equivalent to the condition that $\{X \mid X \leq_K A\}$ is uniformly computable in $\emptyset'$.

This result provides an answer to question (3.6) since a uniform subclass of Δ^0_2 may be considered 'effectively small' while classes that do not admit a uniform parameterization are, in a sense, 'effectively large'. A more precise treatment of this notion of 'largeness' may be obtained via the use of resource bounded measure, an approach that was developed by Jack Lutz in various contexts and is based on the use of effective martingales. In our case we are interested in the size of a set of reals as a subclass of Δ^0_2. An example of such a study is [HT08] where it is shown that the Δ^0_2 measure of the class of oracles that are computable by an incomplete Δ^0_2 set is 0. We may ask the same question with respect to $\leq_K$, $\leq_C$ as a formal version of question (3.6). In other words, to determine the Δ^0_2 measure of the class of oracles with initial segment complexity bounded by the complexity of a given c.e. set.

Since Section 5 is entirely devoted to the proof of Theorem 3.2, we wish to say a few more words on its relevance with the work of other authors. Both directions of the equivalence that it asserts are nontrivial. Chaitin [Cha77] observed that (by a relativization of an argument from [Lov69]) all K-trivial sequences are computable from the halting problem; equivalently, they have a computable approximation. However this proof is not uniform, hence it does not show that the K-trivial sequences can be listed by a single machine operating with oracle $\emptyset'$. One of the consequences of Nies [Nie05] was that the family of K-trivial sequences is indeed uniformly $\emptyset'$-computable. This is the only known proof of this fact and is highly nontrivial, involving the full power of what is now known as the *decanter method*.

We already pointed out that given any c.e. set A, the class of sequences with initial segment prefix-free complexity that is bounded by the complexity of A is contained in Δ^0_2. This is merely an extension of the argument for the case where $A = \emptyset$, so again it is nonuniform. We wanted to know if this class can be uniformly $\emptyset'$-computable in any cases other than the known case where A is K-trivial. A positive answer would have interesting consequences on the local structures of the K degrees. For example, combined with the results in [BV11] it would establish the existence of a pair of Δ^0_2 sets that form a minimal pair in the K degrees. However, Theorem 3.2 shows that this uniformity is a special feature that characterizes the K-trivial computably enumerable sets.

Corollary 3.3. *The following are equivalent for any finite collection of computably enumerable sets A_i, $i < k$.*

(a) *There exists $i < k$ such that A_i is K-trivial;*
(b) *The sets X such that $X \leq_K A_i$ for all $i < k$ are truth-table reducible to $\emptyset'$.*

Moreover (a), (b) *are equivalent to the condition that $\{X \mid \forall i < k \ (X \leq_K A_i)\}$ is uniformly computable in $\emptyset'$.*

This generalization of Theorem 3.2 may be obtained by an application of a result in [Bar11b]. Namely, it was shown that given two c.e. sets B, C such that $\emptyset <_K B$ and $\emptyset <_K C$ there exists a c.e. set A such that $A \leq_K B$, $A \leq_K C$ and $\emptyset <_K A$.[2]

(3.7) What is 'algorithmical independence' for computably enumerable sets?

The notion of algorithmic independence of random sequences is well understood through the concept of relative randomness. In particular, two random sequences may be regarded algorithmically independent if each of them is random relative to the other. Various authors have attempted to extend the notion of algorithmic independence to a wider class of sequences. Such formalizations have been suggested by Levin [Lev74, Lev84, Lev02] through the concept of 'mutual information' of sequences, and by Calude and Zimand [CZ10]. Although the exact relationship between these formalizations is not known, some of them are clearly too crude for the purpose of expressing independence for pairs of computably enumerable sets. For example the definitions in [CZ10] are not sensitive to additive logarithmic factors. In particular, since the complexity of a c.e. set is at most $4 \log n$, every pair of c.e. sets is independent according to [CZ10]. It would be interesting to address question (3.7) by crafting an appropriate formalization which expresses the informal concept of independence for c.e. sets. We note that the concept of minimal pairs with respect to $\leq_K$, $\leq_C$ does express some notion of independence. The existence of minimal pairs of c.e. sets with respect to $\leq_C$ was shown in [MS07] and the nonexistence with respect to $\leq_K$ was shown in [Bar11b].

4. Relative compression power of computably enumerable oracles

A second way to study the c.e. sets with respect to Kolmogorov complexity is to examine their power when they take the place of an oracle in the underlying optimal universal machine. Recall that K^X denotes the prefix-free complexity relative to oracle X. Hirschfeldt and Nies showed in [Nie05] that K^X does not differ from K more than a constant if and only if $X \equiv_K \emptyset$. In other words, the oracles that do not improve the compression of finite programs significantly are exactly the oracles with trivial initial segment prefix-free complexity. A natural way to compare the compression power of oracles was introduced in [Nie05] in the form of the reducibility $\leq_{LK}$.

$$X \leq_{LK} Y \iff \exists c \forall \sigma \ (K^Y(\sigma) \leq K^X(\sigma) + c).$$

In other words $X \leq_{LK} Y$ formalizes the notion that Y can achieve an overall compression of the strings that is at least as good as the compression achieved by X. Moreover by [KHMS12] it coincides with $X \leq_{LR} Y$ which denotes the relation

[2] At this point we would like to draw a parallel between the study of the K degrees of c.e. sets and the K degrees of Martin-Löf random sets that was the object of study in [MY08, MY10]. One of the main open questions in this study was whether there is a maximal element in the K degrees of random reals, which is an analogue of question (3.5). Moreover it was shown that is a pair of random reals X, Y which has no upper bound with respect to $\leq_K$. Finally it was shown that given any finite collection X_i, $i < k$ of random reals, there exists a random real Y such that $Y <_K X_i$ for all $i < k$. The analogue of this result for the c.e. sets was proved in [Bar11b]. This analogy stems from the analogy between the oscillation of the prefix-free complexity of a random sequence between n and $n + K(n)$ and the oscillation of the complexity of a c.e. set between $K(n)$ and $4 \log n$.

that every random sequence relative to Y is also random relative to X. The induced degree structure is known as the LK degrees.

This measure of relative complexity has been studied extensively in the literature, although most of the publications are written in terms of $\leq_{LR}$. Moreover the structure of the c.e. sets under $\leq_{LK}$ has also been studied. It is not hard to see that $\leq_T$ is contained in $\leq_{LK}$ so the c.e. Turing degrees have more in common with the c.e. LK degrees than the c.e. K degrees. In particular, the LK degree of the halting set is complete. Despite this, the two structures are not elementarily equivalent [Bar10b]. In particular, there are no minimal pairs in the structure of LK degrees of c.e. sets. Analogues of various questions that were discussed in Section 3.2 have been addressed for $\leq_{LK}$. Splitting and non-splitting theorems (see questions (3.3) and (3.4)) where obtained in [BLS08a, BM09] and are discussed in Section 6.2.

The size of lower $\leq_{LK}$ cones (see question (3.6)) in terms of cardinality was studied in [BLS08a, BLS08b, Mil10] and was fully determined in [Bar10c] (for the c.e. case) and in [BL11] in general. In [Bar10c] it was shown that for every c.e. set A which is not K-trivial the class $\{X \mid X \leq_{LK} A\}$ contains a perfect Π^0_1 class. In [BB10] it was shown how this argument can be strengthened so that the constructed Π^0_1 class does not have K-trivial members (this is one of the arguments where avoiding the K-trivial paths in the constructed Π^0_1 class is highly non-trivial). On the other hand in [Bar10a] it was shown that every Π^0_1 class with no K-trivial paths contains a $\leq_{LK}$-antichain of size $2^{\aleph_0}$. The combination of these results shows that the lower $\leq_{LK}$-cone below any c.e. set which is not K-trivial is rather large, in the following sense.

Corollary 4.1. *If A is c.e. and not K-trivial then $\{X \mid X \leq_{LK} A\}$ contains a $\leq_{LK}$-antichain of size $2^{\aleph_0}$.*

Recently, Yu Liang showed that it it is not true that every perfect set contains a $\leq_{LK}$-antichain of size $2^{\aleph_0}$. In fact, quite interestingly, he showed that there exists a perfect set of reals which is a chain with respect to $\leq_{LK}$.

We have already mentioned that $\leq_{LK}$ contains $\leq_T$. This fact allows various standard questions to be asked about the structure of the c.e. Turing degrees inside a c.e. LK degree. Such issues have been studied in [BLS08a, BLS08b]. Ever c.e. LK degree contains infinite chains and antichains of Turing degrees. In fact, the following stronger result was shown, where $|_T$ denotes Turing incomparability.

(4.1) If A is a noncomputable incomplete c.e. set then there exist c.e. sets B, C, D in the LK degree of A such that $B <_T A <_T C$ and $D \mid_T A$.

The proof of (4.1) consists of encapsulating a number of basic c.e. Turing degree constructions (often involving the finite and infinite injury priority method) inside an LK degree.

Another question about the structure of the c.e. LK degrees is whether it is dense. In [BLS08b] it was shown that if $A <_{LK} B$ for two c.e. sets whose LK degrees have $\leq_T$-comparable c.e. members then there exists a c.e. set C such that $A <_{LK} C <_{LK} B$. In particular, the LK degrees of c.e. sets are upward and downward dense. However the density of the c.e. LK degrees is an open question, also stated in [MN06, Question 9.12]. A relevant question is if there are c.e. sets A, B such that $A <_{LK} B$ and every c.e. set in the LK degree of A is $\leq_T$-incomparable with every c.e. set in the LK degree of B. When the sets are not required to be c.e.

this question admits a positive answer. This follows from the fact that each LK degree is countable [Nie05] and the fact that there are LK degrees with uncountably many predecessors [BLS08a].

Finally, not much is known about least upper bounds in the LK degrees of c.e. sets. For every sets A, B a natural upper bound in the Turing (and hence the LK) degrees is $A \oplus B$. However although the Turing degree of $A \oplus B$ is the least upper bound of the degrees of A, B the same is not necessarily true for their LK degrees, even when they are computably enumerable. This was first noticed in [Nie05] and various results since show that in some cases the LK degree of $A \oplus B$ is, in a certain sense, very far from being the least upper bound of the degrees of A and B (e.g. [BLS08b, Corollary 12]). Diamondstone [Dia11] showed that with respect to $\leq_{LK}$ every pair of low sets has a low c.e. upper bound. If we consider a pair of low c.e. sets A, B such that $A \oplus B \equiv_T \emptyset'$ then by [Dia11] the LK degree of $A \oplus B$ is, in a certain sense, very far from being the least upper bound of A, B. We do not know of any pair of c.e. sets of incomparable LK degrees which have a least upper bound in the LK degrees of c.e. sets. We also do not know of a pair of c.e. sets whose LK degrees do not have a least upper bound in the LK degrees of c.e. sets.

5. Proof of Theorem 3.2

Let us call (c) the clause that $\{X \mid X \leq_K A\}$ is uniformly computable in $\emptyset'$. Then Theorem 3.2 says that clauses (a), (b), (c) are equivalent. The implication from (a) to (b) is a result from [Nie05]. Since the sets that are truth-table reducible to $\emptyset'$ are uniformly $\emptyset'$-computable, clause (b) implies (c). For the remaining implication assume that A is c.e. and not K-trivial. Moreover let (X_i) be a uniformly $\emptyset'$-computable family of sets. This means that there is a universal computable approximation $(X_i[s])$ such that each $X_i[s]$ converges to X_i as $s \to \infty$. It suffices to construct a computable approximation $B[s]$ converging to set B such that $B \leq_K A$ and $B \neq X_i$ for all i.

For the satisfaction of $B \neq X_i$ we pick a number n_i (a witness) and at each stage s we let $B(n_i)[s] = 1 - X_i(n_i)[s]$. Since $X_i(n_i)$ converges, $B(n_i)[s]$ converges to an appropriate value. For $B \leq_K A$ we will construct a prefix-free machine M such that

$$K_M(B \restriction_k) \leq K(A \restriction_k) \quad \text{for all } k \tag{5.1}$$

where K_M denotes the prefix-free complexity relative to the machine M. Recall that K denotes the prefix-free complexity relative to a fixed universal prefix-free machine U. Without loss of generality we may assume that $\mathtt{wgt}(U) < 2^{-2}$. The enumeration of M is straightforward. At each stage we look for the least k such that (5.1) is not satisfied and we enumerate an M-description of $B[s] \restriction_k$ of length $K(A \restriction_k)[s]$. The condition $B \leq_K A$ is satisfied provided that we manage to keep the weight of the requests that we enumerate into M bounded. This bound may be obtained via an analysis of the requests that are enumerated in M in relation to the descriptions that are produced in U.

5.1. **The model.** Each description that is enumerated in M corresponds to a unique description in the domain of the universal machine U of the same length. Indeed, when the construction requests M to describe some $B[s] \restriction_x$ at stage $s+1$, this is in order to achieve $K_M(B \restriction_x)[s] \leq K(A \restriction_x)[s]$. Hence the new description in M corresponds to the (least) shortest description in U of $A \restriction_x [s]$. Since the

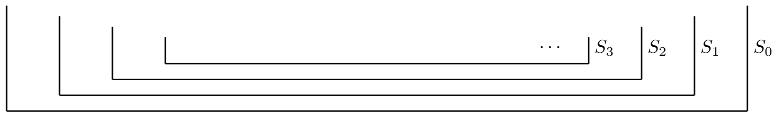

Figure 1. Infinite nested decanter model.

approximation to B changes in the course of the construction, this correspondence is not one-to-one. If a U-description σ corresponds to n distinct M-descriptions we say that σ is used n times.

Let S_0 be the domain of U and for each $k > 0$ let S_k contain the descriptions in the domain of U which are used at least k times. Note that $S_{i+1} \subseteq S_i$ for each i. According to the correspondence that we defined between the domains of U, M a string σ in the domain of U that is used k times incurs weight $k \cdot 2^{-|\sigma|}$ to the domain of M. Hence (5.2) holds.

$$\texttt{wgt}(M) \leq \sum_k k \cdot \texttt{wgt}(S_k). \tag{5.2}$$

A U-description is called *active* at stage s if $U(\sigma)[s] \subseteq A[s]$. By the direct way that M is enumerated, all descriptions that enter S_1 at some stage s are currently active. More generally, only currently active strings may move from S_k to S_{k+1} at any given stage.

The sets S_k may be visualized as the nested containers of the infinite decanter model of Figure 1. As the figure indicates, descriptions may move from S_k to S_{k+1} but they also remain in S_k. If $B(n_i)[s] \neq B(n_i)[s+1]$ at some stage $s+1$ for some witness n_i, some strings move from S_k to S_{k+1} for various $k \in \mathbb{N}$ (i.e. they are used one more time). In this case we say that these strings were reused by n_i.

5.2. **Movable markers and auxiliary machines.** A single witness n_i may use each description in the domain of U at most once, thus contributing to a 2:1 correspondence between the domains of M and U. However t many witnesses may create a $2^t : 1$ correspondence between U, M which may inflate the bound in (5.2). For this reason the witnesses n_i will be movable and obey the following rules (provided that they are defined).

- $n_i[s] < n_{i+1}[s]$ and $n_i[s] \leq n_i[s+1]$;
- If $B(n_i)[s] \neq B(n_i)[s+1]$ then $n_{i+1}[s+1]$ moves to a *large* value.

The witnesses $n_i[s]$ will ultimately reach limits n_i. In order to ensure that the descriptions in U that are reused many times have sufficiently small weight (i.e. they describe sufficiently complex strings), for each witness n_i we enumerate a prefix-free machine N_i during the construction. The purpose of N_i is to achieve $\forall z\, (K_{N_i}(A\restriction_z) \leq K(z) + c_i)$ for some constant c_i. Since A is not K-trivial, this will ultimately fail. However this failure will help to demonstrate that n_{i+1} converges (and, indirectly, that the weight of the strings that are reused by n_i is sufficiently small). Each time n_i moves, the value of c_i increases by 1. Such an event is described as an 'injury' of n_i. In particular, if at some stage s a witness n_i moves while n_j, $j < i$ remain constant this causes n_x, $x \geq i$ to be injured, which has the following consequences:

- n_x, $x > i$ become undefined;

- the values $c_x, x \geq i$ increase by 1.

Each witness will only be injured finitely many times. We let $c_i[s]$ denote the value of c_i at stage s and $c_i[0] = i + 3$.

At each stage s let $t_i[s]$ be the least number t such that $K_{N_i}(A\upharpoonright_t)[s] > K(t)[s] + c_i[s]$. Each witness n_i has the incentive to move its successor n_{i+1} at some stage $s+1$ if it observes a set of descriptions of segments of $A[s]$ that are longer than n_{i+1}, of sufficient weight. This weight is determined by the threshold $q_i[s]$ and is set to $2^{-K(t_i)[s]-c_i[s]}$. Witness n_i may move its successor n_{i+1} either because the above weight exceeds the threshold or because the approximation to $B(n_i)$ changes. Due to this second incentive for movement, a new parameter $p_i[s]$ will tune the threshold to an appropriate value. In particular, witness n_i requires attention at stage $s+1$ if it is defined and one of the following occurs:

(a) $B(n_i)[s] = X_i(n_i)[s]$;
(b) $\sum_{n_{i+1}[s]<j\leq s} 2^{-K(A\upharpoonright_j)[s]} \geq q_i[s] - p_i[s]$.

At each stage $s+1$ the machines N_i will be adjusted according to changes of $K(n)$ for $n < t_i[s]$. This is done by running the following subroutine.

(5.3) For each $i \leq s$ and each $n < t_i[s]$, if $K(n)[s+1] < K(n)[s]$ then enumerate an N_i-description of $A[s]\upharpoonright_n$ of length $K(n)[s+1] + c_i[s]$.

A *large* number at stage $s+1$ is one that is larger than any number that has been the value of any parameter in the construction up to stage s.

5.3. **Construction of B, M, N_r.** At stage 0 place n_0 on 0. At stage $s+1$ run (5.3). If none of the currently defined witnesses requires attention, let k be the largest number with $n_k[s]\downarrow$, let z be the least number that is bounded by s and the current value of some marker such that $K_M(B\upharpoonright_z)[s] > K(A\upharpoonright_z)[s]$ and

- place n_{k+1} on the least *large* number;
- enumerate an M-description of $B[s]\upharpoonright_z$ of length $K(A\upharpoonright_z)[s]$.

Otherwise let x be the least number such that n_x requires attention and define $n_{x+1}[s+1]$ to be a *large* number. Moreover declare $n_i[s+1]$, $i > x+1$ undefined, set $c_j[s+1] = c_j[s]+1$ for each $j > x$ and if (a) applies set $B(n_x)[s] = 1 - X_i(n_x)[s]$. If (b) applies set $p_x[s+1] = 0$ and enumerate an N_x-description of $A\upharpoonright_{t_x}[s]$ of length $K(t_x)[s]$. If (b) does not apply set $p_x[s+1] = p_x[s] + \sum_{m_x[s]<j\leq s} 2^{-K(A\upharpoonright_x)[s]}$.

5.4. **Verification.** When n_{i+1} is first defined at some stage s it takes a *large* value so $t_i[s] < n_{i+1}[s]$. Moreover t_i can only increase when N_i computations are enumerated on strings of length t_i, which happens only when n_{i+1} moves. Hence by induction we have (5.4).

(5.4) For all i, s, if $n_{i+1}[s]$ is defined then $t_i[s] < n_{i+1}[s]$.

If $K(n)$ decreases at some stage $s+1$ for some $n < t_i[s]$, subroutine (5.3) will ensure that $K_{N_i}(A\upharpoonright_n)[s+1] \leq K(n)[s+1] + c_i[s+1]$. Hence t_i may only decrease at $s+1$ if $A[s+1]\upharpoonright_{t_i[s]} \neq A[s]\upharpoonright_{t_i[s]}$, which implies (5.5).

(5.5) If $A[s]\upharpoonright_{t_i[s]} = A[s+1]\upharpoonright_{t_i[s]}$ then $t_i[s] \leq t_i[s+1]$.

The enumeration of descriptions into N_i occurs with overhead c_i, in the sense that at stage s any description of a string of length n that is defined in N_i has length

$K(n)[s] + c_i[s]$. This implies (5.6).

(5.6) At any stage s the weight of the N_i-descriptions that describe initial segments of $A[s]$ is less than $2^{-c_i[s]}$.

For each i there is a machine N_i as prescribed in the construction.

Lemma 5.1. *For each i the weight of the requests in N_i is bounded.*

Proof. The weight of the requests that are enumerated in N_i by subroutine (5.3) is bounded by the weight of the domain of U, which is at most 2^{-2}. In order to calculate the weight of the requests that are enumerated by the main construction, let s_j be the stages where requests are enumerated into N_i. Note that during each interval $[s_j, s_{j+1})$ the successor witness n_{i+1} may move many times, thereby increasing p_i which becomes 0 at s_{j+1}. At each s_j the witness n_{i+1} moves to a *large* value and the weight of the request that is issued in N_i is $q_i[s] \leq \sum_{x \in (n_{i+1}[s_{j-1}], s]} w_x$, where $s_{-1} = 0$ and w_x is the weight of the descriptions in U that describe strings of length x. Hence by induction the weight of the requests that are enumerated in N_i in this way is also bounded by the weight of the domain of U. Hence $\mathtt{wgt}(N_i) \leq 2^{-2} + 2^{-2} = 2^{-1}$. □

In order to calculate a suitable upper bound for each $\mathtt{wgt}(S_k)$ of (5.2) we need (5.7). Recall that a U description σ is reused by n_j at stage $s+1$ if it is the leftmost string describing $A[s] \restriction_z$ via U for some $z \in (n_j[s], n_{j+1}[s]]$ and at stage $s+1$ the construction enumerates an M-description of $B[s] \restriction_z$ of length $K(A \restriction_z)[s]$. Note that every reused U-description is actually reused by some n_i.

(5.7) If during the interval of stages $[s, r]$ the witness n_j is not injured then the weight of the strings that it reuses during this interval which remain active at stage r is at most $2^{-c_j[s]} + p_j[s]$.

Indeed, if at some stage the witness n_{j+1} moves but no enumeration in N_j takes place, the weight of the additional strings that n_j may be called to repay at some later stage equals the increase in p_j. Hence to prove (5.7) it suffices to show that at each stage in $[s, r]$ where an N_j enumeration takes place (and p_j becomes 0) the weight of the strings that have been reused by m_j and remain active is at most $2^{-c_j[s]}$.

At each such stage x the weight of the additional strings that n_j may be called to repay at some later stage at most $q_j[x]$. Moreover at stage x a string of weight $q_j[x]$ is used to describe $A \restriction_{t_j} [x]$ via N_j. By (5.4) if at least one of these additional descriptions in U continues to be active at stage r, then $A[x] \restriction_{t_j[s]} = A[r] \restriction_{t_j[s]}$. Hence by (5.5) we get that $t_j[y] \geq t_j[x]$ for all $y \in [x, r]$. So during the stages in $[x, r]$ the weight of the descriptions in U that n_j repaid and remain active at stage r is bounded by the weight of the descriptions in N_j that describe segments of $A[r]$. By (5.6) this is at most $2^{-c_j[s]}$. This concludes the proof of (5.7).

Lemma 5.2. *The weight of the requests that are enumerated in M is finite.*

Proof. Since only strings in the domain of U are used, $\mathtt{wgt}(S_1) < 2^{-2}$ and since $S_2 \subseteq S_1$ we also have $\mathtt{wgt}(S_2) < 2^{-2}$. Let $k > 1$. Every entry of a string into S_{k+1} is due to some witness n_x which reused it when it was already in S_k. Since $k > 1$, this string entered S_k due to another witness n_y with $y > x \geq 0$ which subsequently moved to a *large* value. Inductively, that string entered S_1 due to a witness n_z with $z \geq k-1$. Fix z, let S_k^z contain the strings in S_k that entered S_1 due to witness n_z

and let (s_i) be the increasing sequence of stages where witness n_z is injured. Note that at this point we do not assume that (s_j) is a finite.

The strings that move from S_k^z to S_{k+1} are currently active and may be divided into the packets of strings that enter S_1 due to n_z during each interval $(s_i, s_{i+1}]$. According to (5.7) the weight of the strings in the ith packet which are still active at s_{i+1} (hence, may move from S_k^z to S_{k+1} later on) is bounded by $2^{-c_z[s_i-1]}$. So the weight of the strings that enter S_{k+1} from S_k^z is bounded by $\sum_j 2^{-c_z[s_j-1]}$. Since $c_z[s_{j+1}-1] = c_z[s_j] < c_z[s_j-1]$ for all j, this weight is bounded by $\sum_j 2^{-c_z[0]-j} = 2^{-c_z[0]+1}$. Since $c_z[0] = z+3$ this bound becomes 2^{-z-2}. Since $S_k = \cup_{z \geq k-1} S_k^z$ the total weight of the strings that enter S_{k+1} from S_k is bounded by $\sum_{z \geq k-1} 2^{-z-2} = 2^{-k}$. Therefore by (5.2) the weight of M is finite. □

By Lemma 5.2 the machine M prescribed in the construction exists. The following proof uses the fact that each N_i is a prefix-free machine, which was established in Lemma 5.1.

Lemma 5.3. *For each i the witness n_i moves only finitely many times i.e. $n_i[s]$ reaches a limit.*

Proof. Assume that this holds for all $i \leq k$. Then n_k stops moving after some stage s_0. The construction will define n_{k+1} at some later stage s_1. In the following we write n_k for the the limit of $n_k[s]$ when $s \to \infty$. Similarly, $c_k[s]$ reaches a limit $c_k := c_k[s_0]$ at s_0. Since A is not K-trivial there is some least j such that $K_{N_k}(A \restriction_j) > K(j) + c_k$. If $s_2 > s_1$ is a stage where the approximations to $A \restriction_j$ and $K(i)$, $i \leq j$ have settled then the approximations to t_k, q_k also reach a limit by this stage.

Let $s_3 > s_2$ be a stage at which the approximation to the membership of n_k in X_k has reached a limit. If n_{k+1} moved after stage s_3 this would be solely due to clause (b) of Section 5.2. Hence at such a stage the construction would enumerate an N_k-description of $A \restriction_j$ of length $K(j) + c_k$ which contradicts the choice of j. Hence n_{k+1} reaches a limit by stage s_3 and this concludes the induction step. □

We may now show that $B \leq_K A$.

Lemma 5.4. *The approximation to B converges and (5.1) is met.*

Proof. If k is not the limit of some witness n_i then $B(k)$ will only change finitely often, since witnesses are defined monotonically and redefined to *large* values. On the other hand for the limit value n_i of the ith witness the construction will stop changing the approximation to $B(n_i)$ once the approximation to $X(n_i)$ stops changing. Hence the approximation to B converges. Moreover the constant enumeration of M descriptions by the construction ensures that (5.1) holds for B. □

Finally, we may conclude that B is not a member of the given uniformly $\emptyset'$-computable family of sets. Given k consider the limit n_k of the kth witness that was established in Lemma 5.3. The construction explicitly ensures that the final value of $B(n_k)$ is $1-X(n_k)$. Hence $B \neq X_k$ for all k. This concludes the verification of the construction and the proof of Theorem 3.2.

6. Computably enumerable splittings and Kolmogorov complexity

A computably enumerable (c.e.) splitting of a c.e. set A is a pair of disjoint c.e. sets B, C such that $A = B \cup C$. This notion has been the subject of interest for many researchers in computability but also in logic in general. For example it plays a special role in the study of the lattice of the c.e. sets under inclusion (see [DS90]), which is a very developed area of computability theory (see [Soa87, Chapter X] for an overview). Moreover the Turing degrees of c.e. splittings have been studied extensively (see [Soa87, Chapter XI] for an overview). For a comprehensive survey of c.e. splittings in computability theory we suggest [DS93].

In this section we discuss c.e. splittings in the context of Kolmogorov complexity. For example, we are interested in the initial segment complexity of the members B, C of the splitting given the complexity of the original set A. In Section 6.1 we show that some of the classical theory of c.e. splittings can be generalized (both in terms of results and in terms of methods) to the context of initial segment complexity. Section 6.2 discusses analogous results when c.e. sets are used as oracles for compressing finite programs. Our presentation has a bias toward the prefix-free version of Kolmogorov complexity, but most of the results and methods that we discuss in this section also hold for plain Kolmogorov complexity.

The results that we present regarding the structure of the K degrees (comparing the initial segment complexity of c.e. sets) and the structure of the LK degrees (comparing the compression strength of c.e. oracles) revolve around the same structural questions and often have the same answers. However in general the two structures are very different, even in the case of c.e. sets. For example, $\leq_{LK}$ is an extension of Turing reducibility but, as many results in [MS07] demonstrate, there is no direct connection between $\leq_K$ and Turing reducibility. In particular, there is a complete c.e. LK degree (i.e. maximum amongst the c.e. LK degrees) but it is not known if there is a maximum c.e. K degree. In our view this is unlikely, and a more interesting open question is whether there exist maximal c.e. K degrees.

The general programme of transferring results and methods from classical computability theory to the study of Kolmogorov complexity (as well as its limitations) is a fascinating topic and a critical discussion of it and its relation with arithmetical definability may be found in [BV11, Section 1].

6.1. Initial segment complexity and c.e. splittings.

Given a computably enumerable set A, we are interested in the initial segment complexity of the members of the various splittings of A. It is not hard to see that the Kolmogorov complexity of $A\restriction_n$ is equal to the Kolmogorov complexity of the last number $< n$ that enters A in a computable enumeration of it. A basic result from [Bar11a, Section 5] and [Ste11, Chapter 2] is that the analogue of the Sacks splitting theorem (e.g. see [Soa87, Theorem 3.1]) holds in the context of plain or prefix-free Kolmogorov complexity.

(6.1) If A is c.e. set and $A >_K \emptyset$ then A is the disjoint union of two c.e. sets A_0, A_1 such that $A_0|_K A_1$ and $A_0, A_1 <_K A$.

In particular, if a c.e. set has nontrivial initial segment complexity then it can be split into two c.e. sets with strictly less initial segment complexity. This fact also holds for the plain complexity C in place of K. Note that the assumption $A >_K \emptyset$ is stronger than $A >_C \emptyset$ which is merely another way to say that A is noncomputable. For various combinations of (6.1) with other splitting theorems (like the classic Sacks splitting theorem) we refer to [Bar11a, Section 5] and [Ste11, Chapter 2].

We wish to combine (6.1) with cone avoidance. First, we demonstrate a cone avoidance argument in the K degrees in isolation, in terms of Π^0_1 classes. A tree is a computable function from strings to strings which preserves the compatibility and incompatibility relations. A real X is a path through a tree T if all of its initial segments belong to the downward closure of the image of T under the prefix relation. We denote the set of infinite paths through a tree T by $[T]$.

Theorem 6.1 (Cone avoidance for $\leq_K$ and Π^0_1 classes)**.** *If A is Δ^0_2 and $A \not\leq_K \emptyset$ then there exists a Π^0_1 class of reals X such that $X \not\leq_K \emptyset$ and $A \not\leq_K X$.*

Proof. The task of avoiding K-trivial members when constructing a Π^0_1 class has been extensively discussed in [KS07, BLS08b, BV11, BB10]. In order to focus on the cone avoidance ideas we only prove a simpler version of Theorem 6.1 which merely requires the Π^0_1 class to be perfect. The combination of this construction with extra requirements guaranteeing that the Π^0_1 class does not have K-trivial members is along the lines of the argument discussed in [KS07]. In particular, it is not as simple as the case in [BV11] but it is much simpler than the argument discussed in [BLS08b] (which is in turn simpler than the one in [BB10]).

Let $A[s]$ be a computable approximation to A. We approximate a perfect Π^0_1 tree $T : 2^{<\omega} \to 2^{<\omega}$ and ensure that $A \not\leq_K X$ for all paths X through T. Let T_σ denote the image of σ under T and let $T_\sigma[s]$ denote its approximation at (the end of) stage s. We will satisfy the following requirements.

$$R_e: \ \forall \sigma \in 2^e \ \exists n \leq |T_\sigma| \ [K(A \restriction_n) > K(T_\sigma \restriction_n) + e].$$

We define the length of agreement ℓ_σ for each string σ of length e in order to monitor the satisfaction of R_e with respect to T_σ. Let $\ell_\sigma[s]$ be the largest number $n \leq s$ such that $K(A \restriction_j)[s] \leq K(X \restriction_j)[s] + e$ for some extension X of $T_\sigma[s]$ in $[T[s]]$ and all $j < n$. We view the approximations to each T_σ as a movable marker. We say that T_σ requires attention at stage $s+1$ if $\ell_\sigma[s]$ is larger than $|T_\sigma[s]|$.

At stage 0 we let $T_\sigma[0] = \sigma$ for all σ. At stage $s+1$ let σ be the least string of length at most s that requires attention (if there is no such string, do nothing). Let $T_\sigma[s+1] = T_\rho[s]$ where ρ is the least extension of σ such that $|T_\rho[s]|$ is larger than $\ell_\sigma[s]$ and $K(A \restriction_j)[s] \leq K(T_\rho[s] \restriction_j)[s] + e$ for all $j < |T_\rho[s]|$. Moreover for each string η let $T_{\sigma * \eta}[s+1] = T_{\rho * \eta}[s]$.

We start the verification of the construction by noting that the reals that are paths through all trees $T[s]$ form a Π^0_1 class. Moreover since the nodes T_σ can only move to existing nodes in $T[s]$ at stage $s+1$ in a monotone fashion, it follows that $[T[s+1]] \subseteq [T[s]]$ for all stages s. Hence if we show that each node T_σ reaches a limit then the paths through the limit tree T form a perfect Π^0_1 class.

It remains to show that $T[s]$ reaches a limit T such that R_e is satisfied for all e. We show this by induction. Assume that by stage s_0 all nodes T_σ with $|\sigma| \leq e$ have reached a (finite) limit such that R_i is satisfied for each $i \leq e$. Fix a string σ of length $e+1$. If T_σ is redefined infinitely often, then $T_\sigma[s]$ converges to a computable real X such that $K(A \restriction_n) \leq K(X \restriction_n) + e + 1$ for all n. This is a contradiction since A is not K-trivial. Hence $T_\sigma[s]$ reaches a finite limit T_σ such that $K(A \restriction_n) > K(T_\sigma \restriction_n) + e + 1$ for some $n \leq |T_\sigma|$. Hence R_{e+1} is satisfied. This completes the induction step and the proof. $\square$

The methods of construction behind (6.1) (from [Bar11a, Section 5] and [Ste11, Chapter 2]) and Theorem 6.1 can be combined in a proof of the following enhanced splitting theorem.

Theorem 6.2 (Splitting with cone avoidance for $\leq_K$)**.** *Let A be a c.e. set such that $A \not\leq_K \emptyset$ and let X be a Δ^0_2 set such that $X \not\leq_K \emptyset$. Then A is the union of two disjoint c.e. sets A_0, A_1 such that $A_0|_K A_1$, $A_i <_K A$ and $X \not\leq_K A_i$ for $i = 0, 1$.*

Proof. In the course of enumerating the elements of A into A_0 and A_1 we satisfy the following requirement for $e \in \mathbb{N}$ and $i = 0, 1$.

$$R_{\langle e,i\rangle} : \exists n \ \big[K(A_{1-i} \restriction_n) > K(A_i \restriction_n) + e\big].$$

Thus we ensure that $A_0 \not\leq_K A_1$ and $A_1 \not\leq_K A_0$. By [Bar11a, Lemma 5.1] we also get $A_0, A_1 <_K A$. Define the *length of agreement* $l(e,i)[s]$ of $R_{\langle e,i\rangle}$ at stage s to be the largest $n \leq s$ such that $\forall j < n \ (K(A_{1-i} \restriction_j)[s] \leq K(A_i \restriction_j)[s] + e)$. We also need to meet the following requirement for all $e \in \mathbb{N}$ and $i = 0, 1$.

$$N_{\langle e,i\rangle} : \ \exists n \ [K(X \restriction_n) > K(A_i \restriction_n) + e].$$

We define the length of agreement $m(e,i)$ in order to monitor the satisfaction of $R_{\langle e,i\rangle}$. Let $m(e,i)[s]$ be the largest number $n \leq s$ with the property that for all $j < n$, $K(X \restriction_j)[s] \leq K(A_i \restriction_j)[s] + e$. Let the restraint imposed by $R_{\langle e,i\rangle}, N_{\langle e,i\rangle}$ on the enumeration of A_i at stage $s+1$ be given by

$$r(e,i)[s] = \max_{t \leq s}\{l(e,i)[t], m(e,i)[t], e\}.$$

Note that by definition the restraint is non-decreasing in the stages s. Let $A_i[0] = \emptyset$ for $i = 0, 1$ and without loss of generality assume that at each stage exactly one element is enumerated in A.

Construction. If $x \in A[s+1] - A[s]$ consider the least $\langle e,i\rangle$ such that $x \leq r(e,i)[s]$ and enumerate x into A_{1-i}.

Verification. By induction we show that for all $\langle e,i\rangle$ requirements $R_{\langle e,i\rangle}, N_{\langle e,i\rangle}$ are met and $r(e,i)$ reaches a limit. Suppose that there is a stage s_0 such that for all $\langle e',i'\rangle < \langle e,i\rangle$ the requirements $R_{\langle e',i'\rangle}, N_{\langle e',i'\rangle}$ are met and $r(e',i')[s]$ remains constant for all $s \geq s_0$. Without loss of generality we may assume that s_0 is large enough so that all numbers enumerated in A after s_0 are larger than the final values of $r(e',i')$, $\langle e',i'\rangle < \langle e,i\rangle$. By the choice of s_0, after that stage all numbers enumerated into A_i will be larger than the current value of $r(e,i)$.

For a contradiction, suppose that $R_{\langle e,i\rangle}$ is not met. Then the length of agreement $l(e,i)$ and the restraint $r(e,i)$ tend to infinity. Since $r(e,i)[s]$ is nondecreasing in s it follows that A_i is computable; hence K-trivial. Since $R_{\langle e,i\rangle}$ is not met, it follows that A_{1-i} is K-trivial. Since A is the disjoint union of A_0 and A_1 we have $A \equiv_T A_0 \oplus A_1$. Then A is K-trivial, given that K-triviality is closed under the join operator. This contradicts the assumption about A. Hence $R_{\langle e,i\rangle}$ is met.

For a second contradiction, assume that $N_{\langle e,i\rangle}$ is not met. Then the length of agreement $m(e,i)$ and the restraint $r(e,i)$ tend to infinity. Since $r(e,i)[s]$ is nondecreasing in s it follows that A_i is computable, hence K-trivial. Since $N_{\langle e,i\rangle}$ is not met, it follows that X is K-trivial. This contradicts our assumption about X. Hence $N_{\langle e,i\rangle}$ is met.

To conclude the induction step (and the proof) it suffices to show that $r(e,i)[s]$ reaches a limit as s tends to infinity. But this is a direct consequence of its definition and the fact that $R_{\langle e,i\rangle}, N_{\langle e,i\rangle}$ are met. $\square$

The proof of Theorem 6.2 can be written for $\leq_C$ instead of $\leq_K$ with no essential changes. This trivial modification gives the following analogue.

Theorem 6.3 (Splitting with cone avoidance for $\leq_C$)**.** *Let A be a c.e. set such that $A \not\leq_C \emptyset$ and let X be a Δ^0_2 set such that $X \not\leq_C \emptyset$. Then A is the union of two c.e. sets A_0, A_1 such that $A_0 \cap A_1 = \emptyset$, $A_0|_C A_1$ and $A_i <_C A$, $X \not\leq_C A_i$ for $i = 0, 1$.*

The above results establish the existence of c.e. splittings of strictly lesser complexity. We would like to know if it is always possible to split a c.e. set into two parts of the same initial segment complexity. It is not hard to see that there exist specially crafted c.e. sets that can be split into two c.e. sets of the same complexity. For example, given any c.e. set A the set $A \oplus A$ has this property. We do not know whether every c.e. set A is the union of two c.e. sets B, C such that $B \cap C = \emptyset$ and $A \equiv_K B \equiv_K \equiv C$. The same question has been studied in the context of Turing degrees. Lachlan [Lac67] showed that there exists a c.e. set that cannot be split into two c.e. sets of the same Turing degree. The c.e. sets which can be split into two c.e. sets of the same Turing degree are called mitotic and were studied in [Lad73b, Lad73a] and [DS89].

A related topic of interest concerns the relationship of $\leq_K, \leq_C$ with $\leq_T$ in the context of c.e. sets. This was studied in [MS07] more generally, but the arguments used there provide facts about the c.e. case as well. For example, the K-degree of the halting set does not contain all the c.e. members of any c.e. Turing degree. Moreover the same holds for the C-degree of the halting set. Further results (along the lines of the theorems in [MS07]) relating $\leq_K, \leq_C$ with $\leq_T$ in the case of the c.e. sets may be obtained by a more careful examination of the methods in [MS07].

6.2. **Compression with c.e. oracles and splittings.** We have already stressed that the LK reducibility (measuring the compressing power of oracles) is quite different to K reducibility (measuring the initial segment complexity of reals) even in the context of c.e. sets. However the two measures are quite related on the random sequences. In fact, it was shown in [MY08] that if X, Y are random sequences then $X \leq_{LK} Y$ if and only if $Y \leq_K X$. In other words, a random sequence X can compress finite programs at least as efficiently as another random sequence Y exactly if its initial segments are at most as complex as those of Y.

A similarity between the two measures also occurs in the case of c.e. oracles, with respect to the various splitting properties that where discussed in Section 6.1. Such properties of LK have mostly been studied in terms of the related reducibility $\leq_{LR}$. As we discussed in Section 4 it coincides with $\leq_{LK}$. Morever it is rather straightforward to translate an argument concerning $\leq_{LR}$ to the analogous argument about $\leq_{LK}$, by replacing effectively open sets with prefix-free machines. The argument in [BL11] may be useful as a guide for such a translation. In the following we only use $\leq_{LK}$, although most of the original proofs in the literature of the results that we discuss refer to $\leq_{LR}$. We start with the following analogue of (6.1) that was proved in [BLS08a] (also see [Bar10b, Footnote 7] and [Ste11, Chapter 2]).

(6.2) If A is c.e. set and $A >_{LK} \emptyset$ then A is the disjoint union of two c.e. sets A_0, A_1 such that $A_0|_{LK} A_1$ and $A_0, A_1 <_{LK} A$.

Cone avoidance works for $\leq_{LK}$ much in the same way as it does for $\leq_K$.

(6.3) Theorems 6.1 and 6.2 hold for $\leq_{LK}$ in place of $\leq_K$.

The proof of (6.3) does not involve any new ideas, other than the ones presented in [BLS08a] and in Section 6.1. For this reason it is left to the motivated reader. Similar cone avoidance arguments have been used in [Mor11].

The study of c.e. sets that cannot be split in the same LK degree was the topic of [BM09, Section 3].

(6.4) There exists a c.e. set A which cannot be split into two c.e. sets B, C such that $A \equiv_{LK} B \equiv_{LK} C$.

Moreover the set A that was constructed was shown to be Turing complete. Recall that the analogue of (6.4) for $\leq_K$ is an open question. The following characterization of triviality with respect to $\leq_{LK}$ was another result from [BM09, Section 3].

(6.5) A c.e. set is K-trivial if and only if it computes a set which cannot be split into two c.e. sets of the same LK degree.

The analogue of (6.5) for the case of Turing degrees (i.e. that every noncomputable c.e. set computes a non-mitotic set) was shown in [Lad73b]. An interesting open question on this topic is the existence of c.e. LK degrees in which all c.e. sets can be split into two c.e. sets of the same LK degree. The analogue of this question for the Turing degrees was answered positively in [Lad73a] and was further studied in [DS89].

References

[ASDFM11] Klaus Ambos-Spies, Decheng Ding, Yun Fan, and Wolfgang Merkle. Maximal pairs of computably enumerable sets in the computable-Lipschitz degrees. Preprint, 2011.

[Bar68] Janis Barzdins. Complexity of programs to determine whether natural numbers not greater than n belong to a recursively enumerable set. *Soviet Mathematics Doklady*, 9:1251–1254, 1968.

[Bar05] George Barmpalias. Computably enumerable sets in the Solovay and the strong weak truth table degrees. In S. Barry Cooper, Benedikt Löwe, and Leen Torenvliet, editors, *CiE*, volume 3526 of *Lecture Notes in Computer Science*, pages 8–17. Springer, 2005.

[Bar10a] George Barmpalias. Compactness arguments with effectively closed sets for the study of relative randomness. *J. Logic Computation*, 2010. In press.

[Bar10b] George Barmpalias. Elementary differences between the degrees of unsolvability and the degrees of compressibility. *Ann. Pure Appl. Logic*, 161(7):923–934, 2010.

[Bar10c] George Barmpalias. Relative randomness and cardinality. *Notre Dame J. Formal Logic*, 51(2), 2010.

[Bar11a] George Barmpalias. On strings with trivial Kolmogorov complexity. *Int J Software Informatics*, 5(4):609–623, 2011.

[Bar11b] George Barmpalias. Universal computably enumerable sets and initial segment prefix-free complexity. Submitted, 2011.

[BB10] Martijn Baartse and George Barmpalias. On the gap between trivial and nontrivial initial segment prefix-free complexity. Submitted, 2010.

[BL06] George Barmpalias and Andrew E. M. Lewis. The ibT degrees of computably enumerable sets are not dense. *Ann. Pure Appl. Logic*, 141(1-2):51–60, 2006.

[BL11] George Barmpalias and Andrew E.M. Lewis. Chaitin's halting probability and the compression of strings using oracles. *Proceedings of the Royal Society A*, 467:2912–2926, 2011.

[BLS08a] George Barmpalias, Andrew E. M. Lewis, and Mariya Soskova. Randomness, Lowness and Degrees. *J. of Symbolic Logic*, 73(2):559–577, 2008.

[BLS08b] George Barmpalias, Andrew E. M. Lewis, and Frank Stephan. Π^0_1 classes, LR degrees and Turing degrees. *Ann. Pure Appl. Logic*, 156(1):21–38, 2008.

[BM09] George Barmpalias and Anthony Morphett. Non-cupping, measure and computably enumerable splittings. *Mathematical Structures in Computer Science*, 19(1):25–43, 2009.

[BV11] George Barmpalias and C.S. Vlek. Kolmogorov complexity of initial segments of sequences and arithmetical definability. *Theoretical Computer Science*, 412(41):5656–5667, 2011.

[Cha75] Gregory J. Chaitin. A theory of program size formally identical to information theory. *J. Assoc. Comput. Mach.*, 22:329–340, 1975.

[Cha76a] G. Chaitin. Information-theoretical characterizations of recursive infinite strings. *Theoretical Computer Science*, 2:45–48, 1976.

[Cha76b] Gregory J. Chaitin. Algorithmic entropy of sets. *Computers & Mathematics with Applications*, 2:233–245, 1976.

[Cha77] Gregory J. Chaitin. Algorithmic information theory. *IBM Journal of Research and Development*, 21(4):350–359, 1977.

[CZ10] Cristian S. Calude and Marius Zimand. Algorithmically independent sequences. *Inf. Comput.*, 208(3):292–308, 2010.

[Day10] Adam R. Day. The computable Lipschitz degrees of computably enumerable sets are not dense. *Ann. Pure Appl. Logic*, 161(12):1588–1602, 2010.

[DH10] Rod Downey and Denis Hirshfeldt. *Algorithmic Randomness and Complexity.* Springer, 2010.

[DHL04] Rod G. Downey, Denis R. Hirschfeldt, and Geoff LaForte. Randomness and reducibility. *J. Comput. System Sci.*, 68(1):96–114, 2004.

[DHNS03] Rod G. Downey, Denis R. Hirschfeldt, André Nies, and Frank Stephan. Trivial reals. In *Proceedings of the 7th and 8th Asian Logic Conferences*, pages 103–131, Singapore, 2003. Singapore Univ. Press.

[Dia11] David Diamondstone. Low upper bounds in the *lr* degrees. *Ann. Pure Appl. Logic*, 2011. In press.

[dLMSS55] K. de Leeuw, E. F. Moore, C. E. Shannon, and N. Shapiro. Computability by probabilistic machines. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 183–212. Princeton University Press, Princeton, NJ, 1955.

[DS89] R. G. Downey and Theodore A. Slaman. Completely mitotic r.e. degrees. *Ann. Pure Appl. Logic*, 41:119–152, 1989.

[DS90] R. G. Downey and M. Stob. Automorphisms and splittings of recursively enumerable sets. In *Proceedings of the Fourth Asian Logic Conference*, pages 75–87. CSSK Centre, 1990.

[DS93] Rod Downey and Michael Stob. Splitting theorems in recursion theory. *Ann. Pure Appl. Logic*, 65(1):106 pp, 1993.

[HKM09] Rupert Hölzl, Thorsten Kräling, and Wolfgang Merkle. Time-bounded kolmogorov complexity and solovay functions. In *Mathematical foundations of computer science, 2009*, volume 5734 of *Lecture Notes in Comput. Sci.*, pages 392–402. Springer, 2009.

[HT08] D. R. Hirschfeldt and S. A. Terwijn. Limit computability and constructive measure. In *Computational Prospects of Infinity, Part II*, volume 15 of *Lecture Notes Series*, pages 131–141, National University of Singapore, 2008. Institute of Mathematical Sciences.

[KHMS12] Bjørn Kjos-Hanssen, Joseph S. Miller, and Reed Solomon. Lowness notions, measure and domination. *J. Lond. Math. Soc.*, 2012. In press.

[Kol65] A. N. Kolmogorov. Three approaches to the definition of the concept "quantity of information". *Problemy Peredači Informacii*, 1(vyp. 1):3–11, 1965.

[KS07] A. Kučera and T. Slaman. Turing incomparability in Scott sets. *Proc. Amer. Math. Soc.*, 135(11):3723–3731, 2007.

[Kum96] Martin Kummer. Kolmogorov complexity and instance complexity of recursively enumerable sets. *SIAM J. Comput.*, 25(6):1123–1143, 1996.

[Lac67] Alistair H. Lachlan. The priority method I. *Z. Math. Logik Grundlag. Math.*, 13:1–10, 1967.

[Lad73a] Richard Ladner. A complete mitotic recursively enumerable degree. *Trans. Amer. Math. Soc.*, 184:497–507, 1973.

[Lad73b] Richard Ladner. Mitotic recursively enumerable sets. *J. Symbolic Logic*, 38:199–211, 1973.

[Lev73] L. A. Levin. The concept of a random sequence. *Dokl. Akad. Nauk SSSR*, 212:548–550, 1973.

[Lev74] Leonid A. Levin. Laws of information conservation (nongrowth) and aspects of the foundation of probability theory. *Problems of Information Transmission*, 10(1):206–210, 1974.

[Lev84] Leonid A. Levin. Randomness conservation inequalities; information and independence in mathematical theories. *Information and Control*, 61(1):15–37, 1984.

[Lev02] Leonid A. Levin. Forbidden information. In *FOCS*, pages 761–. IEEE Computer Society, 2002.

[Lov69] Donald W. Loveland. A variant of the kolmogorov concept of complexity. *Information and Control*, 15(6):510–526, 1969.

[LV08] Ming Li and Paul Vitányi. *An introduction to Kolmogorov complexity and its applications*. Graduate Texts in Computer Science. Springer-Verlag, New York, third edition, 2008.

[Mil10] Joseph S. Miller. The K-degrees, low for K degrees, and weakly low for K sets. *Notre Dame J. Formal Logic*, 50(4):381391, 2010.

[ML66] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[MN06] Joseph S. Miller and André Nies. Randomness and computability: open questions. *Bull. Symbolic Logic*, 12(3):390–410, 2006.

[Mor11] Anthony Morphett. Incompleteness, approximation and relative randomness. Preprint, 2011.

[MS07] Wolfgang Merkle and Frank Stephan. On C-degrees, H-degrees and T-degrees. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC 2007), San Diego, USA, 12–16 June 2007*, pages 60–69, Los Alamitos, CA, USA, 2007. IEEE Computer Society.

[MY08] Joseph S. Miller and Liang Yu. On initial segment complexity and degrees of randomness. *Trans. Amer. Math. Soc.*, 360(6):31933210, 2008.

[MY10] Joseph S. Miller and Liang Yu. Oscillation in the initial segment complexity of random reals. *Adv. Math.*, 2010. In press.

[Nie05] André Nies. Lowness properties and randomness. *Adv. Math.*, 197(1):274–305, 2005.

[Sac63] G. E. Sacks. On the degrees less than $\mathbf{0}'$. *Ann. of Math. (2)*, 77:211–231, 1963.

[Soa87] Robert I. Soare. *Recursively enumerable sets and degrees*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1987. A study of computable functions and computably generated sets.

[Sol75] R. Solovay. Handwritten manuscript related to Chaitin's work. IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 215 pages, 1975.

[Sol77] R. M. Solovay. On random r.e. sets. In *In A. I. Arruda, N. C. A. da Costa, and R. Chuaqui, editors, Non-Classical Logics, Model Theory and Computability. Proceedings of the Third Latin-American Symposium on Mathematical Logic, Campinas, July 1117, 1976*, volume 89 of *Studies in Logic and the Foundations of Mathematics*, pages 283–307, Amsterdam, 1977. North Holland.

[Ste11] Tom Sterkenburg. *Sequences with trivial initial segment complexity.* MSc Dissertation, University of Amsterdam, February 2011.

[Ver07] Nikolai K. Vereshchagin. Kolmogorov complexity of enumerating finite sets. *Inf. Process. Lett.*, 103(1):34–39, 2007.

George Barmpalias: State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190, P.O. Box 8718, China.

E-mail address: barmpalias@gmail.com

URL: http://www.barmpalias.net

Angsheng Li: State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190, P.O. Box 8718, China.

E-mail address: angsheng@ios.ac.cn