The Bulletin of Symbolic Logic
Volume 22, Number 2, June 2016

DEEP Π^0_1 CLASSES

LAURENT BIENVENU AND CHRISTOPHER P. PORTER

Abstract. A set of infinite binary sequences $\mathcal{C} \subseteq 2^{\mathbb{N}}$ is *negligible* if there is no partial probabilistic algorithm that produces an element of this set with positive probability. The study of negligibility is of particular interest in the context of Π^0_1 classes. In this paper, we introduce the notion of depth for Π^0_1 classes, which is a stronger form of negligibility. Whereas a negligible Π^0_1 class $\mathcal{C}$ has the property that one cannot probabilistically compute a member of $\mathcal{C}$ with positive probability, a deep Π^0_1 class $\mathcal{C}$ has the property that one cannot probabilistically compute an *initial segment* of a member of $\mathcal{C}$ with *high* probability. That is, the probability of computing a length n initial segment of a deep Π^0_1 class converges to 0 effectively in n.

We prove a number of basic results about depth, negligibility, and a variant of negligibility that we call *tt*-negligibility. We provide a number of examples of deep Π^0_1 classes that occur naturally in computability theory and algorithmic randomness. We also study deep classes in the context of mass problems, examine the relationship between deep classes and certain lowness notions in algorithmic randomness, and establish a relationship between members of deep classes and the amount of mutual information with Chaitin's Ω.

§1. Introduction.

Much of the work carried out in computability theory since its inception has been concerned with deterministic computation, that is, computational procedures with the property that after any finite number of steps have been carried out, there is one unique step that follows from them. In contrast, the idea of probabilistic computation, which is prominent in computational complexity, rarely comes up in computability theory. One of the reason for this is the early result of De Leeuw, Moore, Shannon and Shapiro [6] (independently proven by Sacks [28]), which states that if an infinite sequence can be generated with positive probability by a probabilistic computation, then it can in fact be generated deterministically. However, there are some *sets* $\mathcal{C}$ of infinite sequences which contain no computable element, but such that a probabilistic computation can produce a member of $\mathcal{C}$ with positive probability. A trivial example is the set $\mathcal{C}$ of noncomputable sequences: it suffices to pick each bit at random (with equal probability for 0 and 1, independently of the other bits), and since there are only countably many computable sequences, the resulting sequence will be noncomputable with probability one.

Received March 8, 2016.
2010 *Mathematics Subject Classification.* 03D32, 68Q30, 03D80.
Key words and phrases. computability theory, algorithmic randomness, Π^0_1 classes, probabilistic computation.

© 2016, Association for Symbolic Logic
1079-8986/16/2202-0005
DOI:10.1017/bsl.2016.9

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

The basic objects of the present investigation are Π^0_1 classes, that is, effectively closed subclasses of $2^{\mathbb{N}}$. Our goal is to study the properties of those Π^0_1 classes whose members cannot be probabilistically obtained with positive probability. For our purposes, the model of probabilistic computation that is most convenient is given by oracle Turing machines equipped with algorithmically random oracles. Thus, for a given Π^0_1 class $\mathcal{C} \subseteq 2^{\mathbb{N}}$, the question

> Can one probabilistically compute some member of $\mathcal{C}$ with positive probability?

amounts to asking

> Does the class of oracles that compute some member of $\mathcal{C}$ have positive Lebesgue measure?

We will discuss two specific types of Π^0_1 classes, *negligible classes* and *deep classes*, focussing in particular on the latter type, which is a special case of the former.

Roughly, a Π^0_1 class $\mathcal{C}$ is *negligible* if the probability of producing a member of $\mathcal{C}$ via any probabilistic algorithm is 0. Rephrased in terms of the model of probabilistic computation we use here, $\mathcal{C}$ is negligible if the collection $\bigcup_{e\in\mathbb{N}} \Phi_e^{-1}(\mathcal{C})$ has Lebesgue measure 0, where $(\Phi_e)_{e\in\mathbb{N}}$ is an effective enumeration of all Turing functionals. The first example of a negligible class appeared in [17], in which it was proved that the Π^0_1 class of consistent completions of Peano arithmetic is negligible. Negligibility was further studied in [21], [23], and [33] (the term *negligible* is explicitly introduced in the latter reference).

Depth can be viewed as a strong form of negligibility. A deep Π^0_1 class $\mathcal{C}$ has the property that producing an initial segment of some member of $\mathcal{C}$ is in some sense maximally difficult: the probability of obtaining such an initial segment not only converges to 0 as we consider longer initial segments of members of $\mathcal{C}$, but this convergence is fast, as we can effectively bound the rate of convergence. That initial segments of members of deep classes are so difficult to produce via probabilistic computation reflects the fact that members of deep classes are highly structured: initial segments of these sequences cannot be successfully produced by any combination of Turing machine and random oracle.

Although the notion of depth is isolated for the first time in the present study, it is implicitly used in the work of both Levin [22] and Stephan [32]. In fact, in each of these papers one can extract a proof that the consistent completions of Peano arithmetic form a deep class, a result that we reprove in Section 7.

The primary goals of this paper are (1) to prove a number of basic results about deep Π^0_1 classes, (2) to determine the exact relationship between negligibility and depth (and a related notion we call *tt-negligibility*), and (3) to provide a number of examples of deep Π^0_1 classes that occur naturally in computability theory and algorithmic randomness.

The outline of this paper is as follows: In Section 2, we provide the technical background for the remainder of the study. The notions of negligibility and *tt*-negligibility are introduced and compared in Section 3, while the

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

notions of depth and *tt*-depth are introduced in Section 4. In Section 4, we also separate depth from negligibility, show the equivalence between *tt*-depth and *tt*-negligibility, and prove several basic facts about these various kinds of classes. Section 5 contains a brief discussion of the levels of randomness that guarantee that a random sequence cannot compute any member of any *tt*-negligible, negligible, and deep class. We consider the notions of depth and negligibility in the context of Medvedev and Muchnik reducibility in Section 6. In Section 7 we present six examples of families of deep classes, each of which is defined in terms of some well-studied notion from computability theory and algorithmic randomness: completions of arithmetic, shift-complex sequences, diagonally noncomputable functions, compression functions, finite sets of maximally incompressible strings, and dominating martingales related to the class $\mathrm{High}(\mathsf{CR}, \mathsf{MLR})$. In Section 8, we establish connections between depth, *tt*-depth, and several lowness notions. Lastly, in Section 9, we apply the notion of mutual information to deep classes, generalizing a result of Levin's from [22].

§2. Background.

2.1. Notation. Let us fix some notation and terminology. We denote by $2^{\mathbb{N}}$ the set of infinite binary sequences (which we often refer to as "sequences"), also known as Cantor space. We denote the set of finite strings by $2^{<\mathbb{N}}$ and the empty string by Λ. $\mathbb{Q}_2$ is the set of dyadic rationals, i.e., multiples of a negative power of 2. Given $X \in 2^{\mathbb{N}}$ and an integer n, $X\upharpoonright n$ is the string that consists of the first n bits of X, and $X(n)$ is the $(n+1)^{\text{st}}$ bit of X (so that $X(0)$ is the first bit of X). For integers n and k, $X\upharpoonright[n, n+k]$ denotes the subword σ of X of length $k+1$ such that for $i \leq k$, $\sigma(i) = X(i+n)$.

If σ is a string and $x \in 2^{<\mathbb{N}} \cup 2^{\mathbb{N}}$, then $\sigma \preceq x$ means that σ is a prefix of x. A prefix-free set of strings is a set of strings such that none of its elements is a strict prefix of another one. For strings $\sigma, \tau \in 2^{<\mathbb{N}}$, $\sigma^\frown\tau$ denotes the string obtained by concatenating σ and τ. Similarly, for $\sigma \in 2^{<\mathbb{N}}$ and $X \in 2^{\mathbb{N}}$, $\sigma^\frown X$ is the sequence obtained by concatenating σ and X. For $X, Y \in 2^{\mathbb{N}}$, $X \oplus Y \in 2^{\mathbb{N}}$ satisfies $X \oplus Y(2n) = X(n)$ and $X \oplus Y(2n+1) = Y(n)$.

Given a string σ, the *cylinder* $[\![\sigma]\!]$ is the set of elements of $2^{\mathbb{N}}$ having σ as a prefix. Moreover, given $S \subseteq 2^{<\mathbb{N}}$, $[\![S]\!]$ is defined to be the set $\bigcup_{\sigma\in S}[\![\sigma]\!]$. When we refer to the topology of the Cantor space, we implicitly mean the product topology, i.e., the topology whose open sets are exactly those of type $[\![S]\!]$ for some S. For this topology, some sets are both open and closed (clopen): these are the sets of type $[\![S]\!]$ when S is finite.

A *tree* is a set of strings that is closed downwards under the prefix relation. A *path* through a tree T is a member of $2^{\mathbb{N}}$ all of whose prefixes are in T. The set of paths of a tree T is denoted by $[T]$. The n^{th} *level of a tree* T, denoted T_n, is the set of members of T of length n. An effectively open set, also called a Σ^0_1 class, is a set of type $[\![S]\!]$ for some c.e. set of strings S. An effectively closed set, or Π^0_1 class, is the complement of some effectively open set. It is well known that a class is Π^0_1 if and only if it is of the form $[T]$ for

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

some co-c.e. (or computable) tree T. Given a Π^0_1 class $\mathcal{C}$, its *canonical co-c.e. tree* is the tree $T = \{\sigma : [\![\sigma]\!] \cap \mathcal{C} \neq \emptyset\}$. The arithmetic hierarchy is defined inductively: a set is Σ^0_{n+1} if it is a uniform union of Π^0_n sets, and a set is Π^0_{n+1} if it is a uniform intersection of Σ^0_n sets.

A real number $r \in [0,1]$ is *left-c.e.* (resp. *right-c.e.*) if it is the limit of a nondecreasing (resp. nonincreasing) computable sequence of rationals.

An *order function* is a function $h : \mathbb{N} \to \mathbb{N}$ that is nondecreasing and unbounded. Given an order function h, the *inverse of* h, denoted h^{-1} is the order function defined as follows: For all k, $h^{-1}(k)$ is the smallest n such that $h(n) \geq k$. Note that h^{-1} is computable if h is computable.

We adopt standard computability notation: $\leq_T$ denotes Turing reducibility, $\leq_{tt}$ denotes *tt*-reducibility, A' is the Turing jump of A (and $\emptyset'$ denotes the jump of the zero sequence). Throughout the paper, when an object A has a computable enumeration or limit approximation, we use the notation $A[s]$ to denote the approximation of the object at stage s. Moreover, if A contains several expressions that have a computable approximation, the notation $A[s]$ means that all of these expressions are approximated up to stage s. For example, if T is a tree and f is a function, both of which have computable approximation, $T_{f(n)}[s]$ is equal to $(T[s])_{f(n)[s]}$.

All logarithms will be taken with respect to the base 2. Finally, we adopt the following asymptotic notation. For two functions $f, g : \mathbb{N} \to \mathbb{N}$, we sometimes write $f \leq^+ g$ to abbreviate $f \leq g + O(1)$ and $f \leq^\times g$ to abbreviate $f = O(g)$.

2.2. Measures on $2^{\mathbb{N}}$. Recall that by Caratheodory's Theorem, a measure μ on $2^{\mathbb{N}}$ is uniquely determined by specifying the values of μ on the basic open sets of $2^{\mathbb{N}}$, where $\mu([\![\sigma]\!]) = \mu([\![\sigma 0]\!]) + \mu([\![\sigma 1]\!])$ for every $\sigma \in 2^{<\mathbb{N}}$. If we further require that $\mu(2^{\mathbb{N}}) = 1$, then μ is a *probability measure*. Hereafter, we will write $\mu([\![\sigma]\!])$ as $\mu(\sigma)$. In addition, for a set $S \subseteq 2^{<\mathbb{N}}$, we will write $\mu(S)$ as shorthand for $\mu([\![S]\!])$. In the case that S is prefix-free, we will have $\mu(S) = \sum_{\sigma \in 2^{<\mathbb{N}}} \mu(\sigma)$.

The *uniform (or Lebesgue) measure* λ is the unique Borel measure such that $\lambda(\sigma) = 2^{-|\sigma|}$ for all strings σ. A measure μ on $2^{\mathbb{N}}$ is *computable* if $\sigma \mapsto \mu(\sigma)$ is computable as a real-valued function, i.e., if there is a computable function $\tilde{\mu} : 2^{<\mathbb{N}} \times \mathbb{N} \to \mathbb{Q}_2$ such that

$$|\mu(\sigma) - \tilde{\mu}(\sigma, i)| \leq 2^{-i}$$

for every $\sigma \in 2^{<\mathbb{N}}$ and $i \in \mathbb{N}$.

One family of examples of computable measures is given by the collection of Dirac measures concentrated on some computable point. That is, if $X \in 2^{\mathbb{N}}$ is a computable sequence, then the *Dirac measure concentrated on* X, denoted δ_X, is defined as follows:

$$\delta_X(\sigma) = \begin{cases} 1 & \text{if } \sigma \prec X, \\ 0 & \text{if } \sigma \not\prec X. \end{cases}$$

More generally, for a measure μ, we say that $X \in 2^{\mathbb{N}}$ is an *atom of* μ or a μ*-atom*, denoted $X \in \mathrm{Atom}_\mu$, if $\mu(\{X\}) > 0$. Kautz proved the following:

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

LEMMA 2.1 (Kautz [18]). *$X \in 2^{\mathbb{N}}$ is computable if and only if X is an atom of some computable measure.*

There is a close connection between computable measures and a certain class of Turing functionals. Recall that a *Turing functional* $\Phi :\subseteq 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ may be defined from a c.e. set S_Φ of pairs of strings (σ, τ) such that if $(\sigma, \tau), (\sigma', \tau') \in S_\Phi$ and $\sigma \preceq \sigma'$, then $\tau \preceq \tau'$ or $\tau' \preceq \tau$. For a Turing functional Φ, we will always include $(\Lambda, \Lambda) \in S_\Phi$.

For each $\sigma \in 2^{<\mathbb{N}}$, we define Φ^σ to be the maximal string in $\{\tau : (\exists \sigma' \preceq \sigma)((\sigma', \tau) \in S_\Phi)\}$ in the order given by $\preceq$. To obtain a map defined on $2^{\mathbb{N}}$ from the c.e. set of pairs Φ, for each $X \in 2^{\mathbb{N}}$, we let Φ^X be the maximal $y \in 2^{<\mathbb{N}} \cup 2^{\mathbb{N}}$ in the order given by $\preceq$ such that $\Phi^{X\restriction n}$ is a prefix of y for all n. We will thus set $\mathsf{dom}(\Phi) = \{X \in 2^{\mathbb{N}} : \Phi^X \in 2^{\mathbb{N}}\}$. When $\Phi^X \in 2^{\mathbb{N}}$, we will often write Φ^X as $\Phi(X)$ to emphasize the functional Φ as a map from $2^{\mathbb{N}}$ to $2^{\mathbb{N}}$. We also use the notation $\Phi^X \restriction n \downarrow$ to emphasize that Φ^X has length at least n. For $\tau \in 2^{<\mathbb{N}}$ let $\Phi^{-1}(\tau)$ be the set $\{\sigma \in 2^{<\mathbb{N}} : \exists \tau' \succeq \tau\colon (\sigma, \tau') \in S_\Phi\}$. In particular, by our above convention, we have $\Lambda \in \Phi^{-1}(\Lambda)$. Similarly, for $S \subseteq 2^{<\mathbb{N}}$ we define $\Phi^{-1}(S) = \bigcup_{\tau \in S} \Phi^{-1}(\tau)$. When $\mathcal{A}$ is a subset of $2^{\mathbb{N}}$, we denote by $\Phi^{-1}(\mathcal{A})$ the set $\{X \in \mathsf{dom}(\Phi) : \Phi(X) \in \mathcal{A}\}$. Note in particular that $\Phi^{-1}([\![\tau]\!]) = [\![\Phi^{-1}(\tau)]\!] \cap \mathsf{dom}(\Phi)$.

REMARK 2.2. The Turing functionals that induce computable measures are precisely the *almost total* Turing functionals, where a Turing functional Φ is almost total if

$$\lambda(\mathsf{dom}(\Phi)) = 1.$$

Given an almost total Turing functional Φ, the measure induced by Φ, denoted λ_Φ, is defined by

$$\lambda_\Phi(\sigma) = \lambda([\![\Phi^{-1}(\sigma)]\!]) = \lambda(\{X : \Phi^X \succeq \sigma\}).$$

It is not difficult to verify that λ_Φ is a computable measure. Moreover, one can easily show that given a computable probability measure μ, there is some almost total functional Φ such that $\mu = \lambda_\Phi$.

2.3. Left-c.e. semi-measures. We consider semi-measures on both $2^{<\mathbb{N}}$ and $2^{\mathbb{N}}$. In general, semi-measures behave like defective probability measures, as they need not be additive.

A *discrete semi-measure* is a map $m : 2^{<\mathbb{N}} \to [0, 1]$ such that $\sum_{\sigma \in 2^{<\mathbb{N}}} m(\sigma) \leq 1$. Moreover, if S is a set of strings, then $m(S)$ is defined to be $\sum_{\sigma \in S} m(\sigma)$. Henceforth, we will restrict our attention to the class of left-c.e. discrete semi-measures, where a discrete semi-measure m is *left-c.e.* if there is a computable function $\widetilde{m} : 2^{<\mathbb{N}} \times \mathbb{N} \to \mathbb{Q}_2$, nondecreasing in its second argument such that for all σ:

$$\lim_{i \to +\infty} \widetilde{m}(\sigma, i) = m(\sigma).$$

Levin showed that there is a *universal* left-c.e. discrete semi-measure $\mathbf{m}$; that is, for every left-c.e. discrete semi-measure m, there is some constant c such that $m \leq c \cdot \mathbf{m}$.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

This universal discrete semi-measure $\mathbf{m}$ is closely related to the notion of prefix-free Kolmogorov complexity. Recall that $\mathrm{K}(\sigma)$ denotes the prefix-free Kolmogorov complexity of σ, i.e.,

$$\mathrm{K}(\sigma) = \min\{|\tau| : U(\tau) = \sigma\},$$

where U is a universal prefix-free Turing machine. Then by the coding theorem (see [8, Theorem 3.9.4]), we have $\mathrm{K}(\sigma) = -\log \mathbf{m}(\sigma) + O(1)$. In particular, since $\mathrm{K}(n) \leq^+ 2\log(n)$, where $\mathrm{K}(n)$ is simply $\mathrm{K}(1^n)$, it follows that $\mathbf{m}(n) \geq^\times n^{-2}$, a fact we will make use of below.

A *continuous semi-measure* is a map $\rho : 2^{<\mathbb{N}} \to [0, 1]$ satisfying:

(i) $\rho(\Lambda) = 1$ and
(ii) $\rho(\sigma) \geq \rho(\sigma 0) + \rho(\sigma 1)$.

This definition is closely related to the definition of a supermartingale (as defined in [25] or [8]).

If S is a set of strings, then $\rho(S)$ denotes the sum $\sum_{\sigma \in S} \rho(\sigma)$. As in the case of discrete semi-measures, we will restrict our attention to the class of left-c.e. continuous semi-measures (the values of which are effectively approximable from below as defined above in the case of discrete left-c.e. semi-measures).

One particularly important property of left-c.e. continuous semi-measures is their connection to Turing functionals. Just as computable measures are precisely the measures that are induced by almost total Turing functionals (as discussed at the end of the previous section), left-c.e. continuous semi-measures are precisely the continuous semi-measures that are induced by Turing functionals:

THEOREM 2.3 (Levin, Zvonkin [24]).

(i) *For every Turing functional* Φ, *the function* $\lambda_\Phi(\sigma) = \lambda([\![\Phi^{-1}(\sigma)]\!]) = \lambda(\{X : \Phi^X \succeq \sigma\})$ *is a left-c.e. continuous semi-measure.*
(ii) *For every left-c.e. continuous semi-measure* ρ, *there is a Turing functional* Φ *such that* $\rho = \lambda_\Phi$.

As there is a universal left-c.e. discrete semi-measure, so too is there a universal left-c.e. continuous semi-measure. That is, there exists a left-c.e. continuous semi-measure $\mathbf{M}$ such that, for every left-c.e. continuous semi-measure ρ, there exists a $c \in \mathbb{N}$ such that $\rho \leq c \cdot \mathbf{M}$. One way to obtain a universal left-c.e. continuous semi-measure is to effectively list all left-c.e. continuous semi-measures $(\rho_e)_{e \in \mathbb{N}}$ (which can be obtained from an effective list of all Turing functionals by appealing to Theorem 2.3) and set $\mathbf{M} = \sum_{e \in \mathbb{N}} 2^{-e-1} \rho_e$. Alternatively, one can induce it by means of a universal Turing functional: Let $(\Phi_i)_{i \in \mathbb{N}}$ be an effective enumeration of all Turing functionals. Then the functional $\widehat{\Phi}$ such that

$$\widehat{\Phi}(1^e 0 X) = \Phi_e(X)$$

for every $e \in \mathbb{N}$ and $X \in 2^{\mathbb{N}}$ is a universal Turing functional and we can set $\mathbf{M} = \lambda_{\widehat{\Phi}}$. One can readily verify that $\mathbf{M}$ is a universal left-c.e. continuous semi-measure, which is sometimes called *a priori probability* (see for example Gács [11] for the basic properties of $\mathbf{M}$).

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Another feature of continuous semi-measures that we will make use of throughout this study is that there is a canonical measure on $2^{\mathbb{N}}$ that can be obtained from a continuous semi-measure. To motivate the definition of this measure, it is helpful to think of a semi-measure as a network flow through the full binary tree $2^{<\mathbb{N}}$ seen as a directed graph (see, for instance, [23] or [33]). First we give the node at the root of the tree flow equal to 1 (corresponding to the condition $\rho(\Lambda) = 1$). Some amount of this flow at each node σ is passed along to the node corresponding to $\sigma 0$, some is passed along to the node corresponding to $\sigma 1$, and, potentially, some of the flow is lost (corresponding to the condition that $\rho(\sigma) \geq \rho(\sigma 0) + \rho(\sigma 1)$). We obtain a measure $\overline{\rho}$ from ρ if we ignore all of the flow that is lost below a given node and just consider the behavior of the flow that never leaves the network below this node. We will refer to $\overline{\rho}$ as the canonical measure derived from ρ. This can be formalized as follows.

DEFINITION 2.4. Let ρ be a semi-measure. The *canonical measure obtained from* ρ is defined to be

$$\overline{\rho}(\sigma) := \inf_{n \geq |\sigma|} \sum_{\tau \succeq \sigma \ \&\ |\tau| = n} \rho(\tau) = \lim_{n \to \infty} \sum_{\tau \succeq \sigma \ \&\ |\tau| = n} \rho(\tau).$$

Several important facts about these canonical measures are the following:

PROPOSITION 2.5. *Let ρ be a semi-measure and let $\overline{\rho}$ be the canonical measure obtained from ρ.*

(i) *$\overline{\rho}$ is the largest measure μ such that $\mu \leq \rho$.*
(ii) *If $\rho(\sigma) = \lambda(\{X : \Phi^X \succeq \sigma\})$, then $\overline{\rho}(\sigma) = \lambda(\{X \in \mathrm{dom}(\Phi) : \Phi^X \succeq \sigma\})$.*

Thus, in replacing ρ with $\overline{\rho}$, this amounts to restricting the Turing functional Φ that induces ρ to those inputs on which Φ is total. The proof of (i) is straightforward; for a proof of (ii), see [2].

REMARK 2.6. Using Proposition 2.5(ii) and the universality of $\mathbf{M}$, one can readily verify that for every left-c.e. continuous semi-measure ρ, there is some $c \in \mathbb{N}$ such that $\overline{\rho} \leq c \cdot \overline{\mathbf{M}}$. Thus $\overline{\mathbf{M}}$ can be seen as a measure that is universal for the class of canonical measures obtained from some left-c.e. continuous semi-measure (a class that contains all computable measures).

2.4. Notions of algorithmic randomness. The primary notion of algorithmic randomness that we will consider here is Martin-Löf randomness. Although the standard definition is given in terms of the Lebesgue measure, we will consider Martin-Löf randomness with respect to any computable measure.

DEFINITION 2.7. Let μ be a computable measure on $2^{\mathbb{N}}$.

(i) A *μ-Martin-Löf test* is a sequence $\{\mathcal{U}_i\}_{i \in \mathbb{N}}$ of uniformly effectively open subsets of $2^{\mathbb{N}}$ such that for each i,
$$\mu(\mathcal{U}_i) \leq 2^{-i}.$$
(ii) $X \in 2^{\mathbb{N}}$ passes the μ-Martin-Löf test $(\mathcal{U}_i)_{i \in \mathbb{N}}$ if $X \notin \bigcap_{i \in \mathbb{N}} \mathcal{U}_i$.
(iii) $X \in 2^{\mathbb{N}}$ is *μ-Martin-Löf random*, denoted $X \in \mathsf{MLR}_\mu$, if X passes every μ-Martin-Löf test.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

When μ is the uniform (or Lebesgue) measure λ, we will simply write MLR. An important feature of Martin-Löf randomness is the existence of a universal test: For every computable measure μ, there is a single μ-Martin-Löf test $\{\hat{\mathcal{U}}_i\}_{i\in\mathbb{N}}$, having the property that $X \in \mathsf{MLR}_\mu$ if and only if $X \notin \bigcap_{i\in\mathbb{N}} \hat{\mathcal{U}}_i$.

REMARK 2.8. As we saw earlier, some computable measures have atoms, such as the Dirac measure δ_X concentrated on some computable sequence X. Moreover, given a computable measure μ, if X is a μ-atom, it immediately follows that $X \in \mathsf{MLR}_\mu$.

Four additional notions of algorithmic randomness that will be considered in this study are difference randomness, Kurtz randomness, weak 2-randomness, and computable randomness.

DEFINITION 2.9. (i) A *difference test* is a computable sequence $\{(\mathcal{U}_i, \mathcal{V}_i)\}_{i\in\mathbb{N}}$ of pairs of Σ^0_1 classes such that for each i,

$$\lambda(\mathcal{U}_i \setminus \mathcal{V}_i) \leq 2^{-i}.$$

(ii) A sequence $X \in 2^{\mathbb{N}}$ *passes a difference test* $\{(\mathcal{U}_i, \mathcal{V}_i)\}_{i\in\mathbb{N}}$ if $X \notin \bigcap_i (\mathcal{U}_i \setminus \mathcal{V}_i)$.

(iii) $X \in 2^{\mathbb{N}}$ is *difference random* if X passes every difference test. We denote by DiffR the class of difference random reals.

Franklin and Ng proved the following remarkable theorem about difference randomness:

THEOREM 2.10 (Franklin and Ng [10]). *A sequence X is difference random if and only if X is Martin-Löf random and $X \not\geq_T \emptyset'$.*

Recall that a sequence X has PA *degree* if X can compute a consistent completion of Peano arithmetic (hereafter, PA). This is equivalent to requiring that X computes a total function extending a universal partial computable $\{0, 1\}$-valued function, a fact that will be useful in Section 7. A related result is the following:

THEOREM 2.11 (Stephan [32]). *A Martin-Löf random sequence X has* PA *degree if and only if $X \geq_T \emptyset'$.*

It follows from the previous two results that a Martin-Löf random sequence is difference random if and only if it does not have PA degree.

Another approach to defining randomness is to require that a random sequence avoid all null sets that are definable at some fixed level of syntactic complexity. For instance, if we take all Π^0_1 definable null sets or all Π^0_2 definable null sets, we have the following two notions of randomness, first introduced by Kurtz in [20].

DEFINITION 2.12. Let $X \in 2^{\mathbb{N}}$.

(i) X is *Kurtz random* (or *weakly 1-random*) if and only if X is not contained in any Π^0_1 class of Lebesgue measure 0 (equivalently, if and only if it is not contained in any Σ^0_2 class of measure 0).

(ii) X is *weakly 2-random* if and only if X is not contained in any Π^0_2 class of Lebesgue measure 0.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Let KR denoted the collection of Kurtz random sequences and W2R denote the collection of weakly 2-random sequences.

The last definition of randomness we will consider in the study is defined in terms of certain effective betting strategies called *martingales*.

DEFINITION 2.13. (i) A *martingale* is a function $d : 2^{<\mathbb{N}} \to \mathbb{R}^{\geq 0}$ such that for every $\sigma \in 2^{<\mathbb{N}}$,

$$2d(\sigma) = d(\sigma 0) + d(\sigma 1).$$

(ii) A martingale d succeeds on $X \in 2^{\mathbb{N}}$ if

$$\limsup_{n\to\infty} d(X\upharpoonright n) = +\infty.$$

(iii) A sequence $X \in 2^{\mathbb{N}}$ is *computably random* if there is no computable martingale d that succeeds on X.

The collection of computably random sequences will be written as CR. The different notions of randomness discussed in this section form a strict hierarchy. Namely, the following relations hold:

$$\mathsf{W2R} \subsetneq \mathsf{DiffR} \subsetneq \mathsf{MLR} \subsetneq \mathsf{CR} \subsetneq \mathsf{KR}.$$

We note that each of these notions of randomness can be defined with respect to a computable measure μ, similarly to how we defined μ-Martin-Löf randomness. Moreover, each of the above notions can be relativized to an oracle $A \in 2^{\mathbb{N}}$ in a straightforward manner.

§3. Negligibility and *tt*-Negligibility. We are now in a position to define negligibility and *tt*-negligibility. As discussed in the introduction, the intuitive idea behind negligibility is that a set $\mathcal{C} \subseteq 2^{\mathbb{N}}$ is negligible if no member of $\mathcal{C}$ can be produced with positive probability by means of any Turing functional with a random oracle. Similarly, a set $\mathcal{C} \subseteq 2^{\mathbb{N}}$ is *tt*-negligible if no member of $\mathcal{C}$ can be produced with positive probability by means of any *total* Turing functional with a random oracle. However, we will primarily work with the following measure-theoretic definition of these two notions.

DEFINITION 3.1. Let $\mathcal{C} \subseteq 2^{\mathbb{N}}$.

(i) $\mathcal{C}$ is *negligible* if $\overline{\mathbf{M}}(\mathcal{C}) = 0$.
(ii) $\mathcal{C}$ is *tt-negligible* if $\mu(\mathcal{C}) = 0$ for every computable measure μ.

The intuitive description of negligibility and *tt*-negligibility given above is justified by the following proposition. For $\mathcal{C} \subseteq 2^{\mathbb{N}}$, we define $\mathcal{C}^{\leq_T}$ to be the set $\{Y \in 2^{\mathbb{N}} : (\exists X \in \mathcal{C})[X \leq_T Y]\}$. Furthermore, $\mathcal{C}^{\leq_{tt}}$ denotes the set $\{Y \in 2^{\mathbb{N}} : (\exists X \in \mathcal{C})[X \leq_{tt} Y]\}$.

PROPOSITION 3.2. *Let $\mathcal{C} \subseteq 2^{\mathbb{N}}$.*

(i) *$\mathcal{C}$ is negligible if and only if $\lambda(\mathcal{C}^{\leq_T}) = 0$.*
(ii) *$\mathcal{C}$ is tt-negligible if and only if $\lambda(\mathcal{C}^{\leq_{tt}}) = 0$.*

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

PROOF. (i) ($\Rightarrow$) If $\mathcal{C}$ is negligible, then $\overline{\rho}(\mathcal{C}) = 0$ for every left-c.e. continuous semi-measure ρ by Remark 2.6. In particular, by Theorem 2.3(ii) and Proposition 2.5(ii), $\overline{\lambda}_\Phi(\mathcal{C}) = 0$ for every Turing functional Φ and hence

$$\lambda(\mathcal{C}^{\leq_T}) = \sum_{i\in\mathbb{N}} \lambda(\{Y \in 2^{\mathbb{N}} : \Phi_i(Y) \in \mathcal{C}\}) = \sum_{i\in\mathbb{N}} \overline{\lambda}_{\Phi_i}(\mathcal{C}) = 0.$$

($\Leftarrow$) Since $\lambda(\mathcal{C}^{\leq_T}) = 0$, it follows that $\overline{\lambda}_\Phi(\mathcal{C}) = 0$, where Φ is a universal Turing functional. Thus $\overline{\mathbf{M}}(\mathcal{C}) = \overline{\lambda}_\Phi(\mathcal{C}) = 0$.

(ii) ($\Rightarrow$) If $\mathcal{C}$ is *tt*-negligible, then $\lambda_\Phi(\mathcal{C}) = 0$ for every *tt*-functional Φ. The result clearly follows.
($\Leftarrow$) Now suppose that $\mu(\mathcal{C}) > 0$ for some computable measure μ. Then by Remark 2.2, there is some almost total Turing functional Φ such that $\lambda_\Phi = \mu$. Since the domain of a Turing functional is Π^0_2, if $X \notin \mathsf{dom}(\Phi)$, it follows that $X \notin \mathsf{MLR}$ (in fact, X is not even Kurtz random). Indeed, the complement of $\mathsf{dom}(\Phi)$ is a Σ^0_2 class of measure 0, so $X \notin \mathsf{dom}(\Phi)$ implies that X is contained in a Π^0_1 class of measure 0.

Let i be the least such that $\mu(\mathcal{C}) > 2^{-i}$. Then $\lambda(\Phi^{-1}(\mathcal{C}) \cap \hat{\mathcal{U}}_i^c) > 0$, where $\hat{\mathcal{U}}_i$ is the i^{th} level of the universal Martin-Löf test (so that $\lambda(\hat{\mathcal{U}}_i^c) > 1 - 2^{-i}$). Now, since Φ is total on $\hat{\mathcal{U}}_i^c$ (because Φ is defined on all Martin-Löf randoms, and $\hat{\mathcal{U}}_i^c$ contains only Martin-Löf randoms), which is a Π^0_1 class, there is a *tt*-functional Ψ which coincides with Φ on $\hat{\mathcal{U}}_i^c$. This holds because for every Π^0_1 class $\mathcal{P}$ and every Turing functional Φ that is total on $\mathcal{P}$, there is a Turing functional Ψ that is total on $2^{\mathbb{N}}$ and coincides with Φ on $\mathcal{P}$.

Since $\Phi^{-1}(\mathcal{C}) \cap \hat{\mathcal{U}}_i^c = \Psi^{-1}(\mathcal{C}) \cap \hat{\mathcal{U}}_i^c$, it follows that $\lambda(\Psi^{-1}(\mathcal{C}) \cap \hat{\mathcal{U}}_i^c) > 0$ and hence $\lambda(\mathcal{C}^{\leq_u}) > 0$. ⊣

By the following proposition, there is a simple characterization of negligible and *tt*-negligible singletons.

PROPOSITION 3.3. *For $X \in 2^{\mathbb{N}}$, the following are equivalent*:

(i) $\{X\}$ *is negligible.*
(ii) $\{X\}$ *is tt-negligible.*
(iii) X *is noncomputable.*

PROOF. (i)⇒(ii) is immediate. (ii)⇒(iii) follows from Lemma 2.1, which tells us that X is computable if $\mu(\{X\}) > 0$ for some computable measure μ. Lastly, (iii)⇒(i) is the theorem of De Leeuw et al. that was mentioned in the introduction. ⊣

Further, it is clear that for a Σ^0_1 class $\mathcal{S} \subseteq 2^{\mathbb{N}}$, $\mathcal{S}$ is negligible if and only if $\mathcal{S}$ is *tt*-negligible if and only if $\mathcal{S}$ is empty. However, the notions of negligibility and *tt*-negligibility are nontrivial for Π^0_1 classes. For instance, if we let $\mathcal{PA}$ be the Π^0_1 class of consistent completions of PA, then as shown by Jockusch and Soare in [17], $\mathcal{PA}$ is negligible (and thus *tt*-negligible).

Although every negligible Π^0_1 class is *tt*-negligible, the converse does not hold.

THEOREM 3.4. *There exists a tt-negligible Π^0_1 class that is not negligible.*

The proof of Theorem 3.4 draws upon a theorem of Downey, Greenberg and Miller [7]: there exists a non-negligible perfect thin Π^0_1 class, where a Π^0_1

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

class $\mathcal{C}$ is *thin* if for every Π^0_1 subclass $\mathcal{C}'$ of $\mathcal{C}$, there exists a clopen set D such that $\mathcal{C}' = \mathcal{C} \cap D$. We argue that any perfect thin class must be *tt*-negligible, which extends a result of Simpson [29] who proved that every thin class must have Lebesgue measure 0. However, our proof strategy is very different from Simpson's. We first need the following lemma, which is folklore.

LEMMA 3.5. *Let $\mathcal{C}$ be a Π^0_1 class. If for some computable probability measure μ the value of $\mu(\mathcal{C})$ is a positive computable real number, then $\mathcal{C}$ contains a computable member.*

PROOF. Consider the measure $\nu : 2^{<\mathbb{N}} \to [0,1]$ defined by $\nu(\sigma) = \mu(\mathcal{C} \cap [\![\sigma]\!])$. Since $\mathcal{C}$ is Π^0_1, the measure ν is right-c.e. (that is, the values $\nu(\sigma)$ for $\sigma \in 2^{<\mathbb{N}}$ are uniformly right-c.e.). However, for every σ, we have

$$\nu(\sigma) = \nu(\Lambda) - \sum_{\substack{|\tau|=|\sigma| \\ \tau \neq \sigma}} \nu(\tau)$$

and since $\nu(\Lambda)$ is computable (it is equal to $\mu(\mathcal{C})$), this shows that $\nu(\sigma)$ is also left-c.e. uniformly in σ. Therefore ν is computable. It is then easy to computably build by induction a sequence of strings $\sigma_0 \prec \sigma_1 \prec \cdots$ with $|\sigma_i| = i$ and such that $\nu(\sigma_i) \geq \nu(\Lambda) \cdot 4^{-i} > 0$. In particular $\mathcal{C} \cap [\![\sigma_i]\!] \neq \emptyset$. Therefore, the sequence X extending all σ_i is computable and must be an element of $\mathcal{C}$. ⊣

LEMMA 3.6. *If $\mathcal{C}$ is a perfect thin Π^0_1 class, then it is tt-negligible.*

PROOF. First, observe that if a thin class $\mathcal{P}$ contains a computable member X, then since $\{X\}$ is a Π^0_1 subclass of $\mathcal{P}$, there is some $\sigma \in 2^{<\mathbb{N}}$ such that $\{X\} = \mathcal{P} \cap [\![\sigma]\!]$. Thus, X is isolated in $\mathcal{P}$. It thus follows that a perfect thin Π^0_1 class contains no computable members.

Now, for the sake of contradiction, suppose that there exists a computable measure μ such that $\mu(\mathcal{C}) > q$ for some positive rational q. Identifying $2^{\mathbb{N}}$ with the unit interval $[0,1]$ in the usual way, let

$$\alpha = \sup\big\{r \in [0,1] \cap \mathbb{Q} \,:\, \mu(\mathcal{C} \cap [0,r]) < q\big\}.$$

Since $\mu(\mathcal{C} \cap [0,r]) < q$ is a Σ^0_1 predicate in r, α is a left-c.e. real. Moreover, since $\mathcal{C}$ has no computable member, it contains no atom of μ, and thus the function $x \mapsto \mu(\mathcal{C} \cap [0,x])$ is continuous. Therefore, by definition of α, $\mu(\mathcal{C} \cap [0,\alpha)) = \mu(\mathcal{C} \cap [0,\alpha]) = q$.

Consider the class $\mathcal{C} \cap [\alpha,1]$, which is a Π^0_1 subclass of $\mathcal{C}$ since α is left-c.e. By the thinness of $\mathcal{C}$, there exists a clopen set D such that $\mathcal{C} \cap [\alpha,1] = \mathcal{C} \cap D$ and thus such that $\mathcal{C} \cap D^c = \mathcal{C} \cap [0,\alpha)$. From this we deduce $\mu(\mathcal{C} \cap D^c) = q$, which is a positive, computable real number. Since D is clopen, $\mathcal{C} \cap D^c$ is Π^0_1, and so by Lemma 3.5, $\mathcal{C} \cap D^c$ contains a computable member. This contradicts our above observation that $\mathcal{C}$ contains no such members. ⊣

§4. Depth and *tt*-depth. Members of negligible and *tt*-negligible classes are difficult to produce, in the sense that their members cannot be computed from random oracles with positive probability. Given a Π^0_1 class $\mathcal{P}$, we can instead consider the probability of producing an *initial segment* of some

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

member of $\mathcal{P}$. By looking at local versions of negligibility and *tt*-negligibility, given in terms of initial segments of members of the classes in question, we will obtain the notions of depth and *tt*-depth, respectively. Since we are interested in strings that are initial segments of some member of $\mathcal{C}$, the natural representation of $\mathcal{C}$ as a tree is via its co-c.e. tree (recall that it is precisely the tree consisting of the strings σ that are prefixes of some path of $\mathcal{C}$, and by the standard compactness argument, this is a co-c.e. set of strings).

Recall our convention from Section 2.3: if S is a set of strings, then $\mathbf{M}(S)$ and $\mathbf{m}(S)$ are defined to be $\sum_{\sigma\in S}\mathbf{M}(\sigma)$ and $\sum_{\sigma\in S}\mathbf{m}(\sigma)$, respectively. Similarly, if μ is a measure, $\mu(S)$ denotes $\sum_{\sigma\in S}\mu(\sigma)$. Furthermore, recall from Section 2.1 that given a tree $T\subseteq 2^{<\mathbb{N}}$, T_n denotes the set of all members of T of length n. We now define the central notion of this paper, namely the notion of a deep Π^0_1 class. Depth strengthens the notion of negligibility. It is easy to see that a class $\mathcal{C}\subseteq 2^{\mathbb{N}}$ of canonical co-c.e. tree T is negligible if and only if $\mathbf{M}(T_n)$ converges to 0 as n grows without bound. When this convergence to 0 is effective, $\mathcal{C}$ is said to be deep.

Definition 4.1. Let $\mathcal{C}\subseteq 2^{\mathbb{N}}$ be a Π^0_1 class and T its associated co-c.e. tree.

(i) $\mathcal{C}$ is *deep* if there is some computable order function h such that for all n,
$$\mathbf{M}(T_n)\leq 2^{-h(n)}.$$

(ii) $\mathcal{C}$ is *tt-deep* if for every computable measure μ there exists a computable order function h such that for all n,
$$\mu(T_n)\leq 2^{-h(n)}.$$

Our choice of the term "deep" is due to the similarity between the notions of a deep class and that of a logically deep sequence, introduced by Bennett in [1]. Roughly, a sequence X is logically deep if there is no computable function that bounds the amount of time to recover each initial segment of X from its shortest description (measured in terms of Kolmogov complexity). Logically deep sequences are highly structured; in particular, it is difficult to produce initial segments of a deep sequence via probabilistic computation, a fact made precise for finite strings by what Bennett calls the slow growth law (see [1, Theorem 1]). Deep classes can thus be seen as an analogue for Π^0_1 classes of logically deep sequences.

Remark 4.2. Note that depth can be equivalently defined as follows: $\mathcal{C}$ is deep if for some computable function f one has $\mathbf{M}(T_{f(k)})\leq 2^{-k}$ for all k. Indeed, if $\mathbf{M}(T_n)\leq 2^{-h(n)}$, then, setting $f=h^{-1}$, we have $\mathbf{M}(T_{f(k)})\leq 2^{-n}$, and conversely, if $\mathbf{M}(T_{f(n)})\leq 2^{-n}$, we can assume that f is increasing and then taking $h=f^{-1}$, we have $\mathbf{M}(T_n)\leq 2^{-h(n)}$ (here we use the fact that $\mathbf{M}(T_n)$ is nonincreasing in n).

The same argument shows that $\mathcal{C}$ is *tt*-deep if and only if for every computable measure μ, there exists a computable function f such that for all n, $\mu(T_{f(n)})\leq 2^{-n}$.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

REMARK 4.3. Another alternative way to define depth is to use $\mathbf{m}$ instead of $\mathbf{M}$: a Π^0_1 class $\mathcal{C}$ with canonical co-c.e. tree T is deep if and only if there is a computable order function h such that $\mathbf{m}(T_n) \leq 2^{-h(n)}$ for all n, if and only if there is a computable function f such that $\mathbf{m}(T_{f(n)}) \leq 2^{-n}$ for all n. Indeed, the following inequality holds for all strings σ (see for example [11]):

$$\mathbf{m}(\sigma) \leq^{\times} \mathbf{M}(\sigma) \leq^{\times} \mathbf{m}(\sigma)/\mathbf{m}(|\sigma|).$$

Thus, for h a computable order function, if $\mathbf{M}(T_n) \leq 2^{-h(n)}$ for all n, then there is some c such that $\mathbf{m}(T_n) \leq 2^{-h(n)+c}$ for all n. Conversely, if $\mathbf{m}(T_n) \leq 2^{-h(n)}$, then $\mathbf{m}(T_{f(n)}) \leq 2^{-n}$ for $f = h^{-1}$, and by the above inequality,

$$\mathbf{M}(T_{f(n)}) \leq^{\times} 2^{-n}/\mathbf{m}(f(n)) \leq^{\times} 2^{-n} \cdot n^2.$$

Thus, taking $g(n) = f(2n + c)$ for some large enough constant c, we get $\mathbf{M}(T_{g(n)}) \leq 2^{-n}$, and thus $\mathcal{C}$ is deep.

Since every Π^0_1 class $\mathcal{C}$ is the set of paths through a computable tree, why can't we simply define depth and *tt*-depth in terms of this tree and not the canonical co-c.e. tree associated to $\mathcal{C}$? In the case of depth, there are two reasons to restrict to the canonical co-c.e. trees associated to Π^0_1 classes.

First, the idea behind a deep class is that it is difficult to produce initial segments of some member of the class. In general, for a Π^0_1 class $\mathcal{C}$, any computable tree T contains nonextendible nodes, and so if we have a procedure that can compute these nonextendible nodes of T with high probability, this tells us nothing about the difficulty of computing the extendible nodes of T.

Second, if we were to use any tree T representing a Π^0_1 class $\mathcal{C}$ in the definition of depth, then depth would become a void notion, by the following proposition.

PROPOSITION 4.4. *If T is an infinite computable tree, then there is no computable order function f such that $\mathbf{m}(T_{f(n)}) \geq 2^{-n}$.*

PROOF. Given an infinite computable tree T, let f be a computable order function. Then there is a computable sequence $(\sigma_n)_{n\in\mathbb{N}}$ of strings such that $\sigma_n \in T_{f(n)}$ for every n. Thus

$$\mathbf{m}(T_{f(n)}) \geq \mathbf{m}(\sigma_n) \geq^{\times} \mathbf{m}(n) \geq^{\times} 1/n^2. \qquad \dashv$$

By contrast with the notion of depth, for the definition of *tt*-depth, it does not matter whether we work with the canonical co-c.e. tree associated to $\mathcal{C}$ or some computable tree T such that $\mathcal{C} = [T]$ (a direct consequence of Theorem 4.5).

An important question concerns the relationship between depth and negligibility (and between *tt*-depth and *tt*-negligibility). First, every deep Π^0_1 class is negligible. Indeed, if $\mathcal{C}$ is deep, T is the canonical co-c.e. tree associated to $\mathcal{C}$, and $\mathbf{M}(T_n) \leq 2^{-h(n)}$ for some computable order function h, then

$$\overline{\mathbf{M}}(\mathcal{C}) = \overline{\mathbf{M}}\Big(\bigcap_{n\in\mathbb{N}} [\![T_n]\!]\Big) = \lim_{n\to\infty} \overline{\mathbf{M}}(T_n) \leq \lim_{n\to\infty} \mathbf{M}(T_n) = 0,$$

where the second equality follows from the continuity of $\overline{\mathbf{M}}$ from above. A similar argument shows that *tt*-depth implies *tt*-negligibility. Does the

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

converse hold? In the case of *tt*-depth and *tt*-negligibility, the answer is positive. We also identify two other equivalent formulations of *tt*-depth.

THEOREM 4.5. *Let $\mathcal{C}$ be a Π^0_1 class. The following are equivalent*:

(i) *$\mathcal{C}$ is tt-deep.*
(ii) *$\mathcal{C}$ is tt-negligible.*
(iii) *For every computable measure μ, $\mathcal{C}$ contains no μ-Kurtz random element.*
(iv) *For every computable measure μ, $\mathcal{C}$ contains no μ-Martin-Löf random element.*

PROOF. (i)⇒(ii). This is shown by a direct modification of the above proof that depth implies negligibility.
(ii)⇒(iii). This follows directly from the definition of Kurtz randomness.
(iii) ⇒ (iv). This follows from the fact that Martin-Löf randomness implies Kurtz randomness.
(iv) ⇒ (ii). If $\mu(\mathcal{C}) > 0$ for some computable measure μ, then $\mathcal{C}$ must contain some μ-Martin-Löf random element since the set of μ-Martin-Löf random sequences has μ-measure 1.
(ii)⇒ (i). Suppose that $\mathcal{C}$ is *tt*-negligible and let μ be a computable measure. Let T be the co-c.e. tree associated to $\mathcal{C}$. By *tt*-negligibility, $\mu(\mathcal{C}) = 0$, or equivalently, $\mu(T_n)$ tends to 0. Since the T_n are co-c.e. sets of strings, $\mu(T_n)$ is right-c.e. uniformly in n. Thus, given k, it is possible to effectively find an n such that $\mu(T_n) \leq 2^{-k}$. Setting $f(k) = n$, we have a computable function f such that $\mu(T_{f(k)}) \leq 2^{-k}$ for all k, therefore $\mathcal{C}$ is *tt*-deep by Remark 4.2. ⊣

REMARK 4.6. Recall that a Π^0_1 class is *special* if it has no computable member. Given a computable sequence X, by Remark 2.8, there is a computable measure μ such that $X \in \mathsf{MLR}_\mu$. It thus follows from Theorem 4.5 that *tt*-negligible classes contain no computable members and are thus special.

Significantly, in contrast to the case with *tt*-negligibility and *tt*-depth, negligibility and depth do not coincide. This is due to a fundamental aspect of depth, namely that it is not invariant under Turing equivalence, unlike negligibility. Suppose that $\mathcal{C}$ and $\mathcal{D}$ are two classes such that for every $X \in \mathcal{C}$ there exists a $Y \in \mathcal{D}$ such that $X \equiv_T Y$, and vice-versa. Then by Proposition 3.2, $\mathcal{C}$ is negligible if and only if $\mathcal{D}$ is negligible. This invariance does not hold in general for deep classes, as the next theorem shows.

THEOREM 4.7. *For any Π^0_1 class $\mathcal{C}$, there is a Π^0_1 class $\mathcal{D}$ such that*:

(a) *the elements of $\mathcal{D}$ are the same as the elements of $\mathcal{C}$, modulo deletion of a finite prefix* (*which in particular guarantees that the elements of $\mathcal{C}$ and $\mathcal{D}$ have the same Turing degrees*); *and*
(b) *$\mathcal{D}$ is not deep.*

PROOF. Let $\mathcal{C}$ be a Π^0_1 class. Let T be a computable tree such that $[T] = \mathcal{C}$, and let S be the canonical co-c.e. tree associated to $\mathcal{C}$. Consider the tree U obtained by appending a copy of S to each terminal node of T. Formally,

$$U = T \cup \{\sigma^\frown\tau : \sigma \in T_{term} \text{ and } \tau \in S\},$$

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

where T_{term} is the set of terminal nodes of T, which is a computable set. One can readily verify that U is the canonical co-c.e. tree associated to some Π^0_1 class. Indeed, suppose ρ is a node of U. Then,

- either $\rho \in T$, in which case ρ is either a prefix of some $X \in [T]$ (and thus $X \in [U]$), or ρ is a prefix of a terminal node ρ', which can then be extended to some $(\rho')^\frown X \in [U]$, where $X \in [S]$; or
- ρ is of the form $\sigma^\frown\tau$, with $\sigma \in T_{term}$ and $\tau \in S$. Since S is a canonical co-c.e. tree, all its nodes extend to an infinite path, and thus $\sigma^\frown\tau$ has an extension $\sigma^\frown X \in [U]$ for some path $X \in [S]$.

The fact that U is co-c.e. follows directly from the fact that T_{term} is computable and S is co-c.e. Let $\mathcal{D}$ be the Π^0_1 class whose canonical co-c.e. tree is U. Then the elements of $\mathcal{D}$ are either elements of $[T] = \mathcal{C}$, or are of the form $\sigma^\frown X$ for some finite string σ and $X \in \mathcal{C}$, which gives us the first part of the conclusion. Finally, the canonical co-c.e. tree of $\mathcal{D}$ contains an infinite computable tree, namely, the tree T, and therefore, by the argument as in the proof of Proposition 4.4, $\mathcal{D}$ is not deep. ⊣

It is now straightforward to get a Π^0_1 class that is negligible but not deep: it suffices to take a negligible class $\mathcal{C}$ and apply the above theorem. The resulting class $\mathcal{D}$ is also negligible as its elements have the same Turing degrees as the elements of $\mathcal{C}$ and negligibility is preserved under Turing equivalence by Proposition 3.2(i). However, $\mathcal{D}$ is not deep.

The following summarizes the implications between the different properties introduced above for Π^0_1 classes.

$$\text{deep} \Rightarrow \text{negligible} \Rightarrow \textit{tt}\text{-negligible} \Leftrightarrow \textit{tt}\text{-deep} \Rightarrow \text{special}.$$

§5. Computational limits of randomness. As we have seen, negligible Π^0_1 classes (and thus deep classes) have the property that one cannot compute a member of them with positive probability. Although some random sequences can compute a member in any negligible Π^0_1 class (namely the Martin-Löf random sequences of PA degree, as sequences of PA degree compute a member of *every* Π^0_1 class), by the definition of negligibility, almost every random sequence fails to compute a member of a negligible class.

Similarly, one cannot *tt*-compute a member of a *tt*-negligible Π^0_1 class with positive probability. The definition of *tt*-negligibility implies that the collection of sequences that can *tt*-compute a member of a *tt*-negligible class form a set of Lebesgue measure zero.

In this section, we specify a precise level of randomness at which computing a member of a *tt*-negligible, negligible, or deep class fails. First, we consider the case for *tt*-negligible classes.

THEOREM 5.1. *If $X \in 2^{\mathbb{N}}$ is Kurtz random, it cannot tt-compute any member of a tt-negligible Π^0_1 class.*

PROOF. Let $\mathcal{C}$ be a *tt*-negligible Π^0_1 class and Φ a *tt*-functional. The set $\Phi^{-1}(\mathcal{C})$ is a Π^0_1 class that, by *tt*-negligibility, has Lebesgue measure 0. Thus, it contains no Kurtz random. ⊣

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

A similar proof can be used to prove the following:

THEOREM 5.2. *If $X \in 2^{\mathbb{N}}$ is weakly 2-random, it cannot compute any member of a negligible Π^0_1 class.*

PROOF. Let $\mathcal{C}$ be negligible Π^0_1 class, Φ be a Turing functional, and T a computable tree such that $\mathcal{C} = [T]$. The set $\Phi^{-1}(\mathcal{C})$ is a Π^0_2 class, since

$$X \in \Phi^{-1}(\mathcal{C}) \text{ if and only if } (\forall k)(\exists \sigma)(\exists n)[\sigma \in T \ \&\ |\sigma| = k \ \&\ \sigma \preceq \Phi^{X\upharpoonright n}].$$

By negligibility, $\Phi^{-1}(\mathcal{C})$ has Lebesgue measure 0 and thus contains no weakly 2-random sequence. ⊣

We do not know whether this last theorem is optimal, i.e., whether it can be extended to randomness notions that are weaker than weak-2-randomness.

Our next result, despite its simplicity, is probably the most interesting of this section. It will help us unify a number of theorems that have appeared in the literature. These are theorems of form

$(*)$ If X is difference random, then it cannot compute an element of $\mathcal{C}$,

where $\mathcal{C}$ is a given Π^0_1 class. Theorem 2.11 is an example of such a theorem, with $\mathcal{C}$ the class of consistent completions of PA. The same result has been obtained with $\mathcal{C}$ the class of shift-complex sequences (Khan [19]), the set of compression functions (Greenberg, Miller, Nies [15]), and the set $\mathcal{DNC}_q$ functions for some computable order functions q (Miller, unpublished). We will give the precise definition of these classes in Section 7 but the important fact is that all of these classes are deep, and indeed, showing the depth of a Π^0_1 class is sufficient to obtain a theorem of the form $(*)$.

THEOREM 5.3. *If a sequence X is difference random, it cannot compute any member of a deep Π^0_1 class.*

PROOF. Let $\mathcal{C}$ be a deep Π^0_1 class with associated co-c.e. tree T and let f be a computable function such that $\mathbf{M}(T_{f(n)}) \leq 2^{-n}$. Let X be a sequence that computes a member of $\mathcal{C}$ via a Turing functional Φ. Let

$$\mathcal{Z}_n = \{Z : \Phi^Z\upharpoonright f(n)\downarrow \in T_{f(n)}\}.$$

The set $\mathcal{Z}_n$ can be written as the difference $\mathcal{U}_n \setminus \mathcal{V}_n$ of two effectively open sets (uniformly in n) with $\mathcal{U}_n = \{Z : \Phi^Z\upharpoonright f(n)\downarrow\}$ and $\mathcal{V}_n = \{Z : \Phi^Z\upharpoonright f(n)\downarrow \notin T_{f(n)}\}$.

Moreover, by definition of the semi-measure induced by Φ,

$$\lambda(\mathcal{Z}_n) \leq \lambda_\Phi(T_{f(n)}) \leq^\times \mathbf{M}(T_{f(n)}) \leq 2^{-n}.$$

The sequence $\{\mathcal{Z}_n\}_{n\in\mathbb{N}}$ thus yields a difference test. Therefore, the sequence X, which by assumption belongs to all $\mathcal{Z}_n$, is not difference random. ⊣

We remark that the converse does not hold: i.e., there is a class $\mathcal{D}$ such that (i) $\mathcal{D}$ is not deep but (ii) no difference random real can compute an element of $\mathcal{D}$. Indeed, take a deep class $\mathcal{C}$ and apply Theorem 4.7 to get a class $\mathcal{D}$ that is not deep but whose members have the same Turing degrees as the elements of $\mathcal{C}$. Thus, no difference random real can compute an element of $\mathcal{D}$. Also, the hypothesis of difference randomness cannot be improved. Indeed, if X is

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

difference random but not Martin-Löf random, then it computes $\emptyset'$, and thus it can compute a member of any nonempty Π^0_1 class.

§6. **Depth, negligibility, and mass problems.** In this section we discuss depth and negligibility in the context of mass problems, i.e., in the context of Muchnik and Medvedev reducibility. Both Muchnik and Medvedev reducibility are generalizations of Turing reducibility. Whereas Turing reducibility is defined in terms of a pair of sequences, both Muchnik and Medevedev reducibility are defined in terms of a pair of collections of sequences. In what follows, we will consider these two reducibilities when restricted to Π^0_1 subclasses of $2^{\mathbb{N}}$. We will follow the notation of the survey [31], to which we refer the reader for a thorough exposition of mass problems in the context of Π^0_1 classes.

Let $\mathcal{C}, \mathcal{D} \subseteq 2^{\mathbb{N}}$ be Π^0_1 classes. We say that $\mathcal{C}$ is *Muchnik reducible* to $\mathcal{D}$, denoted $\mathcal{C} \leq_w \mathcal{D}$, if for every $X \in \mathcal{D}$ there exists a Turing functional Φ such that (i) $X \in \mathsf{dom}(\Phi)$ and $\Phi(X) \in \mathcal{C}$. Moreover, $\mathcal{C}$ is *Medvedev reducible* to $\mathcal{D}$, denoted $\mathcal{C} \leq_s \mathcal{D}$, if $\mathcal{C}$ is Muchnik reducible to $\mathcal{D}$ via a single Turing functional, i.e., there exists a Turing functional Φ such that (i) $\mathcal{D} \subseteq \mathsf{dom}(\Phi)$ and $\Phi(\mathcal{D}) \subseteq \mathcal{C}$.

Just as Turing reducibility gives rise to a degree structure, we can define degree structures from $\leq_w$ and $\leq_s$. We say that $\mathcal{C}$ and $\mathcal{D}$ are *Muchnik equivalent* (resp. *Medvedev equivalent*), denoted $\mathcal{C} \equiv_w \mathcal{D}$ (resp. $\mathcal{C} \equiv_s \mathcal{D}$) if and only if $\mathcal{C} \leq_w \mathcal{D}$ and $\mathcal{D} \leq_w \mathcal{C}$ (resp. $\mathcal{C} \leq_s \mathcal{D}$ and $\mathcal{D} \leq_s \mathcal{C}$). The collections of Muchnik and Medvedev degrees given by the equivalence classes under $\equiv_w$ and $\equiv_s$ are denoted $\mathcal{E}_w$ and $\mathcal{E}_s$, respectively.

Both $\mathcal{E}_w$ and $\mathcal{E}_s$ are lattices, unlike the Turing degrees, which only form an upper semi-lattice. We define the meet and join operations as follows. Given Π^0_1 classes $\mathcal{C}, \mathcal{D} \subseteq 2^{\mathbb{N}}$, $\sup(\mathcal{C}, \mathcal{D})$ is the Π^0_1 class $\{X \oplus Y : X \in \mathcal{C} \ \&\ Y \in \mathcal{D}\}$. Furthermore, we define $\inf(\mathcal{C}, \mathcal{D})$ to be the Π^0_1 class $\{0^\frown X : X \in \mathcal{C}\} \cup \{1^\frown Y : Y \in \mathcal{D}\}$. One can readily check that the least upper bound of $\mathcal{C}$ and $\mathcal{D}$ in $\mathcal{E}_w$ is the Muchnik degree of $\sup(\mathcal{C}, \mathcal{D})$ while their greatest lower bound is the Muchnik degree $\inf(\mathcal{C}, \mathcal{D})$, and similarly for $\mathcal{E}_s$.

Recall that a filter $\mathcal{F}$ in a lattice $(\mathcal{L}, \leq, \inf, \sup)$ is a subset that satisfies the following two conditions: (i) for all $x, y \in \mathcal{L}$, if $x \in \mathcal{F}$ and $x \leq y$, then $y \in \mathcal{F}$, and (ii) for all $x, y \in \mathcal{F}$, $\inf(x, y) \in \mathcal{F}$.

The goal of this section is to study the role of depth and negligibility in the structures $\mathcal{E}_s$ and $\mathcal{E}_w$. Let us start with an easy result.

THEOREM 6.1. *The collection of negligible Π^0_1 classes forms a filter in both $\mathcal{E}_s$ and $\mathcal{E}_w$.*

PROOF. Let $\mathcal{C}$ and $\mathcal{D}$ be negligible Π^0_1 classes. By Proposition 3.2, we have that $\lambda(\mathcal{C}^{\leq_T}) = 0$ and $\lambda(\mathcal{D}^{\leq_T}) = 0$. But since

$$\inf(\mathcal{C}, \mathcal{D})^{\leq_T} = \mathcal{C}^{\leq_T} \cup \mathcal{D}^{\leq_T},$$

it follows that $\lambda(\inf(\mathcal{C}, \mathcal{D})^{\leq_T}) = 0$, which shows that $\inf(\mathcal{C}, \mathcal{D})$ is negligible. Thus, the degrees of negligible classes in both $\mathcal{E}_w$ and $\mathcal{E}_s$ are closed under inf.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Let $\mathcal{D}$ be non-negligible Π^0_1 and $\mathcal{C} \leq_w \mathcal{D}$. For each i, we define

$$\mathcal{D}_i := \{X \in \mathcal{D} : X \in \mathsf{dom}(\Phi_i) \ \&\ \Phi_i(X) \in \mathcal{C}\}.$$

Since $\mathcal{C} \leq_w \mathcal{D}$, it follows that $\mathcal{D} = \bigcup_i \mathcal{D}_i$. Furthermore, we have $\mathcal{D}^{\leq_T} = \left(\bigcup_i \mathcal{D}_i\right)^{\leq_T} = \bigcup_i \mathcal{D}_i^{\leq_T}$. Since $\mathcal{D}$ is non-negligible, we have

$$0 < \lambda(\mathcal{D}^{\leq_T}) \leq \sum_i \lambda(\mathcal{D}_i^{\leq_T}),$$

and thus $\lambda(\mathcal{D}_k^{\leq_T}) > 0$ for some k. But since $\Phi_k(\mathcal{D}_k) \subseteq \mathcal{C}$, it follows that $\lambda(\mathcal{C}^{\leq_T}) > 0$, thus $\mathcal{C}$ is non-negligible. Thus, negligibility is closed upwards under $\leq_w$ (and a fortiori, under $\leq_s$ as well). ⊣

Remark 6.2. Simpson proved in [30] that in $\mathcal{E}_w$, the complement of the filter of negligible classes is in fact a principal ideal, namely the ideal generated by the class $\inf(\mathcal{PA}, 2\mathcal{RAN})$, where $2\mathcal{RAN}$ is the class of 2-random sequences, i.e., the sequences that are Martin-Löf random relative to $\emptyset'$.

For the next two theorems, we need the following fact.

Fact 6.3. *Let $\mathcal{C}$ and $\mathcal{D}$ be Π^0_1 classes such that $\mathcal{C} \leq_s \mathcal{D}$. Then there is a total Turing functional Ψ such that $\Psi(\mathcal{D}) \subseteq \mathcal{C}$.*

Theorem 6.4. *The collection of deep Π^0_1 classes forms a filter in $\mathcal{E}_s$.*

Proof. Let $\mathcal{C}$ and $\mathcal{D}$ be deep Π^0_1 classes with associated co-c.e. trees S and T, respectively. Moreover, let g and h be computable order functions such that $\mathbf{M}(S_n) \leq 2^{-g(n)}$ and $\mathbf{M}(T_n) \leq 2^{-h(n)}$. We define $f(n) = \min\{g(n), h(n)\}$, which is clearly a computable order function. It follows immediately that $\mathbf{M}(S_n) \leq 2^{-f(n)}$ and $\mathbf{M}(T_n) \leq 2^{-f(n)}$.

Now the co-c.e. tree associated with $\inf(\mathcal{C}, \mathcal{D})$ is

$$R = \{0^\frown\sigma : \sigma \in S\} \cup \{1^\frown\tau : \tau \in T\}.$$

Consider the class $\inf(\mathcal{C}, \mathcal{D})$. Setting $0^\frown S_n = \{0^\frown\sigma : \sigma \in S_n\}$ and $1^\frown T_n = \{1^\frown\sigma : \sigma \in T_n\}$ for each n, the co-c.e. tree associated to $\inf(\mathcal{C}, \mathcal{D})$ is $0^\frown S_n \cup 1^\frown T_n$. Moreover,

$$\mathbf{M}(0^\frown S_n \cup 1^\frown T_n) = \mathbf{M}(0^\frown S_n) + \mathbf{M}(1^\frown T_n) \leq^\times 2^{-f(n)} + 2^{-f(n)}$$

which shows that $\inf(\mathcal{C}, \mathcal{D})$ is deep.

Next, suppose that $\mathcal{C}$ is a deep Π^0_1 class and $\mathcal{D}$ is a Π^0_1 class satisfying $\mathcal{C} \leq_s \mathcal{D}$ via the Turing functional Ψ, which we can assume to be total by Fact 6.3. Let S and T be the co-c.e. trees associated to $\mathcal{C}$ and $\mathcal{D}$, respectively, and let g be a computable order function such that $\mathbf{M}(S_n) \leq 2^{-g(n)}$. Since Ψ is total, its use is bounded by some computable function f. It follows that for every $\sigma \in T_{f(n)}$, there is some $\tau \in S_n$ such that $(\sigma, \tau) \in S_\Psi$ (the c.e. set of pairs of strings that generates Ψ). Thus $[\![T_{f(n)}]\!] \subseteq \Psi^{-1}([\![S_n]\!])$. Now let Φ be a universal Turing functional such that $\mathbf{M} = \lambda_\Phi$. Then:

$$\begin{aligned}\mathbf{M}(T_{f(n)}) &= \lambda(\Phi^{-1}(T_{f(n)})) \leq \lambda(\Phi^{-1}([\![T_{f(n)}]\!]) \\ &\leq \lambda(\Phi^{-1}(\Psi^{-1}([\![S_n]\!]))) \leq \lambda_{\Psi\circ\Phi}(S_n),\end{aligned}$$

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

where the last inequality holds because $\Phi^{-1}(\Psi^{-1}([\![S_n]\!])) = (\Psi \circ \Phi)^{-1}([\![S_n]\!]) \subseteq \mathrm{dom}(\Psi\circ\Phi)$ and hence $(\Psi\circ\Phi)^{-1}([\![S_n]\!]) \subseteq (\Psi\circ\Phi)^{-1}(S_n)$. Since $\lambda_{\Psi\circ\Phi} \leq^\times \mathbf{M}$, we have

$$\mathbf{M}(T_{f(n)}) \leq \lambda_{\Psi\circ\Phi}(S_n) \leq^\times \mathbf{M}(S_n) \leq 2^{-g(n)}.$$

Thus, $\mathbf{M}(T_{f\circ g^{-1}(n)}) \leq 2^{-n}$, which shows that $\mathcal{D}$ is deep. ⊣

The invariance of the notion of depth under Medvedev-equivalence is of importance for the next section. There, we prove that certain classes of objects, which are not necessarily infinite binary sequences, are deep. To do so, we fix a certain encoding of these objects by infinite binary sequences and prove the depth of the corresponding encoded class. By the above theorem, the particular choice of encoding is irrelevant: if the class is deep for an encoding, it will be deep for another encoding, as long as switching from the first encoding to the second one can be done computably and uniformly.

One could ask whether we have a similar result in the lattice $\mathcal{E}_w$. However, Theorem 4.7 shows that depth is not invariant under Muchnik equivalence: if we apply this theorem to a deep class $\mathcal{C}$, we get a Π^0_1 class $\mathcal{D}$ which is clearly Muchnik equivalent to $\mathcal{C}$ but is not deep itself. Thus depth is Medvedev invariant but not Muchnik invariant.

More surprisingly, *tt*-depth *is* a both Medvedev invariant and Muchnik invariant, as the *tt*-deep classes form a filter in both lattices.

THEOREM 6.5. *The collection of tt-deep Π^0_1 classes forms a filter in both $\mathcal{E}_s$ and $\mathcal{E}_w$.*

PROOF. Suppose that $\mathcal{C} \leq_w \mathcal{D}$ and $\mathcal{D}$ is not *tt*-deep. By definition, this means that $\mathcal{D}$ has positive μ-measure for some computable probability measure μ. Let $\{\mathcal{U}_k\}_{k\in\mathbb{N}}$ be the universal μ-Martin-Löf test and define $\mathcal{R}_k$ to be the complement of $\mathcal{U}_k$ (since $\mathcal{U}_k$ is Σ^0_1, $\mathcal{R}_k$ is Π^0_1). Since $\mu(\mathcal{R}_k) \geq 1 - 2^{-k}$ for every k, there must be a j such that $\mu(\mathcal{D} \cap \mathcal{R}_j) > 0$ and in particular $\mathcal{D} \cap \mathcal{R}_j \neq \emptyset$. By the hyperimmune-free basis theorem, there is some $X \in \mathcal{D} \cap \mathcal{R}_j$ of hyperimmune-free Turing degree. Since $\mathcal{C} \leq_w \mathcal{D}$, X must compute some element Y of $\mathcal{C}$. But since X is of hyperimmune-free degree, X in fact *tt*-computes Y, i.e., $\Psi(X) = Y$ for some total Turing functional Ψ.

Since X is μ-Martin-Löf random, by the preservation of randomness theorem (see, for instance, Theorem 3.2 in [4]), Y is Martin-Löf random with respect to the computable measure μ_Ψ defined by $\mu_\Psi(\sigma) := \mu(\Psi^{-1}(\sigma))$ for every σ. Thus, by Proposition 4.5, $\mathcal{C}$ is not *tt*-deep. This shows, a fortiori, that if $\mathcal{C} \leq_s \mathcal{D}$ and $\mathcal{D}$ is not *tt*-deep, then $\mathcal{C}$ is not *tt*-deep. Thus *tt*-depth is closed upwards in $\mathcal{E}_w$ and $\mathcal{E}_s$ and in particular is compatible with the equivalence relations $\equiv_s$ and $\equiv_w$.

Next, suppose that $\mathcal{C}$ and $\mathcal{D}$ are *tt*-deep classes but that $\inf(\mathcal{C}, \mathcal{D})$ is not *tt*-deep (recall that the inf operator is the same for both $\mathcal{E}_s$ and $\mathcal{E}_w$). Then by Theorem 4.5, $\inf(\mathcal{C}, \mathcal{D})$ contains a sequence X that is μ-Martin-Löf random for some computable measure μ. Then for $i = 0, 1$ we define computable measures μ_i such that $\mu_i(\sigma) = \mu(i^\frown\sigma)/\mu(i)$ for every σ (where this ratio is equal to 0 if $\mu(i) = 0$). It is routine to check that Y is μ_i-Martin-Löf random if and only if $i^\frown Y$ is μ-Martin-Löf random for $i = 0, 1$. Since $X = i^\frown Z$ for

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

some $i = 0, 1$ and $Z \in 2^{\mathbb{N}}$, it follows that Z is μ_i-Martin-Löf random. But then Z is contained in either $\mathcal{C}$ or $\mathcal{D}$, which contradicts our hypothesis that $\mathcal{C}$ and $\mathcal{D}$ are both tt-deep. ⊣

§7. **Examples of deep Π^0_1 classes.** In this section, we provide a number of examples of deep Π^0_1 classes that naturally occur in computability theory and algorithmic randomness. We give a uniform treatment of all these classes, i.e., we give a general method to prove the depth of Π^0_1 classes.

7.1. Consistent completions of Peano Arithmetic. As mentioned in Section 3, Jockusch and Soare proved in [17] that the Π^0_1 class $\mathcal{PA}$ of consistent completions of PA is negligible. However, as shown implicitly by Levin in [22] and Stephan in [32], $\mathcal{PA}$ is also deep. We will reproduce this result here.

Following both Levin and Stephan, we will use fact that the class of consistent completions of PA is Medvedev equivalent to the class of total extensions of a universal, partial-computable $\{0, 1\}$-valued function. Thus, by showing the latter class is deep, we thereby establish that the former class is deep (via Theorem 6.4).

THEOREM 7.1 (Levin [22], Stephan [32]). *Let $(\phi_e)_{e\in\mathbb{N}}$ be a standard enumeration of all $\{0, 1\}$-valued partial computable functions. Let u be a function that is universal for this collection, e.g., defined by $u(\langle e, x\rangle) = \phi_e(x)$. Then the class $\mathcal{C}$ of total extensions of u is a deep Π^0_1 class.*

PROOF. We build a partial computable function ϕ_e, whose index we know in advance by the recursion theorem. This means that we control the value of $u(\langle e, x\rangle)$ for all x. First, we partition $\mathbb{N}$ into consecutive intervals $I_1, I_2, \ldots$ such that we control 2^{k+1} values of u inside I_k. For each k in parallel, we define ϕ_e on I_k as follows.

> Step 1: Wait for a stage s such that the set $E_k[s] = \{\sigma : \sigma{\restriction}I_k \text{ extends } u[s]{\restriction}I_k\}$ is such that $\mathbf{M}(E_k)[s] \geq 2^{-k}$.
>
> Step 2: Find a $y \in I_k$ that we control and on which $u[s]$ is not defined. Consider the two "halves" $E^0_k[s] = \{\sigma \in E_k[s] : \sigma(y) = 0\}$ and $E^1_k[s] = \{\sigma \in E_k[s] : \sigma(y) = 1\}$ of $E_k[s]$. Note that either $\mathbf{M}(E^0_k[s]) \geq 2^{-k-1}$ or $\mathbf{M}(E^1_k)[s] \geq 2^{-k-1}$. If the first holds, set $u(y)[s + 1] = 1$, otherwise set $u(y)[s + 1] = 0$. Go back to Step 1.

The co-c.e. tree T associated to the class $\mathcal{C}$ is the set of strings σ such that σ is an extension of $u{\restriction}|\sigma|$. The construction works because every time we pass by Step 2, we remove from $T[s]$ a set $E^i_k[s]$ (for some $i \in \{0, 1\}$ and $k, s \in \mathbb{N}$) such that $\mathbf{M}(E^i_k)[s] \geq 2^{-k-1}$. Therefore, Step 2 can be executed at most 2^{k+1} times, and by the definition of I_k we do not run out of values $y \in I_k$ on which we control u. Therefore, the algorithm eventually reaches Step 1 and waits there forever. Setting $f(k) = \max(I_k)$, this implies that the $\mathbf{M}$-weight of the set $\{\sigma : \sigma{\restriction}f(k) \text{ extends } u\}$ is bounded by 2^{-k}, or equivalently,

$$\mathbf{M}(T_{f(k)}) \leq 2^{-k},$$

which proves that $\mathcal{C}$ is deep. ⊣

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

The proof provided here gives us a general template to prove the depth of a Π^0_1 class. First of all, the definition of the Π^0_1 class should allow us to control parts of it in some way, either because we are defining the class ourselves or because, as in the above proof, the definition of the class involves some universal object which we can assume to partially control due to the recursion theorem. All the other examples of deep Π^0_1 class we will see below belong to this second category. Let us take a step back and analyze more closely the structure of the proof of Theorem 7.1. Given a Π^0_1 class $\mathcal{C}$ with canonical co-c.e. tree T, the proof consists of the following steps.

(1) For a given k, we identify a level $N = f(k)$ at which we wish to ensure $\mathbf{M}(T_N) \leq 2^{-k}$. The choice of N will depend on the particular class $\mathcal{C}$.
(2) Next we implement a two-step strategy to ensure that $\mathbf{M}(T_N) \leq 2^{-k}$. Such a strategy will be called a *k-strategy*.
  (2.1) First, we wait for a stage s at which $\mathbf{M}(T_N)[s] \geq 2^{-k}$.
  (2.2) If at some stage s this occurs, we remove one or several of the nodes of $T[s]$ at level N such that the total $\mathbf{M}[s]$-weight of these nodes is at least $\delta(k)$ for a certain function δ, and then go back to Step 2.
(3) Each execution of Step 2 (removing nodes from $T[s]$ for some s) comes at a cost $\gamma(k)$ for some k, and we need to make sure that we do not go over some maximal total cost $\Gamma(k)$ throughout the execution of the k-strategy.

In the above example, $\delta(k) = 2^{-k-1}$ and the only cost for us is to define one value of $u(y)$ out of the 2^{k+1} we control, so we can, for example, set $\gamma(k) = 1$ and $\Gamma(k) = 2^{k+1}$. By definition of $\mathbf{M}$, the k-strategy can only go through Step 2 at most $1/\delta(k)$ times, and thus the total cost of the k-strategy will be at most $\gamma(k)/\delta(k)$. All we have to do is to make sure that $\gamma(k)/\delta(k) \leq \Gamma(k)$ (which is the case in the above example). Again, we have some flexibility on the choice of $N = f(k)$, so it will suffice to choose an appropriate N to ensure $\gamma(k)/\delta(k) \leq \Gamma(k)$. In some cases, there will not be a predefined maximum for each k, but rather a global maximum Γ that the sum of the the costs of the k-strategies should not exceed, i.e., we will want to have $\sum_k \gamma(k)/\delta(k) \leq \Gamma$. Let us now proceed to more examples of Π^0_1 classes.

7.2. Shift-complex sequences.

DEFINITION 7.2. Let $\alpha \in (0, 1)$ and c be a non-negative integer.

(i) $\sigma \in 2^{<\mathbb{N}}$ (resp. $X \in 2^{\mathbb{N}}$) is said to be (α, c)-*shift complex* if $\mathrm{K}(\tau) \geq \alpha|\tau| - c$ for every substring τ of σ (resp. of X).
(ii) $X \in 2^{\mathbb{N}}$ is said to be α-*shift complex* if it is (α, c)-shift complex for some c.

The very existence of α-shift complex sequences is by no means obvious. Such sequences were first constructed by Durand, Levin, and Shen [9] who showed that there exist α-shift complex sequences for all $\alpha \in (0, 1)$. It is easy to see that for every computable pair (α, c), the class of (α, c)-shift complex

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

sequences is a Π^0_1 class. Rumyantsev [27] proved that the class of (α, c)-complex sequences is always negligible, but in fact, his proof essentially shows that it is even deep. The cornerstone of Rumyantsev's theorem is the following lemma. It relies on an ingenious combinatorial argument that we do not reproduce here. We refer the reader to [27] for the full proof.

LEMMA 7.3 (Rumyantsev [27, Lemma 6]). *Let $\beta \in (0,1)$. For every rational $\eta \in (0,1)$, there exist two integers n and N, with $n < N$, such that the following holds. For every probability distribution P on $\{0,1\}^N$, there exist finite sets of strings $A_n, A_{n+1}, \dots, A_N$ such that*

(i) *for all $j \in [n, N]$, A_j contains only strings of length j and has at most $2^{j\beta}$ elements; and*
(ii) *$P(\{\sigma \in \{0,1\}^N : \sigma$ has some substring in $\cup_{i=n}^N A_i\}) \geq 1 - \eta$.*

Moreover, n and N can be effectively computed from η and can be chosen to be arbitrarily large. Once n and N are fixed, the sets A_i can be computed uniformly in P.

Rumyantsev does not explicitly state that the conclusion holds for *all* probability measures, but nothing in his proof makes use of a particular measure.

THEOREM 7.4. *For any computable $\alpha \in (0,1)$ and integer $c \geq 0$, the Π^0_1 class of (α, c)-shift complex sequences is deep.*

PROOF. Let $\mathcal{C}$ be the Π^0_1 class of (α, c)-shift complex sequences and let T be its canonical co-c.e. tree. We shall build a left-c.e. discrete semi-measure m whose coding constant, which we will write in the form 2^e, we know in advance. This means that whenever we will set a string σ to be such that $m(\sigma) > 2^{-\alpha|\sigma|+c+e+1}$, then automatically we will have $\mathbf{m}(\sigma) > 2^{-\alpha|\sigma|+c+1}$, and thus $\mathrm{K}(\sigma) < \alpha|\sigma| - c$. This will *de facto* remove σ from T.

Let us now turn to the construction. First, we pick some β such that $0 < \beta < \alpha$. For each k, we apply the above Lemma 7.3 to β and $\eta = 1/2$ to obtain a pair (n, N) with the above properties as described in the statement of the lemma (we will also make use of the fact that n can be chosen arbitrarily large, see below). Then the two step strategy is the following:

> Step 1: Wait for a stage s at which $\mathbf{M}(T_N)[s] \geq 2^{-k}$. Up to delaying the increase of $\mathbf{M}$, we can assume that when such a stage occurs, one in fact has $\mathbf{M}(T_N)[s] = 2^{-k}$.[1]
>
> Step 2: Let P be the (computable) probability distribution on $\{0,1\}^N$ whose support is contained in $T_N[s]$ and such that for all $\sigma \in T_N[s]$, $P(\sigma) = 2^k \cdot \mathbf{M}(\sigma)[s]$. By Lemma 7.3, we can compute a collection of finite sets A_i such that:
>
> (i) For all $j \in [n, N]$, A_j contains only strings of length j and has at most $2^{j\beta}$ elements.
> (ii) Setting $F = \{\sigma \in \{0,1\}^N : \sigma$ has some substring in $\cup_{i=n}^N A_i\}$, we have $P(F) \geq 1 - \eta = 1/2$. Thus, $\mathbf{M}(T_N \cap F)[s] \geq 2^{-k-1}$.

[1] That is, if s satisfies $\mathbf{M}(T_N)[s-1] < 2^{-k}$ and $\mathbf{M}(T_N)[s] = r > 2^{-k}$, we can modify the enumeration of $\mathbf{M}$ so that $\mathbf{M}(T_N)[s] = 2^{-k}$ by delaying the enumeration of the additional $\mathbf{M}$-measure of size $r - 2^{-k}$ until stage $s+1$.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Then, for each i and each $\sigma \in A_i$, we ensure, increasing m if necessary, that $m(\sigma) > 2^{-\alpha|\sigma|+c+e+1}$. As explained above, this ensures that all strings in F are removed from T_N at that stage, and therefore we have removed a **M**-weight of at least $\delta(k) = 2^{-k-1}$ from T_N. Then we go back to Step 1.

To finish the proof, we need to make sure that this algorithm does not cost us too much. The constraint here is that we need to make sure that $\sum_{\tau \in 2^{<\mathbb{N}}} m(\tau) \leq 1$, so we are trying to stay under a global cost of $\Gamma = 1$. For a given k, our cost at each execution of Step 2 is the increase of m on the strings in F. The total m-weight we add during one such execution of Step 2 is at most

$$\gamma(k) = \sum_{j=n}^{N} |A_j| \cdot 2^{-j\alpha+c+e+1} \leq \sum_{j=n}^{N} 2^{j(\beta-\alpha)+c+e+1} \leq \frac{2^{(\beta-\alpha)n+c+e+1}}{1-2^{\beta-\alpha}}.$$

But since $n = n(k)$ can be chosen arbitrarily large, therefore, with an appropriate choice of n, we can make $\gamma(k) \leq 2^{-2k-2}$. Therefore, since the k-strategy executes Step 2 at most $1/\delta(k)$ times, we have

$$\sum_{\tau \in 2^{<\mathbb{N}}} m(\tau) \leq \sum_{k \in \mathbb{N}} \gamma(k)/\delta(k) \leq \sum_{k \in \mathbb{N}} 2^{-2k-2}/2^{-k-1} \leq 1$$

and therefore m is indeed a discrete semi-measure. ⊣

7.3. DNC_q **functions.** Let $(\phi_e)_{e \in \mathbb{N}}$ be a standard enumeration of partial computable functions from $\mathbb{N}$ to $\mathbb{N}$. Let J be an effectively optimal partial computable function. By this we mean that there exists a total computable function h, which we may assume to be one-to-one, such that $J(h(e,x)) = \phi_e(x)$ and $|h(e,x)| \leq |x| + c_e$ for some constant c_e that depends on e only (where the length of an integer is the length of its binary representation). For example, one could take:

$$J(2^e(2x+1)) = \phi_e(x).$$

The following notion was studied in [14].

DEFINITION 7.5. Let $q : \mathbb{N} \to \mathbb{N}$ be a computable order function. The set $\mathcal{DNC}_q$ is the set of total functions $f : \mathbb{N} \to \mathbb{N}$ such that for all n:

(i) $f(n) \neq J(n)$ (where this condition is trivially satisfied if $J(n)$ is undefined), and
(ii) $f(n) < q(n)$.

We should note that this is a slight variation of the standard definition of diagonal noncomputability (used in [14]), which is usually formulated in terms of the condition

(i′) $f(n) \neq \phi_n(n)$

instead of the condition (i) given above. While this makes no difference for diagonal noncomputability alone, using condition (i') would make the class $\mathcal{DNC}_q$ highly dependent on the particular choice of enumeration $(\phi_e)_{e \in \mathbb{N}}$,

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

while with our definition, this dependence is somewhat reduced. More precisely, we have the following robustness property. Let J_0 and J_1 be two optimal partial computable functions, and for an arbitrary computable order function q, let $\mathcal{DNC}_q^0$ and $\mathcal{DNC}_q^1$ be the corresponding classes of functions given by the above definition. Then there are constants c, d such that, setting $\tilde{q}(n) = q(cn + d)$, we have

$$\mathcal{DNC}_{\tilde{q}}^0 \leq_s \mathcal{DNC}_q^1.$$

Indeed there is a constant e such that $J_1(h(e, x)) = J_0(x)$, and h is bounded by a linear function.

Note that for every fixed computable order function q, the class $\mathcal{DNC}_q$ is a Π_1^0 subset of $\mathbb{N}^{\mathbb{N}}$, but due to the bound q, $\mathcal{DNC}_q$ can be viewed as a Π_1^0 subset of $2^{\mathbb{N}}$. Whether the class $\mathcal{DNC}_q$ is deep turns out to depend on q. Indeed, we have the following interesting dichotomy theorem.

THEOREM 7.6. *Let q be a computable order function.*

(i) *If $\sum_{n\in\mathbb{N}} 1/q(n) < \infty$, then $\mathcal{DNC}_q$ is not tt-negligible.*
(ii) *If $\sum_{n\in\mathbb{N}} 1/q(n) = \infty$, then $\mathcal{DNC}_q$ is deep.*

PROOF. Part (i) is a straightforward adaptation of Kučera's proof that every Martin-Löf random sequence computes a DNC function. Pick an $X \in 2^{\mathbb{N}}$ at random and use it as a source to randomly pick the value of $f(i)$ uniformly among $\{0, \ldots, q(i) - 1\}$, independently of the other values of f.

The details are as follows. First, we can assume that $q(n)$ is a power of 2 for all n. Indeed, for $q' \leq q$ the class $\mathcal{DNC}_{q'}$ is contained in $\mathcal{DNC}_q$, so if we take $q'(n)$ to be the largest power of 2 less than or equal to $q(n)$, we have that q' is computable, $q' \leq q$, and $\sum_{n\in\mathbb{N}} 1/q'(n) < \infty$ because q' is equal to q up to factor 2. Now, set $q(n) = 2^{r(n)}$. Split $\mathbb{N}$ into intervals where interval I_n has length $r(n)$. One can now interpret any infinite binary sequence X as a function $f_X : \mathbb{N} \to \mathbb{N}$, where $X(n)$ is the index of $X \upharpoonright I_n$ in the lexicographic ordering of strings of length $r(n)$. For X taken at random with respect to the uniform measure, the event $f_X(n) = J(n)$ has probability at most $1/q(n)$ and is independent of all such events for $n' \neq n$. Thus the total probability over X such that $f_X(n) \neq J(n)$ for all n is at least

$$\prod_{n\in\mathbb{N}} (1 - 1/q(n)),$$

an expression that is positive if and only if $\sum_{n\in\mathbb{N}} 1/q(n) < \infty$, which is satisfied by hypothesis (the proof of Lemma 5.6.4 in [25] includes a proof of this result). Thus, the class $\mathcal{DNC}_q$, encoded as above, has positive uniform measure, and thus is not *tt*-negligible.

For (ii), let q be a computable order function such that $\sum_{n\in\mathbb{N}} 1/q(n) = \infty$. Let T be the canonical co-c.e. binary tree in which the elements of $\mathcal{DNC}_q$ are encoded. By the recursion theorem, we will build a partial recursive function ϕ_e whose index e we know in advance and therefore will be able to define J on the set of values $D = \{h(e, x) : x \in \mathbb{N}\}$, where h is defined together with J at the beginning of this section. By definition of h and the assumption of injectivity, the set D has positive (lower) density in $\mathbb{N}$. Thus there is

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

a constant d such that for every interval I of type $\{2^{(k-1)d}, \dots, 2^{kd} - 1\}$ with k large enough (say, greater than d), $|D \cap I| \geq 2^{-d}|I|$. This in particular implies that

$$\sum_{n \in D} 1/q(n) = \infty.$$

To show this, we appeal to Cauchy's condensation test, according to which for any positive nonincreasing sequence $(a_n)_{n \in \mathbb{N}}$, $\sum_{n \in \mathbb{N}} a_n < \infty$ if and only if $\sum_{k \in \mathbb{N}} 2^k a_{2^k} < \infty$. A trivial adaptation of the proof of this criterion gives that $\sum_{n \in \mathbb{N}} a_n < \infty$ if and only if $\sum_{k \in \mathbb{N}} m^k a_{m^k} < \infty$ for any integer $m \geq 2$. Taking $m = 2^d$, since the sum $\sum_{n \in \mathbb{N}} 1/q(n)$ diverges, we have that $\sum_{k \in \mathbb{N}} 2^{kd}/q(2^{kd})$ diverges, and thus

$$\sum_{n \in D} \frac{1}{q(n)} \geq \sum_{k > d} \frac{|D \cap \{2^{(k-1)d}, \dots, 2^{kd} - 1\}|}{q(2^{kd})} \geq \sum_{k > d} \frac{2^{-d} \cdot 2^{kd-1}}{q(2^{kd})} = \infty.$$

Now that we have established that $\sum_{n \in D} 1/q(n) = \infty$, we remark that this is equivalent to having $\prod_{n \in D}(1 - 1/q(n)) = 0$. Thus, we can effectively partition D into countably many finite sets D_j such that $\prod_{n \in D_j}(1 - 1/q(n)) < 1/2$. We are ready to describe the construction. For each k, reserve some finite collection $D_{j_1}, D_{j_2}, \dots, D_{j_{k+1}}$ of sets D_j. Let N be a level of the binary tree T sufficiently large such that the encoding of each path f up to length N is enough to recover the values of f on $D_{j_1} \cup D_{j_2} \cup \dots \cup D_{j_{k+1}}$, which can be found effectively in k. The k-strategy then works as follows.

Initialization. Set $i = 1$.

Step 1: Wait for a stage s at which $\mathbf{M}(T_N)[s] \geq 2^{-k}$. Up to delaying the increase of $\mathbf{M}$, we can assume that when such a stage occurs, one in fact has $\mathbf{M}(T_N)[s] = 2^{-k}$.

Step 2: Since each $\sigma \in T_N[s]$ can be viewed as a function from some initial segment of $\mathbb{N}$ that contains $D_{j_1} \cup D_{j_2} \cup \dots \cup D_{j_{k+1}}$, take the first value $x \in D_{j_i}$. For all $\sigma \in T_N$, $\sigma(x) < q(x)$, thus by the pigeonhole principle there must be at least one value $v < q(x)$ such that

$$\mathbf{M}(\{\sigma \in T_N : \sigma(x) = v\})[s] \geq 2^{-k}/q(x).$$

Set $J(x)$ to be the least such value v. This thus gives

$$\mathbf{M}(\{\sigma \in T_N : \sigma(x) \neq v\})[s] \leq 2^{-k}(1 - 1/q(x)).$$

Now take another x' in D_{j_i} on which J has not been defined yet. By the same reasoning, there must be a value $v' < q(x')$ such that

$$\mathbf{M}(\{\sigma \in T_N : \sigma(x) \neq v \wedge \sigma(x') \neq v'\})[s] \leq 2^{-k}(1-1/q(x))(1-1/q(x')).$$

We then set $J(x')$ to be equal to v'. Continuing in this fashion, we can assign all the values of J on D_{j_i} in such a way that

$$\mathbf{M}(\{\sigma \in T_N : \forall x \in D_{j_i},\ \sigma(x) \neq J(x)\})[s] \leq 2^{-k} \prod_{n \in D_j} (1 - 1/q(n)) \leq 2^{-k}/2.$$

Then we increment i by 1 and go back to Step 1.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Once again, at each execution of Step 2 we remove an $\mathbf{M}$-weight of at least $\delta(k) = 2^{-k-1}$ from T_N, since setting $\sigma(x) = J(x)$ for some x immediately ensures that no extension of σ is in $\mathcal{DNC}_q$. Moreover, each execution of Step 2 requires us to define J on all values in some D_{j_i} for $1 \leq j \leq k+1$. That is, each such execution costs us one D_{j_i}, so as in the case of consistent completions of PA, we can set $\gamma(k) = 1$ and $\Gamma(k) = 2^{k+1}$, which ensures that $\gamma(k)/\delta(k) \leq \Gamma(k)$. Therefore, $\mathcal{DNC}_q$ is a deep Π^0_1 class. ⊣

7.4. Finite sets of maximally complex strings. For a constant $c > 0$, generating a long string σ of high Kolmogorov complexity, for example $\mathrm{K}(\sigma) > |\sigma| - c$, can be easily achieved with high probability if one has access to a random source (just repeatedly flip a fair coin and output the raw result). One can even use this technique to generate a sequence of strings $\sigma_1, \sigma_2, \sigma_3, \ldots$, such that $|\sigma_n| = n$ and $\mathrm{K}(\sigma_n) \geq n - c$. Indeed, the measure of the sequences X such that $\mathrm{K}(X\upharpoonright n) \geq n - d$ for all n is at least $1 - 2^{-d}$, and thus the Π^0_1 class of such sequences of strings $(\sigma_1, \sigma_2, \ldots,)$ of high complexity (encoded as elements of $2^{\mathbb{N}}$) is not even *tt*-negligible.

The situation changes dramatically if one wishes to obtain many distinct high complexity strings of a given length. Let $\ell, f, d : \mathbb{N} \to \mathbb{N}$ be any computable functions. Consider the Π^0_1 class $\mathcal{K}_{f,\ell,d}$ whose members are sequences $\vec{F} = (F_1, F_2, F_3, \ldots)$ where for all i, F_i is a finite set of $f(i)$ strings σ of length $\ell(i)$ such that $\mathrm{K}(\sigma) \geq \ell(i) - d(i)$. Note once again that $\mathcal{K}_{f,\ell,d}$ can be viewed, modulo encoding, as a Π^0_1 subclass of $2^{\mathbb{N}}$.

THEOREM 7.7. *For all computable functions f, ℓ, d such that $f(i)/2^{d(i)}$ takes arbitrarily large values and ℓ is increasing, the class $\mathcal{K}_{f,\ell,d}$ is deep.*

PROOF. Let T be the canonical co-c.e. tree associated to the class $\mathcal{K}_{f,\ell,d}$. Just like in the proof of Theorem 7.4, we will build a discrete semi-measure m whose coding constant 2^e we know in advance and thus, by setting $m(\sigma) > 2^{-|\sigma|+d(|\sigma|)+e+1}$, we will force $\mathbf{m}(\sigma) > 2^{-|\sigma|+d(|\sigma|)+1}$, and thus $\mathrm{K}(\sigma) < |\sigma| - d(|\sigma|)$, which will consequently remove from T every sequence of sets $(F_i)_{i\in\mathbb{N}}$ such that σ belongs to some F_i.

Fix a k, and let us pick some well-chosen $i = i(k)$, to be determined later. For readability, let $f = f(i(k))$, $\ell = \ell(i(k))$, and $d = d(i(k))$. Let N be the level of T at which the first i sets $F_1 \ldots F_i$ of the sequence $\vec{F}$ of $\mathcal{K}_{f,\ell,d}$ are encoded. The k-strategy does the following.

> Step 1: Wait for a stage s at which $\mathbf{M}(T_N)[s] \geq 2^{-k}$. Up to delaying the increase of $\mathbf{M}$, we can assume that when such a stage occurs, one in fact has $\mathbf{M}(T_N)[s] = 2^{-k}$.
>
> Step 2: The members of $T_N[s]$ consist of finite sequences $F_1, \ldots, F_i$ of sets of strings where F_i contains f distinct strings of length ℓ. By the pigeonhole principle, there must be a string σ of length ℓ such that
>
> $$\mathbf{M}\big(\{(F_1, \ldots, F_i) \in T_N : \sigma \in F_i\}\big)[s] \geq 2^{-k} \cdot f \cdot 2^{-\ell}.$$

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Indeed, there is a total $\mathbf{M}$-weight 2^{-k} of possible sets F_i, each of which contains f strings of length ℓ, thus $\sum_{|\sigma|=\ell} \mathbf{M}(\{(F_1,\dots,F_i) \in T_N : \sigma \in F_i\})[s] \geq 2^{-k} \cdot f$, and there are 2^ℓ strings of length ℓ in total. Effectively find such a string σ, and set $m(\sigma) > 2^{-\ell+d+e+1}$. By the choice of σ, this causes an $\mathbf{M}$-weight of at least $2^{-k} \cdot f \cdot 2^{-\ell}$ of nodes of $T_N[s]$ to leave the tree. Then go back to Step 1.

As in the previous examples, it remains to conduct the "cost analysis". The resource here again is the weight we are allowed to assign to m, which has to be bounded by $\Gamma = 1$. At each execution of Step 2 of the k-strategy, our cost is the increase of m, which is at most of $\gamma(k) = 2^{-\ell+d+e+1}$, while an $\mathbf{M}$-weight of at least $\delta(k) = 2^{-k} \cdot f \cdot 2^{-\ell}$ leaves the tree T_N. This gives a ratio $\gamma(k)/\delta(k) = 2^{d+e+k+1}/f$, which bounds the total cost of the k-strategy, so we want $\sum_k \gamma(k)/\delta(k) \leq 1$, i.e., we want:

$$\sum_k 2^{d(i(k))+e+k+1}/f(i(k)) \leq 1.$$

By assumption, $f(i)/2^{d(i)}$ takes arbitrarily large values. It thus suffices to choose $i(k)$ such that $2^{d(i(k))+e+1}/f(i(k)) \leq 2^{-2k-1}$. ⊣

7.5. Compression functions. Our next example of a family of deep classes is given in terms of compression functions, which were introduced by Nies, Stephan, and Terwijn in [26] to provide a characterization of 2-randomness (that is, $\emptyset'$-Martin-Löf randomness) in terms of incompressibility.

DEFINITION 7.8. A K-compression function with constant $c > 0$ is a function $g : 2^{<\mathbb{N}} \to \mathbb{N}$ such that $g \leq \mathrm{K} + c$ and $\sum_{\sigma \in 2^{<\mathbb{N}}} 2^{-g(\sigma)} \leq 1$. We denote by $\mathcal{CF}_c$ be the class of compression functions with constant c.

The condition $g(\sigma) \leq \mathrm{K}(\sigma) + c$ implies that $g(\sigma) \leq 2|\sigma| + c + c'$ for some fixed constant c', and therefore $\mathcal{CF}_c$ can be seen as a Π^0_1 subclass of $2^{\mathbb{N}}$, modulo encoding the functions as binary sequences. Of course $\mathcal{CF}_c$ contains the function K itself so it is a nonempty class. We now show the following.

THEOREM 7.9. *For all $c \geq 0$, the class $\mathcal{CF}_c$ is deep.*

PROOF. Although we could give a direct proof following the same template as the previous examples, we will instead show that for all c, we have $\mathcal{K}_{f,\ell,d} \leq_s \mathcal{CF}_c$ for some computable f, ℓ, d (where $\mathcal{K}_{f,\ell,d}$ is the class we defined in the previous section) such that ℓ is increasing and $f/2^d$ takes arbitrarily large values, which by Theorem 6.4 implies the depth of $\mathcal{CF}_c$.

Let $g \in \mathcal{CF}_c$. For all n, since $\sum_{|\sigma|=n} 2^{-g(\sigma)} \leq 1$, there are at most 2^{n-1} strings of length n such that $g(\sigma) < n - 1$, and thus at least 2^{n-1} strings σ of length n such that $g(\sigma) \geq n - 1$. Since $g \in \mathcal{CF}_c$, for each σ such that $g(\sigma) \geq |\sigma| - 1$, we also have $\mathrm{K}(\sigma) \geq |\sigma| - c - 1$. Thus, given $g \in \mathcal{CF}_c$ as an oracle we can find, for each $n \geq 3$, 2^{n-1} strings σ of length n such that $\mathrm{K}(\sigma) \geq |\sigma| - c - 1$. Setting $f(i) = 2^{i+2}$, $\ell(i) = i + 3$ and $d(i) = c + 1$, we have uniformly reduced $\mathcal{K}_{f,\ell,d}$ to $\mathcal{CF}_c$. Since d is a constant function it

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

is obvious that $f(i)/2^d$ is unbounded, thus $\mathcal{K}_{f,\ell,d}$ is deep (by Theorem 7.7) and by Theorem 6.4 so is $\mathcal{CF}_c$. ⊣

The above result is not tight: the proof in fact shows that if d is not a constant function but is such that $2^n/d(n)$ takes arbitrarily large values, then the class of functions g such that $g(\sigma) \leq \mathrm{K}(\sigma) + d(|\sigma|)$ for all σ is a deep class.

7.6. A notion related to High(CR, MLR)**.** In Section 8, we discuss lowness for randomness notions. Here we look at a dual notion, highness for randomness, specifically the class High(CR, MLR), whose precise characterization is an outstanding open question in algorithmic randomness. We shall see that this class is tightly connected to the notion of depth.

DEFINITION 7.10. A sequence $A \in 2^{\mathbb{N}}$ is in the class High(CR, MLR) if $\mathsf{MLR} \supseteq \mathsf{CR}^A$.

The class High(CR, MLR) itself is not Π^0_1 (as it is closed under finite change of prefixes), but all its members have a "deep property". Bienvenu and Miller [3] proved that when A is in High(CR, MLR), then for every c.e. set of strings S such that $\sum_{\sigma\in S} 2^{-|\sigma|} < 1$, A computes a martingale d such that $d(\Lambda) = 1$ and for some fixed rational $r > 0$, $d(\sigma) > 1 + r$ for all $\sigma \in S$. It is straightforward to show that real-valued martingales can be approximated by dyadic-valued martingales with arbitrary precision; in particular one can assume that $d(\sigma)$ is dyadic for all σ (and thus can be coded using $f(|\sigma|)$ bits for some computable function f), still keeping the property $\sigma \in S \Rightarrow d(\sigma) > 1 + r$. For a well-chosen S, such martingales form a deep class.

THEOREM 7.11. *Let S be the set of strings σ such that $\mathrm{K}(\sigma) < |\sigma|$ (so that we have $\sum_{\sigma\in S} 2^{-|\sigma|} < 1$). Let g be a computable function and $r > 0$ be a rational number. Define the class $\mathcal{W}_{g,r}$ to be the set of dyadic-valued martingales d such that $d(\Lambda) = 1$, $d(\sigma) > 1 + r$ for all $\sigma \in S$ and such that for all σ, $d(\sigma)$ can be coded using $g(|\sigma|)$ bits. Then $\mathcal{W}_{g,r}$ can be viewed as a Π^0_1 class of $2^{\mathbb{N}}$, and this class is deep.*

PROOF. Again, we are going to show the depth of $\mathcal{W}_{g,r}$ by a Medvedev reduction from $\mathcal{K}_{f,\ell,h}$ for some increasing computable function ℓ and computable functions f and h such that $f/2^h$ is unbounded. More precisely, we will take $f(n) = n$, $\ell(n) = n + c$ (for a well-chosen c given below) and $h(n) = 0$. Now suppose we are given a martingale $d \in \mathcal{W}_{f,r}$. For any given length n, we have, by the martingale fairness condition,

$$\sum_{\sigma:|\sigma|=n} d(\sigma) = 2^n.$$

It follows that there are at most $2^n/(1 + r)$ strings σ of length n such that $d(\sigma) > 1 + r$, and therefore at least $2^n - 2^n/(1 + r)$ strings such that $d(\sigma) \leq 1 + r$. Having oracle access to d, such strings can be effectively found and listed. Since $2^n - 2^n/(1 + r) > n - c$ for all n and some fixed constant c, for each n we can use d to list n strings $\sigma_1, \ldots, \sigma_n$ of length $n + c$ such that $d(\sigma_i) \leq 1 + r$ for $1 \leq i \leq n$. But by definition of d, $d(\sigma) \leq 1 + r$

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

implies that $\sigma \notin S$, which further implies that $\mathrm{K}(\sigma) \geq |\sigma|$. This shows that $\mathcal{W}_{g,r}$ is above $\mathcal{K}_{f,\ell,d}$ in the Medvedev degrees. By Theorem 7.7, $\mathcal{K}_{f,\ell,h}$ is deep, and hence by Theorem 6.4, $\mathcal{W}_{g,r}$ is deep as well. ⊣

The examples of deep classes provided in this section, combined with Theorem 5.3, give us the results mentioned page 264: If X is a difference random sequence,

- it does not compute any shift-complex sequence (Khan);
- it does not compute any $\mathcal{DNC}_q$ function when q is a computable order function such that $\sum_n 1/q(n) = \infty$ (Miller);
- it does not compute any compression function (Greenberg, Miller, Nies);
- it is not in High(CR, MLR) (Greenberg, Miller, Nies).

The fact that a difference random cannot compute large sets of complex strings (in the sense of Theorem 7.7) appears to be new.

§8. **Lowness and depth.** Various lowness notions have been well-studied in algorithmic randomness, where a lowness notion is given by a collection of sequences that are in some sense computationally weak. Many lowness notions take the following form: For a relativizable collection $\mathcal{S} \subseteq 2^{\mathbb{N}}$, we say that A is *low for* $\mathcal{S}$ if $\mathcal{S} \subseteq \mathcal{S}^A$. For instance, if we let $\mathcal{S} = \mathsf{MLR}$, then the resulting lowness notion consists of the sequences that are *low for Martin-Löf random*, a collection we write as Low(MLR).

Lowness notions need not be given in terms of relativizable classes. For instance, if we let $\mathbf{m}^A$ be a universal A-left-c.e. discrete semi-measure, we can defined A to be *low for* $\mathbf{m}$ if $\mathbf{m}^A(\sigma) \leq^\times \mathbf{m}(\sigma)$ for every σ. In addition, some lowness notions are not given in terms of relativization, such as the notion of K-*triviality*, where $A \in 2^{\mathbb{N}}$ is K-trivial if and only if $\mathrm{K}(A\upharpoonright n) \leq \mathrm{K}(n) + O(1)$ (where we take n to be 1^n). Surprisingly, we have the following result (see [25] for a detailed survey of results in this direction).

THEOREM 8.1. *Let $A \in 2^{\mathbb{N}}$. The following are equivalent*:

(i) $A \in \mathrm{Low}(\mathsf{MLR})$;
(ii) *A is low for* $\mathbf{m}$;
(iii) *A is* K-*trivial*.

The cupping problem, a longstanding open problem in algorithmic randomness involving K-triviality, is to determine whether there exists a K-trivial sequence A and some Martin-Löf random sequence $X \not\geq_T \emptyset'$ such that $X \oplus A \geq_T \emptyset'$. A negative answer was recently provided by Day and Miller [5]:

THEOREM 8.2 (Day-Miller [5]). *A sequence A is K-trivial if and only if for every difference random sequence X, $X \oplus A \not\geq_T \emptyset'$.*

Using the notion of depth, we can strengthen the Day–Miller result.

THEOREM 8.3. *Let X be an incomplete Martin-Löf random sequence and A be K-trivial. Then $X \oplus A$ does not compute any member of any deep Π^0_1 class.*

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

PROOF. First, we need to partially relativize the notion of depth: for $A \in 2^{\mathbb{N}}$, a Π^0_1 class $\mathcal{C} \subseteq 2^{\mathbb{N}}$ is *deep relative to* A if there is some computable order function f such that $\mathbf{m}^A(T_{f(n)}) \leq 2^{-n}$ if and only if there is some A-computable order function g such that $\mathbf{M}^A(T_{g(n)}) \leq 2^{-n}$ (where $\mathbf{M}^A$ is a universal A-left-c.e. continuous semi-measure).

Let $\mathcal{C}$ be a deep Π^0_1 class with canonical co-c.e. tree T and let f be a computable function such that $\mathbf{m}(T_{f(n)}) \leq 2^{-n}$. Since A is low for $\mathbf{m}$, we have also have $\mathbf{m}^A(T_{f(n)}) \leq^\times 2^{-n}$, and thus $\mathcal{C}$ is deep relative to A.

Let X be a sequence such that $X \oplus A$ computes a member of $\mathcal{C}$ via a Turing functional Φ and suppose, for the sake of contradiction, that X is difference random. Let

$$\mathcal{D}_n = \{Z : \Phi^{Z\oplus A}\restriction f(n)\downarrow \in T_{f(n)}\}.$$

The set $\mathcal{D}_n$ can be written as the difference $\mathcal{U}_n \setminus \mathcal{V}_n$ of two A-effectively open sets (uniformly in n) with $\mathcal{U}_n = \{Z : \Phi^{Z\oplus A}\restriction f(n)\downarrow\}$ and $\mathcal{V}_n = \{Z : \Phi^{Z\oplus A}\restriction f(n)\downarrow \notin T_{f(n)}\}$.

We can see the functional $Z \mapsto \Phi^{Z\oplus A}$ as an A-Turing functional Ψ, and thus by the univerality of $\mathbf{M}^A$ for the class of A-left-c.e. continuous semi-measures, we have $\mathbf{M}^A \geq^\times \lambda_\Psi$. By definition of $\mathcal{D}_n$, we therefore obtain:

$$\lambda(\mathcal{D}_n) \leq \lambda_\Psi(T_{f(n)}) \leq^\times \mathbf{M}^A(T_{f(n)}) \leq 2^{-n}.$$

This shows that the sequence X, which by assumption belongs to all $\mathcal{D}_n$, is not A-difference random. It is, however, A-Martin-Löf random as A is low for Martin-Löf randomness. Relativizing Theorem 2.10 to A, this shows that $X \oplus A \geq_T A'$. But this contradicts the Day–Miller theorem (Theorem 8.2). ⊣

As we cannot compute any member of a deep class by joining a Martin-Löf random sequence with a low for Martin-Löf random sequence, it is not unreasonable to ask if there is a notion of randomness $\mathcal{R}$ such that we cannot *tt*-compute any members of a *tt*-deep class by joining an $\mathcal{R}$-random sequence with a low-for-$\mathcal{R}$ sequence. We obtain a partial answer to this question using Kurtz randomness.

From the discussion of lowness at the beginning of this section, we have $A \in \mathrm{Low}(\mathsf{KR})$ if and only if $\mathsf{KR} \subseteq \mathsf{KR}^A$. Moreover, we define the class $\mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$ to be the collection of sequences A such that $\mathsf{MLR} \subseteq \mathsf{KR}^A$. Since $\mathsf{MLR} \subseteq \mathsf{KR}$, it follows that $\mathrm{Low}(\mathsf{KR}) \subseteq \mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$. Miller and Greenberg [13] obtained the following characterization of $\mathrm{Low}(\mathsf{KR})$ and $\mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$. Recall that $A \in 2^{\mathbb{N}}$ is computably dominated if every $f \leq_T A$ is dominated by some computable function.

THEOREM 8.4 (Greenberg-Miller [13]). *Let* $A \in 2^{\mathbb{N}}$.

(i) $A \in \mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$ *if and only if* A *is of non-*DNC *degree.*
(ii) $A \in \mathrm{Low}(\mathsf{KR})$ *if and only if* $A \in \mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$ *and computably dominated.*

For $A \in 2^{\mathbb{N}}$, a Π^0_1 class $\mathcal{C}$ is *tt-negligible relative to* A if $\mu^A(\mathcal{C}) = 0$ for every A-computable measure μ^A. We first prove the following.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

PROPOSITION 8.5. *If* $A \in \mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$, *then every deep* Π^0_1 *class is tt-negligible relative to* A.

PROOF. Let $\mathcal{C}$ be a deep Π^0_1 class and T its associated canonical co-c.e. tree. Let $A \in \mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$, which by Theorem 8.4 (i) is equivalent to being of non-DNC degree. We appeal to a useful characterization of non-DNC degrees due to Hölzl and Merkle [16]: A is of non-DNC degree if and only if it is *infinitely often c.e. traceable* (hereafter, i.o. c.e. traceable). This means that there exists a computable order function h such that the following holds: for every total A-computable function $s : \mathbb{N} \to \mathbb{N}$, there exists a family $(S_n)_{n\in\mathbb{N}}$ of uniformly c.e. finite sets such that $|S_n| < h(n)$ for all n and $s(n) \in S_n$ for infinitely many n.

Let h be a computable order function witnessing the i.o. c.e. traceability of A. Since $\mathcal{C}$ is deep, let f be a computable function such that $\mathbf{M}(T_{f(n)}) \leq 2^{-2n}/h(n)$. Suppose for the sake of contradiction that $\mathcal{C}$ is not tt-negligible relative to A, which means that there exists an A-computable measure μ^A such that $\mu^A(\mathcal{C}) > r$ for some rational $r > 0$. Let $s \leq_T A$ be the function that on input n gives a rational lower-approximation, with precision $1/2$, of the values of μ^A on all strings of length $f(n)$ (encoded as an integer). By this we mean that $s(n)$ gives us for all strings σ of length $f(n)$ a rational value $s(n,\sigma)$ such that $\mu^A(\sigma)/2 \leq s(n,\sigma) \leq \mu^A(\sigma)$. Let $(S_n)_{n\in\mathbb{N}}$ witness the traceability of s, i.e., the S_n's are uniformly c.e., $|S_n| < h(n)$ for every n, and $s(n) \in S_n$ for infinitely many n.

We now build a left-c.e. continuous semi-measure ρ as follows. For all n, enumerate S_n. For each member $z \in S_n$, interpret z as a mass distribution ν on the collection of strings of length $f(n)$. Then, for each string σ of length $f(n)$, increase $\rho(\sigma)$ (as well as strings comparable with σ in any way that ensures that ρ remains a semi-measure) by $2^{-n-1}\nu(\sigma)/h(n)$. This has a total cost of $2^{-n-1}/h(n)$, and since there are at most $h(n)$ elements in S_n, the total cost at level $f(n)$ is at most 2^{-n-1}. Therefore the total cost of the construction of ρ is bounded by 1 and thus ρ is indeed a left-c.e. continuous semi-measure.

Now, for any n such that $s(n) \in S_n$ (as there are infinitely many such n), ρ distributes an amount of at least $(\mu^A(T_{f(n)})/2) \cdot (2^{-n-1}/h(n))$ on $T_{f(n)}$, and since $\mu^A(T_n) > r$, this gives

$$\rho(T_{f(n)}) \geq 2^{-n-O(1)}/h(n).$$

However, we had assumed that

$$\mathbf{M}(T_{f(n)}) \leq 2^{-2n}/h(n),$$

for all n. But since $\rho \leq^{\times} \mathbf{M}$, we get a contradiction. ⊣

The following result involves the notion of relative tt-reducibility. For a fixed $A \in 2^{\mathbb{N}}$, a $tt(A)$-functional is a total A-computable Turing functional. Equivalently, we can define a $tt(A)$-functional Ψ^A in terms of a Turing functional Φ as follows: Let Φ be defined on all inputs of the form $X \oplus A$. Then we set $\Psi^A(X) = \Phi(X \oplus A)$. Furthermore, one can show that there is an A-computable bound on the use of X in the computation (just as there is a computable bound in the use function for unrelativized tt-computations).

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

THEOREM 8.6. *Let X be Kurtz random and $A \in \mathrm{Low}(\mathsf{KR})$. Then X does not $tt(A)$-compute any member of any deep Π^0_1 class.*

PROOF. Let $\mathcal{C}$ be a deep Π^0_1 class and $A \in \mathrm{Low}(\mathsf{KR})$. By Proposition 8.5, $\mathcal{C}$ is also tt-deep relative to A. Let Φ^A be a $tt(A)$-functional. The pre-image $\mathcal{D}$ of $\mathcal{C}$ under Φ^A is a $\Pi^0_1(A)$ class, which must be tt-deep relative to A as well, by Theorem 6.5 relativized to A. Now, applying Proposition 4.5 relativized to A, $\mathcal{D}$ contains no A-Kurtz random sequence. But since A is low for Kurtz randomness, $\mathcal{D}$ contains no Kurtz-random sequence as well. ⊣

We now obtain a partial analogue of Theorem 8.3.

COROLLARY 8.7. *Let X be Kurtz random and $A \in \mathrm{Low}(\mathsf{KR})$. Then $X \oplus A$ does not tt-compute any member of any deep Π^0_1 class.*

PROOF. Let Φ be a tt-functional. Since Φ is total, it is certainly total on all sequences of the form $X \oplus A$ for $X \in 2^{\mathbb{N}}$. Thus $\Psi^A(X) = \Phi(X \oplus A)$ is a $tt(A)$-functional. By Theorem 8.6, it follows that $\Phi(X \oplus A)$ cannot be contained in any deep class. ⊣

QUESTION 8.8. *Does Corollary 8.7 still hold if we replace "deep" with "tt-deep"?*

We can extend Theorem 8.6 to the following result, which proceeds by almost the same proof, the details of which are left to the reader.

THEOREM 8.9. *Let X be Martin-Löf random and A be in $\mathrm{Low}(\mathsf{MLR}, \mathsf{KR})$. Then X does not $tt(A)$-compute any member of any deep Π^0_1 class. (In particular, $X \oplus A$ does not tt-compute any member of any deep Π^0_1 class).*

§9. Depth, mutual information, and the Independence Postulate. In this final section, we introduce the notion of mutual information and apply it to the notion of depth. Roughly, what we prove is that every member of every deep class has infinite mutual information with Chaitin's Ω, a Martin-Löf random sequence that encodes the halting problem. This generalizes a result of Levin's, that every consistent completion of PA has infinite mutual information with Ω. We conclude with a discussion of the Independence Postulate, a principle introduced by Levin to derive the statement that no consistent completion of arithmetic is physically obtainable.

9.1. The definition of mutual information. First we review the definitions of Kolmogorov complexity of a pair and the universal conditional discrete semi-measure $\mathbf{m}(\cdot \mid \cdot)$. Let $\langle \cdot, \cdot \rangle : 2^{<\mathbb{N}} \times 2^{<\mathbb{N}} \to 2^{<\mathbb{N}}$ be a computable bijection. Then we define $\mathrm{K}(\sigma, \tau) := \mathrm{K}(\langle \sigma, \tau \rangle)$. Similarly, we set $\mathbf{m}(\sigma, \tau) := \mathbf{m}(\langle \sigma, \tau \rangle)$. A conditional left-c.e. discrete semi-measure $m(\cdot \mid \cdot) : 2^{<\mathbb{N}} \times 2^{<\mathbb{N}} \to [0, 1]$ is a function satisfying $\sum_{\sigma \in 2^{<\mathbb{N}}} m(\sigma \mid \tau) \leq 1$ for every τ. Then $\mathbf{m}(\cdot \mid \cdot)$ is defined to be a universal conditional left-c.e. discrete semi-measure, so that for every conditional left-c.e. discrete semi-measure, there is some c such that $m(\sigma \mid \tau) \leq c \cdot \mathbf{m}(\sigma \mid \tau)$ for every σ and τ. Lastly, we define the conditional prefix-free Kolmogorov complexity $\mathrm{K}(\sigma \mid \tau)$ to be

$$\mathrm{K}(\sigma \mid \tau) = \min\{|\xi| : U(\langle \xi, \tau \rangle) = \sigma\},$$

where U is a universal prefix-free machine.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

The *mutual information* of two strings σ and τ, denoted by $\mathbf{I}(\sigma : \tau)$, is defined by

$$\mathbf{I}(\sigma : \tau) = \mathrm{K}(\sigma) + \mathrm{K}(\tau) - \mathrm{K}(\sigma, \tau)$$

or equivalently by

$$2^{\mathbf{I}(\sigma:\tau)} =^{\times} \frac{\mathbf{m}(\sigma, \tau)}{\mathbf{m}(\sigma) \cdot \mathbf{m}(\tau)}.$$

By the symmetry of information (see Gács [11]), we also have

$$2^{\mathbf{I}(\sigma:\tau)} =^{\times} \frac{\mathbf{m}(\sigma \mid \tau, \mathrm{K}(\tau))}{\mathbf{m}(\sigma)} =^{\times} \frac{\mathbf{m}(\tau \mid \sigma, \mathrm{K}(\sigma))}{\mathbf{m}(\tau)}. \tag{1}$$

Levin [22] extends mutual information to infinite sequences by setting

$$2^{\mathbf{I}(X:Y)} = \sum_{\sigma,\tau \in 2^{<\mathbb{N}}} \mathbf{m}^X(\sigma) \cdot \mathbf{m}^Y(\tau) \cdot 2^{\mathbf{I}(\sigma:\tau)}.$$

9.2. Mutual information and depth. Recall that Chaitin's Ω can be obtained as the probability that a universal prefix-free machine will halt on a given input, that is, $\Omega = \sum_{U(\sigma)\downarrow} 2^{-|\sigma|}$, where U is a fixed universal prefix-free machine. Generalizing a result of Levin's from [22], we have:

THEOREM 9.1. *Let $\mathcal{C}$ be a Π^0_1 class and T its associated co-c.e. tree. Suppose $\mathcal{C}$ is deep, witnessed by a computable order function f such that $\mathbf{m}(T_{f(n)}) \leq 2^{-n}$. Then for every $Y \in \mathcal{C}$ and all n,*

$$\mathbf{I}\big(\Omega\restriction n \,:\, Y\restriction f(n)\big) \geq n - O(\log n).$$

In particular,

$$\mathbf{I}(\Omega \,:\, Y) = \infty.$$

PROOF. Our proof follows the same idea Levin uses for consistent completions of PA (see [22, Theorem 1]), although some extra care is needed for arbitrary deep classes. Suppose for a given n we have an exact description τ of $T_{f(n)}$; that is, on input τ, the universal machine outputs a code for the finite set $T_{f(n)}$. By the definition of f,

$$\sum_{\sigma \in T_{f(n)}} \mathbf{m}(\sigma) \leq 2^{-n}$$

or equivalently

$$\sum_{\sigma \in T_{f(n)}} \mathbf{m}(\sigma) \cdot 2^n \leq 1.$$

Therefore, the quantity $\mathbf{m}(\sigma) \cdot 2^n \cdot \mathbf{1}_{\sigma \in T_{f(n)}}$ is a discrete semi-measure, but it is not necessarily left-c.e. since $T_{f(n)}$ is merely co-c.e. (and, in general, not c.e. by Proposition 4.4). However, it *is* a left-c.e. semi-measure relative to the exact description τ of $T_{f(n)}$. Thus, for every $\sigma \in T_{f(n)}$, by the universality of $\mathbf{m}(\cdot \mid \tau)$,

$$\mathbf{m}(\sigma \mid \tau) \geq^{\times} \mathbf{m}(\sigma) \cdot 2^n.$$

By Equation 1 above, we have

$$2^{\mathbf{I}(\sigma:\tau)} =^{\times} \frac{\mathbf{m}(\sigma \mid \tau, \mathrm{K}(\tau))}{\mathbf{m}(\sigma)} \geq^{\times} \frac{\mathbf{m}(\sigma \mid \tau)}{\mathbf{m}(\sigma)} \geq^{\times} 2^n,$$

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

and hence $\mathbf{I}(\tau : \sigma) \geq^+ n$. We would like to apply this fact to the case where $\sigma = Y\upharpoonright f(n)$ and $\tau = \Omega\upharpoonright n$. But this is not technically sufficient, as $\Omega\upharpoonright n$ does not necessarily contain enough information to exactly describe $T_{f(n)}$. This is not an obstacle in Levin's argument for completions of PA, but it is for arbitrary deep classes.

However, $\Omega\upharpoonright n$ contains enough information to get a "good enough" approximation of $T_{f(n)}$. Let us refine the idea above: suppose now that τ is no longer an exact description of $T_{f(n)}$, but is a description of a set of strings S of length $f(n)$ such that $T_{f(n)} \subseteq S$ and

$$\sum_{\sigma\in S} \mathbf{m}(\sigma) \leq^\times 2^{-n}n^2.$$

Then, by following the same reasoning as above, we would have $\mathbf{I}(\sigma : \tau) \geq^+ n - 2\log n$ for all $\sigma \in S$ (and thus all $\sigma \in T_{f(n)}$). We shall prove that $\Omega\upharpoonright n$ contains enough information to recover such a set S, thus proving the theorem.

The real number Ω is left-c.e. and Solovay complete (see [8, Section 9.1]). As a consequence, for every other left-c.e. real α, knowing the first k bits of Ω allows us to compute the first $k - O(1)$ bits of α. For all n, define:

$$a_n = \sum_{\substack{|\sigma|=f(n)\\ \sigma\notin T_{f(n)}}} \mathbf{m}(\sigma)$$

and observe that a_n is left-c.e. uniformly in n (because $T_{f(n)}$ is co-c.e. uniformly in n), and belongs to $[0, 1]$. Define now

$$\alpha = \sum_{n\in\mathbb{N}} \frac{a_n}{n^2},$$

which is also a left-c.e. real. Thus, knowing the first n bits of Ω gives us the first $n - O(1)$ bits of α, i.e., an approximation of α with precision $2^{-n}\cdot O(1)$. In particular, one can find a stage t_n such that $\alpha - \alpha[t_n] \leq 2^{-n}\cdot O(1)$, and thus $a_n - a_n[t_n] \leq 2^{-n}\cdot n^2\cdot O(1)$. By definition of a_n,

$$a_n[t_n] = \sum_{\substack{|\sigma|=f(n)\\ \sigma\notin T_{f(n)}[t_n]}} \mathbf{m}(\sigma)[t_n],$$

and since $|a_n - a_n[t_n]| \leq 2^{-n}\cdot n^2\cdot O(1)$, this implies

$$\sum_{\substack{|\sigma|=f(n)\\ \sigma\in T_{f(n)}[t_n]\setminus T_{f(n)}}} \mathbf{m}(\sigma) \leq^\times 2^{-n}\cdot n^2.$$

But recall from above that

$$\sum_{\substack{|\sigma|=f(n)\\ \sigma\in T_{f(n)}}} \mathbf{m}(\sigma) \leq 2^{-n}.$$

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

Combining these two facts, and taking S to be the set $T_{f(n)}[t_n]$, we have $T_{f(n)} \subseteq S$ and

$$\sum_{\sigma \in S} \mathbf{m}(\sigma) \leq^{\times} 2^{-n} \cdot n^2,$$

which establishes the first part of the theorem. To see that the second part of the statement follows from the first, observe that

$$\mathbf{m}^Y(Y\restriction f(n)) =^{\times} \mathbf{m}^Y(n) \geq^{\times} \mathbf{m}(n) \geq^{\times} 1/n^2,$$

$$\mathbf{m}^{\Omega}(\Omega\restriction n) =^{\times} \mathbf{m}^{\Omega}(n) \geq^{\times} \mathbf{m}(n) \geq^{\times} 1/n^2,$$

and $2^{\mathbf{I}(Y\restriction f(n) : \Omega\restriction n)} = 2^n/n^{O(1)}$ (since $\mathbf{I}\big(Y\restriction f(n) \,:\, \Omega\restriction n\big) \geq^{+} n - 2\log n$ as established above). Then we have

$$\begin{aligned} 2^{\mathbf{I}(\Omega : Y)} &= \sum_{\sigma,\tau \in 2^{<\mathbb{N}}} \mathbf{m}^{\Omega}(\sigma) \cdot \mathbf{m}^Y(\tau) \cdot 2^{\mathbf{I}(\sigma : \tau)} \\ &\geq \sum_{n \in \mathbb{N}} \mathbf{m}^{\Omega}(Y\restriction f(n)) \cdot \mathbf{m}^Y(\Omega\restriction n) \cdot 2^{\mathbf{I}(Y\restriction f(n) : \Omega\restriction n)} \\ &\geq \sum_{n \in \mathbb{N}} 2^n/n^{O(1)} = \infty. \end{aligned}$$

⊣

REMARK 9.2. The converse of this theorem does not hold, i.e., there is a Π^0_1 class that is not deep but all of whose elements have infinite mutual information with Ω. This follows from Theorem 4.7 and the fact that having infinite mutual information with Ω is a property that is invariant under addition or deletion of a finite prefix.

Levin's proof of Theorem 9.1 restricted to the particular case of completions of PA is the mathematical part of a more general discussion, the other part of which is philosophical in nature (see [22] and [21]). While Gödel's theorem asserts that no completion of PA can be *computably* obtained, Levin's goal is to show that no completion of PA can be obtained by any physical means whatsoever (computationally or otherwise), thus generalizing Gödel's theorem. Levin does not fully specify what he means by physically obtainable (the exact term he uses is "located in the physical world"), but nonetheless he makes the following postulate, which he dubs the "Independence Postulate": if σ is a string which can be unambiguously defined by an n-bit mathematical formula[2] (say, in ZFC) and τ can be located in the physical world with a k-bit description, then for a fixed small constant c (independent of σ and τ), one has $\mathbf{I}(\sigma \,:\, \tau) < n + k + c$.

In particular, if one admits that some infinite sequences are physically obtainable, the Independence Postulate for infinite sequences says that if X and Y are two infinite sequences with X mathematically definable and Y

[2] A caveat: having an unambiguous mathematical definition does not mean that the string can be computably reconstructed from this description. For example, "the first 2^n bits of the halting problem" (written as a mathematical formula) unambiguously defines an object but does not in any way give us access to the actual value of this object. Thus, an object with an n-bit definition *need not* have Kolmogorov complexity less than or equal to $n + O(1)$.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

physically obtainable, then $\mathbf{I}(X : Y) < \infty$. Being Δ^0_2, Ω is mathematically definable, and, as Levin shows, $\mathbf{I}(\Omega : Y) = \infty$ for any completion Y of PA. Thus, assuming the Independence Postulate, no completion of PA is physically obtainable.

Our Theorem 9.1 extends Levin's theorem and, assuming the Independence Postulate, shows that no member of a deep class (shift-complex sequences, compression functions, etc.) is physically obtainable. Of course, evaluating the validity of the Independence Postulate would require an extended philosophical discussion that would take us well beyond the scope of this paper.

In any case, whether or not the reader accepts the Independence Postulate, Theorem 9.1 is interesting in its own right. In fact, it is quite surprising because it seems to contradict the "basis for randomness theorem" (see [8, Theorem 8.7.2]), which states that if X is a Martin-Löf random sequence and $\mathcal{C}$ is a Π^0_1 class, then there exists a member Y of $\mathcal{C}$ such that X is random relative to Y. If a sequence X is random relative to another sequence Y, the intuition is that Y "knows nothing about X", and thus one could conjecture that $\mathbf{I}(X : Y) < \infty$. However, this cannot always be the case, since by Theorem 9.1, $\mathbf{I}(\Omega : Y) = \infty$ for *all* members Y of a deep Π^0_1 class $\mathcal{C}$, even though Ω is random relative to some $Y \in \mathcal{C}$.

This apparent paradox can be explained by taking a closer look at the definition of mutual information. Let $\mathcal{C}$ be a deep Π^0_1 class, whose canonical co-c.e. tree T satisfies $\mathbf{m}(T_{f(n)}) \leq 2^{-n}$ for some computable function f. By Theorem 9.1 and the symmetry of information, for every $Y \in \mathcal{C}$ we have

$$\begin{aligned} \mathrm{K}(\Omega{\upharpoonright}n) - \mathrm{K}\big(\Omega{\upharpoonright}n \mid Y{\upharpoonright}f(n), k_n\big) \\ = \mathbf{I}\big(\Omega{\upharpoonright}n : Y{\upharpoonright}f(n)\big) - O(1) \geq n - O(\log n), \end{aligned} \quad (2)$$

where k_n stands for $\mathrm{K}(Y{\upharpoonright}f(n))$. Take a $Y \in \mathcal{C}$ such that Ω is random relative to Y. It is well-known that a sequence Z is random if and only if $K(Z{\upharpoonright}n \mid n) \geq n - O(1)$ (see for example Gács [12]). Applying this fact (relativized to Y) to Ω, we have

$$\mathrm{K}^Y(\Omega{\upharpoonright}n \mid n) \geq n - O(1)$$

and thus in particular that

$$\mathrm{K}\big(\Omega{\upharpoonright}n \mid Y{\upharpoonright}f(n)\big) \geq n - O(1).$$

Since $\mathrm{K}(\Omega{\upharpoonright}n) \leq n + O(\log n)$, it follows that

$$\mathrm{K}(\Omega{\upharpoonright}n) - \mathrm{K}\big(\Omega{\upharpoonright}n \mid Y{\upharpoonright}f(n)\big) \leq O(\log n). \quad (3)$$

The only difference between (2) and (3) is the term $k_n = \mathrm{K}(Y{\upharpoonright}f(n))$. But it makes a big difference, as one can verify that

$$\mathrm{K}\big(\Omega{\upharpoonright}n \mid Y{\upharpoonright}f(n)\big) - \mathrm{K}\big(\Omega{\upharpoonright}n \mid Y{\upharpoonright}f(n), k_n\big) \geq n - O(\log n).$$

Informally, while Y "knows nothing" about Ω, the complexity of its initial segments, seen as a function, does. In particular, the change of complexity caused by k_n implies that $\mathrm{K}\big(\mathrm{K}(Y{\upharpoonright}f(n)\big) \geq n - O(\log n)$ and thus $\mathrm{K}(Y{\upharpoonright}f(n)) \geq 2^n/n^{O(1)}$.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

§10. **Acknowledgements.** We would like to thank Noam Greenberg, Rupert Hölzl, Mushfeq Khan, Leonid Levin, Joseph Miller, André Nies, Paul Shafer, and Antoine Taveneaux for many fruitful discussions on the subject. We would also like to thank the anonymous referees for a number of helpful suggestions. Lastly, we are particularly grateful to Steve Simpson, who provided very detailed feedback on the first arXiv version of this paper. Both authors acknowledge the support of the John Templeton Foundation through the grant "Structure and Randomness in the Theory of Computation."

REFERENCES

[1] Charles H. Bennett, *Logical depth and physical complexity*, ***The Universal Turing Machine: A Half-century Survey***, Springer, New York, 1995, pp. 207–235.

[2] Laurent Bienvenu, Rupert Hölzl, Christopher P. Porter, and Paul Shafer, *Randomness and semi-measures*, ***Notre Dame Journal of Formal Logic***, preprint, arXiv:1310.5133, 2014.

[3] Laurent Bienvenu and Joseph S. Miller, *Randomness and lowness notions via open covers*. ***Annals of Pure and Applied Logic***, vol. 163 (2012), no. 5, pp. 506–518.

[4] Laurent Bienvenu and Christopher P. Porter, *Strong reductions in effective randomness*. ***Theoretical Computer Science***, vol. 459 (2012), pp. 55–68.

[5] Adam Day and Joseph S. Miller, *Density, forcing, and the covering problem*, ***Mathematical Research Letters***, vol. 22 (2015), no. 3, submitted.

[6] Karel de Leeuw, Edward F. Moore, Claude Shannon, and Norman Shapiro, *Computability by probabilistic machines*, ***Automata Studies***, Princeton University Press, Princeton, NJ, 1956.

[7] Rodney Downey, Noam Greenberg, and Joseph S. Miller, *The upward closure of a perfect thin class*. ***Annals of Pure and Applied Logic***, vol. 156 (2008), pp. 51–58.

[8] Rodney Downey and Denis Hirschfeldt, ***Algorithmic Randomness and Complexity***, Theory and Applications of Computability, Springer, New York, 2010.

[9] Bruno Durand, Leonid Levin, and Alexander Shen, *Complex tilings*. ***The Journal of Symbolic Logic***, vol. 73 (2008), no. 2, pp. 593–613.

[10] Johanna N.Y. Franklin and Keng Meng Ng, *Difference randomness*. ***Proceedings of the American Mathematical Society***, vol. 139 (2011), pp. 345–360.

[11] Peter Gács, *Lecture notes on descriptional complexity and randomness*, manuscript, available at http://www.cs.bu.edu/fac/gacs/recent-publ.html.

[12] ———, *Exact expressions for some randomness tests*, ***Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik***, vol. 26 (1980), pp. 385–394.

[13] Noam Greenberg and Joseph S. Miller, *Lowness for Kurtz randomness*, this Journal, vol. 74 (2009), no. 2, pp. 665–678.

[14] ———, *Diagonally non-recursive functions and effective Hausdorff dimension*. ***Bulletin of the London Mathematical Society***, vol. 43 (2011), no. 4, pp. 636–654.

[15] Noam Greenberg, Joseph S. Miller, and André Nies, *PA-completeness and its weakenings*, in preparation.

[16] Rupert Hölzl and Wolfgang Merkle, *Traceable sets*, ***IFIP TCS***, IFIP Advances in Information and Communication Technology, no. 323, Springer, Springer Berlin Heidelberg, 2010, pp. 301–315.

[17] Carl Jockusch and Robert Soare, Π^0_1 *classes and degrees of theories*. ***Transaction of the American Mathematical Society***, vol. 173 (1972), pp. 33–56.

[18] Steven M. Kautz, ***Degrees of Random Sequences***, Ph.D. Thesis, Cornell University, 1991.

[19] Mushfeq Khan, *Shift-complex sequences*, this Bulletin, vol. 19 (2013), no. 2, pp. 199–215.

[20] Stuart Kurtz, ***Randomness and genericity in the degrees of unsolvability***, Ph. D dissertation, University of Illinois at Urbana, 1981.

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press

[21] Leonid A. Levin, *Randomness conservation inequalities; information and independence in mathematical theories*. ***Information and Control***, vol. 61 (1984), no. 1, pp. 15–37.

[22] ———, *Forbidden information*. ***Journal of the ACM***, vol. 60 (2013), no. 2, p. 9.

[23] Leonid A. Levin and Vladimir V. V'yugin, *Invariant properties of informational bulks*, ***Lecture Notes in Computer Science***, vol. 53 (1977), pp. 359–364.

[24] Leonid A. Levin and Alexander K. Zvonkin, *The complexity of finite objects and the basing of the concepts of information and randomness on the theory of algorithms*. ***Uspekhi Matematicheskikh Nauk***, vol. 25 (1970), no. 6 (156), pp. 85–127.

[25] André Nies, ***Computability and Randomness***, Oxford Logic Guides, vol. 51, Oxford University Press, Oxford, 2009.

[26] André Nies, Frank Stephan, and Sebastiaan A. Terwijn, *Randomness, relativization and Turing degrees*. ***The Journal of Symbolic Logic***, vol. 70 (2005), no. 2, pp. 515–535.

[27] Andrey Yu. Rumyantsev, *Everywhere complex sequences and the probabilistic method*, ***STACS***, LIPIcs, vol. 9, 2011, pp. 464–471.

[28] Gerald Sacks, ***Degrees of Unsolvability***, Princeton University Press, Princeton, NJ, 1963.

[29] Stephen G. Simpson, *Mass problems and randomness*, this Bulletin, vol. 11 (2005), pp. 1–27.

[30] ———, *An extension of the recursively enumerable Turing degrees*. ***Journal of the London Mathematical Society***, vol. 75 (2006), p. 2007.

[31] ———, *Mass problems associated with effectively closed sets*. ***Tohoku Mathematical Journal***, vol. 63 (2011), pp. 489–517.

[32] Frank Stephan, *Martin-Löf random and PA-complete sets*, ***Proceedings of ASL Logic Colloquium 2002***, ASL Lecture Notes in Logic, vol. 27, 2006, pp. 342–348.

[33] Vladimir V. V'yugin, *Algebra of invariant properties of binary sequences*. ***Problemy Peredachi Informatsii***, vol. 18 (1982), no. 2, pp. 83–100.

LABORATOIRE CNRS J.-V. PONCELET
119002, BOLSHOY VLASYEVSKIY PEREULOK 11
MOSCOW, RUSSIA
E-mail: laurent.bienvenu@computability.fr

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611- 8105, USA
E-mail: cp@cpporter.com

https://doi.org/10.1017/bsl.2016.9 Published online by Cambridge University Press