Randomness Deficiency Overlap (Extended Version)

Samuel Epstein[*]

July 5, 2023

Abstract

In this paper we prove a lower bound on the computable measure of sets with high randomness deficiency with respect to two computable measures. We extend this result to computable metric spaces and univseral uniform tests.

1 Introduction

In a previous paper, a lower bound was proved on the computable measure of sets with high randomness deficiency. The deficiency of randomness of an infinite sequence $\alpha \in \{0,1\}^\infty$ with respect to a computable measure P over $\{0,1\}^\infty$ is defined to be $\mathbf{D}(\alpha|P) = \sup_n -\log P(\alpha[0..n]) - \mathbf{K}(\alpha[0..n])$. The term $\mathbf{K}$ is the prefix free Kolmogorov complexity.

Theorem. [Eps22] *For computable measures μ and nonatomic λ over $\{0,1\}^\infty$ and $n \in \mathbb{N}$, $\lambda\{\alpha : \mathbf{D}(\alpha|\mu) > n\} > 2^{-n-\mathbf{K}(n,\mu,\lambda)-O(1)}$.*

This paper generalizes this result, in the context of overlap between the randomness deficiency function with respect to two different computable probability measures.

Theorem. *For computable measures μ, ρ and nonatomic λ over $\{0,1\}^\infty$ and $n \in \mathbb{N}$, $\lambda\{\alpha : \mathbf{D}(\alpha|\mu) > n$ and $\mathbf{D}(\alpha|\rho) > n\} > 2^{-n-\mathbf{K}(n,\mu,\rho,\lambda)-O(1)}$.*

The $O(1)$ term is dependent solely on the choice of the universal Turing machine. With a bit of work, the theorem can be proved for potentially uncomputable λ that has finite mutual information with the halting sequence. With an $O(\log m)$ loss of precision, the above theorem can be generalized to m probability measures. It is possible to see this theorem being referenced in proofs of more sophisticated theorems. This theorem is of note because it factors out the mutual information with the halting sequence term that is so prevalent in the resultant theorems from similar proofs to the ones found in the paper. Lemma 3 is a reworking of Lemma 2 in [Eps23a], Lemma 5 is a reworking of Lemma 4 in [Eps23b], Theorem 2 is a reworking of Theorem 3 in [Eps22], and Theorem 8 is a reworking of Theorem 8 in [Eps23c]. The tight bounds of the main theorem derived from lemmas with looser bounds is achieved through relativization.

In Appendix A, a more direct, accesible proof that doesn't rely on the mutual information with the halting sequence is given. In addition, it uses uniformly computable probability measures. However the results are incompatible with the generalization to computable metric spaces and uncomputable λ.

[*]JP Theory Group. samepst@jptheorygroup.org

1.1 Universal Uniform Tests

The study of randomness of computable metric spaces can be seen in the works of [HR09, Ǵ21]. These spaces are important because physical random phenomena are modeled using infinite objects, and not the Cantor space. For definitions in this introduction, we use [HR09]. A computable metric space $\mathcal{X}$ is a metric space with a dense set of ideal points on which the distance function is computable. A computable probability is defined by a computable sequence of converging points in the corresponding space of Borel probability measures, $\mathcal{M}(\mathcal{X})$, over $\mathcal{X}$. A uniform test T takes in a description of a probability measure μ and produces a lower computable μ test, with $\int_{\mathcal{X}} T^{\mu} d\mu \leq 1$. There exists a universal test, $\mathbf{t}$, where for any uniform test T there is a $c_T \in \mathbb{N}$ where $c_T\mathbf{t} > T$. We extend the main theorem of this paper to computable metric spaces and universal uniform tests. The $O(1)$ constant is dependent solely on the universal Turing machine. A program for a computable probability measure is an algorithm that produces a fast Cauchy sequence converging to that probability measure in the measure space $\mathcal{M}(\mathcal{X})$. A program for the computable metric space $\mathcal{X}$ is an algorithm can compute the distance measure between its dense ideal points. The $\mathbf{K}(n, \lambda, \mu, \rho, \mathcal{X})$ term is the length of the shortest program that computes n, and a program for λ, μ, ρ, and $\mathcal{X}$.

Theorem. *Given computable probability measures μ, ρ, and λ, non-atomic λ, over a computable metric space $\mathcal{X}$ and universal uniform test $\mathbf{t}$, for all n,*
$\lambda(\{\alpha : \mathbf{t}_\mu(\alpha) > 2^n \text{ and } \mathbf{t}_\rho(\alpha) > 2^n\}) > 2^{-n-\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})-\mathbf{K}(n,\lambda,\mu,\rho,\mathcal{X})-O(1)}$.

2 Conventions

The function $\mathbf{K}(x|y)$ is the conditional prefix Kolmogorov complexity. $\mathbf{m}(x|y)$ is the conditional algortihmic probability. The mutual information between two strings $x, y \in \{0,1\}^*$, is $\mathbf{I}(x : y) = \mathbf{K}(x)+\mathbf{K}(y)-\mathbf{K}(x,y)$. For probability p over $\mathbb{N}$, randomness deficiency is $\mathbf{d}(a|p,b) = \lfloor -\log p(a)\rfloor - \mathbf{K}(a|\langle p\rangle, b)$ and measures the extent of the refutation against the hypothesis p given the result a [Ǵ21]. $\mathbf{d}(a|p) = \mathbf{d}(a|p,\emptyset)$. The amount of information that the halting sequence $\mathcal{H} \in \{0,1\}^\infty$ has about $a \in \{0,1\}^*$, conditional to $y \in \{0,1\}^*$ is $\mathbf{I}(a;\mathcal{H}|y) = \mathbf{K}(a|y) - \mathbf{K}(a|y,\mathcal{H})$. $\mathbf{I}(a;\mathcal{H}) = \mathbf{I}(a;\mathcal{H}|\emptyset)$. We use $<^+ f$ to denote $< f+O(1)$ and $<^{\log} f$ to denote $< f+O(\log(f+1))$. In addition $\overset{*}{<} f$, $\overset{*}{>} f$ denote $< f/O(1)$, $> f/O(1)$. The term $\overset{*}{=} f$ denotes $\overset{*}{<} f$ and $\overset{*}{>} f$. Stochasticity is $\mathbf{Ks}(a|b) = \min\{\mathbf{K}(Q|b) + 3\log\max\{\mathbf{d}(a|Q,b),1\} : Q \text{ has finite support and a range in } \mathbb{Q}\}$. $\mathbf{Ks}(a|b) < \mathbf{Ks}(a) + O(\log \mathbf{K}(b))$. For a mathematical statement A, let $[A] = 1$ if A is true and $[A] = 0$, otherwise. The chain rule gives $\mathbf{K}(x,y) =^+ \mathbf{K}(x|y,\mathbf{K}(y)) + \mathbf{K}(y)$. The randomness deficiency of $\alpha \in \{0,1\}^\infty$ with respect to computable continuous probability measure P is $\mathbf{D}(\alpha|P) = \sup_n -\log P(\alpha[0..n]) - \mathbf{K}(\alpha[0..n])$. The following definition is from [Lev74] .

Definition 1 (Information) *For infinite sequences $\alpha, \beta \in \{0,1\}^\infty$, their mutual information is defined to be $\mathbf{I}(\alpha:\beta) = \log\sum_{x,y\in\{0,1\}^*} 2^{\mathbf{I}(x:y)-\mathbf{K}(x|\alpha)-\mathbf{K}(y|\beta)}$.*

There are many proofs in the literature that non-stochastic numbers have high mutual information with the halting sequence. One such detailed proof is in [Eps21].

Lemma 1 $\mathbf{Ks}(x) < \mathbf{I}(x;\mathcal{H}) + O(\mathbf{K}(\mathbf{I}(x;\mathcal{H})))$.

Lemma 2 ([Eps22]) *For partial computable f, $\mathbf{I}(f(x):\mathcal{H}) <^+ \mathbf{I}(x;\mathcal{H}) + \mathbf{K}(f)$.*

Theorem 1 ([Ver21, Lev74, Gei12]) $\Pr_\mu(\mathbf{I}(\alpha:\mathcal{H}) > n) \overset{*}{<} 2^{-n+\mathbf{K}(\mu)}$.

3 On Exotic Sets of Natural Numbers

Lemma 3 *For computable probabilities p, q over $\mathbb{N}$, $D\subset\mathbb{N}$, $|D| = 2^s$, $s < \max_{a\in D}\min\{\mathbf{d}(a|p), \mathbf{d}(a|q)\}+ \mathbf{I}(D;\mathcal{H}) + O(\mathbf{K}(\mathbf{I}(D;\mathcal{H}), p, q, s))$.*

Proof. We relativize the universal Turing machine to $\langle s, p, q\rangle$. Let Q be a probability measure that realizes $\mathbf{Ks}(D)$, with $d = \max\{\mathbf{d}(D|Q), 1\}$. Let $F \subseteq \mathbb{N}$ be a random set where each element $a \in \mathbb{N}$ is selected independently with probability $cd2^{-s}$, where $c \in \mathbb{N}$ is chosen later. $\mathbf{E}[p(F)] = \mathbf{E}[q(F)] \leq cd2^{-s}$. Furthermore

$$\mathbf{E}[Q(\{G : |G| = 2^s, G \cap F = \emptyset\})] \leq \sum_G Q(G)(1 - cd2^{-s})^{2^s} < e^{-cd}.$$

Thus finite $W \subset \mathbb{N}$ can be chosen such that $p(W) \leq 4cd2^{-s}$, $q(W) \leq 4cd2^{-s}$, and $Q(\{G : |G| = 2^s, G\cap W = \emptyset\}) \leq e^{2-cd}$. $D\cap W \neq \emptyset$, otherwise, using the Q-test, $t(G) = e^{cd-1}$ if $(|G| = 2^s, G\cap W = \emptyset)$ and $t(G) = 0$ otherwise, we have

$$\begin{aligned}\mathbf{K}(D|Q, d, c) &<^+ -\log Q(D) - (\log e)cd\\ (\log e)cd &<^+ -\log Q(D) - \mathbf{K}(D|Q) + \mathbf{K}(d, c)\\ (\log e)cd &<^+ d + \mathbf{K}(d, c),\end{aligned}$$

which is a contradiction for large enough c. Thus there is an $a \in D \cap W$, where

$$\begin{aligned}\mathbf{K}(a) &<^+ \min\{-\log q(a), -\log p(a)\} + \log d - s + \mathbf{K}(d) + \mathbf{K}(Q)\\ s &<^+ \min\{\mathbf{d}(a|p), \mathbf{d}(a|q)\} + \mathbf{Ks}(D).\end{aligned}$$

Making the relativization of $\langle s, p, q\rangle$ explicit, and using Lemma 1 results in

$$\begin{aligned}s &<^+ \min\{\mathbf{d}(a|p), \mathbf{d}(a|q)\} + \mathbf{Ks}(D) + O(\mathbf{K}(s, p, q))\\ s &< \max_{a\in D}\min\{\mathbf{d}(a|p), \mathbf{d}(a|q)\} + \mathbf{Ks}(D) + O(\mathbf{K}(s, p, q))\\ s &< \max_{a\in D}\min\{\mathbf{d}(a|p), \mathbf{d}(a|q)\} + \mathbf{I}(D;\mathcal{H}) + O(\mathbf{K}(\mathbf{I}(D;\mathcal{H}), s, p, q))).\square\end{aligned}$$

4 On Exotic Sets of Reals

Let $\Omega = \sum\{2^{-\|p\|} : U(p) \text{ halts}\}$ be Chaitin's Omega, $\Omega_n \in \mathbb{Q}_{\geq 0}$ be be the rational formed from the first n bits of Ω, and $\Omega^t = \sum\{2^{-\|p\|} : U(p) \text{ halts in time } t\}$. For $n \in \mathbb{N}$, let $\mathbf{bb}(n) = \min\{t : \Omega_n < \Omega^t\}$. $\mathbf{bb}^{-1}(m) = \arg\min_n\{\mathbf{bb}(n-1) < m \leq \mathbf{bb}(n)\}$. Let $\Omega[n] \in \{0,1\}^*$ be the first n bits of Ω.

Lemma 4 *For $n = \mathbf{bb}^{-1}(m)$, $\mathbf{K}(\Omega[n]|m, n) = O(1)$.*

Proof. For a string x, let $BB(x) = \inf\{t : \Omega^t > 0.x\}$. Enumerate strings of length n, starting with 0^n, and return the first string x such that $BB(x) \geq m$. This string x is equal to $\Omega[n]$, otherwise let y be the largest common prefix of x and $\Omega[n]$. Thus $BB(y) = \mathbf{bb}(\|y\|) \geq BB(x) \geq m$, which means $\mathbf{bb}^{-1}(m) \leq \|y\| < n$, causing a contradiction. $\square$

The following lemma, while lengthy, is a series of straightforward application of inequalities.

Lemma 5 *For computable probabilities P, Q, over $\{0,1\}^\infty$, $Z \subset \{0,1\}^\infty$, $|Z| = 2^s$, $s < \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{I}(\langle Z\rangle : \mathcal{H}) + O(\mathbf{K}(s,P,Q) + \log \mathbf{I}(\langle Z\rangle;\mathcal{H}))$.*

Proof. We relativize the universal Turing machine to s, which can be done due to the precision of the theorem. Let $Z_n = \{\alpha[0..n] : \alpha \in Z\}$ and $m = \arg\min_m |Z_m| = |Z|$. Let $n = \mathbf{bb}^{-1}(m)$ and $k = \mathbf{bb}(n)$. Let p and q be probabilities over $\{0,1\}^*$, where $p(x) = [\|x\| = k]P(x)$ and $\langle p\rangle = \langle k, P\rangle$ and let $q(x) = [\|x\| = k]Q(x)$ and $\langle q\rangle = \langle k, P\rangle$. Using $D = Z_k$, Lemma 3, relativized to k, produces $x \in Z_k$, where

$$\begin{aligned}
s &< \min\{\mathbf{d}(x|p), \mathbf{d}(x|q)\} + \mathbf{I}(Z_k;\mathcal{H}|k) + O(\mathbf{K}(\mathbf{I}(Z_k;\mathcal{H}|k), q, p|k)) \\
&< \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{K}(Z_k|k) + \mathbf{K}(k) - \mathbf{K}(Z_k|k,\mathcal{H}) + O(\mathbf{K}(\mathbf{I}(Z_k;\mathcal{H}|k), q, p|k)). \\
&< \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{K}(Z_k|k) + \mathbf{K}(k) - \mathbf{K}(Z_k|k,\mathcal{H}) + O(\mathbf{K}(P,Q) + \log \mathbf{I}(Z_k;\mathcal{H}|k)).
\end{aligned}$$

Since $\mathbf{K}(k) <^+ n + \mathbf{K}(n)$, by the chain rule,

$$\begin{aligned}
&\mathbf{K}(Z_k|k) + \mathbf{K}(k) \\
<^+ &\mathbf{K}(Z_k|k, \mathbf{K}(k)) + \mathbf{K}(\mathbf{K}(k)|k) + \mathbf{K}(k) \\
< &\mathbf{K}(Z_k, k) + O(\log n) \\
< &\mathbf{K}(Z_k) + O(\log n).
\end{aligned}$$

So

$$s < \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{K}(Z_k) - \mathbf{K}(Z_k|k,\mathcal{H}) + O(\log n + \mathbf{K}(P,Q) + \log \mathbf{I}(Z_k;\mathcal{H}|k))).$$

Since $\mathbf{K}(k|n,\mathcal{H}) = O(1)$, $\mathbf{K}(Z_k|\mathcal{H}) <^+ \mathbf{K}(Z_k|k,\mathcal{H}) + \mathbf{K}(n)$,

$$s < \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{I}(Z_k;\mathcal{H}) + O(\log n + \mathbf{K}(P,Q) + \log \mathbf{I}(Z_k;\mathcal{H}|k)).$$

Furthermore since $\mathbf{I}(Z_k;\mathcal{H}|k) + \mathbf{K}(k) < \mathbf{I}(Z_k;\mathcal{H}) + O(\log n)$,

$$s < \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{I}(Z_k;\mathcal{H}) + O(\log n + \mathbf{K}(P,Q)) + O(\log \mathbf{I}(Z_k;\mathcal{H})).$$

By Lemma 4, $\mathbf{K}(\Omega[n]|Z_k) <^+ \mathbf{K}(n)$ so by Lemma by 2,

$$n <^{\log} \mathbf{I}(\Omega[n];\mathcal{H}) <^{\log} \mathbf{I}(Z_k;\mathcal{H}) + \mathbf{K}(n) <^{\log} \mathbf{I}(Z_k;\mathcal{H}).$$

The above equation used the common fact that the first n bits of Ω has $n - O(\log n)$ bits of mutual information with $\mathcal{H}$. So

$$s < \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{I}(Z_k;\mathcal{H}) + O(\mathbf{K}(P,Q) + \log \mathbf{I}(Z_k;\mathcal{H})).$$

By the definition of mutual information $\mathbf{I}$ between infinite sequences

$$\mathbf{I}(Z_k;\mathcal{H}) <^+ \mathbf{I}(Z:\mathcal{H}) + \mathbf{K}(Z_k|Z) <^{\log} \mathbf{I}(Z:\mathcal{H}) + \mathbf{K}(k|Z).$$

Now m is simple relative to Z and by Lemma 4, $\Omega[n]$ is simple relative to m and n. Furthermore k is simple relative to $\Omega[n]$. Therefore $\mathbf{K}(Z_k|Z) <^+ \mathbf{K}(n)$. So

$$\begin{aligned}
s &< \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{I}(Z:\mathcal{H}) + O(\log n) + O(\mathbf{K}(P,Q) + \log \mathbf{I}(Z;\mathcal{H})) \\
s &< \max_{\alpha\in Z} \min\{\mathbf{D}(\alpha|P), \mathbf{D}(\alpha|Q)\} + \mathbf{I}(Z:\mathcal{H}) + O(\mathbf{K}(s,P,Q) + \log \mathbf{I}(Z;\mathcal{H}))).
\end{aligned}$$

□

5 Asymptotic Properties of Randomness Deficiency

Theorem 2 *For computable measures μ,ρ and nonatomic λ over $\{0,1\}^\infty$ and $n\in\mathbb{N}$,* $\lambda\{\alpha:\mathbf{D}(\alpha|\mu)>n$ *and* $\mathbf{D}(\alpha|\rho)>n\}>2^{-n-\mathbf{K}(n,\mu,\rho,\lambda)-O(1)}$.

Proof. We first assume not. For all $c\in\mathbb{N}$, there exist computable nonatomic measures μ, ρ λ, and there exists n, where $\lambda\{\alpha:\mathbf{D}(\alpha|\mu)>n$ and $\mathbf{D}(\alpha|\rho)>n\}\leq 2^{-n-\mathbf{K}(n,\mu,\lambda)-c}$. Sample $2^{n+\mathbf{K}(n,\mu,\rho,\lambda,\rho)+c-1}$ elements $D\subset\{0,1\}^\infty$ according to λ. The probability that all samples $\beta\in D$ has $\mathbf{D}(\beta|\mu)\leq n$ or $\mathbf{D}(\beta|\rho)\leq n$ is

$$\prod_{\beta\in D}\lambda\{\mathbf{D}(\beta|\mu)\leq n \text{ or } \mathbf{D}(\beta|\rho)\leq n\}\geq$$
$$(1-|D|2^{-n-\mathbf{K}(n,\mu,\lambda,\rho)-c})\geq$$
$$(1-2^{n+\mathbf{K}(n,\mu,\lambda,\rho)+c-1}2^{-n-\mathbf{K}(n,\mu,\lambda,\rho)-c})\geq 1/2.$$

Let $\lambda^{n,c}$ be the probability of an encoding of $2^{n+\mathbf{K}(n,\mu,\lambda)+c-1}$ elements each distributed according to λ. Thus

$$\lambda^{n,c}(\text{Encoding of } 2^{n+\mathbf{K}(n,\mu,\lambda,\rho)+c-1} \text{ elements } \beta\text{, each having } \mathbf{D}(\beta|\mu)\leq n \text{ or } \mathbf{D}(\beta|\rho)\leq n)\geq 1/2.$$

Let v be a shortest program to compute $\langle n,\mu,\rho,\lambda\rangle$. By Theorem 1, with the universal Turing machine relativized to v,

$$\lambda^{n,c}(\{\gamma:\mathbf{I}(\gamma:\mathcal{H}|v)>m\})\overset{*}{<}2^{-m+\mathbf{K}(\lambda^{n,c}|v)}\overset{*}{<}2^{-m+\mathbf{K}(n,\mathbf{K}(n,\mu,\lambda,\rho),c,\lambda|v)}\overset{*}{<}2^{-m+\mathbf{K}(c)}.$$

Therefore,

$$\lambda^{n,c}(\{\gamma:\mathbf{I}(\gamma:\mathcal{H}|v)>\mathbf{K}(c)+O(1)\})\leq 1/4.$$

Thus, by probabilistic arguments, there exists $\alpha\in\{0,1\}^\infty$, such that $\alpha=\langle D\rangle$ is an encoding of $2^{n+\mathbf{K}(n,\mu,\rho,\lambda)+c-1}$ elements $\beta\in D\subset\{0,1\}^\infty$, where each β has $\mathbf{D}(\beta|\mu)\leq n$ or $\mathbf{D}(\beta|\rho)\leq n$ and $\mathbf{I}(\alpha:\mathcal{H}|v)<^+\mathbf{K}(c)$. By Lemma 5, relativized to v, there are constants $d,f,g\in\mathbb{N}$ where

$$\begin{aligned}
m=\log|D| &< \max_{\beta\in D}\min\{\mathbf{D}(\beta|\mu,v),\mathbf{D}(\beta|\rho,v)\}+2\mathbf{I}(D:\mathcal{H}|v)+d\mathbf{K}(m|v)+f\mathbf{K}(\mu|v)+g\mathbf{K}(\rho|v)\\
m &< \max_{\beta\in D}\min\{\mathbf{D}(\beta|\mu),\mathbf{D}(\beta|\rho)\}+\mathbf{K}(v)+2\mathbf{I}(D:\mathcal{H}|v)+d\mathbf{K}(m|v)+f\mathbf{K}(\mu|v)+g\mathbf{K}(\rho|v)\\
&<^+ n+\mathbf{K}(n,\mu,\lambda,\rho)+d\mathbf{K}(m|v)+2\mathbf{K}(c)+(f+g)O(1). \qquad (1)
\end{aligned}$$

Therefore:

$$m=n+\mathbf{K}(n,\mu,\rho,\lambda)+c-1$$
$$\mathbf{K}(m|v)<^+\mathbf{K}(c).$$

Plugging the inequality for $\mathbf{K}(m|v)$ back into Equation 1 results in

$$n+\mathbf{K}(n,\mu,\lambda,\rho)+c<^+ n+\mathbf{K}(n,\mu,\lambda,\rho)+2\mathbf{K}(c)+d\mathbf{K}(c)+(f+g)O(1)$$
$$c<^+(2+d)\mathbf{K}(c)+(f+g)O(1).$$

This result is a contradiction for sufficiently large c solely dependent on the universal Turing machine. □

Corollary 1 *For computable measures* $\{\mu_i\}_{i=1}^m$ *and nonatomic* λ *over* $\{0,1\}^\infty$ *and* $n \in \mathbb{N}$, $\lambda\{\alpha : \bigwedge_{i=1}^m \mathbf{D}(\alpha|\mu_i) > n\} > 2^{-n-\mathbf{K}(n,\mu,\rho,\lambda)-O(\log m)}$.

Theorem 2 can be extended to incomputable λ, which can be accomplished using a stronger version[1] of Theorem 1. The term $\langle\lambda\rangle \in \{0,1\}^\infty$ represents any encoding of λ that can compute $\lambda(x\{0,1\}^\infty)$ for $x \in \{0,1\}^*$ up to arbitrary precision. Let $\mathbf{I}(\lambda : \mathcal{H}) = \inf_{\langle\lambda\rangle} \mathbf{I}(\langle\lambda\rangle : \mathcal{H})$.

Corollary 2

- *For computable measures* μ, ρ, *potentially uncomputable nonatomic* λ *over* $\{0,1\}^\infty$ *and* $n \in \mathbb{N}$, $\lambda\{\alpha : \mathbf{D}(\alpha|\mu) > n$ *and* $\mathbf{D}(\alpha|\rho) > n\} > 2^{-n-O(\mathbf{K}(n,\mu,\rho)+\mathbf{I}(\lambda:\mathcal{H}))}$.
- *For measures* μ *and* ρ *over* $\{0,1\}^\infty$, *nonatomic* λ, *computable* μ, ρ, *if for every* $c \in \mathbb{N}$, *there is an* $n \in \mathbb{N}$, *where* $\lambda\{\alpha : \mathbf{D}(\alpha|\mu) > n$ *and* $\mathbf{D}(\alpha|\rho) > n\} < 2^{-n-O(\mathbf{K}(n))-c}$, *then* $\mathbf{I}(\lambda : \mathcal{H}) = \infty$.

6 Computable Probability Spaces

The second main result of this paper uses computable metric spaces and computable probability measures from [HR09]. Some constructs need changes, which we present in the later section. But in this section we show the definitions, lemmas, and theorems that are directly taken from [HR09]. If a theorem or lemma is presented without a proof, then it can be found in [HR09]

Definition 2 *A computable metric space consists of a triple* $(\mathcal{X}, \mathcal{S}, d)$, *where*

- $\mathcal{X}$ *is a separable complete metric space.*
- $\mathcal{S}$ *is an enumerable list of dense ideal points* $\mathcal{S}$ *in* $\mathcal{X}$.
- d *is a distance metric that is uniformly computable over points in* $\mathcal{S}$.

The complexity of a metric space $\mathcal{X}$ *is* $\mathbf{K}(\mathcal{X})$, *the smallest program that computes* d.

For $x \in \mathcal{X}$, $r \in \mathbb{Q}_{>0}$ a ball is $B(x,r) = \{y : d(x,y) < r\}$. The ideal points induce a sequence of enumerable ideal balls $B_i = \{B(s_i, r_j) : s_i \in \mathcal{S}, r_j \in \mathbb{Q}_{>0}\}$. A sequence of ideal points $\{x_n\} \subseteq Y$ is said to be a fast Cauchy sequence if $d(x_n, x_{n+1}) < 2^{-n}$ for all $n \in \mathbb{N}$. A point x is computable there is a computable fast Cauchy sequence converging to x. The complexity of such a point, $\mathbf{K}(x)$ is the length of the smallest program that computes a fast Cauchy sequence converging to it. The complexity of a sequence of uniformly computable points $Y = \{x_i\}$ is $\mathbf{K}(Y)$, the length of the smallest program that maps i to a fast Cauchy sequence converging to x_i. Each computable function f between computable metric spaces $\mathcal{X}$ and $\mathcal{Y}$ has an algorithm $\mathfrak{A}$ such that if $f(x) = y$ then for all fast Cauchy sequences $\overrightarrow{x}$ for x, $\mathfrak{A}(\overrightarrow{x})$ outputs an encoding of a fast Cauchy sequence for y. The complexity of a computable function f is $\mathbf{K}(f)$, the length of the shortest program to compute such an algorithm.

Definition 3 *Lower computable functions* $f \in \mathcal{F}$ *have algorithms that enumerate* $\{(B_i, r_i)\}$, *where* B_i *is an ideal ball and* $r_i \in \mathbb{Q}_{>0}$, *and* $f(x) = \sup\{r_i : x \in B_i\}$.

The computable metric space of all Borel probability measures over $\mathcal{X}$ is $\mathcal{M}(\mathcal{X})$. If $\mathcal{X}$ is separable and compact then so is $\mathcal{M}(\mathcal{X})$. The ideal points of $\mathcal{M}(\mathcal{X})$ are $\mathcal{D}$, the set of probability measures that are concentrated on finitely many points with rational values. The distance metric on $\mathcal{M}(\mathcal{X})$ is the *Prokhorov metric*, defined as follows.

[1] Theorem 1 can be extended to $\Pr_{\alpha\sim\lambda}\left(2^{\mathbf{I}(\langle\alpha,\lambda\rangle:\mathcal{H})}\right) \overset{*}{\lesssim} 2^{\mathbf{I}(\lambda:\mathcal{H})}$.

Definition 4 (Prokhorov metric)

$$\pi(\mu, \nu) = \inf \left\{ \epsilon \in \mathbb{R}^+ : \mu(A) \leq \nu(A^\epsilon) \text{ for Borel set } A \right\},$$

where $A^\epsilon = \{x : d(x, A) < \epsilon\}$.

Theorem 3 *Given a probability measure* $\mu \in \mathcal{M}(\mathcal{X})$*, the following are equivalent.*

1. μ *is computable.*
2. $\mu(B_{i_1} \cup \cdots \cup B_{i_k})$ *is lower semi-computable uniformly in* $\langle i_1, \ldots, i_k \rangle$.
3. $\int d\mu : \mathcal{F} \to \mathbb{R}_{\geq 0}$ *is lower semi-computable.*

Definition 5

1. *A constructive* G_δ*-set* U *is a set of the form* $\bigcap_n U_n$ *where* $\{U_n\}$ *is a sequence of uniformly r.e. open sets. The complexity of* U *is,* $\mathbf{K}(U)$*, the size of the smallest program that uniformly enumerates* $\{U_n\}$.
2. *A computable probability space is a pair* $(\mathcal{X}, \mu)$*, where* $\mathcal{X}$ *is a computable metric space and* μ *is a Borel probability measure on* $\mathcal{X}$.
3. *Let* $(\mathcal{X}, \mu)$ *be a computable probability space and* $\mathcal{Y}$ *a computable metric space. A function* $f : D_f \subset (\mathcal{X}, \mu) \to \mathcal{Y}$ *is almost computable if it is computable on a constructive* G_δ*-set* (D_f) *of* μ*-measure one.*
4. *A morphism of computable probability spaces* $Q : (\mathcal{X}, \mu) \to (\mathcal{Y}, \nu)$ *is an almost computable measure-preserving function* $Q : D_Q \subset \mathcal{X} \to \mathcal{Y}$*, where* $\mu(Q^{-1}(A)) = \nu(A)$ *for all Borel sets* A*. An isomorphism* (Q, R) *is a pair of morphisms such that* $Q \circ R = \mathrm{id}$ *on* $R^{-1}(D_Q)$ *and* $R \circ Q = \mathrm{id}$ *on* $Q^{-1}(D_R)$.
5. *A binary representation of a computable probability space* $(\mathcal{X}, \mu)$ *is a pair* (δ, μ_δ) *where* μ_δ *is a computable probability measure on* $\{0,1\}^\infty$ *and* $\delta : (\{0,1\}^\infty, \mu_\delta) \to (\mathcal{X}, \mu)$ *is a surjective morphism such that, calling* $\delta^{-1}(x)$ *the set of expansions of* $x \in X$:
   - *There is a dense full-measure constructive* G_δ*-set* D *of points having a unique expansion.*
   - $\delta^{-1} : D \to \delta^{-1}(D)$ *is computable.*
   - (δ, δ^{-1}) *is an isomorphism.*

Theorem 4 *Every computable probability space* $(\mathcal{X}, \mu)$ *has a binary representation.*

Definition 6 *Given a probability measure* $\mu \in \mathcal{M}(\mathcal{X})$*, a* μ*-randomness test is a* μ*-constructive function* $T \in \mathcal{F}$*, such that* $\int T d\mu \leq 1$*. A uniform randomness test is a constructive function* T *from* $\mathcal{M}(\mathcal{X})$ *to* $\mathcal{F}$ *such that* $\int T^\mu d\mu \leq 1$.

Theorem 5

1. *Let* μ *be a probability measure. For every* μ*-randomness test* t*, there is a uniform randomness test* $T : \mathcal{M}(\mathcal{X}) \to \mathcal{F}$ *with* $T(\mu) = .5t$.
2. *There is a universal uniform randomness test, that is a uniform test* $\mathbf{t}$ *such that for every uniform test* T*, there is a constant* $c > 0$ *with* $\mathbf{t} > cT$.

7 Multi Binary Representation

This paper introduces a new concept that is needed in the first theorem: dual binary representation. While a binary representation is a mapping from one computable probability space to the Cantor space, a multi binary representation maps three computable probability spaces to Cantor spaces, each sharing the same mapping.

Definition 7 *A set A is almost decidable with respect to probability measures (μ,ν,ρ) if there are two. r.e. open sets U and V such that $U \subset A$, $V \subseteq A^{\mathscr{C}}$, $U \cup V$ is dense and has full μ, ν, and ρ measure. We say the elements of a sequence $\{A_i\}$ are uniformly almost decidable with respect to (μ,ν,ρ) if there are two sequences $\{U_i\}$ and $\{V_i\}$ of uniformly r.e. sets satisfying the above conditions. The complexity of the sequence, $\mathbf{K}(\{A_i\})$ is the length of the smallest program the uniformly computes the two sequences $\mathbf{K}(\{U_i\})$ and $\mathbf{K}(\{V_i\})$.*

The following theorem modifies [HR09] to account for the complexities of the involved terms.

Theorem 6 *On a computable metric space, every dense constructive G_δ-set G has a dense sequence of uniformly computable points Y, where $\mathbf{K}(Y) <^+ \mathbf{K}(G,\mathcal{X})$.*

Proof. Let $A = \cap_i U_i$ where U_i is constructive uniformly in i. Let B be an ideal ball: we construct a sequence of ideal balls $\{B(i)\}_i$ such that $B(i+1) \subset U_i \cap B(i)$. Put $B(0) = B$. If $B(i)$ has been constructed, as U_i is dense $B(i) \cap U_i$ is a non-empty open set, so we can find some ball $B' \subseteq B(i) \cap U_i$. $B(i+1)$ is obtained dividing the radius of B' by 2. By completeness of the space, $\cap_i B(i)$ is non-empty. It is a singleton $\{x\}$ where x is a computable point belonging to $A \cap B$. Everything is uniform in the ideal ball B, the number $\{B_k\}_k$ of ideal balls gives a constructive sequence $\{x_i\}$ of uniformly computable points with $x_k \in A \cap B_k$. The program to uniformly compute $Y = \{x_k\}$ requires a program to compute G_δ and a program to compute $\mathcal{X}$, so $\mathbf{K}(Y) <^+ \mathbf{K}(G,\mathcal{X})$.

Lemma 6 *There is a sequence of $\{r_n\}$ of uniformly computable reals such that $\{B(s_i,r_n)\}_{i,n}$ is a basis of uniformly almost computable decidable balls, relative to (μ,ν,ρ). Furthermore $\mathbf{K}(r_n) <^+ \mathbf{K}(\mu,\nu,\rho,\mathcal{X})$.*

Proof. Define $U_{\langle i,k\rangle} = \{r \in \mathbb{R}_{>0} : \mu(\overline{B}(s_i,r)) < \mu(B(s_i,r)) + 1/k\}$. By computability of μ, this is a r.e. open subset of $\mathbb{R}_{>0}$ uniformly in $\langle i,k\rangle$. Let $W_{\langle i,k\rangle} = \{r \in \mathbb{R}_{>0} : \nu(\overline{B}(s_i,r)) < \nu(B(s_i,r)) + 1/k\}$, which is also an r.e. open subset of $\mathbb{R}_{>0}$. Let $X_{\langle i,k\rangle} = \{r \in \mathbb{R}_{>0} : \rho(\overline{B}(s_i,r)) < \rho(B(s_i,r)) + 1/k\}$, which is also an r.e. open subset of $\mathbb{R}_{>0}$. They both are dense in $\mathbb{R}_{>0}$. The spheres $S_r = \overline{B}(s_i,r) \setminus B(s_i,r)$ are disjoint for different radii and μ, ν, and ρ are finite, so the set of r for which $\mu(S_r) \geq 1/k$, $\nu(S_r) \geq 1/k$, or $\rho(S_r) > 1/k$ is finite. Let $V_{\langle i,j\rangle} = \mathbb{R}_{>0} \setminus \{d(s_i,s_j)\}$ be a dense r.e. open set, uniformly in $\langle i,j\rangle$. Thus we get the dense constructive G_δ-set

$$G = \bigcap_{\langle i,k\rangle} U_{\langle i,k\rangle} \cap \bigcap_{\langle i,k\rangle} W_{\langle i,k\rangle} \cap \bigcap_{\langle i,j\rangle} X_{\langle i,j\rangle} \cap \bigcap_{\langle i,j\rangle} V_{\langle i,j\rangle}$$

This set can be constructed given $\mathcal{X}$, and programs for μ, ν, and ρ, with $\mathbf{K}(G) <^+ \mathbf{K}(\mathcal{X},\mu,\nu,\rho)$. Then by Theorem 6, contains a sequence $\{r_n\}$ of uniformly computable reals numbers which is dense in $\mathbb{R}_{>0}$. This sequence has complexity $\mathbf{K}(\{r_n\}) <^+ \mathbf{K}(\mathcal{X},\mu,\nu,\rho)$. For any s_i and r_n, $B(s_i,r_n)$ is almost decidable, relative to (μ,ν,ρ). □

Definition 8 *A multi probability space $(\mathcal{X},\mu,\nu,\rho)$ is a computable metric space $\mathcal{X}$ and three computable Borel probability measures, μ, ν, and ρ over $\mathcal{X}$.*

Definition 9 *A multi binary representation of a dual probability space* $(\mathcal{X},\mu,\nu,\rho)$ *is a tuple* $(\delta,\mu_\delta,\nu_\delta,\rho_\delta)$ *where* μ_δ, ν_δ, *and* ρ_δ *are computable probability measures on* $\{0,1\}^\infty$ *and* $\delta:(\{0,1\}^\infty,\mu_\delta)\to(\mathcal{X},\mu)$ *,* $\delta:(\{0,1\}^\infty,\nu_\delta)\to(\mathcal{X},\nu)$*, and* $\delta:(\{0,1\}^\infty,\rho_\delta)\to(\mathcal{X},\rho)$ *are surjective morphisms. Denoting* $\delta^{-1}(x)$ *to be the set of expansion of* $x\in X$*:*

- *There is a dense full-measure constructive* G_δ*-set* D *of points have a unique expansion.*
- $\delta^{-1}:D\to\delta^{-1}(D)$ *is computable.*
- (δ,δ^{-1}) *is an isomorphism.*
- $\mathbf{K}(\delta,\mu_\delta,\nu_\delta,\rho_\delta)<^+\mathbf{K}(\mathcal{X},\mu,\nu,\rho)$.

Definition 10 *We fix computable probability measures* μ, ν, *and* ρ *and their computable representations. We denote* $B(s_i,r_n)$ *by* B_k *where* $k=\langle i,n\rangle$ *and* r_n *is the sequence defined in 6. Let* $C_k=X\setminus\overline{B}(s_i,r_n)$*. For* $w\in\{0,1\}^*$*, the cell* $\Gamma(w)$ *is defined by* $\Gamma(\epsilon)=X$, $\Gamma(w0)=\Gamma(w)\cap C_i$ *and* $\Gamma(w1)=\Gamma(w)=\cap B_i$*, where* ϵ *is the empty word and* $i=\|w\|$*. This is an almost decidable set, uniformly in* w*.* $\Gamma(w)$ *can be uniformly computed by a program that uniformly computes* $\{r_n\}$*, and therefore of size* $\mathbf{K}(\Gamma(\cdot))<^+\mathbf{K}(\{r_n\})<^+\mathbf{K}(\mathcal{X},\mu,\nu,\rho)$.

Theorem 7 *Every multi probability space* $(\mathcal{X},\mu,\nu)$ *has a multi binary representation.*

Proof. We construct an encoding function $b:D\to\{0,1\}^\infty$, a decoding function $\delta:D_\delta\to X$, and show that δ is a multi binary representation, with $b=\delta^{-1}$. Let $D=\cap_i B_i\cup C_i$. The set D is a full-measure constructive G_δ-set, with $\mathbf{K}(D)<^+\mathbf{K}(\mathcal{X},\mu,\nu,\rho)$. Define the computable function $b:D\to\{0,1\}^\infty$ with $b(x)_i=1$ if $x\in B_i$ and $b(x)_i=0$ if $x\in C_i$. Let $x\in D$: $\omega=b(x)$ is also characterized by $\{x\}=\cap_i\Gamma(\omega_{0\ldots i-1})$. b can be computed from $\Gamma(\cdot)$, thus $\mathbf{K}(b)<^+\mathbf{K}(\Gamma(\cdot))<^+\mathbf{K}(\mathcal{X},\mu,\nu,\rho)$. Let μ_δ, ν_δ, and ρ_δ be the morphisms $\mu\circ b^{-1}$, $\nu\circ b^{-1}$, and $\rho\circ b^{-1}$, respectively. The complexity of the three morphisms is not more than the complexity of the computable function b, and thus $<^+\mathbf{K}(\mathcal{X},\mu,\nu,\rho)$.

Let D_δ be the set of binary sequences ω such that $\cap_i\overline{\Gamma(\omega_{0\ldots i-1})}$ is a singleton. The decoding function $\delta:D_\delta\to X$ is defined by

$$\delta(\omega)=x \text{ if } \cap_i\overline{\Gamma(\omega_{0\ldots i-1})}=\{x\}.$$

The next steps are to prove the δ is a surjective morphism, and the proof for this follows identically to the proof of Theorem 5.1.1 in [HR09]. The to compute this function this function, one needs to compute $\Gamma(\cdot)$. Thus $\mathbf{K}(\delta)<^+\mathbf{K}(\Gamma(\cdot))<^+\mathbf{K}(\mathcal{X},\mu,\nu,\rho)$. □

8 Universal Uniform Tests

Lemma 7 *Let* $Q:D\subset\mathcal{X}\to\mathcal{Y}$ *be a morphism of computable probability spaces* $(\mathcal{X},\mu)$ *and* $(\mathcal{Y},\nu)$*, with universal tests* $\mathbf{t}_\mu$ *and* $\mathbf{t}_\nu$*. If* $x\in\mathcal{X}$ *and* $\mathbf{t}_\mu(x)<\infty$*, then* $Q(x)$ *is defined and* $\mathbf{t}_\nu(Q(x))\overset{*}{<}\mathbf{t}_\mu(x)2^{\mathbf{K}(Q,\mu)}$.

Proof. The proof is a slight modification to Proposition 6.2.1 in [HR09]. So, assuming $\mathbf{t}_\mu(x)<\infty$, then x is a random point then $x\in D$, because due to Lemma 6.2.1 in [HR09], every random point lies in every r.e. open set of full measure, and D is an intersection of full-measure r.e open sets. Thus $Q(x)$ is defined.

Let $\mathscr{A}$ be any algorithm lower semi-computing the function $\mathbf{t}_\nu \circ Q : D \to \mathbb{R}^{\infty}_{\geq 0}$. This algorithm can be converted into a lower computable function $f_{\mathscr{A}} : \mathcal{X} \to \mathbb{R}^{\infty}_{\geq 0}$ by feeding all finite prefixes of fast Cauchy sequences to Q and enumerating all resultant outputted ideal balls and seeing which outputted ideal balls are in the ideal balls of those enumerated by $\mathbf{t}_\nu$. Since $\mu(D) = 1$, $\int \mathbf{t}_\nu \circ Q d\mu$ equals $\int f_{\mathscr{A}} d\mu$. Thus $\mathbf{K}(f_{\mathscr{A}}) <^+ \mathbf{K}(Q, \mu)$. As Q is measure-preserving, $\int \mathbf{t}_\nu \circ Q du = \int \mathbf{t}_\nu d\nu \leq 1$. Hence $f_{\mathscr{A}}$ is a μ-test, with $f_{\mathscr{A}} \overset{*}{<} \mathbf{t}_\mu 2^{\mathbf{K}(Q,\mu)}$. Thus $\mathbf{t}_\nu(Q(x)) = f_{\mathscr{A}}(x) \overset{*}{<} \mathbf{t}_\mu(x) 2^{\mathbf{K}(Q,\mu)}$. □

Claim 1 *We recall that the deficiency of randomness of an infinite sequence $\alpha \in \{0,1\}^\infty$ with respect to a computable probability measure P over $\{0,1\}^\infty$ is defined to be*

$$\mathbf{D}(\alpha|P, x) = \log \sup_n \mathbf{m}(\alpha[0..n]|x)/P(\alpha[0..n]).$$

We have $\mathbf{D}(\alpha|P) = \mathbf{D}(\alpha|P, \emptyset)$. By [Ǵ21], $2^{\mathbf{D}}$ is a lower-computable P-test, in that $\int_{\{0,1\}^\infty} 2^{\mathbf{D}(\alpha|P)} dP(\alpha) = O(1)$. Thus since $\mathbf{t}_P$ is a universal uniform test that takes the constructive point P as a parameter, $\mathbf{t}_P$ can enumerate lower P-tests that takes a program to compute $P(x)$. Thus $\mathbf{t}_P(\alpha) \overset{*}{>} 2^{\mathbf{D}(\alpha|P)}$.

Theorem 8 *Given computable probability measures μ, ρ, and λ, non-atomic λ, over a computable metric space $\mathcal{X}$ and universal uniform test $\mathbf{t}$, for all n,*
$\lambda(\{\alpha : \mathbf{t}_\mu(\alpha) > 2^n$ *and* $\mathbf{t}_\rho(\alpha) > 2^n\}) > 2^{-n-\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})-\mathbf{K}(n,\lambda,\mu,\rho,\mathcal{X})-O(1)}$.

Proof. We fix the algorithmic descriptions of λ, μ, ρ, and $\mathcal{X}$ to be one that minimizes $\mathbf{K}(\lambda, \mu, \rho, \mathcal{X})$. By Theorem 4, fix a multi binary representation $(\delta, \lambda_\delta, \mu_\delta, \rho_\delta)$ for multi probability space $(\mathcal{X}, \lambda, \mu, \rho)$. Note that δ is a measure-preserving transform, where $\lambda(A) = \lambda_\delta(\delta^{-1}(A))$ for all Borel sets A. Due to Lemma 7 and the fact that $\mathbf{K}(\delta, \lambda_\delta, \mu_\delta, \rho_\delta|\lambda, \mu, \rho, \mathcal{X}) = O(1)$,

$$\begin{aligned}
&\lambda(\{\beta : \mathbf{t}_\mu(\beta) \leq 2^n \text{ or } \mathbf{t}_\rho(\beta) \leq 2^n\}) \\
=&\lambda_\delta(\delta^{-1}(\{\beta : \mathbf{t}_\mu(\beta) \leq 2^n \text{ or } \mathbf{t}_\rho(\beta) \leq 2^n\})) \\
\overset{*}{<}&\lambda_\delta(\delta^{-1}(\{\beta : \mathbf{t}_{\mu_\delta}(\delta^{-1}(\beta)) \leq 2^{n+\mathbf{K}(\delta,\mu_\delta)} \text{ or } \mathbf{t}_{\mu_\rho}(\delta^{-1}(\beta)) \leq 2^{n+\mathbf{K}(\delta,\rho_\delta)}\})) \\
\overset{*}{=}&\lambda_\delta(\delta^{-1}(\{\beta \in \delta(\{\alpha : \mathbf{t}_{\mu_\delta}(\alpha) \leq 2^{n+\mathbf{K}(\delta,\mu_\delta)} \text{ or } \mathbf{t}_{\rho_\delta}(\alpha) \leq 2^{n+\mathbf{K}(\delta,\rho_\delta)}\})\}) \\
\overset{*}{=}&\lambda_\delta(\delta^{-1}(\delta(\{\alpha : \mathbf{t}_{\mu_\delta}(\alpha) \leq 2^{n+\mathbf{K}(\delta,\mu_\delta)} \text{ or } \mathbf{t}_{\rho_\delta}(\alpha) \leq 2^{n+\mathbf{K}(\delta,\mu_\rho)}\})) \\
\overset{*}{<}&\lambda_\delta(\{\alpha : \mathbf{t}_{\mu_\delta}(\alpha) <^+ 2^{n+\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})} \text{ or } \mathbf{t}_{\rho_\delta}(\alpha) \leq 2^{n+\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})}\}).
\end{aligned} \tag{2}$$

The $\mathbf{K}(\delta, \mu_\delta)$ is the size of the smallest program that computes both δ and μ_δ, and similarly for $\mathbf{K}(\delta, \rho_\delta)$. From Equation 2, and Claim 1, we get,

$$\begin{aligned}
&\lambda(\{\beta : \mathbf{t}_\mu(\beta) \leq 2^n \text{ or } \mathbf{t}_\rho(\beta) \leq 2^n\}) \\
\overset{*}{<}&\lambda_\delta\{\alpha : \mathbf{D}(\alpha|\mu_\delta) <^+ n + \mathbf{K}(\lambda, \mu, \rho, \mathcal{X}) \text{ or } \mathbf{D}(\alpha|\rho_\delta) <^+ n + \mathbf{K}(\lambda, \mu, \rho, \mathcal{X})\}.
\end{aligned} \tag{3}$$

From Theorem 2, we get

$$\begin{aligned}
&\lambda(\{\beta : \mathbf{t}_\mu(\beta) \leq 2^n \text{ or } \mathbf{t}_\rho(\beta) \leq 2^n\}) \\
<&1 - 2^{-n-\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})-\mathbf{K}(n+\mathbf{K}(\lambda,\mu,\rho,\mathcal{X}),\lambda,\mu,\rho,\mathcal{X})-O(1)} \\
<&1 - 2^{-n-\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})-\mathbf{K}(n,\mathbf{K}(\lambda,\mu,\rho,\mathcal{X}),\lambda,\mu,\rho,\mathcal{X})-O(1)} \\
<&1 - 2^{-n-\mathbf{K}(\lambda,\mu,\rho,\mathcal{X})-\mathbf{K}(n,\lambda,\mu,\rho,\mathcal{X})-O(1)}.
\end{aligned}$$

A Alternative Proof

In this appendix, a more direct proof is given of Theorem 2. It is a straightforward modification to Theorems 4, 5, and 6 in [Eps22]. The results do not allow a corollary for uncomputable λ with finite mutual information with the halting sequence. In addition, it is incompatible with the application of Theorem 2 to computable metric spaces.

A sampling method A is a probabilistic function that maps an integer N with probability 1 to a set containing N different strings. Let $P = P_1, P_2, \ldots$ be a sequence of measures over strings. For example, one may choose $P_1 = P_2 \ldots$ or choose P_n to be the uniform measure over n-bit strings. A conditional probability bounded P-test is a function $t : \{0,1\}^* \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that for all $n \in \mathbb{N}$ and positive real number r, we have $P_n(\{x : t(x|n) \geq r\}) \leq 1/r$. If $P_1, P_2, \ldots$ is uniformly computable, then there exists a lower-semicomputable such P-test t that is "maximal" (i.e., for which $t' \leq O(t)$ for every other such test t'). We fix such a t, and let $\overline{\mathbf{d}}_n(x|P) = \log t(x|n)$.

Lemma 8 *Let P and Q be two probability measures on strings and let A be a sampling method. For all integers N, there exists a finite set $S \subset \{0,1\}^*$ such that $P(S) \leq 32/N$, $Q(S) \leq 32/N$, and with probability strictly more than 0.99: $A(N)$ intersects S.*

Proof. We show that some possibly infinite set S satisfies the conditions, and thus, some finite subset also satisfies the conditions due to the strict inequality. We use the probabilistic method: we select each string to be in S with probability $8/N$ and show that the three conditions are satisfied with positive probability. The expected value of $P(S)$ and $Q(S)$ is $8/N$. By the Markov inequality, the probability that $P(S) > 32/N$ is at most $1/4$ and the probability that $Q(S) > 32/N$ is at most $1/4$. For any set D containing N strings, the probability that S is disjoint from D is

$$(1 - 8/N)^N < e^{-8}.$$

Let Q be the measure over N-element sets of strings generated by the sampling algorithm $A(N)$. The left-hand side above is equal to the expected value of

$$Q(\{D : D \text{ is disjoint from } S\}).$$

Again by the Markov inequality, with probability greater than $3/4$, this measure is less than $4e^{-8} < 0.01$. By the union bound, the probability that at least one of the conditions is violated is less than $1/4 + 1/4 + 1/4$. Thus, with positive probability a required set is generated, and thus such a set exists. $\square$

Theorem 9 *Let $P = P_1, P_2 \ldots$ and $Q = Q_1, Q_2 \ldots$ be a two uniformly computable sequence of measures on strings and let A be a sampling method. There exists $c \in \mathbb{N}$ such that for all n:*

$$\Pr\left(\max_{a \in A(2^n)} \min\{\overline{\mathbf{d}}_n(a|P), \overline{\mathbf{d}}_n(a|Q)\} > n - c\right) \geq 0.99.$$

Proof. We now fix a search procedure that on input N finds a set S_N that satisfies the conditions of Lemma 8. Let $t'(a|n)$ be the maximal value of $2^n/64$ such that $a \in S_{2^n}$. By construction, t' is a computable probability bounded test for both P and Q, because $P_n(\{x : t'(x|n) = 2^\ell\}) \leq 2^{-\ell-1}$, and thus $P_n(t'(x|n) \geq 2^\ell) \leq 2^{-\ell-1} + 2^{-\ell-2} + \ldots$ and similarly for Q. With probability 0.99, the set $A(2^n)$ intersects S_{2^n}. For any number a in the intersection, we have $t'(x|n) \geq 2^{n-6}$, thus by the optimality of t and definition of $\overline{\mathbf{d}}$, we have $\overline{\mathbf{d}}_n(a|P) > n - O(1)$ and $\overline{\mathbf{d}}_n(a|Q) > n - O(1)$. $\square$

An incomplete sampling method A takes in a natural number N and outputs, with probability $f(N)$, a set of N numbers. Otherwise A outputs $\perp$. f is computable.

Corollary 3 *Let $P = P_1, P_2 \dots$ and $Q = Q_1, Q_2 \dots$ be two uniformly computable sequences of measures on strings and let A be an incomplete sampling method. There exists $c \in \mathbb{N}$ such that for all n:*

$$\Pr_{D=A(n)}\left(D \neq \perp \text{ and } \max_{a \in D} \min\{\overline{\mathbf{d}}_n(a|P), \overline{\mathbf{d}}_n(a|Q)\} \leq n - c\right) < 0.01.$$

Let $\mu = \mu_1, \mu_2, \dots$ be a uniformly computable sequence of measures over infinite sequences. Similar way as for strings in the introduction, the randomness deficiency $\overline{\mathbf{D}}_n(\omega|\mu)$ for sequences ω is defined using lower-semicomputable functions $\{0,1\}^\infty \times \mathbb{N} \to \mathbb{R}_{\geq 0}$. A continuous sampling method C is a probabilistic function that maps, with probability 1, an integer N to an infinite encoding of N different sequences.

Theorem 10 *Let $\mu = \mu_1, \mu_2, \dots$ and $\nu = \nu_1, \nu_2, \dots$ be two uniformly computable sequences of measures over infinite sequences. Let C be a continuous sampling method. There exists $c \in \mathbb{N}$ where for all n:*

$$\Pr\left(\max_{\alpha \in C(2^n)} \min\{\overline{\mathbf{D}}_n(\alpha|\mu), \overline{\mathbf{D}}_n(\alpha|\nu)\} > n - c\right) \geq 0.98.$$

Proof. For $D \subseteq \{0,1\}^\infty$, $D_m = \{\omega[0..m] : \omega \in D\}$. Let $g(n) = \arg\min_m \Pr_{D=C(n)}(|D_m| < n) < 0.01$ be the smallest number m such that the initial m-segment of $C(n)$ are sets of n strings with probability > 0.99. g is computable, because C outputs a set of distinct infinite sequences with probability 1. For probability ψ over $\{0,1\}^\infty$, let $\psi^m(x) = [|x| = m]\psi(\{\omega : x \sqsubset \omega\})$. Let $\mu^g = \mu_1^{g(1)}, \mu_2^{g(2)}, \dots$ and $\nu^g = \nu_1^{g(1)}, \nu_2^{g(2)}, \dots$ be two uniformly computable sequences of discrete probability measures and let A be a discrete incomplete sampling method, where for random seed $\omega \in \{0,1\}^\infty$, $A(n, \omega) = C(n, \omega)_{g(n)}$ if $|C(n, \omega)_{g(n)}| = n$; otherwise $A(n, \omega) = \perp$. So $\Pr[A(n) = \perp] < 0.01$. There exists a constant $c \in \mathbb{N}$ such that,

$$\begin{aligned}
&\Pr\left(\max_{\alpha \in C(2^n)} \min\{\overline{\mathbf{D}}_n(\alpha|\mu), \overline{\mathbf{D}}_n(\alpha|\nu)\} \leq n - c\right) \\
\leq &\Pr_{Z=C(2^n)}\left((|Z_{g(n)}| < 2^n) \text{ or } (|Z_{g(n)}| = 2^n \text{ and } \max_{\alpha \in Z} \min\{\overline{\mathbf{D}}_n(\alpha|\mu), \overline{\mathbf{D}}_n(\alpha|\nu)\} \leq n - c\right) \\
\leq &\Pr_{D=A(2^n)}\left(D = \perp \text{ or } (D \neq \perp \text{ and } \max_{x \in D} \min\{\overline{\mathbf{d}}_n(x|\mu^g), \overline{\mathbf{d}}_n(x|\nu^g)\} \leq n - c)\right) \\
< &0.01 + 0.01 \qquad (4) \\
\leq &0.02,
\end{aligned}$$

where Equation 4 is due to Corollary 3. $\square$

Theorem 11 *Let $\lambda = \lambda_1, \lambda_2, \dots$, $\mu = \mu_1, \mu_2, \dots$, and $\nu = \nu_1, \nu_2, \dots$ be three uniformly computable sequences of measures over infinite sequences. Each λ_n is non-atomic. There is a constant $c \in \mathbb{N}$, dependent on μ, ν and λ, where for all $n \in \mathbb{N}$, $\lambda_n\left\{\alpha : \overline{\mathbf{D}}_n(\alpha|\mu) > n - c \text{ and } \overline{\mathbf{D}}_n(\alpha|\nu) > n - c\right\} > 2^{-n-1}$.*

Proof. We define the continuous sampling method C, where on input n, randomly samples n elements from λ_n. Let $d_n = \lambda_n\{\alpha : \min\{\overline{\mathbf{D}}_n(\alpha|\mu), \overline{\mathbf{D}}_n(\alpha|\nu)\} > n - c\}$, where c is the constant in Theorem 10. By that theorem,

$$\begin{aligned}
\Pr\left(\max_{\alpha\in C(2^n)} \min\{\overline{\mathbf{D}}_n(\alpha|\mu), \overline{\mathbf{D}}_n(\alpha|\nu)\} > n - c\right) &> 0.98 \\
1 - (1 - d_n)^{2^n} &> 0.98 \\
1 - 2^n d_n &< 0.02 \\
d_n &> (0.98)2^{-n} \\
\lambda_n\{\alpha : \min\{\overline{\mathbf{D}}_n(\alpha|\mu), \overline{\mathbf{D}}_n(\alpha|\nu)\} > n - c\} &> 2^{-n-1}.
\end{aligned}$$

□

References

[Eps21] Samuel Epstein. All sampling methods produce outliers. *IEEE Transactions on Information Theory*, 67(11):7568–7578, 2021.

[Eps22] S. Epstein. The outlier theorem revisited. *CoRR*, abs/2203.08733, 2022.

[Eps23a] S. Epstein. On the existence of anomalies. *CoRR*, abs/2302.05972, 2023.

[Eps23b] S. Epstein. On the existence of anomalies, the reals case. *CoRR*, abs/2303.05614, 2023.

[Eps23c] S. Epstein. Uniform tests and algorithmic thermodynamic entropy. *CoRR*, abs/2303.05619, 2023.

[G2́1] Peter Gács. Lecture notes on descriptional complexity and randomness. *CoRR*, abs/2105.04704, 2021.

[Gei12] Philipp Geiger. *Mutual information and Gödel incompleteness*. PhD thesis, Heidelberg University, 10 2012.

[HR09] M. Hoyrup and C. Rojas. Computability of probability measures and martin-löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.

[Lev74] L. A. Levin. Laws of Information Conservation (Non-growth) and Aspects of the Foundations of Probability Theory. *Problemy Peredachi Informatsii*, 10(3):206–210, 1974.

[Ver21] N. Vereshchagin. Proofs of conservation inequalities for levin's notion of mutual information of 1974. *Theoretical Computer Science*, 856, 2021.