Randomness Deficiency Overlap

Samuel Epstein[*]

June 20, 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.

1 Introduction

In a previous paper, a lower bound was proved on the computable measure of sets with high randomness deficiency.

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 revisits 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. 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], and Theorem 2 is a reworking of Theorem 3 in [Eps22]. The tight bounds of the main theorem derived from lemmas with looser bounds is achieved through relativization. The next steps will be to transfer the results over to computable metric spaces and universal uniform tests.

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 [G21]. $\mathbf{d}(a|p) = \mathbf{d}(a|p,\emptyset)$. 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 $< f+O(\log(f+1))$. Stochasticity is $\mathbf{Ks}(a|b) = \min\{\mathbf{K}(Q|b)+3\log\max\{\mathbf{d}(a|Q,b),1\} : Q$ has finite support and a range in $\mathbb{Q}\}$. $\mathbf{Ks}(a|b) < \mathbf{Ks}(a)+$

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

$O(\log \mathbf{K}(b))$. 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 $\mathbf{D}(\alpha|P) = \sup_n -\log P(\alpha[0..n]) - \mathbf{K}(\alpha[0..n]|\langle P\rangle)$. The term $\langle P\rangle$ is a program to compute P. 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. $\mathbf{E}[p(F)] = \mathbf{E}[q(F)] \le cd2^{-s}$. Furthermore

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

Thus finite $W\subset\mathbb{N}$ can be chosen such that $p(W)\le 4cd2^{-s}$, $q(W)\le 4cd2^{-s}$, and $Q(\{G:|G|=2^s, G\cap W=\emptyset\})\le e^{2-cd}$. $D\cap W\neq\emptyset$, 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. □ 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

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

□

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 $\lambda\{\alpha : \mathbf{D}(\alpha|\mu) > n \text{ and } \mathbf{D}(\alpha|\rho) > n\} \le 2^{-n-\mathbf{K}(n,\mu,\lambda)-c}$. Sample $2^{n+\mathbf{K}(n,\mu,\rho,\lambda)+c-1}$ elements $D\subset\{0,1\}^\infty$ according to λ. The probability that all samples $\beta\in D$ has $\mathbf{D}(\beta|\mu)\le n$ or $\mathbf{D}(\beta|\rho)\le n$ is

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

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

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

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,

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

Therefore,

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

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 $\mathbf{D}(\beta|\mu)\le n$ or $\mathbf{D}(\beta|\rho)\le n$ 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

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

Therefore:

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

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

$$n + \mathbf{K}(n, \mu, \lambda) + c <^+ n + \mathbf{K}(n, \mu, \lambda) + 2\mathbf{K}(c) + d\mathbf{K}(c)$$
$$c <^+ (2 + d)\mathbf{K}(c).$$

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

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.

[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.

[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.