THE GAMMA QUESTION FOR MANY-ONE DEGREES

MATTHEW HARRISON-TRAINOR

ABSTRACT. A set A is coarsely computable with density $r \in [0,1]$ if there is an algorithm for deciding membership in A which always gives a (possibly incorrect) answer, and which gives a correct answer with density at least r. To any Turing degree $\mathbf{a}$ we can assign a value $\Gamma_T(\mathbf{a})$: the minimum, over all sets A in $\mathbf{a}$, of the highest density at which A is coarsely computable. The closer $\Gamma_T(\mathbf{a})$ is to 1, the closer $\mathbf{a}$ is to being computable. Andrews, Cai, Diamondstone, Jockusch, and Lempp noted that Γ_T can take on the values 0, 1/2, and 1, but not any values in strictly between 1/2 and 1. They asked whether the value of Γ_T can be strictly between 0 and 1/2. This is the Gamma question.

Replacing Turing degrees by many-one degrees, we get an analogous question, and the same arguments show that Γ_m can take on the values 0, 1/2, and 1, but not any values strictly between 1/2 and 1. We will show that for any $r \in [0,1/2]$, there is an m-degree $\mathbf{a}$ with $\Gamma_m(\mathbf{a}) = r$. Thus the range of Γ_m is $[0,1/2] \cup \{1\}$.

Benoit Monin has recently announced a solution to the Gamma question for Turing degrees. Interestingly, his solution gives the opposite answer: the only possible values of Γ_T are 0, 1/2, and 1.

1. INTRODUCTION

We give a solution to the Gamma question for many-one degrees by showing that for each $r \in [0,1/2]$, there is a many-one degree $\mathbf{a}$ such that $\Gamma_m(\mathbf{a}) = r$.

A set $A \subseteq \omega$ is *coarsely computable* if, roughly speaking, we have an algorithm for deciding membership in A which always gives an answer, and the answer is correct except on a set of density zero. By density, we mean asymptotic lower density.

Definition 1. The *lower density* of a set $Z \subseteq \omega$ is
$$\underline{\rho}(Z) := \liminf_{n \to \infty} \frac{|Z \cap [0,n)|}{n}.$$

More generally, we can talk about algorithms which are correct half the time, or a third of the time, or almost never. To a set $A \subseteq \omega$, we can assign a real number which measures the highest density to which it can be approximated by a computable set.

Definition 2 ([HJMS16])**.** A set $A \subseteq \omega$ is *coarsely computable at density* $r \in [0,1]$ if there is a computable set R such that $\underline{\rho}(A \leftrightarrow R) = r$. Here, $A \leftrightarrow R$ is the set on which A and R agree:
$$A \leftrightarrow R := \{x \mid x \in A \iff x \in R\}.$$

Definition 3 ([HJMS16])**.** The *coarse computability bound* of a set $A \subseteq \omega$ is

$$\gamma(A) := \sup\{r \mid A \text{ is coarsely computable at density } r\}.$$

That is, $\gamma(A)$ is the supremum, over all computable sets R, of $\underline{\rho}(A \leftrightarrow R)$.

It is known that for each $r \in (0, 1]$, there are sets with coarse computability bound r such that the supremum is obtained, and sets where the supremum is not obtained [HJMS16].

Jockusch and Schupp [JS12] have shown that every non-zero Turing degree contains a set which is not coarsely computable. (This follows from the proof of Proposition 6 below.) Thus, if $\Gamma_T(\mathbf{a}) = 1$, then $\mathbf{a} = \mathbf{0}$. Andrews, Cai, Diamondstone, Jockusch, and Lempp suggested assigning to each Turing degree a real number which measures the extent to which all sets computable in that degree can be coarsely computed.

Definition 4 ([ACD$^+$16])**.** The *coarse computability bound* of a Turing degree $\mathbf{a}$ is

$$\Gamma_T(\mathbf{a}) := \inf\{\gamma(A) \mid A \text{ is } \mathbf{a}\text{-computable}\}.$$

It suffices to take the infimum only over sets in $\mathbf{a}$.

Andrews, Cai, Diamondstone, Jockusch, and Lempp showed that $\Gamma_T(\mathbf{a})$ can take on the values 0, 1/2, and 1.[1]

Theorem 5 ([ACD$^+$16])**.** *For a Turing degree $\mathbf{a}$:*

(1) *If $\mathbf{a}$ is computable, $\Gamma_T(\mathbf{a}) = 1$.*
(2) *If $\mathbf{a}$ is computably traceable and non-computable, $\Gamma_T(\mathbf{a}) = 1/2$.*
(3) *If $\mathbf{a}$ is 1-random and hyperimmune-free, $\Gamma_T(\mathbf{a}) = 1/2$.*
(4) *If $\mathbf{a}$ is hyperimmune, $\Gamma_T(\mathbf{a}) = 0$.*
(5) *If $\mathbf{a}$ is PA, $\Gamma_T(\mathbf{a}) = 0$.*

Hirschfeldt, Jockusch, McNicholl, and Schupp showed that $\Gamma_T(\mathbf{a})$ cannot take on any values in the open interval $(1/2, 1)$. We will repeat the proof here because we will reference it later.

Proposition 6 ([HJMS16])**.** *Let $\mathbf{a}$ be a nonzero Turing degree. Then $\Gamma_T(\mathbf{a}) \leq \frac{1}{2}$.*

Proof. Fix $A \in \mathbf{a}$. We will show that there is $B \leq_m A$ such that $\gamma(B) \leq \frac{1}{2}$. The idea is that each bit of A will be copied many times by B, so that if we have a computable approximation to B which is correct more than half the time, we can correctly guess at the bits of A with only finitely many errors.

For each $n \in \omega$, define $I_n = [n!, (n + 1)!)$. Let

$$B = \bigcup_{n \in A} I_n.$$

It is easy to see that $B \leq_m A$. Suppose towards a contradiction that $\gamma(B) > \frac{1}{2}$. Let R be a computable approximation to B, with $\underline{\rho}(B \leftrightarrow R) > \frac{1}{2}$. Fix N and p such that for all $n \geq N$,

$$\frac{|(B \leftrightarrow R) \cap [0, n)|}{n} \geq p > \frac{1}{2}.$$

Increasing N, we may assume that $\frac{1}{2} + \frac{1}{N} < p$.

---

[1] See also [MN15] for a unifying approach to some of these examples.

Given $n \geq N$, we will show how to decide computably whether $n \in A$. We claim that $n \in A$ if and only if more than half of the elements of I_n are in R. Indeed, suppose that $n \in A$, but at most half of the elements of I_n are in R. Then

$$\frac{|(B \leftrightarrow R) \cap [0, (n+1)!)|}{(n+1)!} \leq \frac{n! + \frac{(n+1)!-n!}{2}}{(n+1)!} = \frac{1}{2} + \frac{1}{2(n+1)} < p.$$

This is a contradiction. So if $n \in A$, more than half of the elements of I_n are in R. A similar argument works when $n \notin A$. $\qquad\square$

The Gamma question, from [ACD$^+$16], asks whether the value of Γ_T can be strictly between 0 and 1/2. Monin [Mon] has recently given a solution to the Gamma question: The only possible values of Γ_T are 0, 1/2, and 1.

Our work grew out of an independent attempt to answer the Gamma question. If we replace Turing reducibility by many-one reducibility, we get a Gamma function on many-one degrees:

Definition 7. The *coarse computability bound* of an m-degree $\mathbf{a}$ is

$$\Gamma_m(\mathbf{a}) := \inf\{\gamma(A) \mid A \leq_m \mathbf{a}\}.$$

It suffices to take the infimum only over sets in $\mathbf{a}$.

The proof of Proposition 6 used a many-one reduction, so it still holds for m-degrees. Moreover, the examples in Theorem 5 yield examples of m-degrees with Γ_m being 0, 1/2, and 1. Thus, we can ask the Gamma question for m-degrees: Can the value of Γ_m be strictly between 0 and 1/2? Interestingly, we get the opposite answer from Monin's: Every $p \in [0, 1/2]$ is a possible value of Γ_m.

Theorem 8. *Fix* $0 \leq p \leq \frac{1}{2}$. *There is an m-degree $\mathbf{a}$ with* $\Gamma_m(\mathbf{a}) = p$.

Versions of the Gamma question for weaker reducibilities have already been asked in the literature: In [Hir], Hirschfeldt asked the Gamma question for truth table degrees. (Monin's answer to the Gamma question for Turing degrees also yields the same answer for truth table degrees: The value of Γ_{tt} cannot be strictly between 0 and 1/2.) An interesting question is what happens for intermediate reductions, such as bounded truth table reductions. Do such reductions have enough computational power to apply the theorems from coding theory used by Monin, or are they sufficiently simple to allow a construction such as the one we use for many-one degrees?

2. Background on the Hypergeometric Distribution

The proof of Theorem 8 will make use of a probabilistic argument about random variables following a hypergeometric distribution. We will quickly review this distribution here. (See [HPS71, p. 52].)

The hypergeometric distribution is the discrete probability distribution of the number of successes in K draws, without replacement, from a population of size N which contains n successes. For example, one might think of red and blue marbles in a box; if there are N marbles, n of which are red, and we randomly select K marbles, the number of red marbles we pick will follow a hypergeometric distribution. We denote the hypergeometric distribution by $H(K, N, n)$ and, if $X \sim H(K, N, n)$, we have

$$Pr(X = x) = \frac{\binom{n}{x}\binom{N-n}{K-x}}{\binom{N}{K}}.$$

Our particular application will be to have a set U of size N, with a subset V of size n. If $p > q$ are real numbers in $[0,1]$, we will randomly pick from U a set S consisting of about pN elements. We want to choose, in S, at least qn elements which are also in V. Since $p > q$, it is reasonable to think that we should often get enough elements of V. Intuitively, the larger N and n are, the more likely we are to get want we want. Precise bounds are given by the following theorem:

Theorem 9 ([Hoe63], see also [Chv79]). *Let $X \sim H(K, N, n)$ where $p = K/N > q$. Let $t = p - q$. Then*

$$Pr(X \le qn) \le \exp(-2t^2 n).$$

It will be important that this bound does not depend on N (though of course, as n becomes bigger, N will as well).

3. PROOF OF THE MAIN THEOREM

We will now prove Theorem 8.

Proof of Theorem 8. Fix $0 < p < \frac{1}{2}$. We will find a set A whose m-degree $\mathbf{a}$ has $\Gamma_m(\mathbf{a}) = p$. Fix $(C_\ell)_{\ell \in \omega}$ a non-effective list of the computable sets, in which each set is repeated infinitely many times.

We will ensure that $\gamma(A) \le p$ having, for each computable set C_ℓ,

$$\liminf_{n \to 0} \frac{|(A \leftrightarrow C_\ell) \cap [0, n)|}{n} \le p.$$

We will accomplish this by making sure that for each ℓ, there are infinitely many values of n for which we force A to differ from C_ℓ on a large portion of $[0, n)$. This will force $\Gamma_m(\mathbf{a}) \le p$. In fact, since each computable set appears infinitely many times in the list $(C_\ell)_{\ell \in \omega}$, it suffices to find, for each ℓ, a single $n \ge \ell$ with

$$\frac{|(A \leftrightarrow C_\ell) \cap [0, n)|}{n} \le p + \epsilon_\ell$$

where $\epsilon_\ell \to 0$.

To have $\Gamma_m(\mathbf{a}) \ge p$, we must make sure that for each set B which is m-reducible to A via f, $\gamma(B) \ge p$. (One such B will be A itself via the identity reduction.) We will think of A as being approximated by $A^* = \varnothing$. (So we want the bits of A to be 0's with density at least p.) Thus we might initially try to approximate B by $B^* = \varnothing$, which is what we get by applying the reduction f to A^*. This will not work, for the reason that f could be highly non-injective. For example, if f maps every element to the same element y, then we could have $A = \{y\}$, which is very well approximated by A^*, but applying the reduction f we get $B = \omega$ which is very badly approximated by B^*. This is where we will exploit the fact that $p \le 1/2$. Say $z = f(x) = f(y)$. Then if we put $x \in B^*$ but $y \notin B^*$, we are guaranteed to be right about at least one of the two; for if $z \in A$, then $x, y \in B$ in which case we were right about x, and if $z \notin A$, then $x, y \notin B$, in which case we were right about y. If we manage this in the right way, we will be correct with density $1/2$. (The proof of Proposition 6 shows that we must do something like this.)

More formally, let $(f_e : \omega \to \omega)_{e \in \omega}$ be a (non-effective) list of the total many-one reductions. For each e, let

$$B_e = f_e^{-1}(A) := \{x \mid f_e(x) \in A\}.$$

We define a computable approximation B_e^* to B_e as follows. First, let $I_1, I_2, I_3, \ldots$ be the consecutive intervals of length one, two, three, and so on. (So $I_1 = \{0\}$, $I_2 = \{1, 2\}$, $I_3 = \{3, 4, 5\}$, etc.) For each interval I_n, let $J_{n,e} = f_e(I_n)$ be the multiset image[2] of I_n under f_e, and write $J_{n,e} = J_{n,e}^* \uplus 2J_{n,e}^{**}$ where each element has multiplicity one in $J_{n,e}^*$. So, for example, if $f_e(3) = 0$, $f_e(4) = 0$, and $f_e(5) = 1$, then $f_e(I_3) = \{0, 0, 1\} = \{1\} \uplus 2\{0\}$; if $J_{8,e} = f_e(I_8) = \{0, 0, 0, 0, 0, 1, 1, 2\}$, then $J_{8,e} = \{0, 2\} \uplus 2\{0, 0, 1\}$. We can write I_n as a disjoint union $I_{n,e}^* \cup I_{n,e}^{**,1} \cup I_{n,e}^{**,2}$, where $f_e(I_{n,e}^{**,1}) = f_e(I_{n,e}^{**,2}) = J_{n,e}^{**}$ and $f_e(I_{n,e}^*) = J_{n,e}^*$. Then let $B^* = \bigcup_n I_{n,e}^{**,1}$. The simplicity with which we can describe B^* is where we take advantage of the fact that we are considering many-one reductions rather than Turing reductions. We will have that, for some decreasing positive sequence $\gamma_n \to 0$, and for all n,

$$(\dagger) \qquad \frac{|(B_e \leftrightarrow B_e^*) \cap I_n|}{n} \geq p - \gamma_n.$$

Since the length of the intervals I_n are increasing slowly, this will suffice to get $\gamma(B_e) \geq p$.

Claim 10. *Assuming* $(\dagger)$, $\gamma(B_e) \geq p$.

Proof. Since

$$\frac{|(B_e \leftrightarrow B_e^*) \cap I_m|}{m} \geq p - \gamma_m,$$

for each m, there is K_m such that for all $K \geq K_m$,

$$\frac{|(B_e \leftrightarrow B_e^*) \cap (\bigcup_{n \leq K} I_n)|}{\sum_{n \leq K} n} \geq p - \gamma_m.$$

(Assume that the sequence K_m is strictly increasing in m.) Given m, let $x \in I_{N+1}$ for some $N \geq K_m$, with N sufficiently large that $\frac{N-1}{N+1}(p - \gamma_m) \geq p - 2\gamma_m$. Then

$$\left| (B_e \leftrightarrow B_e^*) \cap \Big(\bigcup_{n < N} I_n \Big) \right| \leq |(B_e \leftrightarrow B_e^*) \cap [0, x]|.$$

So

$$\frac{|(B_e \leftrightarrow B_e^*) \cap [0, x]|}{x + 1} \geq \frac{(\sum_{n<N} n)}{x + 1} \frac{|(B_e \leftrightarrow B_e^*) \cap (\bigcup_{n<N} I_n)|}{(\sum_{n<N} n)} \geq \frac{(\sum_{n<N} n)}{x + 1}(p - \gamma_m).$$

Now

$$\frac{(\sum_{n<N} n)}{x + 1} \geq \frac{(\sum_{n<N} n)}{(\sum_{n \leq N} n)} = \frac{N - 1}{N + 1}.$$

By choice of N, $\frac{N-1}{N+1}(p - \gamma_m) \geq p - 2\gamma_m$, and so

$$\frac{|(B_e \leftrightarrow B_e^*) \cap [0, x]|}{x + 1} \geq p - 2\gamma_m.$$

Thus

$$\gamma(B_e) = \liminf_{x \to \infty} \frac{|(B_e \leftrightarrow B_e^*) \cap [0, x]|}{x} \geq p. \qquad \square$$

---

[2]Recall that multisets are a generalization of sets to allow multiple instances of the same element. By the multiset image, we mean that we want to count the number of pre-images of an element in the range.

Note that one of the reductions f_e is the identity reduction, so for that e, $B_e = A$. Thus $\gamma(A) = p$. Hence $\Gamma_m(A) = p$.

We now construct A as the concatenation of infinitely many finite binary strings: $A = \alpha_1 \widehat{\ } \beta_1 \widehat{\ } \alpha_2 \widehat{\ } \beta_2 \widehat{\ } \alpha_3 \widehat{\ } \ldots$. Each α_i will consist entirely of 0's.[3] Note that we have already defined each of the approximations B_e^* computably; our construction of A can be non-effective. Fix $(\epsilon_\ell)_{\ell \in \omega}$ a decreasing sequence of positive reals converging to zero (and assume that the ϵ_ℓ are small relative to p, so that for example $p - \epsilon_\ell > 0$ and $p + \epsilon_\ell < 1/2$). Given $\alpha_1 \widehat{\ } \beta_1 \widehat{\ } \alpha_2 \widehat{\ } \beta_2 \widehat{\ } \cdots \widehat{\ } \beta_\ell$, we must define $\alpha_{\ell+1}$ and $\beta_{\ell+1}$. Let L be the length of $\alpha_1 \widehat{\ } \beta_1 \widehat{\ } \alpha_2 \widehat{\ } \beta_2 \widehat{\ } \cdots \widehat{\ } \beta_\ell$.

At stage ℓ, we consider only the reductions f_e for $e \leq \ell$. We have two competing desires. First, we want to define $\beta_{\ell+1}$ so that A has very little agreement with C_ℓ on this part of their domain. Second, we want $\beta_{\ell+1}$ to have many 0's, particularly on elements in the ranges of the f_e, so that the sets B_e have many 0's. Our probabilistic argument using the hypergeometric distribution will show that we can satisfy both of these desires at the same time.

At stage ℓ, we will divide the intervals I_n which make up the domain of B_e into three types, depending on their size: small, medium, and large. An interval will be of medium size for exactly one stage ℓ; at earlier stages, it will be large, and at later stages, it will be small. Small intervals are too small to use the bounds for the hypergeometric distribution while keeping any errors under ϵ_ℓ; we will ensure that their images, under the f_e, look only at the values in $\alpha_{\ell+1}$ and earlier. The medium and large intervals are large enough to use the bounds for the hypergeometric distribution. The difference between the medium and large intervals will not show up until the verification. Essentially, large intervals are so big that the values of $\beta_{\ell+1}$ do not affect the agreement of B_e with B_e^*.

Choose M sufficiently large (and bigger than the value of M at the previous stage) so that:

(1) $\ell \sum_{n \geq M} \exp(-\epsilon_\ell^3 n) < 1$ and
(2) if $M' \geq M$ and $M' = 2m_1 + m_2$, then $\frac{m_1 + pm_2 - L}{M'} \geq p - \epsilon_\ell$.

Since $p \leq 1/2$, (2) holds for $M \geq L/\epsilon_\ell$. We can get (1) to hold for sufficiently large M because the infinite series $\ell \sum_{n \geq 1} \exp(-\epsilon_\ell^3 n)$ converges. M is the cutoff between the small and medium intervals. (1) says that M is big enough to apply the tail bounds for the hypergeometric distribution, and (2) says that M is large enough compared to L that we do not have to worry about what we have already chosen for $\alpha_1 \widehat{\ } \beta_1 \widehat{\ } \alpha_2 \widehat{\ } \beta_2 \widehat{\ } \cdots \widehat{\ } \beta_\ell$.

Choose K sufficiently large so that each element of $J_{i,e}$, for $i < M$ and $e \leq \ell$, is less than $L + K$. Set $\alpha_{\ell+1} = 0^K$. Thus every small interval $I_{i,e}$ will see, under the reduction f_e, only $\alpha_1 \widehat{\ } \beta_1 \widehat{\ } \alpha_2 \widehat{\ } \beta_2 \widehat{\ } \cdots \widehat{\ } \beta_\ell \widehat{\ } \alpha_{\ell+1}$. Our choice of $\beta_{\ell+1}$ will not affect the reduction on these intervals.

Then choose N sufficiently large that so that $\frac{L + K + pN + \epsilon_\ell N}{L + K + N} \leq p + 2\epsilon_\ell$ and $\frac{\lfloor (p+\epsilon_\ell) N \rfloor}{N} - p \geq \epsilon_\ell/2$. We will have $|\beta_{\ell+1}| = N$. N is large enough that, by making $\beta_{\ell+1}$ agree with C_ℓ at density at most p, we can make A agree with C_ℓ at density at most $p + \epsilon_\ell$. This N will be the N in our applications of the tail bounds for the hypergeometric distribution. Because K was chosen dependent on M, and N was chosen dependent

---

[3] We should expect to have long sequences of zeros. Since $\gamma(A) = p \leq 1/2$, $A \leftrightarrow \omega$ should have density at most p. But that means that the upper destiny of $A \leftrightarrow \varnothing$ should be at least $1 - p \geq 1/2$, so there should be initial segments of A with many 0's.

on K, it is important here that these tail bounds depend only on n (which will be bounded below by a fixed multiple of M).

As before, write $J_{n,e} = f_e(I_n)$ as $J_{n,e}^* \uplus 2J_{n,e}^{**}$ where each element has multiplicity one in $J_{n,e}^*$. Recall that, by choice of B_e^*, we are guaranteed to have B_e agree with B_e^* on half of the elements in $I_{n,e}$ which map to $J_{n,e}^{**}$. If the size of $J_{n,e}^{**}$ is very close to the size of $J_{n,e}$, then we are guaranteed to get agreement of close to $1/2$ on $I_{n,e}$, and we do not have to worry about $I_{n,e}$. The other option is that $J_{n,e}^*$ is very large, in which case we get small tail bounds for the hypergeometric distribution. More formally, if $n \geq M$, then either $\frac{|J_{n,e}^{**}|}{|J_{n,e}|} \geq \frac{1}{2} - \epsilon_\ell$ or $|J_{n,e}^*| > 2\epsilon_\ell|J_{n,e}|$. Indeed, if we are not in the former case, then

$$2|J_{n,e}^{**}| < |J_{n,e}| - 2\epsilon_\ell|J_{n,e}|.$$

Rearranging, and using the fact that $|J_{n,e}| = |J_{n,e}^*| + 2|J_{n,e}^{**}|$, we get

$$|J_{n,e}^*| > 2\epsilon_\ell|J_{n,e}|.$$

Let Ω index the pairs $n \geq M$ and $e \leq \ell$ for which $|J_{n,e}^*| > 2\epsilon_\ell|J_{n,e}|$.

Claim 11. *There is a set $S \subseteq [L+K, L+K+N)$ with $\frac{|S|}{N} \leq p + \epsilon_\ell$ such that for each $(n,e) \in \Omega$,*

$$\frac{|S \cap J_{n,e}^*| + |J_{n,e}^* \smallsetminus [L+K, L+K+N)|}{|J_{n,e}^*|} \geq p.$$

The set S is a set on which A will be forced to have 0's. On the other elements of $[L+K, L+K+N)$, A will have the freedom to be different from C_ℓ. This claim says that we can choose S to simultaneously have S small enough that A can be sufficiently different from C_ℓ and large enough that the reductions f_e find sufficiently many 0's in their ranges.

Proof. First, note that if we modify $J_{n,e}^*$ by removing an element which is outside of the interval $[L+K, L+K+N)$ and adding a new element which is inside of this interval, for any fixed set $S \subseteq [L+K, L+K+N)$ the quantity

$$\frac{|S \cap J_{n,e}^*| + |J_{n,e}^* \smallsetminus [L+K, L+K+N)|}{|J_{n,e}^*|}$$

can only decrease. Also, if $J_{n,e}^* \supseteq [L+K, L+K+N)$, then for any choice of S with $\frac{|S|}{N} \geq p$ we will have

$$\frac{|S \cap J_{n,e}^*| + |J_{n,e}^* \smallsetminus [L+K, L+K+N)|}{|J_{n,e}^*|} \geq p$$

as desired. So we may assume that, for each (n,e), $J_{n,e}^* \subseteq [L+K, L+K+N)$.

We give a probabilistic argument that the desired set S exists. Imagine that we randomly pick a set S of size $r = \lfloor (p + \epsilon_\ell)N \rfloor$. For each $(n,e) \in \Omega$, let $X_{n,e}$ be the random variable $|S \cap J_{n,e}^*|$; we have $X_{n,e} \sim H(r, N, |J_{n,e}^*|)$. Let $t = \frac{r}{N} - p$. By choice of N, we have $t \geq \epsilon_\ell/2$. So by the tail bounds for the hypergeometric distribution, the probability that for some fixed $(e,n) \in \Omega$, $|S \cap J_{n,e}^*| \leq p|J_{n,e}^*|$ is bounded above by

$$Pr(X_{e,n} \leq p|J_{n,e}^*|) \leq \exp(-2t^2|J_{n,e}^*|) \leq \exp(-\epsilon_\ell^3|J_{n,e}|) = \exp(-\epsilon_\ell^3 n).$$

(Note that this holds even if $p|J^*_{n,e}|$ is not an integer.) So the probability that for all $(e, n) \in \Omega$, $|S \cap J^*_{n,e}| \le p|J^*_{n,e}|$, is bounded above by

$$\sum_{(e,n)\in\Omega} \exp(-\epsilon_\ell^3 n) \le \ell \sum_{n\ge M} \exp(-\epsilon_\ell^3 n) < 1.$$

So there is a non-zero probability that we pick a set S as desired; some such set must exist. □

For $i < N$, when $L + K + i \in S$, set $\beta_{\ell+1}(i) = 0$, and otherwise set $\beta_{\ell+1}(i) \ne C_\ell(L + K + i)$. So for $x \in [L + K, L + K + N)$, if $x \in S$ then $A(x) = 0$ and if $x \notin S$ then $A(x) \ne C_\ell(x)$.

First, we will show that we made A sufficiently different from C_ℓ.

Claim 12. $\gamma(A) \le p$.

Proof. We have, for each ℓ, that

$$\frac{|(A \leftrightarrow C_\ell) \cap [0, L + K + N)|}{L + K + N} \le \frac{L + K + |\{x \in S \mid A(i) = C_\ell(i)\}|}{L + K + N}$$
$$\le \frac{L + K + |S|}{L + K + N}$$
$$\le \frac{L + K + pN + \epsilon_\ell N}{L + K + N}.$$

Here, L, K, and N are the values of those variables at stage ℓ of the construction. By choice of N,

$$\frac{L + K + pN + \epsilon_\ell N}{L + K + N} \le p + 2\epsilon_\ell.$$

So

$$\frac{|(A \leftrightarrow C_\ell) \cap [0, L + K + N)|}{L + K + N} \le p + 2\epsilon_\ell.$$

Then, noting that for each ℓ there are infinitely many ℓ' with $C_\ell = C_{\ell'}$,

$$\gamma(A) = \liminf_{n\to\infty} \frac{|(A \leftrightarrow C_\ell) \cap [0, n)|}{n} \le p. \qquad \square$$

Second, we will show that B_e is sufficiently well approximated by B^*_e. The following claim verifies the hypotheses of Claim 10.

Claim 13. *Fix e. Given n, let $\ell \ge e$ be such that $M_{\ell+1} > n \ge M_\ell$, where M_ℓ is the value of M at stage ℓ. Then*

$$\frac{|(B_e \leftrightarrow B^*_e) \cap I_n|}{n} \ge p - \epsilon_\ell.$$

The n satisfying $M_{\ell+1} > n \ge M_\ell$ are the sizes of intervals of medium size for ℓ, that is, those where β_ℓ is exactly the right length to determine the amount of agreement between B_e and B^*_e on the interval I_n.

Proof. Write $I_e = I^{*,1}_{n,e} \cup I^{*,2}_{n,e} \cup I^{*,3}_{n,e} \cup I^{**,1}_{n,e} \cup I^{**,2}_{n,e}$ where $I^{**,1}_{n,e}$ and $I^{**,2}_{n,e}$ are as before, and $f_e(I^{*,1}_{n,e}) \subseteq [0, L)$, $f_e(I^{*,2}_{n,e}) \subseteq [L + K, L + K + N)$, and $f_e(I^{*,3}_{n,e}) \subseteq [L, L + K) \cup [L + K + N, \infty)$. (Again, L, K, and N are the values at stage ℓ.)

For $x \in I^{*,3}_{n,e}$, $f_e(x) \in [L, L + K) \cup (L + K + N, \infty]$, and so, since $n < M_{\ell+1}$, $B^*_e(x) = 0 = A(f_e(x)) = B_e(x)$. Thus

$$|(B_e \leftrightarrow B^*_e) \cap I^{*,3}_{n,e}| = |I^{*,3}_{n,e}|.$$

For $I_{n,e}^{*,2}$, if $f(x) \in S$, then $B_e(x) = B_e^*(x)$ since $x \notin B_e^*$ and $f_e(x) \notin A$, so that $x \notin B_e$. Thus (recalling that $J_{n,e}^{*,2} = f_e(I_{n,e}^{*,2})$, and that f_e is injective on this set):

$$|(B_e \leftrightarrow B_e^*) \cap I_{n,e}^{*,2}| \geq |S \cap J_{n,e}^{*,2}|.$$

By choice of S, we have

$$\frac{|S \cap J_{n,e}^*| + |J_{n,e}^* \smallsetminus [L+K, L+K+N)|}{|J_{n,e}^*|} \geq p.$$

Thus, noting that $|J_{n,e}^* \smallsetminus [L+K, L+K+N)| = |I_{n,e}^{*,1}| + |I_{n,e}^{*,3}|$,

$(*)$ $$|(B_e \leftrightarrow B_e^*) \cap I_{n,e}^*| + L \geq p|I_{n,e}^*|.$$

By definition of B_e^*,

$(**)$ $$|(B_e \leftrightarrow B_e^*) \cap I_{n,e}^{**}| = |I_{n,e}^{**,1}| = |I_{n,e}^{**,2}|.$$

This is because each $x \in I_{n,e}^{**,1}$ can be paired with a $y \in I_{n,e}^{**,2}$ with $f_e(x) = f_e(y)$; we have $x \in B_e^*$ and $y \notin B_e^*$. Either $f_e(x) = f_e(y) \in A$, in which case $x, y \in B_e$, or $f_e(x) = f_e(y) \notin A$, in which case $x, y \notin B_e$.

So combining equations $(*)$ and $(**)$, we get

$$|(B_e \leftrightarrow B_e^*) \cap I_{n,e}| \geq |I_{n,e}^{**,1}| + p|I_{n,e}^*| - L.$$

By choice of M_ℓ,

$$\frac{|I_{n,e}^{**,1}| + p|I_{n,e}^*| - L}{|I_{n,e}|} \geq p - \epsilon_\ell. \qquad \square$$

This completes the proof of the theorem, as Claim 10 now gives that $\gamma(B_e) \geq p$ for each e. $\qquad \square$

References

[ACD+16] Uri Andrews, Mingzhong Cai, David Diamondstone, Carl Jockusch, and Steffen Lempp. Asymptotic density, computable traceability, and 1-randomness. *Fundamenta Mathematicae*, 2016.

[Chv79] Vašek Chvátal. The tail of the hypergeometric distribution. *Discrete Math.*, 25(3):285–287, 1979.

[Hir] Denis Hirschfeldt. Some questions in computable mathematics. To appear.

[HJKS16] Denis R. Hirschfeldt, Carl G. Jockusch, Rutger Kuyper, and Paul E. Schupp. Coarse reducibility and algorithmic randomness. *J. Symb. Log.*, 81(3):1028–1046, 2016.

[HJMS16] Denis Hirschfeldt, Carl Jockusch, Timothy McNicholl, and Paul Schupp. Asymptotic density and the coarse computability bound. *Computability*, 5(1):13–27, 2016.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[HPS71] Paul Hoel, Sidney Port, and Charles Stone. *Introduction to probability theory*. Houghton Mifflin Co., Boston, Mass., 1971. The Houghton Mifflin Series in Statistics.

[JS12] Carl Jockusch and Paul Schupp. Generic computability, Turing degrees, and asymptotic density. *J. Lond. Math. Soc.*, 85(2):472–490, 2012.

[MN15] Benoit Monin and André Nies. A unifying approach to the gamma question. *Proceedings of the Thirtieth Annual IEEE Symposium on Logic in Computer Science (LICS 2015)*, pages 585–596, July 2015.

[Mon] Benoit Monin. Asymptotic density and error-correcting codes. Preprint.

GROUP IN LOGIC AND THE METHODOLOGY OF SCIENCE, UNIVERSITY OF CALIFORNIA, BERKELEY, USA

E-mail address: matthew.h-t@berkeley.edu

URL: www.math.berkeley.edu/~mattht