

CRAMÉR α -RANDOM PRIMES AND THE FUNDAMENTAL THEOREM OF ARITHMETIC

ERICK ROSS

ABSTRACT. Define the set of Cramér α -random primes, where N is chosen to be a Cramér α -random prime with probability $1/\log^\alpha N$. At $\alpha = 1$, this is the classical set of Cramér random primes, which model the actual primes. Now, the Fundamental Theorem of Arithmetic states that every natural number can be written uniquely as a product of primes (allowing multiplicity). In this paper, we investigate how close the Cramér α -random primes come to satisfying this property. Along the way, we also prove an analog of the Hardy-Ramanujan inequality for the classical Cramér random primes.

1. INTRODUCTION

The Prime Number Theorem states that the number of primes less than x grows like

$$\#\{p \leq x \text{ prime}\} \sim \text{Li}(x) := \int_2^x \frac{1}{\log t} dt.$$

From this asymptotic formula, it is natural to attempt to model the primes as a set of random numbers, where N occurs with probability $1/\log N$. This is called the Cramér random prime model, and it turns out to do a relatively good job of predicting many different statistical properties of the primes (with, of course, several notable exceptions [3], [6]).

Now, in some sense, the defining characteristic of the actual primes is the Fundamental Theorem of Arithmetic: that every natural number can be written uniquely as a product of primes. So a natural question to ask is how close the Cramér random primes come to satisfying this property. More generally, we will consider this question for the Cramér α -random primes, where N is chosen as a Cramér α -random prime with probability $1/\log^\alpha N$.

To be precise, fix $\alpha > 0$ and define the independent Bernoulli random variables ξ_N , where

$$\mathbb{P}[\xi_N = 1] = \nu(N) := \begin{cases} 0 & \text{if } N = 1 \\ 1 & \text{if } N = 2 \\ 1/\log^\alpha N & \text{if } N \geq 3. \end{cases}$$

2020 *Mathematics Subject Classification.* Primary: 11K45; Secondary: 11N05, 11B05.

Key words and phrases. Cramér random primes; Fundamental Theorem of Arithmetic; asymptotic density.

Then the set of Cramér α -random primes is defined to be $A_\alpha := \{N \geq 1 \mid \xi_N = 1\}$. Note that the parameter α in this definition allows us to compare factorization behavior between random sets of natural numbers appearing as likely as the actual primes, slightly more likely than the actual primes, and slightly less likely than the actual primes.

In Sections 2 - 5, we will compute how close the Cramér α -random primes come to satisfying the Fundamental Theorem of Arithmetic. Specifically, we prove the following bounds on how many Cramér-factorizations (i.e. factorizations of N into Cramér α -random primes) one can expect the natural numbers to have.

Theorem 1.1. *Let $b(N)$ denote the expected number of Cramér-factorizations of N . Then the average value of $b(N)$ is*

$$\frac{1}{x} \sum_{N \leq x} b(N) = \begin{cases} \Omega(\log^T x) & \text{for arbitrarily large } T & \text{if } \alpha < 1 \\ \Omega((\log \log x)^T / \log x) & \text{for arbitrarily large } T & \text{if } \alpha = 1 \\ O(1/\log^\alpha x) & & \text{if } \alpha > 1. \end{cases}$$

Moreover, for $\alpha = 1$, the average value of $b(N)$ over squarefree N is

$$\frac{1}{\frac{6}{\pi^2}x} \sum_{\substack{N \leq x \\ \text{squarefree}}} b(N) = O(1) \quad \text{for } \alpha = 1.$$

Here, the Bachmann-Landau notation $f(x) = \Omega(g(x))$ means that there exist constants $C, x_0 > 0$ such that $f(x) \geq Cg(x)$ for all $x \geq x_0$. In other words, big- Ω notation denotes a lower bound, just like big- O notation denotes an upper bound.

We make four remarks about Theorem 1.1. First, we note these results about the average growth of $b(N)$ cannot be improved to results about the growth of $b(N)$ itself; the individual values of $b(N)$ are much too erratic. For example, it is straightforward to verify that for any given α , $b(N) \rightarrow 0$ along primes N , whereas $b(N) \rightarrow \infty$ along N such that $\sigma_0(N) \geq \log^{2\alpha+1} N$. Second, we note that Theorem 1.1 predicts what one would expect intuitively: the average value of $b(N)$ tends to ∞ for Cramér α -random primes appearing slightly more likely than the actual primes, and it tends to 0 for Cramér α -random primes appearing slightly less likely than the actual primes. However, Theorem 1.1 gives explicit growth rates regarding how fast these quantities tend to ∞ and 0. Third, we point out that the above results are all effective. For example, in the last case of Theorem 1.1, we compute explicitly that the average value of $b(N)$ over squarefree N is $\leq \pi^2 e^3 / 3$. Finally, we remark that in order to show the last case of Theorem 1.1, we prove an analog of the Hardy-Ramanujan inequality for the classical Cramér random primes.

In Section 6, we also address an alternative way to interpret how close the Cramér α -random primes come to satisfying the Fundamental Theorem of Arithmetic. In particular, we investigate how many natural numbers can be written as a product of Cramér random primes. Let $\text{mult}(A_\alpha)$ denote the set of such natural numbers (i.e. $\text{mult}(A_\alpha)$ is the multiplicative closure of A_α). Then we show the following result.

Theorem 1.2. *Let A_α denote the set of Cramér α -random primes. Then with probability 1, $\text{mult}(A_\alpha)$ has asymptotic density*

$$d(\text{mult}(A_\alpha)) = \begin{cases} 0 & \text{if } \alpha > 1 \\ 1 & \text{if } \alpha < \frac{1}{2} \log 2. \end{cases}$$

Finally, we discuss our results in Section 7. We take note of some related problems that have been investigated before, and also state a conjecture regarding the asymptotic density of $\text{mult}(A_\alpha)$ for $\frac{1}{2} \log 2 \leq \alpha \leq 1$.

2. A LOWER BOUND WHEN $\alpha < 1$

Fix $\alpha < 1$. In this section, we then show that

$$\frac{1}{x} \sum_{N \leq x} b(N) = \Omega(\log^T x) \quad \text{for arbitrarily large } T.$$

Here, and throughout the rest of the paper, we will use the terminology “factorizations” to refer to Cramér-factorizations. And we will use the terminology “ k -factorizations” to refer to ordered factorizations of length k . Let $b_k(N)$ denote the expected number of k -factorizations of N , and let $B_k(x) := \sum_{N \leq x} b_k(N)$.

Then $B_k(x)$ satisfies the following recurrence relation. For any k -factorization $\leq x$, we can condition on the first element d in the k -factorization. This means that

$$\begin{aligned} B_k(x) &= \sum_{d \leq x} \mathbb{P}[\xi_d = 1] \cdot \mathbb{E}[\#\{(k-1)\text{-factorizations} \leq x/d\} \mid \xi_d = 1] \\ &\geq \sum_{d \leq x} \mathbb{P}[\xi_d = 1] \cdot \mathbb{E}[\#\{(k-1)\text{-factorizations} \leq x/d\}] \\ &= \sum_{d \leq x} \nu(d) B_{k-1}(x/d). \end{aligned} \tag{2.1}$$

Using this recurrence relation, we compute the following bounds on $B_k(x)$.

Lemma 2.1. *Fix $\alpha < 1$, and write $\alpha = 1 - \varepsilon$. Then for all $k \geq 1$, $B_k(x) = \Omega(x \log^{-1+k\varepsilon} x)$.*

Proof. We proceed by induction on k . The base case of $k = 1$ is immediate:

$$B_1(x) = \sum_{N \leq x} b_1(N) = \sum_{N \leq x} \nu(N) \geq \sum_{3 \leq N \leq x} \frac{1}{\log^{1-\varepsilon} N} \geq (x-3) \frac{1}{\log^{1-\varepsilon} x} = \Omega(x \log^{-1+\varepsilon} x).$$

For the inductive step, assume that $B_{k-1}(x) = \Omega(x \log^{-1+k\varepsilon} x)$. So there exist $C, x_0 > 0$ such that $B_{k-1}(x) \geq Cx \log^{-1+k\varepsilon} x$ for all $x \geq x_0$. Then by the recursive formula,

$$\begin{aligned} B_k(x) &\geq \sum_{d \leq x} \nu(d) B_{k-1}(x/d) \\ &\geq \sum_{3 \leq d \leq x/x_0} \nu(d) B_{k-1}(x/d) \\ &\geq \sum_{3 \leq d \leq x/x_0} \frac{1}{\log^{1-\varepsilon} d} C (x/d) \log^{-1+(k-1)\varepsilon}(x/d) \\ &= Cx \sum_{3 \leq d \leq x/x_0} \frac{\log^{(k-2)\varepsilon} x/d}{d((\log x - \log d) \log d)^{1-\varepsilon}} \\ &\geq \frac{Cx}{(\frac{1}{4} \log^2 x)^{1-\varepsilon}} \sum_{3 \leq d \leq x/x_0} \frac{\log^{(k-2)\varepsilon} x/d}{d} \\ &\geq 4^{1-\varepsilon} Cx \log^{-2+2\varepsilon}(x) \sum_{3 \leq d \leq \sqrt{x}} \frac{\log^{(k-2)\varepsilon} x/d}{d} \quad \text{for sufficiently large } x \\ &\geq 4^{1-\varepsilon} Cx \log^{-2+2\varepsilon}(x) \log^{(k-2)\varepsilon}(\sqrt{x}) \sum_{3 \leq d \leq \sqrt{x}} \frac{1}{d} \\ &\geq \frac{4^{1-\varepsilon}}{2^{(k-2)\varepsilon}} Cx \log^{-2+k\varepsilon}(x) \left(\log \lfloor \sqrt{x} \rfloor - \frac{3}{2} \right) \\ &= \Omega(x \log^{-1+k\varepsilon} x), \end{aligned}$$

as desired. □

Now, we can easily show the desired result. For any given $T \geq 1$, choose k large enough that $-1 + k\varepsilon \geq T$. Also, note that there are at most $k!$ ways to rearrange a k -factorization of N , which means that $b(N) \geq \frac{1}{k!} b_k(N)$. Thus by Lemma 2.1,

$$\frac{1}{x} \sum_{N \leq x} b(N) \geq \frac{1}{x} \sum_{N \leq x} \frac{1}{k!} b_k(N) = \frac{1}{x} \frac{1}{k!} B_k(N) = \Omega(\log^T x),$$

proving the first case of Theorem 1.1.

We note here that it is tempting to use Lemma 2.1 to bound $b(N)$ by all the k , instead of just one particular value of k . Specifically, it is tempting to compute that

$$\frac{1}{x} \sum_{N \leq x} b(N) = \frac{1}{x} \sum_{N \leq x} \sum_{k \geq 1} \frac{1}{k!} b_k(N) = \frac{1}{x} \sum_{k \geq 1} \frac{1}{k!} B_k(x) = \frac{1}{x} \sum_{k \geq 1} \frac{1}{k!} \Omega(x \log^{-1+k\varepsilon} x),$$

and then hope that this is equal to

$$\Omega \left(\frac{1}{x} \sum_{k \geq 1} \frac{1}{k!} x \log^{-1+k\varepsilon} x \right) = \Omega \left(\frac{\exp(\log^\varepsilon x)}{\log x} \right).$$

However, this argument does not work because the $B_k(x) = \Omega(x \log^{-1+k\varepsilon} x)$ bound is not uniform in k ; for any fixed x_0 , $x_0 \log^{-1+k\varepsilon} x_0$ is unbounded in k , whereas $B_k(x_0) = 0$ for sufficiently large k .

3. A LOWER BOUND WHEN $\alpha = 1$

Fix $\alpha = 1$. In this section, we then show that

$$\frac{1}{x} \sum_{N \leq x} b(N) = \Omega((\log \log x)^T / \log x) \quad \text{for arbitrarily large } T.$$

We will use the same recurrence relation as in the previous section:

$$B_k(x) \geq \sum_{d \leq x} \nu(d) B_{k-1}(x/d).$$

We can then compute the following bounds on $B_k(x)$.

Lemma 3.1. *Fix $\alpha = 1$. Then for all $k \geq 1$, $B_k(x) = \Omega(x (\log \log x)^{k-1} / \log x)$.*

Proof. We proceed by induction on k . The base case of $k = 1$ is immediate:

$$B_1(x) = \sum_{N \leq x} b_1(N) = \sum_{N \leq x} \nu(N) \geq \sum_{3 \leq N \leq x} \frac{1}{\log N} \geq \frac{x-3}{\log x} = \Omega \left(\frac{x}{\log x} \right).$$

For the inductive step, assume that $B_{k-1}(x) = \Omega(x (\log \log x)^{k-2} / \log x)$. So there exist $C, x_0 > 0$ such that $B_{k-1}(x) \geq Cx (\log \log x)^{k-2} / \log x$ for all $x \geq x_0$. Then by the recursive formula,

$$\begin{aligned} B_k(x) &\geq \sum_{d \leq x} \nu(d) B_{k-1}(x/d) \\ &\geq \sum_{3 \leq d \leq x/x_0} \frac{1}{\log d} \frac{C (x/d) (\log \log x/d)^{k-2}}{\log x/d} \\ &\geq Cx \sum_{3 \leq d \leq \sqrt{x}} \frac{(\log \log x/d)^{k-2}}{d \log d (\log x - \log d)} \quad \text{for sufficiently large } x \end{aligned}$$

$$\begin{aligned}
&\geq Cx(\log \log \sqrt{x})^{k-2} \sum_{3 \leq d \leq \sqrt{x}} \frac{1}{d \log d (\log x - \log d)} \\
&\geq \frac{Cx(\log \log \sqrt{x})^{k-2}}{\log x} \sum_{3 \leq d \leq \sqrt{x}} \frac{1}{d \log d} \\
&\geq \frac{Cx(\log \log \sqrt{x})^{k-2}}{\log x} \int_3^{\sqrt{x}} \frac{1}{t \log t} dt \quad \text{since } t \mapsto \frac{1}{t \log t} \text{ is decreasing} \\
&= \frac{Cx(\log \log \sqrt{x})^{k-2}}{\log x} (\log \log \sqrt{x} - \log \log 3) \\
&= \Omega(x(\log \log x)^{k-1} / \log x),
\end{aligned}$$

as desired. \square

Now, for any given $T \geq 1$, choose k large enough that $k - 1 \geq T$. Then by Lemma 3.1,

$$\frac{1}{x} \sum_{N \leq x} b(N) \geq \frac{1}{x} \sum_{N \leq x} \frac{1}{k!} b_k(N) = \frac{1}{x} \frac{1}{k!} B_k(N) = \Omega((\log \log x)^T / \log x),$$

proving the second case of Theorem 1.1.

4. AN UPPER BOUND WHEN $\alpha = 1$

Fix $\alpha = 1$. In this section, we then show that

$$\frac{1}{\frac{6}{\pi^2}x} \sum_{\substack{N \leq x \\ \text{squarefree}}} b(N) = O(1) \quad \text{for } \alpha = 1.$$

Let $B'_k(x) := \sum_{N \leq x, \text{ sqfr.}} b_k(N)$. Then similarly to (2.1), we have the following recurrence relation for $B'_k(x)$:

$$\begin{aligned}
B'_k(x) &= \sum_{d \leq x} \nu(d) \mathbb{E}[\#\{\text{sqfr. } (k-1)\text{-factorizations } \leq x/d, \text{ coprime to } d\} \mid \xi_d = 1] \\
&= \sum_{d \leq x} \nu(d) \mathbb{E}[\#\{\text{sqfr. } (k-1)\text{-factorizations } \leq x/d, \text{ coprime to } d\}] \\
&\leq \sum_{d \leq x} \nu(d) B'_{k-1}(x/d).
\end{aligned}$$

Note we are restricting to squarefree N in this section because the corresponding recurrence relation becomes an upper bound (as opposed to the lower bound of (2.1)).

With this recurrence relation, we can then show the following bounds on $B'_k(N)$. Note that this result is an analog of the Hardy-Ramanujan inequality for the Cramér random

primes [4, 7]. Here, we have an extra factor of $k!$ appearing (compared to the classical Hardy-Ramanujan inequality) just because we are considering *ordered* factorizations.

Lemma 4.1. *Fix $\alpha = 1$. Then for all $x \geq 3$ and $k \geq 1$, $B'_k(x) \leq 2kx(3 + \log \log x)^{k-1} / \log x$.*

Proof. We proceed by induction on k . The base case of $k = 1$ is immediate;

$$B'_1(x) = \sum_{d \leq x} \nu(d) = 1 + \sum_{3 \leq d \leq x} \frac{1}{\log d} \leq 1 + \int_2^x \frac{1}{\log t} dt \leq \frac{2x}{\log x},$$

as desired.

For the inductive step, take $k \geq 2$ and assume that $B'_{k-1}(x) \leq 2(k-1)x(3 + \log \log x)^{k-2} / \log x$ for all $x \geq 3$. Then

$$\begin{aligned} B'_k(x) &\leq \sum_{d \leq x} \nu(d) B'_{k-1}(x/d) \\ &= B'_{k-1}(x/2) + \frac{1}{\log 3} B'_{k-1}(x/3) + \sum_{4 \leq d \leq x} \nu(d) B'_{k-1}(x/d) \\ &\leq 2B'_{k-1}(x) + \sum_{4 \leq d \leq x/3} \nu(d) B'_{k-1}(x/d) + \sum_{x/3 < d \leq x} \nu(d) B'_{k-1}(x/d) \end{aligned}$$

In each of the terms in the final summation, $d > x/3$ implies that $\nu(d) \leq 2/\log x$, and $B'_{k-1}(x/d) \leq 1$. Thus this final summation is $\leq 2x/\log x$. This yields

$$\begin{aligned} B'_k(x) &\leq 2B'_{k-1}(x) + \frac{2x}{\log x} + \sum_{4 \leq d \leq x/3} \nu(d) B'_{k-1}(x/d) \\ &\leq 3 \frac{2(k-1)x(3 + \log \log x)^{k-2}}{\log x} + \sum_{4 \leq d \leq x/3} \frac{1}{\log d} \frac{2(k-1)(x/d)(3 + \log \log x/d)^{k-2}}{\log x/d} \\ &= 2(k-1)x \left[3 \frac{(3 + \log \log x)^{k-2}}{\log x} + \sum_{4 \leq d \leq x/3} \frac{(3 + \log \log x/d)^{k-2}}{d \log d \log x/d} \right]. \end{aligned} \tag{4.1}$$

Now, observe that

$$\begin{aligned} t &\mapsto t \log x/t && \text{is increasing over } (1, x/e) \\ \text{hence } t &\mapsto \frac{(3 + \log \log x/t)^{k-2}}{t \log t \log x/t} && \text{is decreasing over } (1, x/e). \end{aligned}$$

This means that the above summation can be bounded by the integral:

$$\begin{aligned} &\sum_{4 \leq d \leq x/3} \frac{(3 + \log \log x/d)^{k-2}}{d \log d \log x/d} \\ &\leq \int_3^{x/3} \frac{(3 + \log \log x/t)^{k-2}}{t \log t \log x/t} dt \end{aligned}$$

$$\begin{aligned}
&= \int_{\log 3}^{\log x/3} \frac{(3 + \log(\log x - u))^{k-2}}{u(\log x - u)} du \\
&= \frac{1}{\log x} \left[\int_{\log 3}^{\log x/3} \frac{(3 + \log(\log x - u))^{k-2}}{u} + \frac{(3 + \log(\log x - u))^{k-2}}{\log x - u} du \right] \\
&\leq \frac{1}{\log x} \left[(3 + \log \log x)^{k-2} \int_{\log 3}^{\log x/3} \frac{1}{u} du + \int_{\log 3}^{\log x/3} \frac{(3 + \log u)^{k-2}}{u} du \right] \\
&\leq \frac{1}{\log x} \left[(3 + \log \log x)^{k-2} \log \log x + \frac{(3 + \log \log x)^{k-1}}{k-1} \right] \\
&= \frac{(3 + \log \log x)^{k-2}}{\log x} \left[\log \log x + \frac{(3 + \log \log x)}{k-1} \right].
\end{aligned}$$

Substituting this bound into (4.1) yields

$$\begin{aligned}
B'_k(x) &\leq \frac{2(k-1)x(3 + \log \log x)^{k-2}}{\log x} \left[3 + \log \log x + \frac{3 + \log \log x}{k-1} \right] \\
&= \frac{2(k-1)x(3 + \log \log x)^{k-1}}{\log x} \left[1 + \frac{1}{k-1} \right] \\
&= \frac{2kx(3 + \log \log x)^{k-1}}{\log x},
\end{aligned}$$

as desired. \square

We can then use this lemma to show that the average value of $b(N)$ over squarefree N is $O(1)$. Since N is squarefree, any factorization of N will have no repeated factors, so $b(N) = \sum_{k=1}^{\infty} b_k(N)/k!$. This then means that

$$\begin{aligned}
\frac{1}{\frac{6}{\pi^2}x} \sum_{\substack{N \leq x \\ \text{squarefree}}} b(N) &= \frac{1}{\frac{6}{\pi^2}x} \sum_{\substack{N \leq x \\ \text{squarefree}}} \sum_{k=1}^{\infty} \frac{b_k(N)}{k!} \\
&= \frac{1}{\frac{6}{\pi^2}x} \sum_{k=1}^{\infty} \frac{B'_k(x)}{k!} \\
&\leq \frac{1}{\frac{6}{\pi^2}x} \sum_{k=1}^{\infty} \frac{2kx(3 + \log \log x)^{k-1}}{k! \log x} \\
&= \frac{\pi^2}{3 \log x} \sum_{k=1}^{\infty} \frac{(3 + \log \log x)^{k-1}}{(k-1)!} \\
&= \frac{\pi^2}{3 \log x} \exp(3 + \log \log x) \\
&= \frac{\pi^2 e^3}{3} = O(1),
\end{aligned}$$

as desired.

5. AN UPPER BOUND WHEN $\alpha > 1$

Fix $\alpha > 1$. In this section, we then show that

$$\frac{1}{x} \sum_{N \leq x} b(N) = O(1/\log^\alpha x)$$

We would again like to use a recursive formula to obtain this result. However, the recursive relation (2.1) for $B_k(x)$ has its inequality in the wrong direction. So rather than inducting on the length of a factorization k , we will instead induct on the size of the factors.

Let $b_m(N)$ denote the expected number of (unordered) factorizations of N , using only factors that are $\leq m$. Also, let $B_m(x) := \sum_{N \leq x} b_m(N)$. For any factorization $\leq x$, we can condition on the largest factor d (say, with multiplicity j). This then yields the recurrence relation

$$B_m(x) = \sum_{d \leq m} \nu(d) \sum_{\substack{j \geq 1 \\ d^j \leq x}} B_{d-1}(x/d^j)$$

Using this recurrence relation, we compute the following bounds on $B_m(x)$. Note that the constants T , then x_0 , then C are chosen to satisfy a list of constraints. We write out each of these constraints explicitly to make it clear why they can all be achieved simultaneously.

Lemma 5.1. *Fix $\alpha > 1$. Choose T , then x_0 , then C large enough such that:*

$$\left\{ \begin{array}{l} T \geq 3, \quad \int_{\log(T-1)}^{\infty} \frac{1}{u^\alpha} du \leq \frac{1}{5 \cdot 2^\alpha}, \quad \sum_{d \geq T} \frac{1}{d \log^\alpha d} \leq \frac{1}{5 \cdot 2^\alpha}; \end{array} \right. \quad (*T)$$

$$\left\{ \begin{array}{l} x_0 \geq T^2, \quad \frac{x_0}{T-1} \geq \frac{x_0}{T} + 1, \quad (2\alpha + 4) \log \log x_0 \leq \frac{1}{2} \log x_0, \\ \frac{1}{5} \frac{x}{\log^\alpha x} \geq (T-2) \log_2(x) (\log_2(x) + 1)^{T-3} \quad \forall x \geq x_0; \end{array} \right. \quad (*x_0)$$

$$\left\{ \begin{array}{l} C \geq 1, \quad \frac{Cx}{\log^\alpha x} \geq 1 \quad \forall x \geq 1, \\ \frac{1}{5} \frac{Cx}{\log^\alpha x} \geq B_T(T) \left[\frac{\sqrt{x}}{\log^\alpha T} + \frac{x}{\log^\alpha \sqrt{x}} \right] \quad \forall x \geq x_0, \\ \frac{Cx}{\log^\alpha x} \geq B_{[x]}(x) \quad \forall 1 \leq x \leq x_0. \end{array} \right. \quad (*C)$$

Then for each $m \in \mathbb{N}$, $B_m(x) \leq Cx/\log^\alpha x$ for all $x \geq 1$. (Here, the right hand side can be interpreted as ∞ at $x = 1$.)

Proof. We proceed by strong induction on m . The base case of $m = 1$ is immediate: $B_1(x) = 1$ for all $x \geq 1$, which satisfies the desired bound by $(*C)$.

For the inductive step, assume that $B_{m'}(x) \leq Cx/\log^\alpha x$ for each $m' \leq m - 1$. By $(*C)$, we have that $B_m(x) \leq B_{\lfloor x \rfloor}(x) \leq Cx/\log^\alpha x$ for all $1 \leq x \leq x_0$. So assume that $x \geq x_0$. Then by the recursive formula,

$$B_m(x) = \sum_{d \leq m} \nu(d) \sum_{\substack{j \geq 1 \\ d^j \leq x}} B_{d-1}(x/d^j).$$

We will break this summation over d, j into five parts:

- (1) where $2 \leq d \leq T - 1$;
- (2) where $d \geq T, d^j > x/T$;
- (3) where $d \geq T, d^j \leq x/T, j = 1$;
- (4) where $d \geq T, d^j \leq x/T, j \geq 2, d^{j-1} < j^2 \log^\alpha x$;
- (5) where $d \geq T, d^j \leq x/T, j \geq 2, d^{j-1} \geq j^2 \log^\alpha x$;

and show that each part is $\leq \frac{1}{5}Cx/\log^\alpha x$.

First, consider

$$(\text{part 1}) = \sum_{\substack{2 \leq d \leq T-1 \\ d \leq m}} \nu(d) \sum_{\substack{j \geq 1 \\ d^j \leq x}} B_{d-1}(x/d^j).$$

Note that each $B_{d-1}(x/d^j)$ term in this summation is $\leq B_{T-2}(x)$. And $B_{T-2}(x)$ counts the expected number of factorizations $\leq x$ using only the factors $\{2, \dots, T-2\}$. To obtain a product $\leq x$, each of these $T-3$ factors could be used at most $\log_2(x)$ times. Thus the total number of products that could be formed is $\leq (\log_2(x) + 1)^{T-3}$. Hence $B_{d-1}(x/d^j) \leq B_{T-2}(x) \leq (\log_2(x) + 1)^{T-3}$, and so

$$\begin{aligned} (\text{part 1}) &= \sum_{\substack{2 \leq d \leq T-1 \\ d \leq m}} \nu(d) \sum_{\substack{j \geq 1 \\ d^j \leq x}} B_{d-1}(x/d^j) \\ &\leq \sum_{\substack{2 \leq d \leq T-1 \\ d \leq m}} \nu(d) \sum_{\substack{j \geq 1 \\ d^j \leq x}} (\log_2(x) + 1)^{T-3} \\ &\leq \sum_{\substack{2 \leq d \leq T-1 \\ d \leq m}} \nu(d) \log_2(x) (\log_2(x) + 1)^{T-3} \\ &\leq (T-2) \log_2(x) (\log_2(x) + 1)^{T-3} \\ &\leq \frac{1}{5} \frac{Cx}{\log^\alpha x}, \quad \text{by } (*x_0), (*C) \end{aligned}$$

as desired.

Second, we have

$$\begin{aligned}
(\text{part 2}) &= \sum_{\substack{d \geq T \\ d \leq m}} \nu(d) \sum_{\substack{j \geq 1 \\ x/T < d^j \leq x}} B_{d-1}(x/d^j) \\
&\leq \sum_{T \leq d \leq x} \nu(d) B_{d-1}(T) \quad (\text{at most one inner term since } d \geq T) \\
&\leq B_T(T) \sum_{T \leq d \leq x} \frac{1}{\log^\alpha d} \\
&= B_T(T) \left[\sum_{T \leq d < \sqrt{x}} \frac{1}{\log^\alpha d} + \sum_{\sqrt{x} \leq d \leq x} \frac{1}{\log^\alpha d} \right] \quad \text{using } (*x_0) \\
&\leq B_T(T) \left[\frac{\sqrt{x}}{\log^\alpha T} + \frac{x}{\log^\alpha \sqrt{x}} \right] \\
&\leq \frac{1}{5} \frac{Cx}{\log^\alpha x}, \quad \text{by } (*C)
\end{aligned}$$

as desired.

Third, consider

$$\begin{aligned}
(\text{part 3}) &= \sum_{\substack{T \leq d \leq x/T \\ d \leq m}} \nu(d) B_{d-1}(x/d) \\
&\leq \sum_{\substack{T \leq d \leq x/T \\ d \leq m}} \frac{1}{\log^\alpha d} \frac{C(x/d)}{\log^\alpha x/d} \\
&\leq Cx \sum_{T \leq d \leq x/T} \frac{1}{d \log^\alpha d \log^\alpha x/d}.
\end{aligned}$$

Observe that

$$\begin{array}{ll}
u \mapsto e^{u/\alpha} u (\log x - u) & \text{is unimodal up over } (0, \log x) \\
\text{hence } t \mapsto t^{1/\alpha} \log t (\log x - \log t) & \text{is unimodal up over } (1, x) \\
\text{hence } t \mapsto \frac{1}{t \log^\alpha t \log^\alpha x/t} & \text{is unimodal down over } (1, x).
\end{array}$$

This means that

$$\begin{aligned}
(\text{part 3}) &\leq Cx \sum_{T \leq d \leq x/T} \frac{1}{d \log^\alpha d \log^\alpha x/d} \\
&\leq Cx \int_{T-1}^{x/T+1} \frac{1}{t \log^\alpha t \log^\alpha x/t} dt
\end{aligned}$$

$$\begin{aligned}
&\leq Cx \int_{T-1}^{x/(T-1)} \frac{1}{t \log^\alpha t \log^\alpha x/t} dt \quad \text{by } (*x_0) \\
&= Cx \int_{\log(T-1)}^{\log(x/(T-1))} \frac{1}{u^\alpha (\log x - u)^\alpha} du \\
&= \frac{Cx}{\log^\alpha x} \int_{\log(T-1)}^{\log(x/(T-1))} \left(\frac{1}{u} + \frac{1}{\log x - u} \right)^\alpha du \\
&\leq \frac{Cx}{\log^\alpha x} \frac{2^\alpha}{2} \int_{\log(T-1)}^{\log(x/(T-1))} \frac{1}{u^\alpha} + \frac{1}{(\log x - u)^\alpha} du \quad \text{by Jensen's Inequality} \\
&= \frac{Cx}{\log^\alpha x} 2^\alpha \int_{\log(T-1)}^{\log(x/(T-1))} \frac{1}{u^\alpha} du \\
&\leq \frac{1}{5} \frac{Cx}{\log^\alpha x}, \quad \text{by } (*T)
\end{aligned}$$

as desired.

Fourth, consider

$$\begin{aligned}
(\text{part 4}) &= \sum_{\substack{d \geq T \\ d \leq m}} \nu(d) \sum_{\substack{j \geq 2 \\ d^j \leq x/T, \, d^{j-1} \geq j^2 \log^\alpha x}} B_{d-1}(x/d^j) \\
&= \sum_{2 \leq j \leq \log_T(x/T)} \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d \leq m, \, d^{j-1} \geq j^2 \log^\alpha x}} \nu(d) B_{d-1}(x/d^j) \\
&\leq \sum_{2 \leq j \leq \log x} \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d^{j-1} \geq j^2 \log^\alpha x}} \frac{1}{\log^\alpha d} \frac{C(x/d^j)}{\log^\alpha x/d^j} \\
&= \sum_{2 \leq j \leq \log x} Cx \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d^{j-1} \geq j^2 \log^\alpha x}} \frac{1}{d^j \log^\alpha d \log^\alpha x/d^j} \tag{5.1}
\end{aligned}$$

Then using the facts that $\log^\alpha x/d^j \geq \log^\alpha T \geq 1$ and $d^{j-1} \geq j^2 \log^\alpha x$,

$$\begin{aligned}
(\text{part 4}) &\leq \sum_{2 \leq j \leq \log x} Cx \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d^{j-1} \geq j^2 \log^\alpha x}} \frac{1}{d^j \log^\alpha d} \\
&\leq \sum_{2 \leq j \leq \log x} Cx \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d^{j-1} \geq j^2 \log^\alpha x}} \frac{1}{d j^2 \log^\alpha x \log^\alpha d} \\
&\leq \sum_{2 \leq j \leq \log x} \frac{Cx}{\log^\alpha x} \frac{1}{j^2} \sum_{d \geq T} \frac{1}{d \log^\alpha d}
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{2 \leq j \leq \log x} \frac{Cx}{\log^\alpha x} \frac{1}{j^2} \frac{1}{5} \quad \text{by } (*T) \\
&\leq \frac{1}{5} \frac{Cx}{\log^\alpha x},
\end{aligned}$$

as desired.

Fifth, by an identical argument as in (5.1), we have

$$(\text{part 5}) \leq \sum_{2 \leq j \leq \log x} Cx \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d^{j-1} < j^2 \log^\alpha x}} \frac{1}{d^j \log^\alpha d \log^\alpha x / d^j}. \quad (5.2)$$

Note that since $d^{j-1} < j^2 \log^\alpha x$, we have

$$\begin{aligned}
j \log d &\leq 2(j-1) \log d \leq 2 \log(j^2 \log^\alpha x) \leq 2 \log((\log x)^2 \log^\alpha x) \\
&\leq (2\alpha + 4) \log \log x \leq \frac{1}{2} \log x, \quad \text{by } (*x_0)
\end{aligned}$$

$$\text{so } \log^\alpha x / d^j = (\log x - j \log d)^\alpha \geq \left(\frac{1}{2} \log x \right)^\alpha = \frac{\log^\alpha x}{2^\alpha}.$$

Plugging this bound into (5.2) then yields

$$\begin{aligned}
(\text{part 5}) &\leq \sum_{2 \leq j \leq \log x} 2^\alpha \frac{Cx}{\log^\alpha x} \sum_{\substack{T \leq d \leq (x/T)^{1/j} \\ d^{j-1} < j^2 \log^\alpha x}} \frac{1}{d^j \log^\alpha d} \\
&\leq \sum_{2 \leq j \leq \log x} \frac{2^\alpha}{T^{j-1}} \frac{Cx}{\log^\alpha x} \sum_{d \geq T} \frac{1}{d \log^\alpha d} \\
&\leq \sum_{2 \leq j \leq \log x} \frac{1}{T^{j-1}} \frac{1}{5} \frac{Cx}{\log^\alpha x} \quad \text{by } (*T) \\
&\leq \frac{1}{5} \frac{Cx}{\log^\alpha x},
\end{aligned}$$

completing the proof. □

This lemma immediately yields the desired bound:

$$\frac{1}{x} \sum_{N \leq x} b(N) = \frac{1}{x} B_{[x]}(x) = O(1/\log^\alpha x),$$

proving the third case of Theorem 1.1.

6. ASYMPTOTIC DENSITY OF $\text{mult}(A_\alpha)$

Recall that $\text{mult}(A_\alpha)$ denotes the set of all natural numbers which can be written as products of Cramér α -random primes. We then show the following result on the asymptotic density of $\text{mult}(A_\alpha)$.

Theorem 1.2. *Let A_α denote the set of Cramér α -random primes. Then with probability 1, $\text{mult}(A_\alpha)$ has asymptotic density*

$$d(\text{mult}(A_\alpha)) = \begin{cases} 0 & \text{if } \alpha > 1 \\ 1 & \text{if } \alpha < \frac{1}{2} \log 2. \end{cases}$$

Proof. The first half of this theorem follows directly from Theorem 1.1. In particular, Theorem 1.1 implies that the expected asymptotic density of $\text{mult}(A_\alpha)$ is 0:

$$\begin{aligned} \mathbb{E}[d(\text{mult}(A_\alpha))] &= \mathbb{E} \left[\lim_{x \rightarrow \infty} \frac{1}{x} \sum_{N \leq x} \mathbb{1}_{N \in \text{mult}(A_\alpha)} \right] \\ &= \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{N \leq x} \mathbb{E} [\mathbb{1}_{N \in \text{mult}(A_\alpha)}] \\ &\leq \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{N \leq x} b(N) \\ &= 0 \quad \text{when } \alpha > 1, \end{aligned}$$

which yields the desired result.

For the second half of the theorem, we use a probability argument. In particular, for a density 1 set of N , we show that $\mathbb{P}[N \notin A_\alpha A_\alpha]$ tends to 0 as $N \rightarrow \infty$. This is sufficient to show the desired result since then,

$$\mathbb{E}[d(A_\alpha A_\alpha)] = \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{N \leq x} \mathbb{E} [\mathbb{1}_{N \in A_\alpha A_\alpha}] = \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{N \leq x} \mathbb{P}[N \in A_\alpha A_\alpha] = 1,$$

which means that $d(A_\alpha A_\alpha) = 1$ (and hence $d(\text{mult}(A_\alpha)) = 1$) with probability 1.

The probability that $N \notin A_\alpha A_\alpha$ is given by

$$\begin{aligned} \mathbb{P}[N \notin A_\alpha A_\alpha] &= \mathbb{P}[d \notin A_\alpha \text{ or } N/d \notin A_\alpha \quad \forall d \mid N] \\ &\leq \mathbb{P}[d \notin A_\alpha \text{ or } N/d \notin A_\alpha \quad \forall d \mid N, 3 \leq d < \sqrt{N}] \\ &= \prod_{\substack{d \mid N, \\ 3 \leq d < \sqrt{N}}} \mathbb{P}[d \notin A_\alpha \text{ or } N/d \notin A_\alpha] \\ &= \prod_{\substack{d \mid N, \\ 3 \leq d < \sqrt{N}}} \left(1 - \frac{1}{\log^\alpha d} \frac{1}{\log^\alpha N/d} \right), \end{aligned}$$

so $-\log \mathbb{P}[N \notin A_\alpha A_\alpha] \geq \sum_{\substack{d \mid N, \\ 3 \leq d < \sqrt{N}}} -\log \left(1 - \frac{1}{\log^\alpha d} \frac{1}{\log^\alpha N/d} \right)$

$$\begin{aligned}
&\geq \sum_{\substack{d|N, \\ 3 \leq d < \sqrt{N}}} \frac{1}{\log^\alpha d} \frac{1}{\log^\alpha N/d} \\
&\geq \sum_{\substack{d|N, \\ 3 \leq d < \sqrt{N}}} \left(\frac{4}{\log^2 N} \right)^\alpha \\
&\geq \frac{4^\alpha \left(\frac{1}{2} \sigma_0(N) - 3 \right)}{\log^{2\alpha} N}.
\end{aligned}$$

Now, for $\alpha < \frac{1}{2} \log 2$, choose $\varepsilon > 0$ such that $2\alpha < \log 2 - \varepsilon$. It is well-known that the normal order of $\log \sigma_0(N)$ is $(\log 2) \log \log N$ [5, Theorem 432]. So in particular, for a density 1 set of N , we have $\sigma_0(N) \geq \log^{\log 2 - \varepsilon} N$ and so

$$-\log \mathbb{P}[N \notin A_\alpha A_\alpha] \geq \frac{4^\alpha \left(\frac{1}{2} \log^{\log 2 - \varepsilon} N - 3 \right)}{\log^{2\alpha} N} \rightarrow \infty,$$

which means that $\mathbb{P}[N \notin A_\alpha A_\alpha] \rightarrow 0$, as desired. \square

7. DISCUSSION

In this section, we discuss our results. First, we would like to point out a (non-probabilistic) related problem has been investigated before. Define the Beurling generalized primes to be a multiset \mathcal{P} of real numbers $0 < p_1 \leq p_2 \leq \dots$. Then the Beurling generalized integers are defined to be the multiset \mathcal{N} of real numbers of the form $p_1^{a_1} p_2^{a_2} \dots$ for $a_i \in \mathbb{N}_0$. In the past, several works have investigated how the behavior of these Beurling generalized integers compares to that of the classical integers. Specifically, [2, Theorem 2] shows that if, for example, $\#\{\beta \in \mathcal{P} : \beta \leq x\} = \frac{x}{\log x} + O\left(\frac{x}{\log^{1+\varepsilon} x}\right)$ (i.e. the Beurling generalized primes grow like the actual primes) then $\#\{\beta \in \mathcal{N} : \beta \leq x\} = cx + o(x)$ for some constant $c > 0$ (i.e. the Beurling generalized numbers grow like a sequence of natural numbers with a nonzero density).

One critical difference between the Beurling generalized integers and $\text{mult}(A_\alpha)$ studied in this paper is that the Beurling generalized integers are counted with multiplicity. In some sense, the fundamental difficulty in our setting is accounting for natural numbers that can be written as a product of Cramér α -random primes in multiple ways. If one did not have to account for multiplicity (i.e. if one instead considered $\text{mult}(A_\alpha)$ to be a multiset), the results of this paper could be made stronger.

Next, we make the following conjecture on the asymptotic density of $\text{mult}(A_\alpha)$. Only the cases of $\frac{1}{2} \log 2 \leq \alpha \leq 1$ remain to be proven.

Conjecture 7.1. *Let A_α denote the set of Cramér α -random primes. Then with probability 1, $\text{mult}(A_\alpha)$ has asymptotic density*

$$d(\text{mult}(A_\alpha)) = \begin{cases} 0 & \text{if } \alpha \geq 1 \\ 1 & \text{if } \alpha < 1. \end{cases}$$

We suspect that the case of $\alpha = 1$ in Conjecture 7.1 will be a much more difficult problem than the case of $\frac{1}{2} \log 2 \leq \alpha < 1$.

For $\alpha < 1$, Theorem 1.1 shows that the average number of Cramér-factorizations of N tends to infinity rather quickly. But this in itself is not enough to show that $\text{mult}(A_\alpha)$ has asymptotic density 1. This conclusion does seem likely to be true, however. Additionally, note that for the case of $\alpha < \frac{1}{2} \log 2$ in Theorem 1.2, we only utilized a bound on $d(A_\alpha A_\alpha)$ (i.e. only considering length 2 factorizations). This is a very crude estimate of $d(\text{mult}(A_\alpha))$, and we suspect that more accurate estimates would yield the desired result for all $\alpha < 1$.

For $\alpha = 1$, on the other hand, the problem of determining the behavior of $\text{mult}(A_1)$ remains wide open. Recall that every natural number has exactly one prime-factorization. So since A_1 is distributed like the primes, one would expect that the average number of Cramér factorizations of the natural numbers should also be roughly 1. This means that if any natural numbers have multiple Cramér factorizations, other natural numbers would be forced to have zero Cramér factorizations. For this reason, we guessed in Conjecture 7.1 that $\text{mult}(A_1)$ should have asymptotic density 0. However, this guess is mainly just based on intuition, and it is quite possible that the asymptotic density is positive with non-zero probability. In this case, it would also be an interesting problem to compute the expected asymptotic density of $\text{mult}(A_1)$ (assuming it exists).

Finally, we take note of one technicality in the above problem. In Theorem 1.2, we proved that for various α , $\text{mult}(A_\alpha)$ will have a particular asymptotic density with probability 1. Motivated by this result and the Davenport-Erdős theorem [1], we had conjectured in an earlier preprint of this paper that in fact *all* multiplicatively closed sets have some type of density (specifically, logarithmic density). Surprisingly, however, this turns out to not be true; in [8], we managed to construct a multiplicatively closed set without an asymptotic density, logarithmic density, or multiplicative density.

DECLARATIONS

No funding was received to assist with the preparation of this manuscript. The author has no competing interests.

This manuscript has no associated data.

REFERENCES

- [1] H. Davenport and P. Erdős. On sequences of positive integers. *Acta Arith.*, 2, 1937.
- [2] Harold G. Diamond. When do Beurling generalized integers have a density? *J. Reine Angew. Math.*, 295:22–39, 1977.
- [3] Andrew Granville. Harald Cramér and the distribution of prime numbers. *Scand. Actuar. J.*, pages 12–28, 1995. Harald Cramér Symposium (Stockholm, 1993).
- [4] G. H. Hardy and S. Ramanujan. The normal number of prime factors of a number n . *Quart. J. Math.*, 48, 1917.
- [5] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. The Clarendon Press, Oxford University Press, New York, fifth edition, 1979.
- [6] János Pintz. Cramér vs. Cramér. On Cramér’s probabilistic model for primes. *Funct. Approx. Comment. Math.*, 37:361–376, 2007.
- [7] Paul Pollack. A generalization of the Hardy-Ramanujan inequality and applications. *J. Number Theory*, 210:171–182, 2020.
- [8] Erick Ross. Are sets with a given multiplicative structure guaranteed to have a density? Submitted, 2026.

(E. Ross) SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES, CLEMSON UNIVERSITY, CLEMSON, SC

Email address: `erickjohnross@gmail.com`