Primes below $n^{2-\delta}$ in every residue class, for a majority of moduli

Charles O'Neill, Max Kirkby, Astra and Fable

24 November 2026

Abstract

Let $P(a,n)$ denote the least prime congruent to a modulo n. Since the generalised Riemann hypothesis for Dirichlet L-functions is now known, $P(a,n) \ll (\varphi(n)\log n)^2$, but a bound $\max_{(a,n)=1} P(a,n) \ll n^{2-\delta}$ with a fixed $\delta > 0$ is not known for all large n. We show, without using the generalised Riemann hypothesis, that there are absolute constants $C, \delta > 0$ such that, in every sufficiently large dyadic interval, more than 58.7% of the integers n satisfy $\max_{(a,n)=1} P(a,n) \le Cn^{2-\delta}$. The proof combines Maynard's theorem on primes in all residue classes to large moduli that have a conveniently sized factor with a computer-assisted lower bound, more than 58.71%, for the proportion of integers n having a divisor d with $\log d/\log n$ in $[0.142858, 0.199999]$, a compact subinterval of $(1/7, 1/5)$. That bound is obtained from the Donnelly–Grimmett stick-breaking description of the prime factors of a random integer, and an upper bound of 59.56% for the proportion of integers with a divisor in $[n^{1/7}, n^{1/5}]$ shows that the divisor condition itself cannot give much more. Both numerical bounds have also been proved in Lean 4 with Mathlib, and so have both theorems, from the theorems of Maynard and of Donnelly and Grimmett taken as hypotheses.

2020 Mathematics Subject Classification. 11N13; 11N25, 60C05
Keywords. least prime in an arithmetic progression, Linnik's theorem, divisors in intervals, stick-breaking

AI use and author contributions

The mathematical arguments and computer programs in this paper were developed by Astra (GPT-6.2 Astra, OpenAI). Fable (Fable 5.2, Anthropic) worked throughout as a second model: it proposed alternative approaches, reviewed arguments and raised objections. The manuscript was prepared from the project's research records with Claude (Anthropic). All computations were re-run before release, and the programs and certificates are available at https://github.com/charlieoneill11/maths-releases.

Charles O'Neill proposed the problem because the generalised Riemann hypothesis for Dirichlet L-functions, now a theorem [1], gives $P(a,n) \ll$

$(\varphi(n)\log n)^2$, and asked whether the quadratic exponent could be beaten. He also proposed aiming first for averaged statements with a quantified exceptional set, which is the form of Theorem 1. The argument that succeeded for most moduli turned out not to need the generalised Riemann hypothesis. Charles O'Neill has checked the arguments and takes full responsibility for the correctness of the results.

1 Introduction

For $n \geq 2$ and $(a, n) = 1$ let $P(a, n)$ be the least prime $p \equiv a \pmod{n}$. Linnik's theorem gives $P(a, n) \ll n^L$ for an absolute constant L, and the best explicit exponent stated before 2026 without the generalised Riemann hypothesis was $L = 5$ [12, Thm. 2.1]. Alpöge's proof of the generalised Riemann hypothesis for Dirichlet L-functions [1] makes the classical conditional bound $P(a, n) \leq (1 + o(1))(\varphi(n)\log n)^2$ [2] unconditional. The exponent 2 has not been improved for all moduli. For the weighted count $\theta(x; n, a) = \sum_{p \leq x,\, p \equiv a \pmod{n}} \log p$, the generalised Riemann hypothesis, character orthogonality and the explicit formula give an error of order $x^{1/2}(\log nx)^2$ in each class against the main term $x/\varphi(n)$, and the mean square of the errors over the classes is compatible with a single class that contains no prime up to $n^{2-\delta}$.

Uniformity in the residue class for large moduli is exactly what Maynard's third paper on primes in progressions to large moduli provides [10]: for almost all moduli $q \leq x^{1/2+\eta}$ that have a divisor in a suitable range, every reduced class modulo q contains $\gg \pi(x)/\varphi(q)$ primes up to x. Taking q close to $x^{1/2+\eta}$, this gives $P(a, q) \leq q^{2-\delta}$ for every class, provided that q has a divisor d with $\log d/\log q$ slightly inside $(1/7, 1/5)$ (Lemma 4). The question is then how often this happens.

The proportion of integers with a divisor in such a window has been studied before. Tenenbaum [11] showed that, for fixed $0 < \alpha < \beta < 1$, the proportion of integers n having a divisor in $[n^\alpha, n^\beta]$ has a limit. Haddad [6] formulated this limit in terms of the Poisson–Dirichlet distribution, the limiting law of the normalised logarithms of the prime factors of a random integer; the stick-breaking description that we use in Section 3 is the size-biased form of the same distribution. Drappeau and Mounier [5] computed this limiting density earlier, on a grid of windows. The windows of Theorem 2 are not on that grid, and we do not use the values of [5]; Section 5 compares our bounds with three windows of their grid. Our program shows that a majority of integers have a divisor in $[n^{0.15}, n^{0.19}]$. Since $[0.15, 0.19]$ is

strictly contained in $(1/7, 1/5)$, a majority on that window already gives a majority of moduli with a divisor of the kind Lemma 4 needs. What is new here is, first, two verified quantitative bounds: the lower bound 0.587142 for the inward window $[0.142858, 0.199999]$ and the upper bound 0.595573 for the outward window $[0.142857, 0.2]$, which contains $[1/7, 1/5]$; the upper one shows how far this route can go (Theorem 2). Second, the application of Maynard's theorem in every reduced residue class simultaneously turns the divisor condition into the bound for the least prime in Theorem 1. Neither theorem uses the Riemann hypothesis or the generalised Riemann hypothesis; the latter [1] enters only in the comparisons in the first paragraph and in remark (2) of Section 7.

Theorem 1. *There are absolute constants $C, \delta > 0$ such that, for all sufficiently large Q, more than 58.7% of the integers $n \in (Q/2, Q]$ satisfy*

$$\max_{\substack{1 \le a \le n \\ (a,n)=1}} P(a, n) \le C n^{2-\delta}.$$

One may take $C = 4$; the constant δ depends only on the range of the parameter in Proposition 3. The second theorem concerns the divisor condition alone.

Theorem 2. *For all sufficiently large Q, among the integers $n \in (Q/2, Q]$,*

1. *more than 58.71% have a divisor d with $n^{0.142858} \le d \le n^{0.199999}$;*
2. *fewer than 59.56% have a divisor d with $n^{1/7} \le d \le n^{1/5}$.*

Part (2) shows that Proposition 3, used in this way, cannot cover much more than 59.56% of the moduli: as $\delta \to 0$ the divisor range it requires tends to $[n^{4/5}, n^{6/7}]$, whose complementary divisors are those in $[n^{1/7}, n^{1/5}]$ (see Section 7).

The proportions in Theorem 2 are computed from the limiting distribution of the prime factors of a random integer, listed in size-biased order: by the theorem of Donnelly and Grimmett [4], their normalised logarithms behave like the pieces of a stick of length one broken successively at uniform fractions. A divisor corresponds to a set of pieces, and we bound the probability that some set of the first few pieces, together with or without the remainder of the stick, has total length in the target range. This probability satisfies an integral recursion, which we evaluate with rigorous directed rounding (Sections 4 and 5).

Notation. $\pi(x; q, a)$ is the number of primes $p \le x$ with $p \equiv a \pmod q$. For a set $T \subseteq \mathbb{R}$ we write $1 - T = \{1 - t : t \in T\}$ and $|T|$ for its Lebesgue measure. All intervals in targets are closed unless stated otherwise.

2 Primes in every class for moduli with a suitable divisor

Proposition 3 (Maynard [10, Cor. 1.4]). *There is $\eta_0 > 0$ with the following property. Let $0 < \eta \le \eta_0$ and $A > 0$. For $x \ge x_0(\eta, A)$ there is a set $\mathcal{B}(x)$ of integers in $[1, x^{1/2+\eta}]$ with $\#\mathcal{B}(x) \le x^{1/2+\eta}/(\log x)^A$ such that every integer $q \le x^{1/2+\eta}$ with $q \notin \mathcal{B}(x)$ that has a divisor in $[x^{2/5+5\eta}, x^{3/7}]$ satisfies*

$$\pi(x; q, a) \gg_\eta \frac{\pi(x)}{\varphi(q)} \qquad \text{for every } a \text{ with } (a, q) = 1.$$

Lemma 4. *Let $0 < \eta \le \min(\eta_0, 10^{-8})$ and $\delta = 4\eta/(1+2\eta)$. For all sufficiently large Q, all but at most $Q/\log Q$ of the integers $n \in (Q/2, Q]$ that have a divisor d with $n^{0.142858} \le d \le n^{0.199999}$ satisfy $P(a, n) \le 4n^{2-\delta}$ for every a coprime to n.*

Proof. Let $x = Q^{2/(1+2\eta)}$, so that $x^{1/2+\eta} = Q$, and apply Proposition 3 with $A = 1$; we may assume $x \ge x_0(\eta, 1)$ and $\log x \ge \log Q$. Let $n \in (Q/2, Q]$ have a divisor d as stated, and put $e = n/d$, so that $n^{0.800001} \le e \le n^{0.857142}$. We check that $e \in [x^{2/5+5\eta}, x^{3/7}]$. Since $\log x = 2\log Q/(1+2\eta)$,

$$\Big(\frac{2}{5} + 5\eta\Big)\log x = \frac{4/5 + 10\eta}{1+2\eta}\log Q \le (0.8 + 10\eta)\log Q \le 0.8000001\log Q,$$

while $\log e \ge 0.800001 \log n \ge 0.800001(\log Q - \log 2)$, which exceeds $0.8000001\log Q$ once $9 \cdot 10^{-7}\log Q > 0.800001\log 2$. Similarly

$$\begin{aligned}\frac{3}{7}\log x = \frac{6/7}{1+2\eta}\log Q &\ge \Big(\frac{6}{7} - 2\eta\Big)\log Q \\ &\ge (0.857142857 - 2\cdot 10^{-8})\log Q > 0.857142\log Q \ge \log e.\end{aligned}$$

Finally $n \le Q = x^{1/2+\eta}$. Hence, if $n \notin \mathcal{B}(x)$, then $\pi(x; n, a) > 0$ for every a coprime to n, that is $P(a, n) \le x$. Since $Q < 2n$,

$$x = Q^{2/(1+2\eta)} < (2n)^{2/(1+2\eta)} = (2n)^{2-\delta} \le 4n^{2-\delta}.$$

The number of exceptions is at most $\#\mathcal{B}(x) \le Q/\log x \le Q/\log Q$. □

3 Prime factors of a random integer

Let $N \ge 2$ be an integer with prime factorisation $N = a_1 a_2 \cdots a_M$, where $a_1 \ge a_2 \ge \cdots \ge a_M$ are primes, repeated according to multiplicity. A

size-biased ordering of the factors is the random permutation $(\beta_1, \dots, \beta_M)$ of $\{1, \dots, M\}$ in which, given $\beta_1, \dots, \beta_{i-1}$, the index β_i is chosen among the remaining indices k with probability proportional to $\log a_k$. Put

$$\Lambda_i(N) = \frac{\log a_{\beta_i}}{\log N} \quad (1 \le i \le M), \qquad \Lambda_i(N) = 0 \quad (i > M),$$

so that $\sum_i \Lambda_i(N) = 1$; for $N = 1$ put $\Lambda_i(1) = 0$. We write $\Lambda^{(j)}(N) = (\Lambda_1(N), \dots, \Lambda_j(N))$. Let $U_1, U_2, \dots$ be independent and uniform on $[0,1]$, and define the *stick-breaking* variables

$$V_1 = U_1, \qquad V_i = U_i \prod_{k<i}(1-U_k) \quad (i \ge 2), \qquad R_j = \prod_{k \le j}(1-U_k) = 1 - \sum_{i \le j} V_i.$$

Theorem 5 (Donnelly–Grimmett [4])**.** *Let N_X be uniform on $\{1, \dots, X\}$, and let $\Lambda(N_X)$ be formed with a size-biased ordering chosen independently given N_X. For every fixed $j \ge 1$, $\Lambda^{(j)}(N_X)$ converges in distribution to $(V_1, \dots, V_j)$ as $X \to \infty$.*

For $j \ge 1$ let $\Delta_j = \{\lambda \in [0,1]^j : \lambda_1 + \cdots + \lambda_j \le 1\}$, and for $\lambda \in \Delta_j$ and $S \subseteq \{1, \dots, j\}$ put $\sigma_S(\lambda) = \sum_{i \in S} \lambda_i$ and $\rho(\lambda) = 1 - \sum_{i \le j} \lambda_i$. For a set $T \subseteq [0,1]$ define

$$\mathcal{E}_j(T) = \{\lambda \in \Delta_j : \ \sigma_S(\lambda) + \epsilon\rho(\lambda) \in T \text{ for some } S \text{ and some } \epsilon \in \{0,1\}\},$$
$$\mathcal{F}_j(T) = \{\lambda \in \Delta_j : \ \sigma_S(\lambda) + y \in T \text{ for some } S \text{ and some } y \in [0, \rho(\lambda)]\},$$

where S ranges over the subsets of $\{1, \dots, j\}$. Clearly $\mathcal{E}_j(T) \subseteq \mathcal{F}_j(T)$, and both sets increase with T. When T is a finite union of intervals, $\mathcal{E}_j(T)$ is a finite union of preimages of T under affine maps, intersected with Δ_j, and $\mathcal{F}_j(T)$ is a finite union of sets defined by finitely many affine inequalities (see the proof of Lemma 7), so both are Borel sets. For such T only we put, with $V = (V_1, \dots, V_j)$,

$$p_j(T) = \mathbb{P}\big(V \in \mathcal{E}_j(T)\big), \qquad q_j(T) = \mathbb{P}\big(V \in \mathcal{F}_j(T)\big).$$

All targets used below are finite unions of intervals.

Lemma 6. *Let $N \ge 2$, let $\Lambda = \Lambda(N)$ be formed with any size-biased ordering, and let $T \subseteq [0,1]$.*

1. *If $\Lambda^{(j)}(N) \in \mathcal{E}_j(T)$, then N has a divisor d with $\log d / \log N \in T$.*
2. *If N has a divisor d with $\log d / \log N \in T$, then $\Lambda^{(j)}(N) \in \mathcal{F}_j(T)$.*

Proof. Write $\lambda = \Lambda^{(j)}(N)$. The factors a_{β_i} with $i \le \min(j, M)$ are the *revealed* factors, and the product r of the other factors a_k satisfies $\log r / \log N = \rho(\lambda)$. (1) If $\sigma_S(\lambda) + \epsilon\rho(\lambda) \in T$, let d be the product of the revealed factors a_{β_i} with $i \in S$, $i \le M$, times r if $\epsilon = 1$. Then $d \mid N$ and $\log d / \log N = \sigma_S(\lambda) + \epsilon\rho(\lambda)$, since $\Lambda_i = 0$ for $i > M$. (2) Write $d = \prod_{k \in D} a_k$ for a set D of indices. Let $S = \{i \le \min(j, M) : \beta_i \in D\}$ and let $y \log N$ be the logarithm of the product of the a_k with $k \in D$ not revealed. Then $0 \le y \le \rho(\lambda)$ and $\log d / \log N = \sigma_S(\lambda) + y$. □

Lemma 7. *Let $T \subseteq [0, 1]$ be a finite union of intervals whose endpoints lie in $(0, 1)$. Then the boundaries of $\mathcal{E}_j(T)$ and $\mathcal{F}_j(T)$ in $\mathbb{R}^j$ are null sets for the distribution of $(V_1, \dots, V_j)$.*

Proof. The map $(u_1, \dots, u_j) \mapsto (v_1, \dots, v_j)$ with $v_i = u_i \prod_{k<i}(1 - u_k)$ is a smooth bijection from $(0, 1)^j$ onto the interior of Δ_j with nonvanishing Jacobian $\prod_{i \le j}(1 - u_1) \cdots (1 - u_{i-1})$, so $(V_1, \dots, V_j)$ has a density and every hyperplane is a null set. In particular the boundary $\partial\Delta_j$ of the simplex, which lies in the $j + 1$ hyperplanes $\{\lambda_i = 0\}$ and $\{\lambda_1 + \cdots + \lambda_j = 1\}$, is a null set. The set $\mathcal{E}_j(T)$ is the union over (S, ϵ) of the sets $\Delta_j \cap f_{S,\epsilon}^{-1}(T)$, where $f_{S,\epsilon} = \sigma_S + \epsilon\rho$ is affine on $\mathbb{R}^j$, so its boundary in $\mathbb{R}^j$ lies in $\partial\Delta_j$ together with the union of the sets $\{f_{S,\epsilon} = t\}$ with t an endpoint of an interval of T. The map $f_{S,\epsilon}$ is constant only for $(S, \epsilon) = (\emptyset, 0)$, where it is 0, and for $(S, \epsilon) = (\{1, \dots, j\}, 1)$, where it is 1; as $t \in (0, 1)$, the corresponding sets are empty. All other sets $\{f_{S,\epsilon} = t\}$ are hyperplanes. Similarly $\lambda \in \mathcal{F}_j(T)$ if and only if $[\sigma_S(\lambda), \sigma_S(\lambda) + \rho(\lambda)]$ meets an interval I of T for some S. For I with endpoints $l \le u$ this is a condition of the form $\sigma_S(\lambda) \le u$ (or $< u$) and $\sigma_S(\lambda) + \rho(\lambda) \ge l$ (or $> l$), so the boundary of $\mathcal{F}_j(T)$ in $\mathbb{R}^j$ lies in $\partial\Delta_j$ together with the union of the sets $\{\sigma_S = u\}$ and $\{\sigma_S + \rho = l\}$. Since $\sigma_\emptyset = 0$ and $\sigma_S + \rho = 1$ for $S = \{1, \dots, j\}$, these are again empty or hyperplanes. In both cases the boundary is contained in finitely many hyperplanes and is therefore a null set. (The set $\partial\Delta_j$ cannot be omitted: for $j = 1$ and $T = [0.4, 0.6]$ we have $\mathcal{F}_1(T) = [0, 0.6]$, and its boundary point $\lambda = 0$ satisfies none of the equalities $\sigma_S = u$, $\sigma_S + \rho = l$.) □

Proposition 8. *Let $T \subseteq [0, 1]$ be a finite union of intervals with endpoints in $(0, 1)$, and let $j \ge 1$. Let $D_T(Q)$ be the proportion of the integers $n \in (Q/2, Q]$ that have a divisor d with $\log d / \log n \in T$. Then*

$$p_j(T) \le \liminf_{Q \to \infty} D_T(Q) \le \limsup_{Q \to \infty} D_T(Q) \le q_j(T).$$

Proof. For an integer $X \geq 1$ let μ_X be the distribution of $\Lambda^{(j)}(N_X)$, as in Theorem 5. By Theorem 5, Lemma 7 and the portmanteau theorem [3, Thm. 2.1], $\mu_X(\mathcal{E}_j(T)) \to p_j(T)$ and $\mu_X(\mathcal{F}_j(T)) \to q_j(T)$. Let $X = \lfloor Q \rfloor$ and $Y = \lfloor Q/2 \rfloor$; the integers in $(Q/2, Q]$ are those in $(Y, X]$. Let ν_Q be the distribution of $\Lambda^{(j)}(N)$ for N uniform on $(Y, X]$, with an independent size-biased ordering. Then $(X-Y)\nu_Q = X\mu_X - Y\mu_Y$, and as $X/(X-Y) \to 2$ and $Y/(X-Y) \to 1$,

$$\nu_Q(\mathcal{E}_j(T)) \to p_j(T), \qquad \nu_Q(\mathcal{F}_j(T)) \to q_j(T).$$

By Lemma 6(1), every n for which $\Lambda^{(j)}(n) \in \mathcal{E}_j(T)$ with positive probability over the ordering has a divisor as stated, so $D_T(Q) \geq \nu_Q(\mathcal{E}_j(T))$. By Lemma 6(2), every n with such a divisor has $\Lambda^{(j)}(n) \in \mathcal{F}_j(T)$ for every ordering, so $D_T(Q) \leq \nu_Q(\mathcal{F}_j(T))$. □

4 The stick-breaking recursion

In this section $T \subseteq [0,1]$ is a finite union of closed intervals. We call T *symmetric* if $T = 1 - T$. For $0 < r \leq 1$ put

$$A_r(T) = \frac{1}{r}\big(T \cap [0,r]\big) \subseteq [0,1], \qquad T_r = A_r(T) \cup \big(1 - A_r(T)\big).$$

Lemma 9 (Complements)**.** *For every $j \geq 0$, $p_j(T) = p_j(T \cup (1-T))$ and $q_j(T) = q_j(T \cup (1-T))$.*

Proof. For $j = 0$ the events are $\{0,1\} \cap T \neq \emptyset$ and $[0,1] \cap T \neq \emptyset$, which are unchanged. For $j \geq 1$, if $\sigma_S(\lambda) + \epsilon\rho(\lambda) = t$ then $\sigma_{S'}(\lambda) + (1-\epsilon)\rho(\lambda) = 1 - t$ for the complement S' of S in $\{1, \ldots, j\}$, because all coordinates and ρ add up to 1; and if $\sigma_S(\lambda) + y = t$ with $0 \leq y \leq \rho(\lambda)$ then $\sigma_{S'}(\lambda) + (\rho(\lambda) - y) = 1 - t$. Hence $\mathcal{E}_j(1-T) = \mathcal{E}_j(T)$ and $\mathcal{F}_j(1-T) = \mathcal{F}_j(T)$, and the sets for $T \cup (1-T)$ are the unions. □

Lemma 10 (Recursion)**.** *Let T be symmetric and $j \geq 1$. Then*

$$p_j(T) = \int_0^1 p_{j-1}(T_r)\,\mathrm{d}r, \qquad q_j(T) = \int_0^1 q_{j-1}(T_r)\,\mathrm{d}r.$$

If $r \in T$, then $p_{j-1}(T_r) = q_{j-1}(T_r) = 1$.

Proof. Put $r = 1 - U_1$, which is uniform on $[0,1]$, and let V_i', R_i' be the stick-breaking variables formed from $U_2, U_3, \ldots$, which are independent of

r. Then $V_{i+1} = rV'_i$ for $i \geq 1$ and $R_j = rR'_{j-1}$. Suppose that $V \in \mathcal{E}_j(T)$, witnessed by (S, ϵ). By the proof of Lemma 9 and the symmetry of T, the complementary pair is also a witness, so we may assume $1 \notin S$. Then the value $\sigma_S(V) + \epsilon R_j = r\big(\sigma_{S-1}(V') + \epsilon R'_{j-1}\big)$ lies in $[0, r]$, where $S - 1 = \{i - 1 : i \in S\}$; so it lies in T if and only if $\sigma_{S-1}(V') + \epsilon R'_{j-1} \in A_r(T)$. Conversely, a witness for $A_r(T)$ gives a witness for T not containing 1. Hence, given r, the event $V \in \mathcal{E}_j(T)$ is the event $(V'_1, \ldots, V'_{j-1}) \in \mathcal{E}_{j-1}(A_r(T))$, whose conditional probability is $p_{j-1}(A_r(T)) = p_{j-1}(T_r)$ by Lemma 9. The same argument, with witnesses (S, y) and $(S', \rho - y)$, gives the formula for q_j. If $r \in T$ then $1 \in A_r(T)$ and $0 \in T_r$; the empty set with $\epsilon = 0$ (or $y = 0$) is a witness, so both probabilities are 1. □

Since T is a finite union of closed intervals, its boundary ∂T is finite, so $|\partial T| = 0$, and the recursion reads $p_j(T) = |T| + \int_{(0,1]\setminus T} p_{j-1}(T_r)\,\mathrm{d}r$, and likewise for q_j; the point $r = 0$, where T_r is not defined, has measure zero. For r in a *gap* of T, that is a component of $[0, 1] \setminus T$, the set $T \cap [0, r]$ is the union W of the intervals of T lying below the gap, and it does not depend on r.

Lemma 11 (The last two pieces)**.** *Let T be symmetric with $0 \notin T$, let $0 < r \leq 1$ lie in a gap of T, and let $W = T \cap [0, r]$, with least element l and greatest element u if $W \neq \emptyset$. Then*

$$p_1(T_r) = \frac{|W \cup (r - W)|}{r}, \qquad q_1(T_r) = \frac{\max(r - l, u)}{r},$$

and both vanish if $W = \emptyset$. Consequently

$$p_2(T) = |T| + \int_{(0,1]\setminus T} \frac{g^-(r)}{r}\,\mathrm{d}r, \qquad q_2(T) = |T| + \int_{(0,1]\setminus T} \frac{g^+(r)}{r}\,\mathrm{d}r,$$

with $g^-(r) = |W \cup (r - W)|$ and $g^+(r) = \max(r - l, u)$ (and $g^\pm(r) = 0$ if $W = \emptyset$).

Proof. Put $B = T_r$, a symmetric set with $0, 1 \notin B$, since $0 \notin T$ and $r \notin T$. For $j = 1$ the values $\sigma_S + \epsilon\rho$ are 0, V_1, $1 - V_1$ and 1, so $V_1 \in \mathcal{E}_1(B)$ if and only if $V_1 \in B \cup (1 - B) = B$, which has probability $|B| = |A_r \cup (1 - A_r)| = |W \cup (r - W)|/r$. The values $\sigma_S + y$ fill $[0, 1 - V_1] \cup [V_1, 1]$. As B is symmetric, $[0, 1 - V_1]$ meets B if and only if $[V_1, 1]$ meets B, that is, if and only if $V_1 \leq \max B$. Now $\max B = \max(u/r, 1 - l/r)$, which gives $q_1(B)$. The formulas for p_2 and q_2 follow from Lemma 10. □

Lemma 12 (Cells). *Let T be symmetric with $0 \notin T$, and let $0 < a < b \le 1$ with $[a, b]$ in the closure of a gap of T. Let $[l_1, u_1], \dots, [l_m, u_m]$ be the intervals of T lying in $[0, a]$. For every $r \in (a, b)$ and every $k \ge 0$,*

$$p_k(T_r) \ge p_k\Big(\bigcup_{l_i/a \le u_i/b} [l_i/a, u_i/b] \Big), \qquad q_k(T_r) \le q_k\Big(\bigcup_i [l_i/b, u_i/a] \Big).$$

Proof. For $r \in (a, b)$ we have $A_r(T) = \bigcup_i [l_i/r, u_i/r]$, and $[l_i/a, u_i/b] \subseteq [l_i/r, u_i/r] \subseteq [l_i/b, u_i/a]$. The claim follows from the monotonicity of p_k, q_k and Lemma 9. □

Lemma 13. *Let $0 < a < b$ and $f(r) = \alpha + \beta/r$. If $\beta \ge 0$ then $(b-a) f(\frac{a+b}{2}) \le \int_a^b f \le (b-a)\frac{f(a)+f(b)}{2}$, and if $\beta \le 0$ the inequalities are reversed. Moreover $\beta \ge 0$ if and only if $f(a) \ge f(b)$.*

Proof. f is convex for $\beta \ge 0$ and concave for $\beta \le 0$; these are the Hermite–Hadamard inequalities. The last statement holds as $f(a) - f(b) = \beta(b-a)/(ab)$. □

On a gap of T the function g^+ is affine on each side of $l + u$. The function g^- is affine between consecutive points of the form $l_i + l_k$, $l_i + u_k$, $u_i + l_k$, $u_i + u_k$, since

$$g^-(r) = 2|W| - \sum_{i,k} \big| [l_i, u_i] \cap [r - u_k, r - l_k] \big|$$

when $W = \bigcup_i [l_i, u_i]$ with disjoint intervals, and each summand is a piecewise affine function of r with breakpoints at these points. On each piece $g^{\pm}(r)/r$ has the form $\alpha + \beta/r$ of Lemma 13.

Proposition 14. *Let $T_{\text{in}} = [0.142858, 0.199999] \cup [0.800001, 0.857142]$ and $T_{\text{out}} = [0.142857, 0.2] \cup [0.8, 0.857143]$. Then*

$$p_6(T_{\text{in}}) \ge 0.587142, \qquad q_5(T_{\text{out}}) \le 0.595573.$$

The proof is a computation, described in Section 5.

5 The computation

All numbers are integers in units of 10^{-6}; the targets T_{in} and T_{out} have endpoints on this grid. A program computes a lower bound $\underline{p}_k(T)$ for $p_k(T)$ and an upper bound $\overline{q}_k(T)$ for $q_k(T)$, for symmetric unions T of grid

intervals, by the following recursion. It is written in C++17 and uses integer arithmetic only, with 64-bit products and accumulators. Grid points, lengths, rounded probabilities in units of 10^{-6} and their sums and differences have absolute value at most $2 \cdot 10^6 + 2$ and are held in 32-bit integers; every product and accumulated sum is a 64-bit signed integer of absolute value below $3 \cdot 10^{12}$, far inside the 64-bit range.

Trivial cases. If $T = \emptyset$ the value is 0; if $0 \in T$ it is 1.

Levels $k \geq 3$. The contribution of T itself is $|T|$. Each gap (s, e) of T lying above the least element of T is cut into cells $[a, b]$ of length $\lceil 10^6/m_k \rceil$ units starting from s, the last one ending at e; gaps below the least element of T contribute 0, since there $W = \emptyset$. For the lower bound, each interval $[l_i, u_i]$ of T below a gives the grid interval $[\lceil 10^6 l_i/a \rceil, \lfloor 10^6 u_i/b \rfloor]$, rounded inward, which is kept if it is nondegenerate; these intervals and their reflections are merged into a symmetric set L, and the cell contributes $(b-a)\,\underline{p}_{k-1}(L)$. For the upper bound one uses $[\lfloor 10^6 l_i/b \rfloor, \lceil 10^6 u_i/a \rceil]$, rounded outward, giving a symmetric set M, and the cell contributes $(b-a)\,\overline{q}_{k-1}(M)$. The total is divided by 10^6 and rounded down (lower bound) or up (upper bound). By Lemmas 10 and 12 and monotonicity, the results are bounds in the stated direction.

Level $k = 2$. Let $g = g^-$ for the lower bound and $g = g^+$ for the upper bound. On each gap of T the program forms the breakpoints of g listed after Lemma 13, evaluates g exactly at consecutive breakpoints $a < b$, decides whether $g(r)/r$ is convex on $[a, b]$ by comparing $g(a)/a$ with $g(b)/b$, and bounds the integral of $g(r)/r$ over $[a, b]$ by $(b-a)$ times the midpoint value $(g(a) + g(b))/(a + b)$ or the mean of the endpoint values $g(a)/a$ and $g(b)/b$, whichever Lemma 13 allows, each rounded in the safe direction. Together with $|T|$ this gives $\underline{p}_2(T)$ and $\overline{q}_2(T)$, by Lemma 11.

The cell counts are $m_6 = 450$, $m_5 = 320$, $m_4 = 220$, $m_3 = 100$ for the lower bound, starting from $\underline{p}_6(T_{\mathrm{in}})$, and $m_5 = 600$, $m_4 = 500$, $m_3 = 400$ for the upper bound, starting from $\overline{q}_5(T_{\mathrm{out}})$. The program prints

$$\underline{p}_6(T_{\mathrm{in}}) = 587142 \cdot 10^{-6}, \qquad \overline{q}_5(T_{\mathrm{out}}) = 595573 \cdot 10^{-6},$$

which proves Proposition 14. The program itself checks that the lower bound is at least 587142 and the upper bound at most 595573, so a successful run certifies exactly the numbers stated in Proposition 14. It ran in 31 seconds on one core of a 2.1 GHz Intel Xeon (compiled with GCC 13.3) on 24 November 2026. The program is available at `https://github.com/charlieoneill11/maths-releases`.

Comparison with earlier computations. Drappeau and Mounier [5] computed the density on a grid of windows, which does not contain

$[0.142858, 0.199999]$ or $[1/7, 1/5]$. Their values therefore give no proof of either threshold in Theorem 2: by monotonicity, a window of the grid gives a lower bound for the inward window only if it lies inside it, such as $[0.15, 0.19]$, and an upper bound for $[1/7, 1/5]$ only if it contains it, such as $[0.14, 0.20]$, and Table 1 shows that neither of these two reaches the threshold. For the comparison we ran the program with only the target intervals changed, to $T = [\alpha, \beta] \cup [1-\beta, 1-\alpha]$ for three windows $[\alpha, \beta]$ of that grid, with the same cell counts. By Proposition 8, the resulting $\underline{p}_6(T)$ and $\overline{q}_5(T)$ are rigorous lower and upper bounds for the lower and upper limits of the proportion of integers n with a divisor in $[n^\alpha, n^\beta]$. As a

Window $[\alpha, \beta]$	$\underline{p}_6(T)$	$\overline{q}_5(T)$
$[0.14, 0.20]$	0.596755	0.604658
$[0.15, 0.19]$	0.530857	0.544968
$[0.15, 0.20]$	0.561986	0.572090

Table 1: Rigorous bounds from the program of Section 5 on three windows of the grid of [5], with the target intervals changed.

heuristic check, which no result of this paper uses, we estimated the density on these windows by Monte Carlo in the Poisson–Dirichlet formulation of the limit [6], with $2 \cdot 10^7$ samples per window; the program is available at `https://github.com/charlieoneill11/maths-releases`. The estimates 0.6014, 0.5394 and 0.5683, each with standard error about $1.1 \cdot 10^{-4}$, lie inside our rigorous intervals. The same program gives 0.5920 for $[1/7, 1/5]$, between the two bounds of Theorem 2, and 0.6572 for the constant c of [5, Corollary 1.3], inside their published enclosure $0.651 \le c \le 0.666$. Our lower bound 0.530857 on $[0.15, 0.19]$ proves that a majority of integers have a divisor in that window.

6 Proofs of the theorems

Proof of Theorem 2. (1) Let $T = T_{\text{in}}$. By Proposition 8 with $j = 6$ and Proposition 14, $\liminf_Q D_T(Q) \ge p_6(T_{\text{in}}) \ge 0.587142 > 0.5871$. A divisor d with $\log d / \log n \in [0.800001, 0.857142]$ gives the divisor n/d with $\log(n/d)/\log n \in [0.142858, 0.199999]$, so $D_T(Q)$ is the proportion in part (1).

(2) Let $T = [1/7, 1/5] \cup [4/5, 6/7]$, so that $D_T(Q)$ is the proportion in part (2), by the same complementation. Since $T \subseteq T_{\text{out}}$, Proposition 8 with $j = 5$

and Proposition 14 give $\limsup_Q D_T(Q) \le q_5(T) \le q_5(T_{\mathrm{out}}) \le 0.595573 < 0.5956$. □

Proof of Theorem 1. Fix $\eta = \min(\eta_0, 10^{-8})$ and let $\delta = 4\eta/(1+2\eta)$. Let $T = T_{\mathrm{in}}$. The proof of Theorem 2(1) gives $\liminf_Q D_T(Q) \ge 0.587142$, so for all large Q more than $0.58714\, m_Q$ of the $m_Q = \lfloor Q \rfloor - \lfloor Q/2 \rfloor$ integers $n \in (Q/2, Q]$ have a divisor d with $n^{0.142858} \le d \le n^{0.199999}$. By Lemma 4, all of them except at most $Q/\log Q = o(m_Q)$ satisfy $\max_a P(a, n) \le 4n^{2-\delta}$. Hence for all large Q more than $0.587\, m_Q$ integers $n \in (Q/2, Q]$ satisfy this bound, which is Theorem 1 with $C = 4$. □

7 Remarks

1. *The limit of the method.* Suppose Proposition 3 is applied to a modulus n with $x \le n^{2-\delta}$. Since $n \le x^{1/2+\eta}$, we have $n^{2/(1+2\eta)} \le x \le n^{2-\delta}$, and the required divisor $e \in [x^{2/5+5\eta}, x^{3/7}]$ then has $\log e/\log n$ in an interval that tends to $[4/5, 6/7]$ as $\eta, \delta \to 0$; the complementary divisor n/e has $\log(n/e)/\log n$ in an interval tending to $[1/7, 1/5]$. By Theorem 2(2), fewer than 59.56% of the moduli have a divisor in $[n^{1/7}, n^{1/5}]$ itself, so this route cannot go much beyond Theorem 1 without new input. Since $T_{\mathrm{in}} \subseteq [1/7, 1/5] \cup [4/5, 6/7]$, the lower and upper limits of the proportion of integers with a divisor in $[n^{1/7}, n^{1/5}]$ both lie in $[0.5871, 0.5956]$.

2. *Almost all classes.* For every modulus n, the generalised Riemann hypothesis for Dirichlet L-functions [1], the explicit formula for each character and Parseval's identity show that all but $O\big(\varphi(n)^2(\log nx)^4/x\big)$ reduced classes contain a prime up to x. This is far from every class when $x = n^{2-\delta}$.

8 Formal verification in Lean

Proposition 14 has also been proved in the Lean 4 proof assistant [7] with the Mathlib library [8], using version 4.29.1 of Lean and the matching release of Mathlib. So have Theorems 1 and 2, from Proposition 3 and Theorem 5, which enter as explicit hypotheses. The development is available at `https://github.com/charlieoneill11/maths-releases`.

Statements. The formal statements define the integers in $(Q/2, Q]$, the law of $(V_1, \dots, V_j)$ as the image of the uniform measure on $[0, 1]^j$, the sets

$\mathcal{E}_j(T)$ and $\mathcal{F}_j(T)$, and the targets T_{in} and T_{out}, as in Section 3 and Proposition 14, and state Theorems 1 and 2 as written. The two inequalities of Proposition 14 are proved without hypotheses. The two cited inputs are stated as hypotheses. The first is Proposition 3 with $A = 1$, in the weaker form that every reduced class modulo q contains a prime up to x, which is all that Lemma 4 uses. The second is Theorem 5, with the size-biased ordering of Section 3. From these, Lean proves Theorem 1 with $C = 4$ and both parts of Theorem 2, which need only the second input.

Proof. Every result of Sections 2 to 4 and 6 has a Lean counterpart, and all of them are proved. The computation of Section 5 is reimplemented in Lean as a checker for symmetric unions of grid intervals, with a proof that an accepted run bounds the corresponding probability. Run with the cell counts of Section 5, it returns exactly 587142 and 595573. Lean also proves the six bounds of Table 1. On $4{,}248$ random and edge-case targets at levels up to six, the Lean checker and the program of Section 5 returned the same integers.

Trust. The checkers are evaluated with Lean's `native_decide` tactic, which runs compiled code, so accepting the formal proof means trusting the Lean compiler for these 18 evaluations as well as the Lean kernel. Apart from them, the proofs use only the three standard axioms of Lean: propositional extensionality, the axiom of choice and the soundness of quotients. They contain no unproved steps and declare no further axioms. The longest evaluation, for the lower bound of Proposition 14, takes about six minutes.

References

[1] L. Alpöge, *Remark*, 20 November 2026, https://alpo.ge/remark.pdf. Formal Lean proof of the Riemann hypothesis and of GRH for Dirichlet L-functions: https://github.com/leanprover-community/riemann-hypothesis. Independent verification: A. V. Sutherland, https://math.mit.edu/~drew/grh/; M. Rubinstein, https://www.math.uwaterloo.ca/~mrubinst/alpoge_claude_grh/.

[2] E. Bach and J. Sorenson, *Explicit bounds for primes in residue classes*, Math. Comp. **65** (1996), 1717–1735.

[3] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., Wiley, New York, 1999.

[4] P. Donnelly and G. Grimmett, *On the asymptotic distribution of large prime factors*, J. London Math. Soc. (2) **47** (1993), 395–404.

[5] S. Drappeau and A. Mounier, *Computing sieve integrals using LattE, and the density of integers with a localized divisor*, arXiv:2606.30428, 2026.

[6] T. Haddad, *Poisson–Dirichlet approximation for counting integers with divisors in an interval*, arXiv:2512.13669, 2025.

[7] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction – CADE 28, Lecture Notes in Comput. Sci. **12699**, Springer, 2021, 625–635.

[8] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, 367–381.

[9] J. Maynard, *Dense clusters of primes in subsets*, Compos. Math. **152** (2016), 1517–1554.

[10] J. Maynard, *Primes in arithmetic progressions to large moduli III: uniform residue classes*, Mem. Amer. Math. Soc., doi:10.1090/memo/1544; arXiv:2006.08250.

[11] G. Tenenbaum, *Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné*, Compos. Math. **51** (1984), 243–263.

[12] T. Xylouris, *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression*, Bonner Math. Schriften **404** (2011).