Erdős Problem #783: sharp asymptotic value and a stability program

Przemyslaw Chojecki

February 25, 2026

Abstract

Fix $C > 0$ and N large. Among pairwise coprime sets $A \subseteq \{2, \ldots, N\}$ with harmonic budget $\sum_{a \in A} \frac{1}{a} \leq C$, Erdős asked which A minimises the number of integers $m \leq N$ not divisible by any $a \in A$. Tao recently proved that for all such A one has $\sigma_N(A) \geq \rho(e^{\mu(A)}) + o(1)$, where $\sigma_N(A)$ is the unsieved density, $\mu(A) = \sum_{a \in A} \frac{1}{a}$, and ρ is the Dickman–de Bruijn function. In particular, the minimum possible unsieved density with budget $\leq C$ is $\rho(e^C) + o(1)$, matching the classical "prime tail" construction and thereby resolving a natural asymptotic-value form of the problem.

The original problem is structural: classify the minimisers (or near-minimisers). This note collects the unconditional asymptotic-value results and formulates a concrete stability program, motivated by Tao's reduction to primes and the Granville–Soundararajan integral-equation framework in the prime case. The main missing step for a structural theorem is a quantitative stability refinement of the Granville–Soundararajan equality mechanism, together with near-extremal bookkeeping in Tao's reduction. We prove these missing statements in the last section.

1 Problem, notation, and the prime-tail heuristic

Definition 1 (pairwise coprime, harmonic budget, unsieved density)**.** A finite set $A \subset \mathbb{N}$ is *pairwise coprime* if $\gcd(a, b) = 1$ for all distinct $a, b \in A$. Define its *logarithmic size* (harmonic budget)

$$\mu(A) := \sum_{a \in A} \frac{1}{a}.$$

For a natural number N, define the *unsieved density*

$$\sigma_N(A) \coloneqq \frac{1}{N}\Big|\Big\{n \in \{1, \dots, N\} : \ a \nmid n \ \forall a \in A\Big\}\Big|.$$

We always exclude $1 \in A$, since then $\sigma_N(A) = 0$.

Remark 2 (Erdős Problem #783). Fix $C > 0$. For large N, consider pairwise coprime $A \subseteq \{2, \dots, N\}$ with $\mu(A) \leq C$. Erdős asked: *which such A minimises the number of $n \leq N$ divisible by none of the $a \in A$?* Equivalently, minimise $\sigma_N(A)$. See [1, 3] and the discussion around the "prime tail" construction.

1.1 Prime tails and smooth numbers

Write

$$\mathcal{P}(y, N] \coloneqq \{p \leq N : \ p \text{ prime and } p > y\}.$$

Let $P^+(n)$ denote the largest prime factor of n, with $P^+(1) = 1$. Define the smooth-number counting function

$$\Psi(x, y) \coloneqq \#\{n \leq x : \ P^+(n) \leq y\}.$$

If $A = \mathcal{P}(y, N]$, then an integer $n \leq N$ avoids all primes $> y$ iff it is y-smooth, so

$$\sigma_N(\mathcal{P}(y, N]) = \frac{\Psi(N, y)}{N}. \tag{1.1}$$

The Dickman–de Bruijn theorem asserts that if $y = x^{1/u}$ with u fixed, then

$$\Psi(x, y) = x\,\rho(u)\,(1 + o(1)) \qquad (x \to \infty), \tag{1.2}$$

where ρ is the Dickman function. (See e.g. [8, Ch. III] for a modern account.)

Separately, Mertens' theorem gives

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + B_1 + o(1) \qquad (x \to \infty), \tag{1.3}$$

and therefore the harmonic measure of a prime tail is asymptotic to a log-ratio of log's:

$$\sum_{y < p \leq N} \frac{1}{p} = \log\log N - \log\log y + o(1).$$

Choosing y so that $\sum_{y<p\le N} \frac{1}{p} \approx C$ gives $\log\log y \approx \log\log N - C$, hence $\log y \approx e^{-C}\log N$ and so $y \approx N^{e^{-C}}$. Then $u = \frac{\log N}{\log y} \approx e^{C}$, and (1.1)–(1.2) suggest

$$\sigma_N(\mathcal{P}(y,N]) \approx \rho(e^{C}).$$

Erdős suggested that (up to boundary effects) this construction should be extremal.

2 Tao's theorem and the asymptotic minimum value

2.1 Tao's lower bound

Theorem 3 (Tao [7])**.** *Let $A \subseteq \{2,\dots,N\}$ be pairwise coprime with $\mu(A) = O(1)$. Then*

$$\sigma_N(A) \ \ge\ \rho\big(e^{\mu(A)}\big) + o(1) \qquad (N \to \infty),$$

where the $o(1)$ depends on the upper bound for $\mu(A)$ but is uniform in the choice of A.

Remark 4 (Prime case and Tao's strategy). The prime-only case (when A consists entirely of primes) was established earlier by Hildebrand [6]; see also Granville–Soundararajan [5] for a streamlined approach. Tao deduces Theorem 3 from the prime case by showing that: (i) the "large" elements of a pairwise coprime set are mostly prime; and (ii) the effect of "small" elements can be controlled using a splitting/Bonferroni argument and the log-concavity of ρ.

2.2 The minimum unsieved density is $\rho(e^{C}) + o(1)$

Define the extremal density

$$m_C(N) := \inf\Big\{\sigma_N(A) :\ A \subseteq \{2,\dots,N\} \text{ pairwise coprime},\ \mu(A) \le C\Big\}.$$

Lemma 5 (A feasible prime-tail competitor)**.** *Fix $C > 0$ and $N \ge 3$. Let $y = y_N(C) \in [2,N]$ be maximal such that*

$$\sum_{\substack{y<p\le N\\ p \text{ prime}}} \frac{1}{p} \ \le\ C.$$

Let

$$A_N^{\star}(C) := \mathcal{P}(y,N].$$

Then $A^\star_N(C)$ is admissible (pairwise coprime, $\mu(A^\star_N(C)) \le C$). Moreover,

$$\mu(A^\star_N(C)) = C + o(1) \qquad \text{and} \qquad y = N^{e^{-C}+o(1)} \qquad (N \to \infty).$$

Sketch. Admissibility is by construction. For asymptotics, let $S(x) := \sum_{p \le x} \frac{1}{p}$. From (1.3),

$$\mu(A^\star_N(C)) = S(N) - S(y) = \log\log N - \log\log y + o(1).$$

Maximality of y forces $S(N) - S(y) \in (C - 1/y, C]$, hence $\mu(A^\star_N(C)) = C + o(1)$ since $y \to \infty$. Solving $\log\log N - \log\log y = C + o(1)$ gives $\log y = e^{-C+o(1)} \log N$, i.e. $y = N^{e^{-C}+o(1)}$. $\square$

Corollary 6 (Asymptotic minimum value)**.** *For every fixed $C > 0$,*

$$m_C(N) = \rho(e^C) + o(1) \qquad (N \to \infty).$$

Equivalently,

$$\min_A \#\{n \le N : \ a \nmid n \ \forall a \in A\} = \big(\rho(e^C) + o(1)\big)\, N,$$

where the minimum is over pairwise coprime $A \subseteq \{2, \dots, N\}$ with $\mu(A) \le C$.

Proof. Lower bound: If $\mu(A) \le C$, then Theorem 3 gives $\sigma_N(A) \ge \rho(e^{\mu(A)}) + o(1) \ge \rho(e^C) + o(1)$ since ρ is decreasing.

Upper bound: Take $A^\star_N(C)$ from Lemma 5. Then $n \le N$ avoids all primes in $A^\star_N(C)$ iff $P^+(n) \le y$, so by (1.1)

$$\sigma_N(A^\star_N(C)) = \frac{\Psi(N, y)}{N}.$$

Here $y = N^{e^{-C}+o(1)}$, so $u := \log N / \log y = e^C + o(1)$, and Dickman–de Bruijn (1.2) plus continuity of ρ yield $\sigma_N(A^\star_N(C)) = \rho(e^C) + o(1)$. $\square$

Remark 7. Corollary 6 resolves a natural "asymptotic-value" version of Erdős #783: it pins down the leading constant in the minimum uncovered count. The original question is stronger: it asks for the *choice(s) of A* that minimise, and for a structural description of (near-)minimisers.

3 The regime $C \le \log 2$

Tao records a simple Lipschitz inequality, which already yields the correct lower bound when $\mu(A) \le \log 2$.

Lemma 8 (Lipschitz bound [7])**.** *For any finite* $A, A' \subset \mathbb{N}$ *and any* N*,*

$$\left|\sigma_N(A') - \sigma_N(A)\right| \le \mu(A \triangle A').$$

Proof. If an integer $n \le N$ contributes to exactly one of $\sigma_N(A)$ or $\sigma_N(A')$, then n is divisible by some element of $A \triangle A'$. By the union bound, this happens for at most $\sum_{q \in A \triangle A'} \frac{1}{q} = \mu(A \triangle A')$ of the integers up to N. ☐

Lemma 9 (Dickman on [1,2)**.** *] For* $1 \le u \le 2$*, one has* $\rho(u) = 1 - \log u$*.*

Proof. For $1 < u \le 2$ the delay differential equation reads $u\rho'(u) + \rho(u-1) = 0$ with $\rho(u-1) = 1$, hence $\rho'(u) = -1/u$. Integrate from 1 to u using $\rho(1) = 1$. ☐

Corollary 10 (Clean value bound for $C \le \log 2$)**.** *If* $\mu(A) \le C \le \log 2$*, then for all* N*,*

$$\sigma_N(A) \ge 1 - \mu(A) = \rho(e^{\mu(A)}) \ge \rho(e^C) = 1 - C.$$

Proof. Apply Lemma 8 with $A' = \emptyset$, giving $\sigma_N(A) \ge 1 - \mu(A)$. If $\mu(A) \le \log 2$, then $e^{\mu(A)} \le 2$ and Lemma 9 gives $\rho(e^{\mu(A)}) = 1 - \mu(A)$. Monotonicity yields $\rho(e^{\mu(A)}) \ge \rho(e^C)$. ☐

Remark 11 (Extremisers for $C \le \log 2$)*.* Beyond the lower bound, the *structural* extremal problem is easier in the small budget regime: for $C \le \log 2$ the relevant prime-tail threshold satisfies $y \gtrsim \sqrt{N}$, so the sets of multiples $\{n \le N : \ p \mid n\}$ for distinct large primes p are essentially disjoint, making the union bound sharp. See [2] and the discussion on [1]. In this range, Erdős's "consecutive largest primes" construction is known to be extremal (Chojecki; see [1]).

4 What remains open: structural classification for $C > \log 2$

4.1 A stability formulation

Since exact minimisers can depend on boundary effects (prime gaps, floors $\lfloor N/a \rfloor$, etc.), a natural goal is an *asymptotic stability* statement, allowing $o(1)$ perturbations in harmonic weight.

Definition 12 (μ-distance)**.** For finite sets $A, B \subseteq \mathbb{N}$ define

$$d_\mu(A,B) := \mu(A \triangle B) = \sum_{n \in A \triangle B} \frac{1}{n}.$$

Definition 13 (Near-minimiser (one convenient formulation))**.** Fix $C > 0$. A sequence $A = A(N) \subseteq \{2, \ldots, N\}$ is a *near-minimiser at budget* C if $A(N)$ is pairwise coprime, $\mu(A(N)) \le C + o(1)$, and

$$\sigma_N(A(N)) \le \rho(e^C) + o(1) \qquad (N \to \infty).$$

Conjecture 14 (Prime-tail stability for $C > \log 2$)**.** Fix $C > \log 2$ and write $w := e^C > 2$. If $A = A(N)$ is a near-minimiser at budget C, then there exists a threshold $y = y(N) = N^{1/w+o(1)}$ such that

$$d_\mu\big(A,\ \mathcal{P}(y, N]\big) = o(1).$$

In words: any asymptotically extremal family is μ-close to a tail of primes.

Remark 15. Conjecture 14 is compatible with known phenomena: it does *not* assert a unique minimiser, only that minimisers lie in a narrow "boundary layer" around a prime tail, measured in harmonic weight. The Erdős problems page [1] notes that small perturbations can beat naive exact prime-tail rules at finite N without changing the leading constant.

5 A strictness lemma from the Dickman log-concavity

Tao uses the log-concavity of ρ (Hildebrand's inequality) to combine contributions from different scales. For stability, it is useful to isolate a quantitative strictness away from the boundary.

Lemma 16 (Log-concavity and supermultiplicativity [7, 6])**.** *The Dickman function ρ is log-concave on $[1, \infty)$. In particular, for all $u_1, u_2 \ge 1$,*

$$\rho(u_1)\rho(u_2) \ \ge \ \rho(u_1 u_2).$$

Lemma 17 (Strict concavity of $t \mapsto \log \rho(e^t)$)**.** *Let $F : (0, \infty) \to \mathbb{R}$ be defined by $F(t) := \log \rho(e^t)$. Then F is strictly concave on $(0, \infty)$.*

Proof. Write $u = e^t > 1$. Then

$$F'(t) = \frac{u\rho'(u)}{\rho(u)}, \qquad F''(t) = \frac{u\rho'(u)}{\rho(u)} + u^2(\log \rho)''(u).$$

For $u > 1$, the Dickman equation gives $\rho'(u) = -\rho(u-1)/u < 0$, so the first term is strictly negative. Log-concavity yields $(\log \rho)''(u) \le 0$, so the second term is nonpositive. Hence $F''(t) < 0$. ☐

Lemma 18 (Uniform strictness on compacta)**.** *Fix $C > 0$ and $\delta \in (0, C/2]$. There exists $\kappa = \kappa(C, \delta) > 0$ such that for all $x, y \in [\delta, C]$ with $x + y \leq C$,*

$$\rho(e^x)\,\rho(e^y) \;\geq\; e^{\kappa}\,\rho(e^{x+y}).$$

Proof. Let $D(x, y) := F(x) + F(y) - F(x + y)$ with F as in Lemma 17. By strict concavity, $D(x, y) > 0$ for all $x, y > 0$. On the compact set $K = \{(x, y) : x, y \in [\delta, C],\; x + y \leq C\}$, continuity gives $\min_K D = \kappa > 0$. Exponentiate. ☐

Corollary 19 (Near equality forces one parameter to be small)**.** *Fix $C > 0$. Let (x_N, y_N) satisfy $x_N, y_N \geq 0$ and $x_N + y_N \leq C + o(1)$. If*

$$\rho(e^{x_N})\,\rho(e^{y_N}) \leq (1 + o(1))\,\rho(e^{x_N + y_N}),$$

then $\min(x_N, y_N) = o(1)$*.*

Proof. Otherwise, along a subsequence $x_N, y_N \geq \delta > 0$. For N large $x_N + y_N \leq C$ and Lemma 18 yields a fixed multiplicative gap $e^{\kappa} > 1$, contradiction. ☐

Remark 20 (How this enters stability)*.* Tao's proof effectively decomposes A into pieces whose contributions are combined using Lemma 16. Corollary 19 shows that *if* such a product inequality is nearly sharp, then one of the pieces must carry $o(1)$ of the relevant "parameter" (harmonic mass). Turning this into an actual structural theorem requires showing that, for near-minimisers with $C > \log 2$, the piece that must be small is the "small-moduli" part rather than the "large-prime" part; this in turn forces one to invoke the prime-case extremal mechanism.

6 Prime-case structure via Granville–Soundararajan

In the prime case, Granville–Soundararajan recast Hildebrand's inequality in terms of an integral equation.

Definition 21 (Integral equation model [5])**.** Let $\chi : [0, \infty) \to [0, 1]$ be measurable with $\chi(t) = 1$ for $t \leq 1$. Let $\sigma = \sigma_\chi$ be the unique continuous solution to

$$u\sigma(u) = \int_0^u \chi(t)\sigma(u - t)\,\mathrm{d}t, \qquad \sigma(u) = 1 \;\text{ for } u \leq 1, \tag{6.1}$$

and define

$$E_\chi(u) := \exp\Big(\int_0^u \frac{1 - \chi(t)}{t}\,\mathrm{d}t\Big). \tag{6.2}$$

Lemma 22 (Feasibility: $E_\chi(u) \le u$)**.** *For any* χ *as in Definition 21 and* $u \ge 1$*, one has* $1 \le E_\chi(u) \le u$*.*

Proof. Since $0 \le 1 - \chi(t) \le 1$ for $t \ge 1$ and $1 - \chi(t) = 0$ for $t \le 1$,

$$\int_0^u \frac{1-\chi(t)}{t}\,\mathrm{d}t = \int_1^u \frac{1-\chi(t)}{t}\,\mathrm{d}t \le \int_1^u \frac{1}{t}\,\mathrm{d}t = \log u,$$

so $E_\chi(u) \le u$. Also $E_\chi(u) \ge 1$ since the integrand is nonnegative. □

Theorem 23 (Granville–Soundararajan, Theorem 5 [5])**.** *Let* χ, σ, E *be as in Definition 21. Then for all* $u \ge 0$*,*

$$\sigma(u) \;\ge\; \rho\big(E(u)\big).$$

Moreover, if $E(u) \ge 2$ *and* $\sigma(u) = \rho(E(u))$*, then* $E(u/E(u)) = 1$*, equivalently,*

$$\chi(t) = 1 \quad \textit{for } t \le \frac{u}{E(u)}, \qquad \chi(t) = 0 \quad \textit{for } \frac{u}{E(u)} \le t \le u$$

except possibly on a null set.

Remark 24 (Equality forces a step function)*.* The key qualitative consequence is that, in the prime model, *equality forces a sharp cutoff.* In the multiplicative-function interpretation, this corresponds to sieving by an interval of primes (i.e. excluding all primes above some threshold).

6.1 A quantitative stability problem

To upgrade equality-structure into a classification of near-minimisers, one needs a stability statement in the integral-equation model.

Hypothesis 25 (GS stability at $E(u) > 2$)**.** Fix parameters $u \ge 1$ and $w \in (2, u]$. For every $\delta > 0$ there exists $\eta = \eta(u, w, \delta) > 0$ such that: whenever χ satisfies Definition 21 with $E_\chi(u) = w$ and

$$\sigma_\chi(u) \;\le\; \rho(w) + \eta,$$

then χ is close to the step extremiser in the weighted sense

$$\int_0^{u/w} (1-\chi(t))\,\frac{\mathrm{d}t}{t} \le \delta, \qquad \int_{u/w}^{u} \chi(t)\,\frac{\mathrm{d}t}{t} \le \delta.$$

Remark 26. Hypothesis 25 is the natural quantitative analogue of the equality clause in Theorem 23. A soft compactness argument suggests it should hold on compact subsets of the parameter region $\{(u,w) : u \geq w > 2\}$, but implementing this requires: (i) a topology in which the admissible χ are compact under the $E_\chi(u) = w$ constraint; and (ii) a continuity (or lower-semicontinuity plus rigidity) statement for the map $\chi \mapsto \sigma_\chi(u)$. Making this quantitative is a central analytic task in a structural theory.

6.2 Discretisation back to primes

Granville–Soundararajan relate the integral equation to mean values of multiplicative functions, and hence to prime sets, via a dictionary (see [5, §2]). For structural statements in Erdős #783, one needs a stability transfer from "χ close to step" to "prime set close to a tail" in harmonic measure.

Hypothesis 27 (Stability transfer from χ to prime tails)**.** Fix $w > 2$ and a compact interval of u containing $u = w$. There exists a function $\omega(\delta) \to 0$ as $\delta \to 0$ such that the following holds.

Let x be large and let f be completely multiplicative with $0 \leq f(n) \leq 1$ and $f(n) = 1$ for $n \leq x^{1/u}$. Let χ_f be the associated function in [5, §2]. Assume that χ_f satisfies the weighted step-closeness bounds from Hypothesis 25 with parameter δ. Then there exists a threshold $Y = x^{1/w+o(1)}$ such that the prime support

$$A_f(x) := \{p \leq x : \ f(p) = 0\}$$

satisfies the harmonic closeness estimate

$$\sum_{p \leq x} \frac{\left|\mathbf{1}_{p>Y} - \mathbf{1}_{p \in A_f(x)}\right|}{p} \ \leq \ \omega(\delta) + o(1) \qquad (x \to \infty).$$

Remark 28. Hypothesis 27 packages the analytic number theory required to pass from t-space to primes: partial summation, uniform prime number theorem input, and control of discretisation errors in the map $t \mapsto x^{t/u}$.

7 A conditional structural theorem for Erdős #783

Finally, one must control Tao's reduction in the near-extremal regime. Tao's proof for the value theorem proceeds by discarding certain parts of A at small *harmonic* cost and bounding σ_N from below. For structural results, one needs a near-extremal "no slack" form: if $\sigma_N(A)$ is close to the minimum, then the discarded parts must have $o(1)$ harmonic mass.

Hypothesis 29 (Near-extremal reduction to primes)**.** Fix $C > \log 2$ and write $w = e^C$. If $A \subseteq \{2, \ldots, N\}$ is pairwise coprime with $\mu(A) \le C + o(1)$ and $\sigma_N(A) \le \rho(w) + o(1)$, then there exists a prime set $P = P_N \subseteq \{2, \ldots, N\}$ such that

$$\mu(A \triangle P) = o(1) \qquad \text{and} \qquad \sigma_N(P) \le \rho(w) + o(1),$$

as $N \to \infty$.

Theorem 30 (Conditional prime-tail stability)**.** *Fix $C > \log 2$ and write $w = e^C > 2$. Assume Hypotheses 25, 27, and 29 in the parameter range relevant to w. Then Conjecture 14 holds: every near-minimiser $A = A(N)$ is μ-close to a prime tail, i.e. there exists $y = y(N) = N^{1/w + o(1)}$ such that*

$$d_\mu\big(A,\ \mathcal{P}(y, N]\big) = o(1).$$

Outline. Use Hypothesis 29 to replace A by a near-minimising prime set P with $\mu(A \triangle P) = o(1)$. Apply the prime-case Hildebrand/Granville–Soundararajan mechanism to encode P by a function χ with $E_\chi(u) = w$ and with $\sigma_\chi(u)$ close to $\rho(w)$ at the relevant parameter u. Hypothesis 25 forces χ to be close to a step function, and Hypothesis 27 transfers this to harmonic closeness of P to a prime tail. Finally, $\mu(A \triangle P) = o(1)$ transfers the conclusion from P back to A. □

8 Further remarks: boundary effects and the Buchstab obstruction

Remark 31 (Boundary layers and finite-N optimisation)*.* Even if prime-tail stability holds in the sense of Conjecture 14, it does not by itself determine the *exact* minimiser for each N: near the threshold $y \sim N^{1/w}$, small modifications of which primes are included can change $\mu(A)$ by $o(1)$ and $\sigma_N(A)$ by $o(1)$, and can improve lower-order terms. Understanding these boundary layers appears to require finer asymptotics for smooth numbers and/or quantitative versions of the sieve inequalities (e.g. power-saving errors), as discussed in [7, 1].

Remark 32 (Why composites create a multi-state recursion)*.* For a prime tail $A = \mathcal{P}(y, N]$, Buchstab's identity yields a closed recursion for $\Psi(N, y)$ and hence Dickman asymptotics. For general pairwise coprime A, attempting the same "largest prime factor" decomposition introduces a state change: if p divides some modulus $a \in A$, then the condition $a \nmid (pt)$ becomes

$(a/p) \nmid t$, so one must replace A by a modified set A_p in the recursion. This produces a multi-state Buchstab recursion rather than a closed one, and is a concrete obstruction to naively porting the smooth-number derivation to general coprime sets. See [2] for a clear discussion.

9 Stability and prime-tail rigidity (proving 3 Hypotheses)

Throughout this section we fix a constant $w > 2$ and write

$$C := \log w \qquad (\text{so } w = e^C).$$

All $o(1)$ terms are as the relevant size parameter tends to ∞ (typically $y \to \infty$ or $N \to \infty$), with w (and any fixed u) held constant.

We now verify the three hypotheses from §7: the GS stability hypothesis (Hypothesis 25), the discretisation hypothesis (Hypothesis 27), and the near-extremal reduction-to-primes hypothesis (Hypothesis 29). Combining these with Theorem 30 yields an unconditional asymptotic rigidity statement.

9.1 Proving GS stability by compactness

We record a compactness proof of Hypothesis 25.

Proposition 33 (GS stability)**.** *Fix $u \geq 1$ and $w \in (2, u]$. For every $\delta > 0$ there exists $\eta = \eta(u, w, \delta) > 0$ such that whenever $\chi : [0, u] \to [0, 1]$ is measurable with $\chi(t) = 1$ for $t \leq 1$ and*

$$E_\chi(u) = w, \qquad \sigma_\chi(u) \leq \rho(w) + \eta,$$

one has

$$\int_0^{u/w} \frac{1 - \chi(t)}{t}\, dt \leq \delta \qquad \textit{and} \qquad \int_{u/w}^{u} \frac{\chi(t)}{t}\, dt \leq \delta. \tag{9.1}$$

Proof. Set $u_0 := u/w$ (so $u_0 \geq 1$). Let λ denote the measure $\lambda(dt) := \frac{dt}{t}$ on $[1, u]$. For such a χ, define the absolutely continuous measure

$$\nu_\chi(B) := \int_B \frac{1 - \chi(t)}{t}\, dt \qquad (B \subseteq [1, u] \text{ Borel}).$$

Then $0 \leq \nu_\chi \leq \lambda$ (as measures) and the constraint $E_\chi(u) = w$ is exactly

$$\nu_\chi([1, u]) = \int_1^u \frac{1 - \chi(t)}{t}\, dt = \log w. \tag{9.2}$$

Moreover, since $\chi(t) = 1$ for $t \le 1$ and $u_0 \ge 1$,

$$\int_0^{u_0} \frac{1-\chi(t)}{t}\,dt = \nu_\chi([1,u_0]).$$

Also, using (9.2) and $\int_{u_0}^{u} \frac{dt}{t} = \log w$,

$$\int_{u_0}^{u} \frac{\chi(t)}{t}\,dt = \int_{u_0}^{u} \frac{1}{t}\,dt - \int_{u_0}^{u} \frac{1-\chi(t)}{t}\,dt = \log w - \nu_\chi([u_0,u]) = \nu_\chi([1,u_0]).$$

So the two inequalities in (9.1) are in fact equivalent under $E_\chi(u) = w$, and it suffices to prove $\nu_\chi([1,u_0]) \le \delta$.

We now re-express $\sigma_\chi(u)$ as a continuous functional of ν_χ. For $j \ge 0$ let

$$\Delta_{u,j} := \Big\{(t_1,\dots,t_j) \in [1,u]^j :\ t_1 + \cdots + t_j \le u\Big\},$$

(with $\Delta_{u,0}$ a singleton). As in Granville–Soundararajan [5, Eq. (2.7)], one may write

$$\sigma_\chi(u) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!}\, I_j(u;\chi), \qquad I_j(u;\chi) := \int_{\Delta_{u,j}} d\nu_\chi^{\otimes j}. \tag{9.3}$$

Since $\nu_\chi([1,u]) = \log w$, we have the uniform bound $0 \le I_j(u;\chi) \le (\log w)^j$ for all j.

Let $\mathcal{K}$ be the compact set of all finite Borel measures ν on $[1,u]$ such that $0 \le \nu \le \lambda$ and $\nu([1,u]) = \log w$, equipped with the weak topology. (Compactness follows, e.g., from Banach–Alaoglu/Prokhorov on the compact space $[1,u]$, using the uniform domination $\nu \le \lambda$.) For $\nu \in \mathcal{K}$ define

$$\sigma(u;\nu) := \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_{\Delta_{u,j}} d\nu^{\otimes j}.$$

By the uniform tail bound $\sum_{j>J} (\log w)^j/j! \to 0$ as $J \to \infty$, it suffices to check that each map $\nu \mapsto \nu^{\otimes j}(\Delta_{u,j})$ is continuous on $\mathcal{K}$. But $\partial\Delta_{u,j}$ lies inside the hyperplane $t_1 + \cdots + t_j = u$, which has $\lambda^{\otimes j}$-measure 0, hence also $\nu^{\otimes j}$-measure 0 for every $\nu \le \lambda$. Thus $\Delta_{u,j}$ is a continuity set for all $\nu^{\otimes j}$, and therefore $\nu_n \Rightarrow \nu$ implies $\nu_n^{\otimes j}(\Delta_{u,j}) \to \nu^{\otimes j}(\Delta_{u,j})$. Consequently $\sigma(u;\nu)$ is continuous on $\mathcal{K}$.

Let

$$\mathcal{K}_\delta := \{\nu \in \mathcal{K} :\ \nu([1,u_0]) \ge \delta\},$$

which is closed in the weak topology (since $[1, u_0]$ is closed), hence compact. Define

$$m_\delta := \min_{\nu \in \mathcal{K}_\delta} \sigma(u; \nu),$$

which exists by compactness and continuity.

We claim that $m_\delta > \rho(w)$. Indeed, if $m_\delta = \rho(w)$ then there exists $\nu^\star \in \mathcal{K}_\delta$ with $\sigma(u; \nu^\star) = \rho(w)$. Let $\chi^\star$ be the function on $[0, u]$ given by $\chi^\star(t) = 1$ for $t \le 1$ and, for $t \in [1, u]$, $\chi^\star(t) = 1 - \frac{d\nu^\star}{d\lambda}(t)$. Then $E_{\chi^\star}(u) = w$ by (9.2) and $\sigma_{\chi^\star}(u) = \sigma(u; \nu^\star) = \rho(w)$. By Granville–Soundararajan [5, Theorem 5] (classification of equality cases for $w > 2$), this forces $\chi^\star$ to be the step extremiser

$$\chi^\star(t) = \mathbf{1}_{t \le u_0},$$

which corresponds to the measure $\nu^\star = \lambda\restriction_{[u_0, u]}$ and in particular $\nu^\star([1, u_0]) = 0$, contradicting $\nu^\star \in \mathcal{K}_\delta$.

Thus $m_\delta > \rho(w)$. Set $\eta := \frac{1}{2}(m_\delta - \rho(w)) > 0$. If χ satisfies $E_\chi(u) = w$ and $\sigma_\chi(u) \le \rho(w) + \eta$, then $\nu_\chi \notin \mathcal{K}_\delta$, so $\nu_\chi([1, u_0]) < \delta$, which is exactly (9.1). $\square$

9.2 A discretisation lemma for the GS dictionary

Fix $u \ge 1$ and write $x = y^u$ with $y \to \infty$. If $f : \mathbb{N} \to [0, 1]$ is completely multiplicative, Granville–Soundararajan [5] associate to f the function

$$\chi_f(t) := \frac{1}{\vartheta(y^t)} \sum_{p \le y^t} f(p) \log p, \qquad 0 \le t \le u, \tag{9.4}$$

where $\vartheta(X) := \sum_{p \le X} \log p$.

We need a simple lemma comparing weighted integrals of χ_f with harmonic prime sums.

Lemma 34 (Integral/harmonic comparison)**.** *Fix $1 \le a < b \le u$ and let $x = y^u$ with $y \to \infty$. For any completely multiplicative $f : \mathbb{N} \to [0, 1]$,*

$$\int_a^b \frac{\chi_f(t)}{t}\, dt = \sum_{y^a < p \le y^b} \frac{f(p)}{p} + o(1),$$

where the $o(1)$ depends on u, a, b but is uniform in f.

Proof. Make the change of variables $X := y^t$ (so $t = \frac{\log X}{\log y}$), giving $dt/t = dX/(X \log X)$ and

$$\int_a^b \frac{\chi_f(t)}{t}\, dt = \int_{y^a}^{y^b} \frac{1}{X \log X} \cdot \frac{1}{\vartheta(X)} \Big(\sum_{p \le X} f(p) \log p \Big)\, dX.$$

Write $F(X) := \sum_{p \le X} f(p) \log p$. Since $0 \le f(p) \le 1$, we have $0 \le F(X) \le \vartheta(X)$. By the prime number theorem, $\vartheta(X) = X(1 + o(1))$ uniformly for $X \in [y^a, y^b]$ as $y \to \infty$, so $1/\vartheta(X) = (1 + o(1))/X$ uniformly on this range. Therefore

$$\int_a^b \frac{\chi_f(t)}{t}\,dt = \int_{y^a}^{y^b} \frac{F(X)}{X^2 \log X}\,dX + o(1) \cdot \int_{y^a}^{y^b} \frac{dX}{X \log X}. \tag{9.5}$$

Since $\int_{y^a}^{y^b} \frac{dX}{X \log X} = \log \frac{b}{a} + o(1) = O_{a,b}(1)$, the error term in (9.5) is $o(1)$.

It remains to show that

$$\int_{y^a}^{y^b} \frac{F(X)}{X^2 \log X}\,dX = \sum_{y^a < p \le y^b} \frac{f(p)}{p} + o(1).$$

Apply partial summation in the form

$$\sum_{y^a < p \le y^b} \frac{f(p)}{p} = \sum_{y^a < p \le y^b} f(p) \log p \cdot \frac{1}{p \log p} = \Big[\frac{F(X)}{X \log X}\Big]_{X=y^a}^{y^b} - \int_{y^a}^{y^b} F(X)\, d\Big(\frac{1}{X \log X}\Big).$$

Since $F(X) \le \vartheta(X) \ll X$, the boundary terms are $O(1/\log y^a) = o(1)$. Also

$$-d\Big(\frac{1}{X \log X}\Big) = \Big(\frac{1}{X^2 \log X} + \frac{1}{X^2 \log^2 X}\Big)\,dX,$$

so

$$\sum_{y^a < p \le y^b} \frac{f(p)}{p} = \int_{y^a}^{y^b} \frac{F(X)}{X^2 \log X}\,dX + \int_{y^a}^{y^b} \frac{F(X)}{X^2 \log^2 X}\,dX + o(1).$$

Finally, $\int_{y^a}^{y^b} \frac{F(X)}{X^2 \log^2 X}\,dX \ll \int_{y^a}^{y^b} \frac{dX}{X \log^2 X} = O(1/\log y^a) = o(1)$. This completes the proof. □

A convenient corollary is that "weighted closeness" of χ_f to a step function forces "harmonic closeness" of the prime support of f to a prime tail.

Corollary 35 (Step-to-harmonic transfer)**.** *Fix $u \ge 1$ and let $x = y^u$ with $y \to \infty$. Let $f : \mathbb{N} \to \{0, 1\}$ be completely multiplicative, and define χ_f by (9.4). Suppose $u_0 \in [1, u]$ and*

$$\int_1^{u_0} \frac{1 - \chi_f(t)}{t}\,dt \le \delta, \qquad \int_{u_0}^{u} \frac{\chi_f(t)}{t}\,dt \le \delta.$$

Then there exists a threshold $Y = y^{u_0 + o(1)} = x^{u_0/u + o(1)}$ such that

$$\sum_{\substack{p \le Y \\ f(p)=0}} \frac{1}{p} + \sum_{\substack{Y < p \le x \\ f(p)=1}} \frac{1}{p} \le \delta + o(1).$$

Proof. Apply Lemma 34 with $(a, b) = (1, u_0)$ to obtain

$$\int_1^{u_0} \frac{1 - \chi_f(t)}{t}\, dt = \sum_{y < p \le y^{u_0}} \frac{1 - f(p)}{p} + o(1),$$

and with $(a, b) = (u_0, u)$ to obtain

$$\int_{u_0}^{u} \frac{\chi_f(t)}{t}\, dt = \sum_{y^{u_0} < p \le y^{u}} \frac{f(p)}{p} + o(1).$$

Since $y^{u_0} = x^{u_0/u}$, taking $Y = y^{u_0}$ (or any nearby threshold $y^{u_0+o(1)}$) yields the claim. ☐

9.3 Prime-tail rigidity for near-extremal prime sieves

The GS stability Proposition 33 was stated for a fixed parameter w. In applications one often only knows $E_\chi(u) = w + o(1)$. The following "robust" variant (which is still proved by compactness) lets w vary in a small neighbourhood.

Lemma 36 (Robust GS stability for varying w)**.** *Fix $u \ge 1$ and $w_0 \in (2, u]$. For every $\delta > 0$ there exist $\varepsilon_0 = \varepsilon_0(u, w_0, \delta) > 0$ and $\eta = \eta(u, w_0, \delta) > 0$ with the following property.*

If $w \in [w_0 - \varepsilon_0, w_0 + \varepsilon_0]$ and $\chi : [0, u] \to [0, 1]$ is measurable, $\chi(t) = 1$ for $t \le 1$, such that

$$E_\chi(u) = w \qquad \textit{and} \qquad \sigma_\chi(u) \le \rho(w) + \eta,$$

then

$$\int_0^{u/w} \frac{1 - \chi(t)}{t}\, dt \le \delta \qquad \textit{and} \qquad \int_{u/w}^{u} \frac{\chi(t)}{t}\, dt \le \delta. \tag{9.6}$$

Proof. Suppose not. Then there exists $\delta > 0$ and sequences $w_n \to w_0$ with $w_n \in (2, u]$ and measurable $\chi_n : [0, u] \to [0, 1]$ with $\chi_n(t) = 1$ for $t \le 1$ such that

$$E_{\chi_n}(u) = w_n, \qquad \sigma_{\chi_n}(u) \le \rho(w_n) + \frac{1}{n},$$

but such that for every n at least one inequality in (9.6) fails. Since $E_{\chi_n}(u) = w_n$, the two inequalities in (9.6) are equivalent (cf. the proof of Proposition 33), so after passing to a subsequence we may assume

$$\int_0^{u/w_n} \frac{1 - \chi_n(t)}{t}\, dt \ge \delta \qquad \text{for all } n. \tag{9.7}$$

Define ν_n on $[1,u]$ by $d\nu_n(t) = \frac{1-\chi_n(t)}{t}\,dt$. Then $0 \le \nu_n \le \lambda := dt/t$ and $\nu_n([1,u]) = \log w_n$. Since $\log w_n$ is bounded and $[1,u]$ is compact, by sequential compactness we may pass to a subsequence (still denoted n) such that $\nu_n \Rightarrow \nu$ weakly for some ν with $0 \le \nu \le \lambda$ and

$$\nu([1,u]) = \lim_{n\to\infty} \nu_n([1,u]) = \lim_{n\to\infty} \log w_n = \log w_0.$$

As in the proof of Proposition 33, $\sigma_{\chi_n}(u) = \sigma(u;\nu_n)$ and the map $\nu \mapsto \sigma(u;\nu)$ is continuous on $\{\nu :\ 0 \le \nu \le \lambda,\ \nu([1,u]) \le \log(u)\}$. Hence

$$\sigma(u;\nu) = \lim_{n\to\infty} \sigma(u;\nu_n) = \lim_{n\to\infty} \sigma_{\chi_n}(u) = \lim_{n\to\infty} \rho(w_n) = \rho(w_0).$$

Let χ be defined from ν by $\chi(t) = 1$ for $t \le 1$ and $\chi(t) = 1 - \frac{d\nu}{d\lambda}(t)$ for $t \in [1,u]$. Then $E_\chi(u) = w_0$ and $\sigma_\chi(u) = \rho(w_0)$, so by Granville–Soundararajan's equality case classification (for $w_0 > 2$) one has $\chi(t) = \mathbf{1}_{t \le u/w_0}$ almost everywhere. Equivalently,

$$\nu([1,u/w_0]) = 0. \tag{9.8}$$

On the other hand, $u/w_n \to u/w_0$. Hence

$$\lambda\big([1,u/w_n] \triangle [1,u/w_0]\big) = \left|\log \frac{u/w_n}{u/w_0}\right| \to 0.$$

Since $\nu_n \le \lambda$, it follows that

$$\big|\nu_n([1,u/w_n]) - \nu_n([1,u/w_0])\big| \le \nu_n\big([1,u/w_n] \triangle [1,u/w_0]\big) \le \lambda\big([1,u/w_n] \triangle [1,u/w_0]\big) = o(1).$$

Combining with (9.7), we obtain $\nu_n([1,u/w_0]) \ge \delta/2$ for all sufficiently large n. By Portmanteau applied to the closed set $[1,u/w_0]$, we get

$$\nu([1,u/w_0]) \ \ge\ \limsup_{n\to\infty} \nu_n([1,u/w_0]) \ \ge\ \delta/2,$$

contradicting (9.8). This contradiction proves the lemma. □

We can now prove prime-tail rigidity.

Theorem 37 (Prime-tail rigidity)**.** *Fix $u \ge w > 2$ and write $x = y^u$ with $y \to \infty$. Let $P \subset (y,x]$ be a set of primes such that*

$$\sum_{\substack{p \le x \\ p \in P}} \frac{1}{p} = \log w + o(1) \qquad \textit{and} \qquad \sigma_x(P) = \rho(w) + o(1). \tag{9.9}$$

Then for every $\delta > 0$ *there exists a threshold* $Y = x^{1/w+o(1)}$ *such that*

$$\sum_{\substack{p\le Y\\ p\in P}} \frac{1}{p} + \sum_{\substack{Y<p\le x\\ p\notin P}} \frac{1}{p} \le \delta.$$

Proof. Define a completely multiplicative $f : \mathbb{N} \to \{0,1\}$ by

$$f(p) := \mathbf{1}_{p\notin P} \quad (p \text{ prime}),$$

extended multiplicatively. Since $P \subset (y, x]$, one has $f(n) = 1$ for all $n \le y$. Let χ_f be as in (9.4).

Step 1: translate $\sigma_x(P)$ to $\sigma_{\chi_f}(u)$. By Granville–Soundararajan's mean-value dictionary [5, Proposition 2.1], for fixed u we have

$$\frac{1}{x}\sum_{n\le x} f(n) = \sigma_{\chi_f}(u) + O\Big(\frac{u}{\log y}\Big) = \sigma_{\chi_f}(u) + o(1).$$

The left-hand side equals $\sigma_x(P)$, hence $\sigma_{\chi_f}(u) = \rho(w) + o(1)$.

Step 2: translate $\sum_{p\in P} 1/p$ to $E_{\chi_f}(u)$. By Lemma 34 with $(a, b) = (1, u)$,

$$\int_1^u \frac{1-\chi_f(t)}{t}\,dt = \sum_{y<p\le y^u} \frac{1-f(p)}{p} + o(1) = \sum_{\substack{p\le x\\ p\in P}} \frac{1}{p} + o(1) = \log w + o(1).$$

Since $\chi_f(t) = 1$ for $t \le 1$, we have

$$\int_0^u \frac{1-\chi_f(t)}{t}\,dt = \int_1^u \frac{1-\chi_f(t)}{t}\,dt,$$

and therefore

$$w_x \;:=\; E_{\chi_f}(u) \;=\; \exp\Big(\int_0^u \frac{1-\chi_f(t)}{t}\,dt\Big) = \exp(\log w + o(1)) = w + o(1).$$

By continuity of ρ, also $\rho(w_x) = \rho(w) + o(1)$, hence

$$\sigma_{\chi_f}(u) = \rho(w) + o(1) = \rho(w_x) + o(1).$$

Step 3: apply robust GS stability at w_x. Fix $\delta > 0$. Apply Lemma 36 with $w_0 = w$ and this δ to obtain $\varepsilon_0, \eta > 0$. For x large we have $w_x \in [w - \varepsilon_0, w + \varepsilon_0]$ and $\sigma_{\chi_f}(u) \le \rho(w_x) + \eta$, so the lemma yields

$$\int_0^{u/w_x} \frac{1-\chi_f(t)}{t}\,dt \le \delta, \qquad \int_{u/w_x}^{u} \frac{\chi_f(t)}{t}\,dt \le \delta.$$

Since $w_x \to w$ and $u \geq w > 2$, we have $u/w_x \geq 1$ for all sufficiently large x.

Step 4: discretise back to primes. Apply Corollary 35 with $u_0 = u/w_x$ to obtain a threshold

$$Y = y^{u_0+o(1)} = y^{u/w_x+o(1)} = x^{1/w_x+o(1)} = x^{1/w+o(1)}$$

such that

$$\sum_{\substack{p \leq Y \\ f(p)=0}} \frac{1}{p} + \sum_{\substack{Y<p\leq x \\ f(p)=1}} \frac{1}{p} \leq \delta + o(1).$$

Since $f(p) = 0$ iff $p \in P$, this is exactly the desired conclusion (after absorbing $o(1)$ into δ). ☐

9.4 Near-extremal reduction to a prime sieve

We now prove the stability version of Tao's reduction-to-primes mechanism (Hypothesis 29), using (i) Tao's factorisation scheme from [7] together with (ii) strictness of the ρ log-concavity (Corollary 19).

Lemma 38 (Quantitative strictness of log-concavity)**.** *Fix $C_0 > 0$. For every $\tau > 0$ there exists $\varepsilon = \varepsilon(C_0, \tau) > 0$ such that whenever $a, b \in [0, C_0]$ satisfy*

$$\frac{\rho(e^a)\rho(e^b)}{\rho(e^{a+b})} \leq 1 + \varepsilon,$$

then $\min\{a, b\} \leq \tau$.

Proof. Define $\Psi(a, b) := \log \rho(e^a) + \log \rho(e^b) - \log \rho(e^{a+b})$ on $[0, C_0]^2$. Then $\Psi \geq 0$ by log-concavity (supermultiplicativity), and $\Psi(a, b) = 0$ iff $ab = 0$ by strict log-concavity for $w > 2$. The compact set

$$K_\tau := \{(a, b) \in [0, C_0]^2 : \ a \geq \tau, \ b \geq \tau\}$$

is disjoint from the zero set, hence $\min_{K_\tau} \Psi =: m_\tau > 0$. If $\rho(e^a)\rho(e^b) \leq (1+\varepsilon)\rho(e^{a+b})$, then $\Psi(a, b) \leq \log(1+\varepsilon)$. Choose $\varepsilon > 0$ so that $\log(1+\varepsilon) < m_\tau$, forcing $(a, b) \notin K_\tau$ and thus $\min\{a, b\} < \tau$. ☐

We first record two elementary estimates used to force the "large-prime part" of Tao's factorisation to have non-negligible harmonic mass when $w > 2$.

Lemma 39 (A crude lower bound for σ_∞)**.** *For any set B of pairwise coprime integers ≥ 2,*

$$\sigma_\infty(B) = \prod_{q \in B} \Big(1 - \frac{1}{q}\Big) \geq \exp\Big(- \frac{3}{2}\mu(B)\Big).$$

Proof. For $q \geq 2$, using $\log(1-x) \geq -x - x^2$ for $0 \leq x \leq 1/2$ and $x = 1/q$ gives

$$\log\Big(1 - \frac{1}{q}\Big) \geq -\frac{1}{q} - \frac{1}{q^2} \geq -\frac{1}{q} - \frac{1}{2q} = -\frac{3}{2q}.$$

Summing over $q \in B$ and exponentiating yields the claim. $\square$

Lemma 40 (A crude power bound for Dickman)**.** *For all $u \geq 2$ one has $\rho(u) < u^{-3/2}$.*

Proof. Set $g(u) := u^{3/2}\rho(u)$. For $u > 1$ we have $u\rho'(u) = -\rho(u-1)$ and ρ is decreasing, so

$$g'(u) = \frac{3}{2}u^{1/2}\rho(u) + u^{3/2}\rho'(u) = u^{1/2}\Big(\frac{3}{2}\rho(u) - u\rho(u-1)\Big) \leq u^{1/2}\rho(u)\Big(\frac{3}{2} - u\Big) < 0$$

for all $u > 3/2$. Hence g is decreasing on $[2,\infty)$, so $g(u) \leq g(2) = 2^{3/2}\rho(2) = 2^{3/2}(1 - \log 2)$. Since $\log 2 > 2/3$ (because $e^{2/3} < 2$), we have $1 - \log 2 < 1/3$, and thus

$$g(2) < \frac{2^{3/2}}{3} < 1 \qquad (\text{as } 2^{3/2} = 2\sqrt{2} < 3).$$

Therefore $g(u) < 1$ for all $u \geq 2$, i.e. $\rho(u) < u^{-3/2}$. $\square$

Proposition 41 (Near-extremal reduction to primes)**.** *Fix $C > \log 2$ and write $w := e^C > 2$. Let $A = A_N \subseteq \{2, \dots, N\}$ be pairwise coprime with*

$$\mu(A) \leq C + o(1) \qquad \text{and} \qquad \sigma_N(A) \leq \rho(w) + o(1) \qquad (N \to \infty).$$

Then for every $\delta > 0$ and all sufficiently large N there exists a set of primes $P = P_{N,\delta} \subseteq \{p \leq N\}$ such that

$$\mu(A \triangle P) \leq \delta \qquad \text{and} \qquad \sigma_N(P) = \rho(w) + o(1),$$

and moreover one may arrange $P \subset (N^{1/u_\star}, N]$ for some fixed $u_\star = u_\star(w,\delta) > w$.

Proof. Fix $\delta > 0$. Since $\rho(e^t)$ is strictly decreasing for $t > \log 2$ and $\sigma_N(A) \geq \rho(e^{\mu(A)}) + o(1)$ (Tao's lower bound), the hypothesis $\sigma_N(A) \leq \rho(w) + o(1)$ forces $\mu(A) = C + o(1)$; we use this freely.

Step 1: choose sieve parameters depending on δ. Choose a small $\varepsilon = \varepsilon(\delta) > 0$ (to be fixed at the end). Choose an integer $r = r(w, \varepsilon)$ large enough that

$$\frac{C}{r} \leq \varepsilon, \qquad \sum_{j=r+1}^{\infty} \frac{(2C)^j}{j!} \leq \varepsilon, \qquad \frac{e^C}{r} \leq \varepsilon. \tag{9.10}$$

(Existence is clear as $r \to \infty$.) Define

$$u_\star := w\, r^{2r+2}, \qquad z_0 := N^{1/u_\star}, \qquad z_j := z_0^{r^{2j}} = N^{r^{2j}/u_\star} \quad (j = 0, 1, \ldots, r). \tag{9.11}$$

Then $z_0 \to \infty$ and $z_r = N^{r^{2r}/u_\star} = N^{1/(wr^2)} < N$.

Step 2: delete large composites above z_0. Let $A_{\mathrm{comp}} := \{q \in A : \ q > z_0 \text{ and } q \text{ composite}\}$. By Tao's Lemma 2.1 and dyadic decomposition one has $\mu(A_{\mathrm{comp}}) \ll 1/\sqrt{z_0} = o(1)$. By Lipschitz (Lemma 8), deleting A_{comp} changes $\sigma_N(\cdot)$ by at most $\mu(A_{\mathrm{comp}}) = o(1)$. Thus, after deleting A_{comp} and relabelling, we may assume:

All elements of A exceeding z_0 are prime.

Step 3: choose a sparse intermediate interval and split $A = A_1 \cup A_{\mathrm{mid}} \cup A_2$. By pigeonhole,

$$\sum_{j=1}^{r} \mu\big(A \cap (z_{j-1}, z_j]\big) \leq \mu(A) = C + o(1),$$

so for all sufficiently large N there exists $1 \leq j \leq r$ with

$$\mu\big(A \cap (z_{j-1}, z_j]\big) \leq \frac{C + o(1)}{r} \leq 2\varepsilon.$$

Fix such a j and set $z_- := z_{j-1}$, $z_+ := z_j$, and

$$A_{\mathrm{mid}} := A \cap (z_-, z_+], \qquad A_1 := A \cap [2, z_-], \qquad A_2 := A \cap (z_+, N].$$

Then $\mu(A_{\mathrm{mid}}) \leq 2\varepsilon$ and A_2 consists entirely of primes.

By Lipschitz again,

$$\sigma_N(A) = \sigma_N(A_1 \cup A_2) + O(\mu(A_{\mathrm{mid}})) = \sigma_N(A_1 \cup A_2) + O(\varepsilon), \tag{9.12}$$

and also

$$\mu(A_1) + \mu(A_2) = \mu(A) - \mu(A_{\mathrm{mid}}) = C + O(\varepsilon) + o(1). \tag{9.13}$$

Step 4: Tao's factorisation (with tracked errors). Recall the truncated Bonferroni approximants (Tao's notation)

$$\sigma_{N,r}(B) := \sum_{j=0}^{r} \sum_{\substack{q_1<\cdots<q_j\in B \\ q_1\cdots q_j\le N}} \frac{(-1)^j}{q_1\cdots q_j}.$$

Applying Tao's Bonferroni inequality [7, Lemma 2.3] to $A_1 \cup A_2$ gives

$$\sigma_N(A_1\cup A_2) = \sigma_{N,r}(A_1\cup A_2) + O\Big(\frac{C^{r+1}}{(r+1)!}\Big) + O\Big(\frac{re^C}{\log N}\Big). \tag{9.14}$$

By our choice of r and then taking N large, the right-hand error is $O(\varepsilon) + o(1)$.

Next, since $A_1 \subset (1, z_-]$ and $A_2 \subset (z_+, N]$ with A_2 prime, Tao's splitting lemma [7, Lemma 2.4] gives

$$\sigma_{N,r}(A_1\cup A_2) = \sigma_{N,r}(A_1)\sigma_{N,r}(A_2) + O\Big(\sum_{j=r+1}^{2r} \frac{C^j}{j!}\Big) + O\Big(r\frac{\log z_-}{\log z_+}e^C\Big). \tag{9.15}$$

Here $\frac{\log z_-}{\log z_+} = r^{-2}$ by construction, so the last term is $O(e^C/r) = O(\varepsilon)$, and the sum over j is $O(\varepsilon)$ by (9.10). Hence

$$\sigma_{N,r}(A_1\cup A_2) = \sigma_{N,r}(A_1)\sigma_{N,r}(A_2) + O(\varepsilon). \tag{9.16}$$

Finally, since $A_1 \subset [2, z_-]$ and $z_-^r \le N$ (which holds because $z_- \le z_{r-1} = N^{1/(wr^4)}$), Tao's pure Brun sieve [7, Lemma 2.5] gives

$$\sigma_{N,r}(A_1) = \sigma_\infty(A_1) + O\Big(\sum_{j=r+1}^{\infty} \frac{C^j}{j!}\Big) = \sigma_\infty(A_1) + O(\varepsilon). \tag{9.17}$$

Similarly, applying [7, Lemma 2.3] to A_2 gives

$$\sigma_{N,r}(A_2) = \sigma_N(A_2) + O(\varepsilon) + o(1). \tag{9.18}$$

Combining (9.14), (9.16), (9.17), and (9.18) yields

$$\sigma_N(A_1\cup A_2) = \sigma_\infty(A_1)\sigma_N(A_2) + O(\varepsilon) + o(1). \tag{9.19}$$

Together with (9.12), we conclude

$$\sigma_N(A) = \sigma_\infty(A_1)\sigma_N(A_2) + O(\varepsilon) + o(1). \tag{9.20}$$

Step 5: $\mu(A_2)$ is bounded away from 0 (since $w > 2$). Since $\mu(A_1) \le \mu(A) = C + o(1)$, Lemma 39 gives $\sigma_\infty(A_1) \ge w^{-3/2} + o(1)$. If $\mu(A_2) \le c$ then $\sigma_N(A_2) \ge 1 - \mu(A_2) \ge 1 - c$ by the union bound. Inserting into (9.20) gives

$$\sigma_N(A) \;\ge\; (w^{-3/2} + o(1))(1 - c) - O(\varepsilon).$$

Since $\rho(w) < w^{-3/2}$ by Lemma 40, we may choose $c_0 = c_0(w) > 0$ and then $\varepsilon > 0$ small enough (in terms of w) so that the right-hand side is $> \rho(w) + \delta$ for all large N, contradicting $\sigma_N(A) \le \rho(w) + o(1)$. Thus

$$\mu(A_2) \ge c_0(w) \qquad \text{for all sufficiently large } N. \tag{9.21}$$

Step 6: strictness forces $\mu(A_1)$ small. Because A_2 is a prime set with $\mu(A_2) = O(1)$, the prime-moduli theorem (Hildebrand/Tao Theorem 1.2) gives

$$\sigma_N(A_2) \ge \rho(e^{\mu(A_2)}) + o(1). \tag{9.22}$$

Also, for $q \ge 2$ we have $1 - \frac{1}{q} = \rho(e^{1/q})$, so

$$\sigma_\infty(A_1) = \prod_{q \in A_1} \rho(e^{1/q}) \;\ge\; \rho\Big(\prod_{q \in A_1} e^{1/q}\Big) = \rho(e^{\mu(A_1)})$$

by log-concavity of ρ. Therefore (9.20) and (9.22) imply

$$\sigma_N(A) \;\ge\; \rho(e^{\mu(A_1)})\,\rho(e^{\mu(A_2)}) \;-\; O(\varepsilon) \;+\; o(1).$$

Since $\sigma_N(A) \le \rho(w) + o(1)$ and $\mu(A_1) + \mu(A_2) = C + O(\varepsilon) + o(1)$ by (9.13), we obtain

$$\frac{\rho(e^{\mu(A_1)})\,\rho(e^{\mu(A_2)})}{\rho(e^{\mu(A_1)+\mu(A_2)})} \;\le\; 1 + O(\varepsilon) + o(1).$$

Fix a compactness bound $C_0 := C + 1$ and apply Lemma 38: for any $\tau > 0$, if ε is small enough (depending on C_0, τ) and N large enough, then

$$\min\{\mu(A_1), \mu(A_2)\} \le \tau.$$

Using (9.21), we conclude $\mu(A_1) \le \tau$ for all large N. Choose $\tau = \varepsilon$ (say) and then choose $\varepsilon = \varepsilon(\delta)$ small enough so that

$$\mu(A_1) \le \delta/3 \qquad \text{for all sufficiently large } N. \tag{9.23}$$

Step 7: define the prime set P and conclude. Set $P := A_2$, which is a prime set and satisfies $P \subset (z_+, N] \subset (z_0, N] = (N^{1/u_\star}, N]$. Moreover,

$$\mu(A \triangle P) \;\le\; \mu(A_1) + \mu(A_{\mathrm{mid}}) + \mu(A_{\mathrm{comp}}) \;\le\; \delta/3 + 2\varepsilon + o(1) \;\le\; \delta$$

for all sufficiently large N, by (9.23), $\mu(A_{\mathrm{mid}}) \le 2\varepsilon$, and $\mu(A_{\mathrm{comp}}) = o(1)$, provided $\varepsilon(\delta) \le \delta/10$.

Finally, $\mu(P) = \mu(A_2) = C + o(1)$ by (9.13) and $\mu(A_1), \mu(A_{\mathrm{mid}}) = O(\varepsilon)$. Thus by the prime-moduli theorem (9.22),

$$\sigma_N(P) = \sigma_N(A_2) \ge \rho(e^{\mu(P)}) + o(1) = \rho(w) + o(1),$$

while (9.20) and $\sigma_\infty(A_1) \le 1$ give $\sigma_N(P) \le \sigma_N(A) + O(\varepsilon) + o(1) = \rho(w) + o(1)$. Hence $\sigma_N(P) = \rho(w) + o(1)$ as claimed. ☐

9.5 Coprime tail rigidity

We can now complete the proof of the tail rigidity statement.

Theorem 42 (Coprime tail rigidity)**.** *Fix $C > \log 2$ and set $w = e^C > 2$. Let $A \subseteq \{2, \dots, N\}$ be pairwise coprime with*

$$\mu(A) \le C + o(1) \qquad \text{and} \qquad \sigma_N(A) \le \rho(w) + o(1).$$

Then for every $\delta > 0$ there exists a threshold $Y = N^{1/w+o(1)}$ such that

$$\sum_{\substack{p \le Y \\ p \text{ prime} \\ p \in A}} \frac{1}{p} + \sum_{\substack{Y < p \le N \\ p \text{ prime} \\ p \notin A}} \frac{1}{p} + \sum_{\substack{q \in A \\ q \text{ composite}}} \frac{1}{q} \le \delta. \tag{9.24}$$

Equivalently, $d_\mu\big(A, \mathcal{P}(Y, N]\big) \le \delta$.

Proof. Apply Proposition 41 to obtain a prime set $P \subseteq (N^{1/u_\star}, N]$ (for some fixed $u_\star = u_\star(w) > w$) such that

$$\mu(A \triangle P) = o(1), \qquad \mu(P) = \log w + o(1), \qquad \sigma_N(P) = \rho(w) + o(1).$$

Now set $x := N$ and $y := N^{1/u_\star}$, so $x = y^{u_\star}$ with $y \to \infty$. Since $P \subset (y, x]$, Theorem 37 applies to P (with this fixed $u_\star$), and gives: for the given $\delta > 0$ there exists $Y = x^{1/w+o(1)} = N^{1/w+o(1)}$ such that

$$\mu\big(P \triangle \mathcal{P}(Y, N]\big) = \sum_{\substack{p \le Y \\ p \in P}} \frac{1}{p} + \sum_{\substack{Y < p \le N \\ p \notin P}} \frac{1}{p} \le \delta/2$$

for all sufficiently large N (say, after adjusting the $o(1)$ and shrinking $\delta/2$ if needed).

Finally, by the triangle inequality for d_μ,

$$d_\mu\big(A, \mathcal{P}(Y, N]\big) \le d_\mu(A, P) + d_\mu\big(P, \mathcal{P}(Y, N]\big) = \mu(A \triangle P) + \mu\big(P \triangle \mathcal{P}(Y, N]\big) \le \delta$$

for all large N. Expanding $d_\mu\big(A, \mathcal{P}(Y, N]\big)$ gives (9.24). ☐

Remark 43. The conclusion (9.24) is a "harmonic" rigidity statement: it controls the symmetric difference between A and a prime tail in the $\sum 1/p$ metric. It does *not* in general control the *counting* symmetric difference $|A\Delta\mathcal{P}(Y,N]|$, since primes near N have very small reciprocal weights.

9.6 Cardinality of the canonical prime-tail extremiser

For the specific extremal model set $A_N^\star(C) := \{p \le N : \ p \text{ prime and } p > N^{1/w}\}$, one can record its size explicitly from the prime number theorem.

Corollary 44 (Size of the prime tail)**.** *Fix $C > 0$ and set $w = e^C$. Let*

$$A_N^\star(C) = \{p \le N : \ p \textit{ prime and } p > N^{1/w}\}.$$

Then

$$|A_N^\star(C)| = \pi(N) - \pi(N^{1/w}) = \Big(1 + o(1)\Big)\frac{N}{\log N} \qquad (N \to \infty).$$

Proof. By the prime number theorem, $\pi(N) \sim N/\log N$ and

$$\pi(N^{1/w}) \sim \frac{N^{1/w}}{(1/w)\log N} = \frac{w\,N^{1/w}}{\log N} = o\Big(\frac{N}{\log N}\Big),$$

since $N^{1/w} = o(N)$. The claim follows. $\square$

References

[1] T. F. Bloom (ed.), *Erdős Problems, Problem #783* (online database page), `https://www.erdosproblems.com/783` (accessed 2026-02-23).

[2] P. Chojecki, *A Buchstab–Dickman Note and an Attempted Port to an Extremal Coprime Sieve Problem*, `https://www.ulam.ai/research/erdos783.pdf`.

[3] P. Erdős, *Problems and results on combinatorial number theory*, in *A survey of combinatorial theory* (Proc. Internat. Sympos., Colorado State Univ., Fort Collins, Colo., 1971), North-Holland, Amsterdam, 1973, pp. 117–138.

[4] P. Erdős and I. Z. Ruzsa, *On the small sieve. I. Sifting by primes*, J. Number Theory **12** (1980), no. 4, 385–394.

[5] A. Granville and K. Soundararajan, *The number of unsieved integers up to x*, Acta Arith. **115** (2004), no. 4, 305–328. (Preprint: `https://arxiv.org/abs/math/0308009`.)

[6] A. Hildebrand, *Quantitative mean value theorems for nonnegative multiplicative functions. II*, Acta Arith. **49** (1987), no. 3, 209–260.

[7] T. Tao, *Sieving by coprime numbers* (Feb. 22, 2026), `https://terrytao.wordpress.com/wp-content/uploads/2026/02/erdos783-3.pdf`.

[8] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, 3rd ed., Graduate Studies in Mathematics 163, AMS, 2015.