A Proposed Complete Solution to Erdős Problem 536

Shouqiao Wang

Columbia University   Multiscalar Intelligence

Abstract

Let $f(N)$ be the largest size of a set $A \subseteq \{1, \ldots, N\}$ containing no three distinct integers whose three pairwise least common multiples are equal. We prove that $f(N) = o(N)$, answering a question of Erdős. The proof first reduces positive density to a weighted finite-prime extremal problem and then deletes all prime exponents equal to one. The resulting squarefree problem is controlled by balanced pair-product cubes and the cap-set theorem. A five-state prime-band model, an exact factorial insertion identity, and matching first- and second-moment estimates produce cube laws whose word marginals are asymptotically flat. The only external mathematical inputs are standard prime-number estimates, the Brun–Titchmarsh inequality, and the polynomial-method cap-set bound. Finite algebraic and combinatorial checks are reproduced by a companion Python verifier.

1 Introduction

Call a set of positive integers *safe* if it contains no three distinct members a, b, c satisfying

$$\mathrm{lcm}(a, b) = \mathrm{lcm}(a, c) = \mathrm{lcm}(b, c). \tag{1.1}$$

Let $f(N)$ be the maximum size of a safe subset of $[N] = \{1, \ldots, N\}$. Erdős asked whether every set of positive upper density must contain a triple of the form (1.1) [4, p. 646]; see also the modern formulation [2]. Abbott and Gardner gave an early lower-bound construction for the associated extremal problem [1].

Our main result is the following.

Theorem 1.1. *As $N \to \infty$,*

$$f(N) = o(N).$$

Here is the proof strategy. Equal pairwise least common multiples have an exact pair-product form

$$txy, \qquad txz, \qquad tyz,$$

with x, y, z pairwise coprime. This simple observation is used at four different scales. The logical chain is

$$\begin{aligned} &\text{positive density} \Longrightarrow \text{a finite-prime envelope} \Longrightarrow \text{a squarefree moving-prefix capacity} \\ &\qquad\qquad \Longrightarrow \text{balanced pair-product cubes} \Longrightarrow \text{a cap-set saving.} \end{aligned} \tag{1.2}$$

The first two arrows are deterministic. The difficult point is to build balanced cubes whose individual word marginals are close to the relevant product measure even though the extremizing prefix is allowed to move.

For that purpose we use many disjoint prime bands. On one band, a five-state coupling creates a common part and three disjoint petals. If B_T denotes the event that the three petal weights nearly agree and all four active labels satisfy lower prefix profiles, then

$$\mathbb{P}(B_T) \gg w^2, \qquad \mathbb{E}[\mathbb{P}(B_T \mid S)^2] \ll w^4.$$

The first estimate is proved by a whole-band Poisson coupling and two adaptive anchors. The second is a fully discrete calculation: two canonical root pivots give two factors of w, and two missing-petal pivots give two more. Consequently the conditioned root density has a uniform L^2 bound. Averaging the active band among many alternatives converts this L^2 control into the L^1 flatness needed by the moving-prefix argument.

The paper is organized in dependency order. Section 2 proves the elementary pair-product description and records the external inputs. Section 3 reduces the theorem to a normalized squarefree capacity. Section 4 proves the balanced-cube transference principle. Section 5 constructs one prime-band coordinate and proves its first-moment lower bound. Section 6 proves the exact pivot estimates and the matching second-moment upper bound. Section 7 flattens the word marginals by alternative bands, and Section 8 places those bands and completes the proof. Appendix A describes the companion verifier and its precise scope.

2 Structure and external inputs

2.1 The pair-product form

Lemma 2.1 (Pair-product form)**.** *Three distinct positive integers a, b, c satisfy (1.1) if and only if, after possibly permuting them, there are positive integers t, x, y, z such that*

$$a = txy, \qquad b = txz, \qquad c = tyz, \tag{2.1}$$

and x, y, z are pairwise coprime.

Proof. If (2.1) holds, pairwise coprimality gives

$$\operatorname{lcm}(a, b) = \operatorname{lcm}(a, c) = \operatorname{lcm}(b, c) = txyz$$

prime by prime.

Conversely, let L be the common value in (1.1). Fix a prime p and write

$$A = v_p(a), \qquad B = v_p(b), \qquad C = v_p(c), \qquad M = v_p(L).$$

The maximum of each pair among A, B, C is M. Therefore at least two of A, B, C equal M. Define

$$x = \frac{L}{c}, \qquad y = \frac{L}{b}, \qquad z = \frac{L}{a}.$$

If, for example, p divided both x and y, then $C < M$ and $B < M$, which would contradict $\max(B, C) = M$. The same argument for the other two pairs shows that x, y, z are pairwise coprime.

Finally put $t = abc/L^2$. At the prime p its exponent is

$$A + B + C - 2M \geq 0,$$

because at least two of A, B, C equal M. Hence t is an integer. A direct valuation calculation now gives

$$v_p(txy) = A, \qquad v_p(txz) = B, \qquad v_p(tyz) = C.$$

Since this holds for every prime, (2.1) follows. □

For squarefree integers, least common multiple corresponds to union of prime supports. Thus Lemma 2.1 is the bridge between the integer problem and the set systems used below.

2.2 Prime-number estimates

We use the following standard consequences of the classical zero-free region and the Brun–Titchmarsh inequality. The quoted forms follow by partial summation from the prime number theorem with de la Vallée Poussin error; see [6, Chapters 2 and 6] and [5, Chapters 5 and 6]. The interval estimate is the classical Brun–Titchmarsh theorem; see [7].

Proposition 2.2 (Analytic prime inputs)**.** *There are constants $B_0, B_1 \in \mathbb{R}$ and $c > 0$ such that, for $x \geq 3$,*

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + B_0 + O\left(\mathrm{e}^{-c\sqrt{\log x}}\right), \tag{2.2}$$

$$\sum_{p \leq x} \frac{1}{p+1} = \log\log x + B_1 + O\left(\mathrm{e}^{-c\sqrt{\log x}}\right), \tag{2.3}$$

$$\sum_{p \leq x} \frac{\log p}{p-1} = (1 + o(1)) \log x. \tag{2.4}$$

Moreover, whenever $2 \leq h \leq x$,

$$\pi(x+h) - \pi(x) \leq \frac{2h}{\log h}. \tag{2.5}$$

If $h > x$, the ordinary prime number theorem gives the weaker estimates needed below.

The replacement of $1/p$ by $1/(p+1)$ in (2.3) changes only the constant, because $\sum_p p^{-2} < \infty$ and its tail is $O(1/x)$.

2.3 The cap-set estimate

Proposition 2.3 (Cap-set bound)**.** *Let $A \subseteq \mathbb{F}_3^H$ contain no nonconstant affine line. Then*

$$|A| \leq 3\kappa_{\mathrm{cap}}^H 3^H, \qquad \kappa_{\mathrm{cap}} := \frac{7}{12} 2^{2/3} < 1. \tag{2.6}$$

Proof. The polynomial-method bound of Ellenberg and Gijswijt [3] gives

$$|A| \leq 3 \min_{0<t<1} \left((1 + t + t^2) t^{-2/3}\right)^H.$$

Taking $t = 1/2$ gives $(1 + t + t^2)t^{-2/3} = 3\kappa_{\mathrm{cap}}$. Also

$$\kappa_{\mathrm{cap}}^3 = 4 \left(\frac{7}{12}\right)^3 = \frac{343}{432} < 1,$$

which proves the displayed strict inequality. The same exact calculation is reproduced by the companion verifier. □

3 Reduction to a squarefree moving-prefix capacity

3.1 The finite-prime envelope

Let P be a finite set of primes. Write $\mathcal{M}_P$ for the multiplicative monoid generated by P, and put

$$\delta_P := \prod_{p \in P} \left(1 - \frac{1}{p}\right).$$

Let $b_P(T)$ be the largest size of a safe subset of $\mathcal{M}_P \cap [1, T]$, and define

$$C(P) := \delta_P \int_1^\infty b_P(T) \frac{\mathrm{d}T}{T^2}. \tag{3.1}$$

This integral converges. Indeed, a member of $\mathcal{M}_P \cap [1, T]$ is determined by $|P|$ nonnegative exponents, each $O_P(1 + \log T)$, and hence

$$b_P(T) \le |\mathcal{M}_P \cap [1, T]| = O_P((1 + \log T)^{|P|}). \tag{3.2}$$

Proposition 3.1 (Finite-prime envelope)**.** *For every fixed finite set P of primes,*

$$\limsup_{N\to\infty} \frac{f(N)}{N} \le C(P). \tag{3.3}$$

Proof. Fix P throughout the proof. Every positive integer has a unique factorization $n = mq$ with $q \in \mathcal{M}_P$ and $(m, \prod_{p\in P} p) = 1$. If a safe subset of $[N]$ is restricted to one fixed m-fibre, the resulting set of q's is safe: here m is coprime to every q, so

$$\operatorname{lcm}(mq_i, mq_j) = m \operatorname{lcm}(q_i, q_j).$$

It follows that

$$f(N) \le \sum_{\substack{m\le N \\ (m,\prod_{p\in P} p)=1}} b_P(N/m). \tag{3.4}$$

Finite inclusion–exclusion gives

$$A_P(x) := \#\{m \le x : (m, \textstyle\prod_{p\in P} p) = 1\} = \delta_P x + E_P(x), \qquad E_P(x) = O_P(1).$$

The function $t \mapsto b_P(N/t)$ is nonincreasing, and its total variation on $[1, N]$ is at most $b_P(N)$. Stieltjes summation therefore gives

$$\begin{aligned} \sum_{\substack{m\le N \\ (m,\prod_{p\in P} p)=1}} b_P(N/m) &= \int_{1^-}^N b_P(N/t)\ \mathrm{d}A_P(t) \\ &= \delta_P \int_1^N b_P(N/t)\ \mathrm{d}t + O_P(b_P(N)). \end{aligned} \tag{3.5}$$

For completeness, the error term follows by integration by parts: the two endpoint terms are $O_P(b_P(N))$, and the remaining integral is bounded by $\|E_P\|_\infty$ times the total variation of $b_P(N/t)$.

In the main integral set $u = N/t$. Then

$$\delta_P \int_1^N b_P(N/t)\ \mathrm{d}t = \delta_P N \int_1^N b_P(u) \frac{\mathrm{d}u}{u^2}. \tag{3.6}$$

By (3.2), $b_P(N) = o(N)$ for fixed P. Divide (3.4)–(3.6) by N and let $N \to \infty$ to obtain (3.3). □

3.2 Deleting all unit exponents

For a finite prime set $\mathcal{R}$ and $S \subseteq \mathcal{R}$, write

$$d(S) := \prod_{p\in S} p, \qquad Z_\mathcal{R} := \prod_{p\in\mathcal{R}} (1 + p^{-1}), \qquad \mu_\mathcal{R}(S) := \frac{1}{Z_\mathcal{R} d(S)}.$$

Thus $\mu_\mathcal{R}$ is the product law in which p is selected with probability $1/(p + 1)$.

A family $\mathcal{H} \subseteq 2^{\mathcal{R}}$ is *admissible* if it contains no three distinct members S_1, S_2, S_3 for which

$$S_1 \cup S_2 = S_1 \cup S_3 = S_2 \cup S_3. \tag{3.7}$$

Let $A^{\mathrm{sf}}_{\mathcal{R}}(x)$ be the largest size of an admissible subfamily of

$$\{S \subseteq \mathcal{R} : d(S) \le x\},$$

and adopt the convention

$$A^{\mathrm{sf}}_{\mathcal{R}}(x) := 0 \qquad (0 < x < 1). \tag{3.8}$$

Thus $A^{\mathrm{sf}}_{\mathcal{R}}$ is defined for every $x > 0$. We then define

$$I^{\mathrm{sf}}_{\mathcal{R}} := \int_1^\infty A^{\mathrm{sf}}_{\mathcal{R}}(x) \frac{\mathrm{d}x}{x^2}. \tag{3.9}$$

The elementary bound

$$0 \le \frac{I^{\mathrm{sf}}_{\mathcal{R}}}{Z_{\mathcal{R}}} \le 1 \tag{3.10}$$

follows from

$$I^{\mathrm{sf}}_{\mathcal{R}} \le \sum_{S \subseteq \mathcal{R}} \int_{d(S)}^\infty \frac{\mathrm{d}x}{x^2} = \sum_{S \subseteq \mathcal{R}} \frac{1}{d(S)} = Z_{\mathcal{R}}.$$

For $E \subseteq P$ and $m \in \mathcal{M}_P$, set

$$U_E(m) := \{p \in E : v_p(m) = 1\}.$$

Proposition 3.2 (All-unit deletion)**.** *For finite prime sets $E \subseteq P$,*

$$C(P) \le \delta_P \sum_{\substack{m \in \mathcal{M}_P \\ U_E(m) = \varnothing}} \frac{1}{m} I^{\mathrm{sf}}_{E \setminus \mathrm{supp}(m)}. \tag{3.11}$$

Moreover,

$$C(P) \le \frac{I^{\mathrm{sf}}_E}{Z_E} + \sum_{p \in E} \frac{1}{p^2}. \tag{3.12}$$

Proof. Every $q \in \mathcal{M}_P$ has a unique decomposition

$$q = m\, d(S), \qquad S = U_E(q), \qquad U_E(m) = \varnothing, \qquad S \subseteq E \setminus \mathrm{supp}(m). \tag{3.13}$$

Indeed, $d(S)$ removes exactly one copy of each prime whose exponent in q is one. Hence every prime of E has exponent either zero or at least two in m.

Fix m in (3.13) and consider a safe family in $\mathcal{M}_P \cap [1, T]$. Its m-section is admissible. Otherwise three supports satisfying (3.7) would give

$$\mathrm{lcm}(md(S_i), md(S_j)) = m\, d(S_i \cup S_j)$$

for each pair, contradicting safety. Therefore

$$b_P(T) \le \sum_{\substack{m \in \mathcal{M}_P \\ U_E(m) = \varnothing}} A^{\mathrm{sf}}_{E \setminus \mathrm{supp}(m)}(T/m).$$

Insert this inequality into (3.1). Positivity allows termwise integration. For each eligible m, the substitution $T = mx$ and the convention (3.8) give

$$\int_1^\infty A^{\mathrm{sf}}_{E \setminus \mathrm{supp}(m)}(T/m) \frac{\mathrm{d}T}{T^2} = \frac{1}{m} \int_{1/m}^\infty A^{\mathrm{sf}}_{E \setminus \mathrm{supp}(m)}(x) \frac{\mathrm{d}x}{x^2}$$

$$= \frac{1}{m} \int_1^\infty A^{\mathrm{sf}}_{E \setminus \mathrm{supp}(m)}(x) \frac{\mathrm{d}x}{x^2} = \frac{1}{m} I^{\mathrm{sf}}_{E \setminus \mathrm{supp}(m)}.$$

This proves (3.11).

It remains to normalize the right side. For an eligible m, put $R = E \setminus \mathrm{supp}(m)$ and define

$$\mathcal{W}(m) := \delta_P \frac{Z_R}{m}. \tag{3.14}$$

We verify prime by prime that $\mathcal{W}$ is a probability law. If $p \notin E$, summing all valuations gives

$$(1 - p^{-1}) \sum_{k \geq 0} p^{-k} = 1.$$

If $p \in E$, valuation zero contributes

$$(1 - p^{-1})(1 + p^{-1}) = 1 - p^{-2},$$

while all permitted positive valuations, namely $k \geq 2$, contribute

$$(1 - p^{-1}) \sum_{k \geq 2} p^{-k} = p^{-2}.$$

Thus the local masses sum to one. They also show that the events $\{p \notin R\}$ are independent and satisfy

$$\mathbb{P}_{\mathcal{W}}(p \notin R) = p^{-2} \qquad (p \in E). \tag{3.15}$$

The right side of (3.11) is exactly

$$\mathbb{E}_{\mathcal{W}} \left[\frac{I^{\mathrm{sf}}_R}{Z_R} \right].$$

On $\{R = E\}$ the random ratio equals I^{sf}_E / Z_E; off that event it lies in $[0, 1]$ by (3.10). Hence

$$C(P) \leq \frac{I^{\mathrm{sf}}_E}{Z_E} + \mathbb{P}_{\mathcal{W}}(R \neq E) \leq \frac{I^{\mathrm{sf}}_E}{Z_E} + \sum_{p \in E} p^{-2},$$

which is (3.12). □

Corollary 3.3 (The remaining target)**.** *Let $P_y = \{p : p \leq y\}$. Suppose that $E_y \subseteq P_y$ satisfies*

$$\min E_y \longrightarrow \infty, \qquad \frac{I^{\mathrm{sf}}_{E_y}}{Z_{E_y}} \longrightarrow 0. \tag{3.16}$$

Then $f(N) = o(N)$.

Proof. Apply Proposition 3.2 with $P = P_y$ and $E = E_y$. Since $\min E_y \to \infty$,

$$\sum_{p \in E_y} p^{-2} \leq \sum_{n \geq \min E_y} n^{-2} \longrightarrow 0.$$

Thus $C(P_y) \to 0$. Proposition 3.1 gives, for each fixed y,

$$\limsup_{N \to \infty} \frac{f(N)}{N} \leq C(P_y).$$

Letting $y \to \infty$ proves the corollary. □

4 Balanced cubes and moving prefixes

4.1 Pair-product cubes

A dimension-H *pair-product cube* consists of a common prime support and, for each coordinate $1 \le i \le H$, three nonempty prime supports x_i, y_i, z_i. All $3H$ petals are mutually disjoint, and the common support is disjoint from every petal. A word $\omega = (\omega_1, \ldots, \omega_H) \in \mathbb{F}_3^H$ takes one of the following three contributions in coordinate i:

state ω_i	chosen petals	omitted petal
0	$y_i \cup z_i$	x_i
1	$x_i \cup z_i$	y_i
2	$x_i \cup y_i$	z_i

The support of the word is the union of its H coordinate contributions and the common support.

Lemma 4.1 (Cap-set saving on one cube)**.** *An admissible family occupies at most a proportion $3\kappa_{\mathrm{cap}}^H$ of the 3^H words of a pair-product cube.*

Proof. The word map is injective: if two words first differ in coordinate i, then one of the nonempty petals x_i, y_i, z_i occurs in exactly one of the two supports.

Consider a nonconstant affine line

$$\omega, \qquad \omega + v, \qquad \omega + 2v \quad \text{in } \mathbb{F}_3^H.$$

In a coordinate where $v_i = 0$, all three words make the same choice. In a coordinate where $v_i \neq 0$, the three states are $0, 1, 2$ in some order. The pairwise union of any two corresponding coordinate contributions is then $x_i \cup y_i \cup z_i$. It follows coordinate by coordinate that the three full supports have identical pairwise unions. Consequently the preimage of an admissible family contains no nonconstant affine line. Proposition 2.3 bounds that preimage by $3\kappa_{\mathrm{cap}}^H 3^H$, proving the lemma. □

4.2 A joint law for a moving prefix

The extremal family in the definition of $A_{\mathcal{R}}^{\mathrm{sf}}(x)$ depends on x. It is therefore not enough to construct a good cube at one fixed cutoff. The next proposition couples the cube and the cutoff so that all words are tested against the same moving prefix.

Proposition 4.2 (Joint-prefix transference)**.** *Let $\mathcal{R}$ be finite, and let λ be a probability law on dimension-H pair-product cubes in $2^{\mathcal{R}}$. For a word ω, let ν_ω be its support marginal. Suppose that*

$$\frac{\mathrm{d}\nu_\omega}{\mathrm{d}\mu_{\mathcal{R}}} = G_\omega, \qquad \mathbb{E}_{\mu_{\mathcal{R}}} |G_\omega - 1| \le \varepsilon \tag{4.1}$$

for every $\omega \in \mathbb{F}_3^H$. Suppose also that $0 \le \zeta \le 1$ and, for every sampled cube c and every word ω,

$$|\log d(S_\omega(c)) - \log \mathcal{G}(c)| \le \zeta, \qquad \mathcal{G}(c) := \left(\prod_{\tau \in \mathbb{F}_3^H} d(S_\tau(c)) \right)^{1/3^H}. \tag{4.2}$$

Then

$$\frac{I_{\mathcal{R}}^{\mathrm{sf}}}{Z_{\mathcal{R}}} \le 3\kappa_{\mathrm{cap}}^H + O(\varepsilon + \zeta), \tag{4.3}$$

where the implied constant is absolute.

Proof. We separate the construction into four steps.
Step 1: a joint cube–cutoff measure. For a cube c, put

$$D(c) := \max_{\omega\in\mathbb{F}_3^H} d(S_\omega(c)).$$

Define a finite measure on the cube space times $[1,\infty)$ by

$$\mathrm{d}\Lambda(c,x) := \mathcal{G}(c)\,\mathbf{1}_{\{x\geq D(c)\}}\frac{\mathrm{d}x}{x^2}\;\mathrm{d}\lambda(c). \tag{4.4}$$

Since $D(c)\geq\mathcal{G}(c)$, while (4.2) gives $D(c)\leq \mathrm{e}^{\zeta}\mathcal{G}(c)$,

$$\mathrm{e}^{-\zeta}\leq\|\Lambda\|=\mathbb{E}_\lambda\frac{\mathcal{G}(c)}{D(c)}\leq 1. \tag{4.5}$$

Step 2: the weighted mass of one fibre. Fix a word ω and a support $S\subseteq\mathcal{R}$. Whenever $S_\omega(c)=S$, the balance assumption implies

$$\mathrm{e}^{-\zeta}d(S)\leq\mathcal{G}(c)\leq\mathrm{e}^{\zeta}d(S),\qquad D(c)\leq\mathrm{e}^{2\zeta}d(S). \tag{4.6}$$

The second inequality follows because every word product is at most $\mathrm{e}^{\zeta}\mathcal{G}(c)$ and $\mathcal{G}(c)\leq\mathrm{e}^{\zeta}d(S)$. By (4.1),

$$\nu_\omega(S)=\frac{G_\omega(S)}{Z_{\mathcal{R}}d(S)}.$$

Consequently the weighted fibre mass

$$M_\omega(S) := \int_{\{c:S_\omega(c)=S\}}\mathcal{G}(c)\;\mathrm{d}\lambda(c)$$

satisfies the pointwise bounds

$$\frac{\mathrm{e}^{-\zeta}G_\omega(S)}{Z_{\mathcal{R}}}\leq M_\omega(S)\leq\frac{\mathrm{e}^{\zeta}G_\omega(S)}{Z_{\mathcal{R}}}. \tag{4.7}$$

Step 3: comparison with the canonical prefix law. We use the following convention. If σ is a finite signed measure, then $\|\sigma\|_{\mathrm{var}}:=|\sigma|(\Omega)$ denotes its total variation norm. For probability measures P,Q we write

$$d_{\mathrm{TV}}(P,Q):=\sup_E|P(E)-Q(E)|=\frac{1}{2}\|P-Q\|_{\mathrm{var}}.$$

Let Λ_ω be the pushforward of Λ under $(c,x)\mapsto(S_\omega(c),x)$. With respect to

$$\rho:=\sum_{S\subseteq\mathcal{R}}\delta_S\otimes\frac{\mathrm{d}x}{x^2}\quad\text{on }2^{\mathcal{R}}\times[1,\infty),$$

its density is

$$L_\omega(S,x):=\int_{\{c:S_\omega(c)=S\}}\mathcal{G}(c)\mathbf{1}_{\{D(c)\leq x\}}\;\mathrm{d}\lambda(c).$$

Let Π be the probability measure whose ρ-density is

$$\frac{\mathrm{d}\Pi}{\mathrm{d}\rho}(S,x):=\frac{1}{Z_{\mathcal{R}}}\mathbf{1}_{\{x\geq d(S)\}}. \tag{4.8}$$

Indeed,

$$\Pi(2^{\mathcal{R}}\times[1,\infty))=\frac{1}{Z_{\mathcal{R}}}\sum_{S\subseteq\mathcal{R}}\frac{1}{d(S)}=1.$$

Fix $S \subseteq \mathcal{R}$ and abbreviate $G = G_\omega(S)$ and $M = M_\omega(S)$. If $G > 0$, (4.7) permits us to write

$$M = \frac{\theta_S G}{Z_{\mathcal{R}}}, \qquad \mathrm{e}^{-\zeta} \le \theta_S \le \mathrm{e}^{\zeta}.$$

If $G = 0$, then (4.7) gives $M = 0$, and the same formula holds after setting $\theta_S = 1$. Since $0 \le \zeta \le 1$,

$$|\theta_S - 1| \le \mathrm{e}^{\zeta} - 1 \le 2\zeta.$$

On the fibre $S_\omega(c) = S$, one has $D(c) \ge d(S)$ and, by (4.6), $D(c) \le \mathrm{e}^{2\zeta} d(S)$. Consequently $L_\omega(S, x) = 0$ for $x < d(S)$ and $L_\omega(S, x) = M_\omega(S)$ for $x \ge \mathrm{e}^{2\zeta} d(S)$. On the latter region,

$$\begin{aligned}
&\sum_{S \subseteq \mathcal{R}} \int_{\mathrm{e}^{2\zeta} d(S)}^{\infty} \left| L_\omega(S, x) - \frac{1}{Z_{\mathcal{R}}} \right| \frac{\mathrm{d}x}{x^2} \\
&= \mathrm{e}^{-2\zeta} \sum_{S \subseteq \mathcal{R}} \frac{|M_\omega(S) - Z_{\mathcal{R}}^{-1}|}{d(S)} \\
&\le \sum_{S \subseteq \mathcal{R}} \frac{|G_\omega(S) - 1| + 2\zeta G_\omega(S)}{Z_{\mathcal{R}} d(S)} \le \varepsilon + 2\zeta.
\end{aligned} \tag{4.9}$$

Here the last line uses $\mathbb{E}_{\mu_{\mathcal{R}}} |G_\omega - 1| \le \varepsilon$ and $\mathbb{E}_{\mu_{\mathcal{R}}} G_\omega = 1$.

It remains to control the transition strip

$$\mathcal{T} := \{(S, x) : d(S) \le x < \mathrm{e}^{2\zeta} d(S)\}.$$

Its Π-mass is exactly

$$\Pi(\mathcal{T}) = (1 - \mathrm{e}^{-2\zeta}) \sum_{S \subseteq \mathcal{R}} \frac{1}{Z_{\mathcal{R}} d(S)} = 1 - \mathrm{e}^{-2\zeta} \le 2\zeta. \tag{4.10}$$

For the other marginal, Tonelli's theorem and nonnegativity give

$$\begin{aligned}
\Lambda_\omega(\mathcal{T}) &= \int \mathcal{G}(c) \int_{d(S_\omega(c))}^{\mathrm{e}^{2\zeta} d(S_\omega(c))} \mathbf{1}_{\{D(c) \le x\}} \frac{\mathrm{d}x}{x^2} \, \mathrm{d}\lambda(c) \\
&\le (1 - \mathrm{e}^{-2\zeta}) \int \frac{\mathcal{G}(c)}{d(S_\omega(c))} \, \mathrm{d}\lambda(c) \\
&\le \mathrm{e}^{\zeta} (1 - \mathrm{e}^{-2\zeta}) \le 6\zeta.
\end{aligned} \tag{4.11}$$

The penultimate inequality follows from (4.2); the last uses $\mathrm{e}^{\zeta} \le \mathrm{e} < 3$ and $1 - \mathrm{e}^{-2\zeta} \le 2\zeta$. On $\mathcal{T}$ the variation norm of the difference is at most the sum of the two masses, while below $x = d(S)$ both measures vanish. Combining (4.9)–(4.11) yields

$$\|\Lambda_\omega - \Pi\|_{\mathrm{var}} \le \varepsilon + 10\zeta. \tag{4.12}$$

Put $a_\Lambda := \|\Lambda\|$. Every Λ_ω has mass a_Λ, and (4.5) gives $\mathrm{e}^{-\zeta} \le a_\Lambda \le 1$. The word marginal $\overline{\Lambda}_\omega$ of $\overline{\Lambda} := a_\Lambda^{-1} \Lambda$ is $a_\Lambda^{-1} \Lambda_\omega$. Therefore

$$\begin{aligned}
\|\overline{\Lambda}_\omega - \Pi\|_{\mathrm{var}} &\le \|a_\Lambda^{-1} \Lambda_\omega - \Lambda_\omega\|_{\mathrm{var}} + \|\Lambda_\omega - \Pi\|_{\mathrm{var}} \\
&= (1 - a_\Lambda) + \|\Lambda_\omega - \Pi\|_{\mathrm{var}} \le \varepsilon + 11\zeta,
\end{aligned} \tag{4.13}$$

because $1 - a_\Lambda \le 1 - \mathrm{e}^{-\zeta} \le \zeta$. Thus, uniformly in ω,

$$d_{\mathrm{TV}}(\overline{\Lambda}_\omega, \Pi) \le \frac{\varepsilon + 11\zeta}{2}. \tag{4.14}$$

Step 4: the moving extremizer. Fix a lexicographic order on the finite set $2^{\mathcal{R}}$, and use the induced lexicographic order on its subfamilies. For each $x \ge 1$, choose the first admissible family $\mathcal{F}_x \subseteq \{S : d(S) \le x\}$ of maximum size. Only finitely many prefixes occur, so $x \mapsto \mathcal{F}_x$ is piecewise constant and hence measurable. Set

$$\mathcal{E} := \{(S, x) : S \in \mathcal{F}_x\}.$$

Because $\mathcal{F}_x$ is contained in the x-prefix, Tonelli's theorem gives

$$\begin{aligned} \Pi(\mathcal{E}) &= \frac{1}{Z_{\mathcal{R}}} \sum_{S \subseteq \mathcal{R}} \int_{d(S)}^{\infty} \mathbf{1}_{\{S \in \mathcal{F}_x\}} \frac{\mathrm{d}x}{x^2} \\ &= \frac{1}{Z_{\mathcal{R}}} \int_1^{\infty} |\mathcal{F}_x| \frac{\mathrm{d}x}{x^2} = \frac{I_{\mathcal{R}}^{\mathrm{sf}}}{Z_{\mathcal{R}}}. \end{aligned} \tag{4.15}$$

For $\overline{\Lambda}$-almost every (c, x), one has $x \ge D(c)$, so all 3^H word supports lie in the same x-prefix. Since $\mathcal{F}_x$ is admissible, Lemma 4.1 implies

$$\frac{1}{3^H} \sum_{\omega \in \mathbb{F}_3^H} \mathbf{1}_{\{S_\omega(c) \in \mathcal{F}_x\}} \le 3\kappa_{\mathrm{cap}}^H.$$

Integrating this pointwise inequality and using that $\overline{\Lambda}_\omega$ is the (S_ω, x)-pushforward of $\overline{\Lambda}$, we obtain

$$\begin{aligned} \frac{1}{3^H} \sum_{\omega \in \mathbb{F}_3^H} \overline{\Lambda}_\omega(\mathcal{E}) &= \int \frac{1}{3^H} \sum_{\omega \in \mathbb{F}_3^H} \mathbf{1}_{\{S_\omega(c) \in \mathcal{F}_x\}} \, \mathrm{d}\overline{\Lambda}(c, x) \\ &\le 3\kappa_{\mathrm{cap}}^H. \end{aligned} \tag{4.16}$$

On the other hand, (4.14) and (4.15) imply, for every ω,

$$\overline{\Lambda}_\omega(\mathcal{E}) \ge \frac{I_{\mathcal{R}}^{\mathrm{sf}}}{Z_{\mathcal{R}}} - \frac{\varepsilon + 11\zeta}{2}.$$

Averaging this inequality over all words and comparing with (4.16) gives

$$\frac{I_{\mathcal{R}}^{\mathrm{sf}}}{Z_{\mathcal{R}}} \le 3\kappa_{\mathrm{cap}}^H + \frac{\varepsilon + 11\zeta}{2},$$

which proves (4.3). □

5 One prime band: the model and its first moment

5.1 Explicit parameters and normalized prime weights

We make all structural constants explicit:

$$a = \frac{23}{25}, \qquad a_0 = \frac{24}{25}, \qquad \vartheta = \frac{1}{200}, \qquad \alpha = \frac{8}{25}, \qquad q := 3\alpha = \frac{24}{25}. \tag{5.1}$$

These choices satisfy

$$\frac{1}{\log 3} < a < a_0 < 1, \qquad a(1-\vartheta)\log 3 > 1. \tag{5.2}$$

Here is an exact certificate. The positive series

$$\log 3 = 2 \sum_{n \ge 0} \frac{(1/2)^{2n+1}}{2n+1}$$

gives, on retaining its first five terms,

$$\log 3 > 1 + \frac{1}{12} + \frac{1}{80} + \frac{1}{448} + \frac{1}{2304} > \frac{549}{500}.$$

Consequently

$$a(1-\vartheta)\log 3 > \frac{23}{25}\frac{199}{200}\frac{549}{500} = \frac{2512773}{2500000} > 1.$$

This also implies $a\log 3 > 1$. The companion verifier reproduces these comparisons with exact rational arithmetic.

Let $T \to \infty$, and define

$$\mathcal{B}(T) := \{p : \mathrm{e}^{T^\vartheta} < p \le \mathrm{e}^T\}, \qquad u_p := \frac{\log p}{T}, \qquad s_p := -\log u_p. \tag{5.3}$$

Thus $u_p \in (T^{\vartheta-1}, 1]$. The largest possible depth is

$$\lambda_T := (1-\vartheta)\log T. \tag{5.4}$$

Let $\eta = \eta(T)$ satisfy

$$\eta \longrightarrow 0, \qquad \log(1/\eta) = o(\log T), \qquad w := \frac{\eta}{T}. \tag{5.5}$$

Equivalently, $\eta = T^{-o(1)}$. All statements in this section hold for every such function η once T is sufficiently large; the threshold may depend on the function, but all displayed constants are uniform in T and η.

5.2 Reciprocal-prime mass on a normalized band

The upper bound below is expressed in terms of w, not the actual length of the interval. This is important because reciprocal-prime measure is atomic.

Lemma 5.1 (Local reciprocal-prime mass)**.** *Fix $C_0 > 0$. For every sufficiently large T, every interval $I \subseteq [T^{\vartheta-1}, 1]$ of length at most $C_0 w$ satisfies*

$$\sum_{\substack{p\in\mathcal{B}(T)\\ u_p\in I}} \frac{1}{p} \ll_{C_0} \frac{w}{\inf I}. \tag{5.6}$$

This includes degenerate intervals and intervals containing only one prime.

If $0 < r_0 < r_1 < 1$ and $c_0 > 0$ are fixed, then, uniformly for $t \in [r_0, r_1]$,

$$\sum_{\substack{p\in\mathcal{B}(T)\\ t\le u_p\le t+c_0 w}} \frac{1}{p} = \log\left(1 + \frac{c_0 w}{t}\right) + o(w) = (1+o(1))\frac{c_0 w}{t}. \tag{5.7}$$

Both statements remain valid with $1/p$ replaced by $1/(p+1)$.

Proof. Put $\epsilon_T = T^{\vartheta-1}$. Since

$$\frac{w}{\epsilon_T} = \frac{\eta}{T^\vartheta} = o(1),$$

we may enlarge I, only in an available direction at a band endpoint, to an interval $[r,s] \subseteq [\epsilon_T, 1]$ such that

$$I \subseteq [r,s], \qquad w \le s-r \le (C_0+2)w, \qquad r \ge \frac{1}{2}\inf I. \tag{5.8}$$

Set $x = \mathrm{e}^{Tr}$ and $\Delta = \mathrm{e}^{Ts} - \mathrm{e}^{Tr}$. Because $T(s-r) \asymp_{C_0} \eta$,

$$\Delta \asymp_{C_0} x\eta, \qquad \Delta \le x, \qquad \log\Delta = Tr + \log\eta + O_{C_0}(1).$$

Now $Tr \geq T^\vartheta/2$, while $|\log \eta| = o(\log T) = o(T^\vartheta)$. Hence, for large T,

$$\log \Delta \geq \frac{1}{4} T \inf I \quad \text{and} \quad \Delta \geq 2.$$

Brun–Titchmarsh, Proposition 2.2, gives

$$\pi(x + \Delta) - \pi(x) \ll_{C_0} \frac{x\eta}{T \inf I}.$$

Allowing one possible prime at a closed lower endpoint,

$$\begin{aligned}\sum_{\substack{p \in \mathcal{B}(T) \\ u_p \in I}} \frac{1}{p} &\leq \frac{\pi(x + \Delta) - \pi(x) + 1}{x} \\ &\ll_{C_0} \frac{\eta}{T \inf I} + \mathrm{e}^{-T^\vartheta} \ll_{C_0} \frac{w}{\inf I}.\end{aligned}$$

The last absorption uses $\mathrm{e}^{-T^\vartheta} = o(w)$. The enlargement in (5.8) was made for every I, including a degenerate interval. Equivalently, a degenerate I contains at most one normalized prime weight, whose reciprocal contribution is at most $\mathrm{e}^{-T^\vartheta} = o(w)$; thus single-atom intervals cause no exception.

For (5.7), subtract (2.2) at e^{Tt} and $\mathrm{e}^{T(t+c_0 w)}$. Uniformly for $t \in [r_0, r_1]$ this gives

$$\sum_{\substack{p \in \mathcal{B}(T) \\ t \leq u_p \leq t + c_0 w}} \frac{1}{p} = \log \frac{t + c_0 w}{t} + O(\mathrm{e}^{-c\sqrt{Tr_0}}) + O(\mathrm{e}^{-Tr_0}).$$

Both error terms are $o(w)$ because $\eta = T^{-o(1)}$. Finally,

$$0 \leq \sum_{p \in \mathcal{B}(T)} \left(\frac{1}{p} - \frac{1}{p+1} \right) \leq \sum_{n > \mathrm{e}^{T^\vartheta}} \frac{1}{n^2} \ll \mathrm{e}^{-T^\vartheta} = o(w),$$

which proves the versions with $1/(p+1)$. □

5.3 The five-state coupling and the target event

For $p \in \mathcal{B}(T)$ put

$$r_p := \frac{1}{p+1}, \qquad h_p := \frac{r_p}{3} = \frac{1}{3(p+1)}.$$

Independently at each prime, choose one of the five labels

$$\varnothing, \quad c, \quad x, \quad y, \quad z$$

with probabilities

$$\mathbb{P}(c) = \mathbb{P}(x) = \mathbb{P}(y) = \mathbb{P}(z) = h_p, \qquad \mathbb{P}(\varnothing) = 1 - 4h_p. \tag{5.9}$$

These probabilities are nonnegative because $r_p \leq 1/3$. From one labelled configuration form the following three state supports:

state	support on the band	missing petal
$S^{(x)}$	$c \cup y \cup z$	x
$S^{(y)}$	$c \cup x \cup z$	y
$S^{(z)}$	$c \cup x \cup y$	z

A fixed prime belongs to any one state with probability $3h_p = r_p$. Independence across primes therefore shows that each state has exactly the product law $\mu_{\mathcal{B}(T)}$.

For a labelled configuration define the normalized petal weights

$$X := \sum_{p\colon x(p)} u_p, \qquad Y := \sum_{p\colon y(p)} u_p, \qquad Z := \sum_{p\colon z(p)} u_p \tag{5.10}$$

and the four prefix counts

$$N_\ell(s) := \#\{p : \ell(p) = \ell,\ s_p \le s\}, \qquad \ell \in \{c, x, y, z\}. \tag{5.11}$$

We now fix, once and for all,

$$J := \left[\frac{9}{20}, \frac{11}{20}\right], \qquad R_0 := 75. \tag{5.12}$$

The parameter choice gives, for every real $s \ge R_0$,

$$3\lfloor \alpha s \rfloor \ge 3\alpha s - 3 = a_0 s - 3 \ge a s, \tag{5.13}$$

because $(a_0 - a)R_0 = 3$. Also $\lfloor \alpha R_0 \rfloor = 24 \ge 1$.

Define the symmetric event

$$\begin{aligned} B_T := \{X, Y, Z \in J, \quad & \max(X,Y,Z) - \min(X,Y,Z) \le w; \\ & N_\ell(s) \ge \lfloor \alpha s \rfloor \text{ for } \ell \in \{c,x,y,z\} \text{ and } R_0 \le s \le \lambda_T\}. \end{aligned} \tag{5.14}$$

Fix the first state $S = S^{(x)}$ and set

$$g_T(S) := \mathbb{P}(B_T \mid S), \qquad \beta_T := \mathbb{P}(B_T). \tag{5.15}$$

The event B_T is invariant under every permutation of x, y, z. Consequently the same function g_T describes all three states. Bayes' formula gives the exact Radon–Nikodym identity

$$\frac{\mathrm{d}\nu_T}{\mathrm{d}\mu_{\mathcal{B}(T)}}(S) = \frac{g_T(S)}{\beta_T}, \tag{5.16}$$

where ν_T is the law of any state conditioned on B_T.

The next three lemmas prepare the lower bound $\beta_T \gg w^2$.

Lemma 5.2 (Whole-band Poisson coupling)**.** *Let Q_T be the categorical law* (5.9). *Regard it as a law on four nonnegative integer arrays, one for each active label. Let $\widetilde{\mathsf{Q}}_T$ be the law of independent random variables*

$$K_{\ell,p} \sim \mathrm{Poisson}(h_p), \qquad \ell \in \{c,x,y,z\}, \quad p \in \mathcal{B}(T).$$

Then

$$d_{\mathrm{TV}}(\mathsf{Q}_T, \widetilde{\mathsf{Q}}_T) \le 16 \sum_{p \in \mathcal{B}(T)} h_p^2 \ll \mathrm{e}^{-T^{\vartheta}} = o(w^A) \tag{5.17}$$

for every fixed $A > 0$.

Proof. Fix p and write $h = h_p$. Under the Poisson law, the total number of points at p is Poisson with mean $4h$. The empty configuration and each of the four singleton configurations have probabilities e^{-4h} and $h\mathrm{e}^{-4h}$, respectively. Therefore

$$2d_{\mathrm{TV}}(\mathsf{Q}_p, \widetilde{\mathsf{Q}}_p) \le |1 - 4h - \mathrm{e}^{-4h}| + 4h(1 - \mathrm{e}^{-4h}) + \mathbb{P}(\mathrm{Poisson}(4h) \ge 2).$$

The elementary inequalities

$$1-x\le \mathrm{e}^{-x}\le 1-x+\frac{x^2}{2},\qquad 1-\mathrm{e}^{-x}\le x\quad (x\ge 0)$$

bound the first two terms by $8h^2$ and $16h^2$. If $N\sim\mathrm{Poisson}(4h)$, then

$$\mathbb{P}(N\ge 2)\le \frac{\mathbb{E}[N(N-1)]}{2}=8h^2.$$

Thus the local total-variation distance is at most $16h^2$.

Couple the arrays prime by prime. The probability that any local coupling fails is at most the sum of the local distances, which proves the first inequality in (5.17). Since $h_p\le 1/(3p)$,

$$\sum_{p\in\mathcal{B}(T)} h_p^2\le \frac{1}{9}\sum_{n>\mathrm{e}^{T^\vartheta}}\frac{1}{n^2}\ll \mathrm{e}^{-T^\vartheta}.$$

Finally,

$$\log(1/w)=\log T+\log(1/\eta)=(1+o(1))\log T,$$

whereas $T^\vartheta/\log T\to\infty$. Hence $\mathrm{e}^{-T^\vartheta}=o(w^A)$ for every fixed A. $\square$

On the common array space define

$$X_\ell^{\mathrm{P}}:=\sum_{p\in\mathcal{B}(T)}u_pK_{\ell,p},\qquad N_\ell^{\mathrm{P}}(s):=\sum_{\substack{p\in\mathcal{B}(T)\\ s_p\le s}}K_{\ell,p}. \tag{5.18}$$

Let B_T^{P} be the event obtained from (5.14) by replacing X,Y,Z with $X_x^{\mathrm{P}},X_y^{\mathrm{P}},X_z^{\mathrm{P}}$ and every N_ℓ with N_ℓ^{P}. Multiplicities are counted under the Poisson law. On the categorical subspace, where each entry is zero or one and at most one label is nonzero at a prime, B_T^{P} is literally B_T. Thus

$$\mathsf{Q}_T(B_T^{\mathrm{P}})=\mathsf{Q}_T(B_T)=\beta_T. \tag{5.19}$$

For a depth interval D, write

$$\Lambda_T(D):=\sum_{\substack{p\in\mathcal{B}(T)\\ s_p\in D}}h_p.$$

Lemma 5.3 (Uniform prime time change and small tails)**.** *Uniformly for $0\le t\le\lambda_T-R_0$,*

$$\Lambda_T((R_0,R_0+t])=\frac{t}{3}+\varepsilon_T(t),\qquad \sup_t|\varepsilon_T(t)|=o(1). \tag{5.20}$$

For $\ell\in\{c,x,y,z\}$, put

$$V_{\ell,T}:=\sum_{\substack{p\in\mathcal{B}(T)\\ s_p>R_0}}u_pK_{\ell,p}. \tag{5.21}$$

For all sufficiently large T,

$$\mathbb{E}_{\widetilde{\mathsf{Q}}_T}V_{\ell,T}\le \mathrm{e}^{-R_0}. \tag{5.22}$$

Finally, for every fixed interval $D\subseteq[0,R_0]$,

$$\Lambda_T(D)\longrightarrow\frac{|D|}{3}. \tag{5.23}$$

All estimates hold simultaneously for the four labels.

Proof. Let

$$H_+(x) := \sum_{p \le x} \frac{1}{p+1}.$$

For $0 \le t \le \lambda_T - R_0$,

$$3\Lambda_T((R_0, R_0+t]) = H_+(\mathrm{e}^{T\mathrm{e}^{-R_0}}) - H_+(\mathrm{e}^{T\mathrm{e}^{-(R_0+t)}}) + O(\mathrm{e}^{-T^{\vartheta}}). \tag{5.24}$$

The logarithms of both prime cutoffs are at least T^{ϑ}. By (2.3), the error in (5.24) is $O(\mathrm{e}^{-cT^{\vartheta/2}})$ uniformly in t. Since

$$\log\log(\mathrm{e}^{T\mathrm{e}^{-s}}) = \log T - s,$$

the main term in (5.24) is t. This proves (5.20). The same subtraction over a fixed depth interval proves (5.23).

For the weighted tail, positivity and (2.4) give

$$\begin{aligned}\mathbb{E}V_{\ell,T} &= \frac{1}{3T} \sum_{\mathrm{e}^{T^{\vartheta}} < p < \mathrm{e}^{T\mathrm{e}^{-R_0}}} \frac{\log p}{p+1} \\ &\le \frac{1}{3T} \sum_{p \le \mathrm{e}^{T\mathrm{e}^{-R_0}}} \frac{\log p}{p-1} = \left(\frac{1}{3} + o(1)\right) \mathrm{e}^{-R_0}.\end{aligned}$$

The last expression is at most e^{-R_0} for all sufficiently large T, proving (5.22). □

Lemma 5.4 (A uniform Poisson lower profile)**.** *Let $\Pi(v)$ be a rate-one Poisson process. With*

$$q = \frac{24}{25}, \qquad \gamma := \frac{1}{50},$$

one has, for every $b \ge 0$,

$$\mathbb{P}\big(\Pi(v) + b \ge qv \text{ for every } v \ge 0\big) \ge 1 - \mathrm{e}^{-\gamma b}. \tag{5.25}$$

Proof. The inequality $\mathrm{e}^{-x} \le 1 - x + x^2/2$ gives

$$q\gamma + \mathrm{e}^{-\gamma} - 1 \le (q-1)\gamma + \frac{\gamma^2}{2} = -\frac{3}{5000} < 0.$$

It follows from independent increments that

$$M(v) := \exp\big(\gamma(qv - \Pi(v))\big)$$

is a nonnegative supermartingale.

We include the short maximal argument, including the harmless strictness at the boundary. Fix $L < \infty$ and $\epsilon > 0$. On a finite time grid in $[0, L]$, stop M at the first grid point at which it reaches $\mathrm{e}^{\gamma(b+\epsilon)}$. Conditioning on whether the stopping time has already occurred shows directly that the stopped process is still a nonnegative supermartingale. Its expectation is therefore at most $M(0) = 1$, so the grid crossing probability is at most $\mathrm{e}^{-\gamma(b+\epsilon)}$. Apply this to nested dyadic grids. Right continuity detects every strict crossing, and hence

$$\mathbb{P}\big(\exists v \in [0, L] : qv - \Pi(v) > b + \epsilon\big) \le \mathrm{e}^{-\gamma(b+\epsilon)}.$$

Let first $\epsilon \downarrow 0$ and then $L \to \infty$. We obtain

$$\mathbb{P}\big(\exists v \ge 0 : qv - \Pi(v) > b\big) \le \mathrm{e}^{-\gamma b}.$$

The complement is exactly the event in (5.25). □

Proposition 5.5 (Profiled anchor lower bound)**.** *There is an absolute constant $c > 0$ such that, for every function $\eta(T)$ satisfying (5.5) and all sufficiently large T,*

$$\beta_T = \mathbb{P}(B_T) \geq cw^2. \tag{5.26}$$

The constant c is independent of T and η.

Proof. We construct under $\widetilde{\mathbb{Q}}_T$ an event contained in B_T^{P} whose probability is at least a fixed multiple of w^2. The whole-band coupling will then transfer it to the categorical law.

Step 1: fixed numerical choices. Set

$$\delta := \frac{1}{100}, \qquad B := 350, \qquad m := \lceil \alpha R_0 \rceil + B = 374. \tag{5.27}$$

These constants satisfy the three inequalities needed below:

$$4\mathrm{e}^{-\gamma B/2} < \frac{1}{8}, \qquad m\mathrm{e}^{1-R_0} < \frac{\delta}{2}, \qquad \frac{8\mathrm{e}^{-R_0}}{\delta} < \frac{1}{8}. \tag{5.28}$$

We record exact witnesses. First,

$$\mathrm{e}^{7/2} > \sum_{n=0}^{7} \frac{(7/2)^n}{n!} = \frac{66007}{2048} > 32,$$

which proves the first inequality. Next,

$$374\mathrm{e}^{-74} < \frac{374 \cdot 4!}{74^4} = \frac{561}{1874161} < \frac{1}{200}.$$

Finally,

$$800\mathrm{e}^{-75} < \frac{800 \cdot 3!}{75^3} = \frac{64}{5625} < \frac{1}{8}.$$

Only positive terms of the exponential series were used. These calculations are also checked exactly in the companion verifier.

Step 2: finite-depth buffers. Let

$$\mathcal{K} := \left[\frac{9}{20}, \frac{11}{20}\right), \qquad \mathcal{A} := \left[\frac{12}{25}, \frac{13}{25}\right).$$

For a half-open weight interval $\mathcal{I}$, write $D_{\mathcal{I}} := \{-\log u : u \in \mathcal{I}\}$; its image is again taken with the corresponding half-open endpoint convention. Choose the four explicit, pairwise disjoint depth intervals

$$Q_c = (R_0 - 1, R_0 - 7/8], \qquad Q_x = (R_0 - 3/4, R_0 - 5/8],$$
$$Q_y = (R_0 - 1/2, R_0 - 3/8], \qquad Q_z = (R_0 - 1/4, R_0 - 1/8].$$

All lie in $(R_0 - 1, R_0)$.

Let F_{fix} be the following finite-depth event:

(i) label c has exactly m points in Q_c and no other point of depth at most R_0;

(ii) label z has exactly m points in Q_z, exactly one point in $D_{\mathcal{A}}$, and no other point of depth at most R_0;

(iii) each label $\ell \in \{x, y\}$ has exactly m points in Q_ℓ, no point in $[0, R_0] \setminus (Q_\ell \cup D_{\mathcal{K}})$, and no condition imposed on its process in $D_{\mathcal{K}}$.

Since $D_{\mathcal{A}}, D_{\mathcal{K}} \subset (0,1)$ whereas every Q_ℓ lies in $(R_0-1, R_0) = (74, 75)$, the anchor-depth regions are disjoint from all four buffer intervals. We now specify the finite atoms used to evaluate this event. For label c, take Q_c itself and the nonempty interval components of $[0, R_0] \setminus Q_c$; prescribe counts m and 0, respectively. For label z, take Q_z, $D_{\mathcal{A}}$, and the nonempty interval components of their complement in $[0, R_0]$; prescribe counts m, 1, and 0. For each $\ell \in \{x, y\}$, take Q_ℓ and the nonempty interval components of

$$[0, R_0] \setminus (Q_\ell \cup D_{\mathcal{K}});$$

prescribe count m on Q_ℓ and count 0 on every listed component. The region $D_{\mathcal{K}}$ is deliberately absent from the constrained atom family for these two labels. Endpoint assignments are inherited from the half-open conventions defining the intervals, so within each label the listed atoms are disjoint and describe exactly the conditions in (i)–(iii). Let $k_{\ell,D}$ denote the prescribed count. When $k_{\ell,D} = 0$, the corresponding factor below is simply $\mathrm{e}^{-\Lambda_T(D)}$. Poisson independent increments give the exact product

$$\widetilde{\mathsf{Q}}_T(F_{\text{fix}}) = \prod_{(\ell,D)\ \text{constrained}} \mathrm{e}^{-\Lambda_T(D)} \frac{\Lambda_T(D)^{k_{\ell,D}}}{k_{\ell,D}!}. \tag{5.29}$$

By finite additivity and (5.23), every intensity in this finite product has a finite limit. Each atom on which a positive count is required has positive length and hence a strictly positive limiting intensity. It follows that, for all sufficiently large T,

$$\widetilde{\mathsf{Q}}_T(F_{\text{fix}}) \geq c_{\text{fix}} > 0, \tag{5.30}$$

where c_{fix} is independent of T.

Step 3: survival through the entire remaining depth. For a label ℓ, define

$$M_{\ell,T}(t) := \sum_{\substack{p \in \mathcal{B}(T) \\ R_0 < s_p \leq R_0 + t}} K_{\ell,p}, \qquad 0 \leq t \leq \lambda_T - R_0.$$

This independent-increment process can be realized as

$$M_{\ell,T}(t) = \Pi_\ell(\Lambda_T((R_0, R_0 + t])), \tag{5.31}$$

where the Π_ℓ are independent rate-one Poisson processes. Put $\epsilon_T := \sup_t |\varepsilon_T(t)|$. For all sufficiently large T, $q\epsilon_T \leq B/2$.

Suppose that

$$\Pi_\ell(v) + B/2 \geq qv \quad \text{for every } v \geq 0. \tag{5.32}$$

With $v = \Lambda_T((R_0, R_0 + t])$, the time-change estimate gives $t/3 \leq v + \epsilon_T$. Therefore

$$M_{\ell,T}(t) + B \geq qv + B/2 \geq \alpha t, \tag{5.33}$$

$$m + M_{\ell,T}(t) \geq \alpha(R_0 + t). \tag{5.34}$$

The second line uses $m \geq \alpha R_0 + B$. By Lemma 5.4 and the first inequality in (5.28), the probability that (5.34) fails for at least one of the four labels is less than $1/8$.

The weighted tails are also small at fixed positive cost. By (5.22), Markov's inequality, and the last inequality in (5.28),

$$\mathbb{P}\left(\max_{\ell \in \{c,x,y,z\}} V_{\ell,T} > \delta/2 \right) \leq \frac{8\mathrm{e}^{-R_0}}{\delta} < \frac{1}{8}. \tag{5.35}$$

All depth-$> R_0$ processes are independent of $F_{\rm fix}$. Let $\mathcal{E}_T$ be the event that (5.34) holds for all four labels and all $0 \le t \le \lambda_T - R_0$, and that

$$V_{\ell,T} \le \delta/2 \qquad (\ell = c, x, y, z) \tag{5.36}$$

holds. The preceding two union bounds show that

$$\widetilde{\mathsf{Q}}_T(\mathcal{E}_T \mid F_{\rm fix}) \ge \frac{3}{4}. \tag{5.37}$$

No error depending on w has been subtracted.

Step 4: two adaptive correcting anchors. Let $\mathcal{H}_T$ be the σ-field generated by the restrictions of all four labelled Poisson arrays except for the restrictions of the x- and y-arrays to $D_{\mathcal{K}}$. The event $F_{\rm fix}$ is $\mathcal{H}_T$-measurable because it imposes no condition on those two omitted restrictions. The event $\mathcal{E}_T$ is also $\mathcal{H}_T$-measurable: it is determined entirely by the four depth-$> R_0$ processes. By independence across labels and independent increments on disjoint depth sets, conditionally on $\mathcal{H}_T$ the two omitted restrictions remain independent Poisson processes, each with its original intensity Λ_T restricted to $D_{\mathcal{K}}$.

Let $U_z \in \mathcal{A}$ be the normalized weight of the unique z-point in $D_{\mathcal{A}}$. For $\ell \in \{x, y, z\}$, let E_ℓ be the total weight of the m points in Q_ℓ plus $V_{\ell,T}$. Every point in Q_ℓ has weight at most e^{1-R_0}, so (5.28) and (5.36) imply

$$0 \le E_\ell \le \delta. \tag{5.38}$$

Work conditionally on $\mathcal{H}_T$, on $F_{\rm fix} \cap \mathcal{E}_T$. The variables U_z, E_x, E_y, E_z are then fixed. Define the $\mathcal{H}_T$-measurable centers

$$t_x := U_z + E_z - E_x, \qquad t_y := U_z + E_z - E_y \tag{5.39}$$

and target intervals

$$I_x := \left[t_x - \frac{w}{8}, t_x + \frac{w}{8}\right], \qquad I_y := \left[t_y - \frac{w}{8}, t_y + \frac{w}{8}\right]. \tag{5.40}$$

From $U_z \in [12/25, 13/25)$ and (5.38),

$$t_x, t_y \in \left[\frac{47}{100}, \frac{53}{100}\right].$$

Since $w \to 0$, both intervals in (5.40) lie in $\mathcal{K}$ for all sufficiently large T.

Require the x-process to have exactly one point in I_x and no other point in $D_{\mathcal{K}}$, and impose the analogous requirement on the y-process. Let $\mathcal{A}_T$ denote this two-anchor event on $F_{\rm fix} \cap \mathcal{E}_T$, and declare it empty off that event. Conditionally on $\mathcal{H}_T$, the probability of the x-requirement is exactly

$$\left(\sum_{\substack{p \in \mathcal{B}(T)\\ u_p \in I_x}} h_p\right) \exp\left(-\sum_{\substack{p \in \mathcal{B}(T)\\ u_p \in \mathcal{K}}} h_p\right). \tag{5.41}$$

The interval I_x has the form $[t, t + w/4]$ with $t = t_x - w/8$; for large T, this left endpoint lies in the fixed compact interval $[23/50, 27/50] \subset (0,1)$. Thus Lemma 5.1, with $c_0 = 1/4$, gives, almost surely on $F_{\rm fix} \cap \mathcal{E}_T$ and uniformly over every possible conditional realization,

$$\sum_{\substack{p \in \mathcal{B}(T)\\ u_p \in I_x}} h_p = \frac{1}{3} \sum_{\substack{p \in \mathcal{B}(T)\\ u_p \in I_x}} \frac{1}{p+1} \ge c_1 w.$$

Although I_x is random before conditioning, it is $\mathcal{H}_T$-measurable and hence is a fixed deterministic interval under each conditional realization. The lower bound in Lemma 5.1 is uniform over all left endpoints in $[23/50, 27/50]$, so the same constant c_1 applies simultaneously to every such realization. The identical observation applies to I_y. Moreover, (2.3) gives

$$\sum_{\substack{p\in\mathcal{B}(T)\\ u_p\in\mathcal{K}}} h_p \longrightarrow \frac{1}{3}\log\frac{11}{9}.$$

Consequently the exponential factor in (5.41) is at least a fixed $c_2 > 0$. The same bounds hold for I_y. Independence of the two omitted processes therefore gives both correcting anchors with conditional probability at least

$$c_3 w^2. \tag{5.42}$$

Step 5: verification of the target event and transfer. Let U_x, U_y be the weights of the two correcting anchors. On the event just constructed,

$$X = U_x + E_x, \qquad Y = U_y + E_y, \qquad Z = U_z + E_z.$$

By (5.39)–(5.40),

$$|X - Z| \le w/8, \qquad |Y - Z| \le w/8. \tag{5.43}$$

Thus the range of X, Y, Z is at most $w/4$. Also $Z \in [48/100, 53/100]$ by (5.38); since $w \to 0$, (5.43) places all three totals in the fixed closed interval J for large T.

At depth R_0, every label has at least m points. For $s = R_0 + t$, (5.34) gives

$$N_\ell^{\mathrm{P}}(s) \ge \alpha s \ge \lfloor \alpha s \rfloor \qquad (\ell = c, x, y, z).$$

The constructed event is therefore contained in B_T^{P}. The conditional estimate (5.42) is uniform on $F_{\mathrm{fix}} \cap \mathcal{E}_T$. The tower property, (5.30), and (5.37) therefore give

$$\begin{aligned}\widetilde{\mathsf{Q}}_T(F_{\mathrm{fix}} \cap \mathcal{E}_T \cap \mathcal{A}_T) &= \mathbb{E}_{\widetilde{\mathsf{Q}}_T}\Big[\mathbf{1}_{F_{\mathrm{fix}}\cap\mathcal{E}_T}\,\widetilde{\mathsf{Q}}_T(\mathcal{A}_T \mid \mathcal{H}_T)\Big]\\ &\ge c_3 w^2\, \widetilde{\mathsf{Q}}_T(F_{\mathrm{fix}} \cap \mathcal{E}_T) \ge \frac{3}{4} c_{\mathrm{fix}} c_3 w^2.\end{aligned}$$

The event on the left is contained in B_T^{P}, so

$$\widetilde{\mathsf{Q}}_T(B_T^{\mathrm{P}}) \ge \frac{3}{4} c_{\mathrm{fix}} c_3 w^2 =: c_4 w^2. \tag{5.44}$$

Finally, the defining property of total-variation distance gives

$$\Big|\mathsf{Q}_T(B_T^{\mathrm{P}}) - \widetilde{\mathsf{Q}}_T(B_T^{\mathrm{P}})\Big| \le d_{\mathrm{TV}}(\mathsf{Q}_T, \widetilde{\mathsf{Q}}_T).$$

Lemma 5.2 with $A = 2$ makes the right side $o(w^2)$. Hence (5.19) and (5.44) give

$$\beta_T = \mathsf{Q}_T(B_T^{\mathrm{P}}) \ge \widetilde{\mathsf{Q}}_T(B_T^{\mathrm{P}}) - o(w^2) \ge c_4 w^2 - o(w^2) \ge \frac{c_4}{2} w^2$$

for all sufficiently large T. This proves (5.26). $\square$

6 The exact rooted second moment

The first-moment estimate shows that conditioning on B_T is not too expensive. We now show that the conditioned root density does not concentrate. The proof is annealed: the root support and its two independent colourings are summed together. This is essential, because no comparable anti-concentration statement is true uniformly for every fixed root.

6.1 The conditional root law and factorial insertion

Fix the first state $S = S^{(x)} = c \cup y \cup z$. Since its marginal is $\mu_{\mathcal{B}(T)}$, the primes of S are independent with probabilities $r_p = 1/(p+1)$ and inclusion odds

$$\frac{r_p}{1-r_p} = \frac{1}{p}. \tag{6.1}$$

Conditional on S, every $p \in S$ is independently and uniformly coloured c, y, or z. Every $p \notin S$ is independently assigned to the missing x-petal with probability

$$\frac{h_p}{1-r_p} = \frac{1}{3p}. \tag{6.2}$$

Indeed, absence from S means that the original label is either $\varnothing$ or x.

Take two independent completions conditional on the same root. Put

$$\sigma(c) = 0, \qquad \sigma(y) = 1, \qquad \sigma(z) = -1$$

and, for $p \in S$, define the nine-valued mark

$$\xi_p := (\sigma(\ell_1(p)), \sigma(\ell_2(p))) \in \mathcal{M} := \{-1, 0, 1\}^2. \tag{6.3}$$

All nine marks are equiprobable. If

$$D_h := Y_h - Z_h \qquad (h = 1, 2), \tag{6.4}$$

then the vector $D = (D_1, D_2)$ is exactly

$$D = \sum_{p \in S} u_p \xi_p. \tag{6.5}$$

Both pivot estimates use the following exact identity.

Lemma 6.1 (Factorial deletion and insertion)**.** *Let $\mathcal{P}$ be finite, and let μ be the product law*

$$\mu(S) = \prod_{p \in S} r_p \prod_{p \notin S} (1 - r_p), \qquad \omega_p := \frac{r_p}{1 - r_p}, \qquad 0 < r_p < 1.$$

For $d \geq 1$, let $F(S; p_1, \ldots, p_d)$ be nonnegative and defined when $p_1, \ldots, p_d$ are distinct members of S. Then

$$\begin{aligned} &\mathbb{E}_{S \sim \mu} \sum_{p_1, \ldots, p_d \in S}^{\neq} F(S; p_1, \ldots, p_d) \\ &\quad = \sum_{A \subseteq \mathcal{P}} \mu(A) \sum_{\substack{p_1, \ldots, p_d \in \mathcal{P} \setminus A \\ p_1, \ldots, p_d \text{ distinct}}} \left(\prod_{j=1}^{d} \omega_{p_j} \right) F(A \cup \{p_1, \ldots, p_d\}; p_1, \ldots, p_d). \end{aligned} \tag{6.6}$$

Proof. Expand the expectation on the left. For every ordered tuple appearing in the expansion, make the bijective substitution

$$A = S \setminus \{p_1, \ldots, p_d\}.$$

The deleted points are pairwise distinct and absent from A, so

$$\mu(A \cup \{p_1, \ldots, p_d\}) = \mu(A) \prod_{j=1}^{d} \frac{r_{p_j}}{1 - r_{p_j}}.$$

Substitution gives (6.6). Every sum is finite, so no limiting interchange is involved. $\square$

6.2 Profiles and ordered weights

Order a nonempty support S by decreasing normalized weight:

$$u_{(1)} > u_{(2)} > \cdots > u_{(K)}.$$

There are no ties because distinct primes have distinct logarithms.

Lemma 6.2 (Ordered weights forced by a profile)**.** *Suppose that*

$$N_S(s) := \#\{p \in S : s_p \le s\} \ge as \qquad (R_0 \le s \le \lambda_T). \tag{6.7}$$

Then $K = |S| \ge a\lambda_T$, and for $1 \le k \le K$,

$$u_{(k)} \ge \ell_k, \qquad \ell_k := \exp\left(-\max\left\{R_0, \frac{k+1}{a}\right\}\right). \tag{6.8}$$

Proof. The inequality $K \ge a\lambda_T$ is (6.7) at $s = \lambda_T$. Put $s_{(k)} = -\log u_{(k)}$. If $s_{(k)} \le R_0$, then $u_{(k)} \ge \mathrm{e}^{-R_0} \ge \ell_k$. Otherwise take $R_0 \le s < s_{(k)}$ and let $s \uparrow s_{(k)}$. At depth s there are at most $k-1$ points, so $as \le k-1$. Hence $s_{(k)} \le (k-1)/a$, which is stronger than (6.8). $\square$

We record the two convergent endpoint estimates before using them. From (5.2),

$$\frac{3^{-a\lambda_T}}{w} = \frac{T^{\,1-a(1-\vartheta)\log 3}}{\eta} = o(1). \tag{6.9}$$

Indeed the exponent of T is a fixed negative number and $1/\eta = T^{o(1)}$. Moreover,

$$\sum_{k\ge 1} \frac{3^{-k}}{\ell_k} < \infty. \tag{6.10}$$

To see this, discard finitely many indices controlled by R_0; the remaining ratio is $\mathrm{e}^{1/a}/3 < 1$, which is equivalent to $a\log 3 > 1$.

Lemma 6.3 (Annealed two-pivot small-ball bound)**.** *Let $S \sim \mu_{\mathcal{B}(T)}$, give its primes independent uniform marks in $\mathcal{M} = \{-1,0,1\}^2$, and define D by (6.5). For every axis-parallel square $Q \subseteq \mathbb{R}^2$ of side at most $C_Q w$, uniformly in its position,*

$$\mathbb{P}\big(N_S(s) \ge as \text{ for } R_0 \le s \le \lambda_T,\ D \in Q\big) \ll_{C_Q} w^2. \tag{6.11}$$

Proof. Write

$$\mathcal{P}_T(S) := \{N_S(s) \ge as \text{ for every } R_0 \le s \le \lambda_T\}.$$

For brevity, write $\mu := \mu_{\mathcal{B}(T)}$. For each realization put $K := |S|$. All ranks below refer to decreasing order of the weights u_p. For a marked realization $\xi = (\xi_p)_{p\in S}$, define

$$r(\xi) := \dim_{\mathbb{R}} \operatorname{span}\{\xi_p : p \in S\}.$$

The span of the empty set is understood to be $\{0\}$. Thus the cases $r(\xi) = 0, 1, 2$ are disjoint and exhaustive.

Mark-counting sublemma. For $v \in \mathcal{M} \setminus \{0\}$, put

$$L(v) := \mathcal{M} \cap \mathbb{R}v = \{-v, 0, v\}.$$

Thus $|L(v)| = 3$ and $|\mathcal{M} \setminus L(v)| = 6$. On a support of size K, the total probability of all mark sequences whose first nonzero mark has rank i and whose first subsequent mark outside its real span has rank $j > i$ is

$$9^{-K} \cdot 8 \cdot 3^{j-i-1} \cdot 6 \cdot 9^{K-j} = 16\, 3^{-(i+j)}. \tag{6.12}$$

The total probability of all rank-one mark sequences whose first nonzero mark has rank i is

$$9^{-K} \cdot 8 \cdot 3^{K-i} = 8\,3^{-(K+i)}. \tag{6.13}$$

Indeed, there are eight choices for the mark at rank i. For (6.12), the ranks $i+1, \dots, j-1$ have the three choices in $L(v)$, rank j has the six choices outside $L(v)$, and every later rank has nine choices. For (6.13), every later mark must belong to $L(v)$. Multiplication and division by the 9^K equally likely mark sequences prove both formulas.

Rank two. For a rank-two marked support, let i be the rank of its first nonzero mark and let $j > i$ be the rank of its first mark outside the span of the mark at rank i. Let p and q be the primes at these ranks. This choice is canonical.

We spell out the data remaining after deletion of p and q. Put $A = S \setminus \{p, q\}$ and $K = |A| + 2$. Order A as $r_1, \dots, r_{K-2}$ with

$$b_1 := u_{r_1} > \cdots > b_{K-2} := u_{r_{K-2}}, \qquad b_0 := +\infty, \quad b_{K-1} := 0.$$

When $A = \varnothing$, the displayed ordered list is empty and only the two endpoint conventions apply. For $1 \le i < j \le K$, let $\mathcal{G}_{ij}(A)$ be the set of ordered pairs of distinct primes $p, q \in \mathcal{B}(T) \setminus A$ such that

$$\begin{cases} b_{i-1} > u_p > u_q > b_i, & j = i+1, \\ b_{i-1} > u_p > b_i \ \text{ and } \ b_{j-2} > u_q > b_{j-1}, & j \ge i+2. \end{cases} \tag{6.14}$$

These inequalities say exactly that p and q occupy ranks i and j after insertion. For $v \neq 0$, let $\mathcal{C}_{ij}(A; v)$ be the set of mark vectors $\zeta = (\zeta_1, \dots, \zeta_{K-2})$ on this ordered copy of A satisfying

$$\zeta_k = 0 \ (k < i), \qquad \zeta_k \in L(v) \ (i \le k \le j-2), \qquad \zeta_k \in \mathcal{M} \ (k \ge j-1), \tag{6.15}$$

where an empty index range imposes no condition. Put

$$D_A(\zeta) := \sum_{k=1}^{K-2} b_k \zeta_k.$$

We now verify explicitly that deletion and insertion preserve every canonical index. Let $t_1, \dots, t_K$ denote the decreasing-weight ordering after p, q are inserted. For $(p, q) \in \mathcal{G}_{ij}(A)$ this ordering is, and can only be,

$$(t_1, \dots, t_K) = (r_1, \dots, r_{i-1}, p, r_i, \dots, r_{j-2}, q, r_{j-1}, \dots, r_{K-2}), \tag{6.16}$$

where every sublist with lower index larger than its upper index is empty. If $j \ge i+2$, the two insertion conditions in (6.16) are precisely

$$b_{i-1} > u_p > b_i, \qquad b_{j-2} > u_q > b_{j-1}.$$

If $j = i+1$, both points lie in the same residual gap and their required order is precisely

$$b_{i-1} > u_p > u_q > b_i.$$

Thus (6.16) is equivalent to $(p, q) \in \mathcal{G}_{ij}(A)$, including the endpoint ranks under the sentinel conventions.

For an ordered marked support (S, ξ), define $\mathcal{E}_{ij}^{v,v'}(S, \xi; p, q)$ to be the event

$$\begin{gathered} p = t_i, \qquad q = t_j, \qquad \xi_{t_k} = 0 \quad (k < i), \\ \xi_p = v \neq 0, \qquad \xi_{t_k} \in L(v) \quad (i < k < j), \qquad \xi_q = v' \notin L(v). \end{gathered}$$

There is no restriction on marks after rank j. Pointwise in (S,ξ),

$$\mathbf{1}_{\{r(\xi)=2\}} = \sum_{\substack{p,q\in S\\ p\neq q}} \sum_{1\le i<j\le K} \sum_{v\in\mathcal{M}\setminus\{0\}} \sum_{v'\in\mathcal{M}\setminus L(v)} \mathbf{1}_{\mathcal{E}_{ij}^{v,v'}(S,\xi;p,q)}. \tag{6.17}$$

Indeed, if $r(\xi)=2$, the sole nonzero summand uses the prime carrying the first nonzero mark and the prime carrying the first later mark outside its span. If $r(\xi)<2$, no summand can occur. In particular, exchanging p and q never creates a second summand: their canonical ranks are different, and (6.16) assigns each rank to one specific insertion gap.

Under (6.16), deletion of the two pivot marks turns the full-rank conditions in $\mathcal{E}_{ij}^{v,v'}$ into exactly (6.15). Namely, residual indices $k<i$ retain full rank k; indices $i\le k\le j-2$ retain full rank $k+1$ and lie strictly between the pivots; and indices $k\ge j-1$ retain full rank $k+2$ and occur after the second pivot.

To apply Lemma 6.1 without suppressing the mark average, for distinct $p,q\in S$ define

$$\begin{aligned} F_2(S;p,q) := 9^{-|S|} &\sum_{\xi\in\mathcal{M}^S} \mathbf{1}_{\mathcal{P}_T(S)} \mathbf{1}_{\{\sum_{s\in S} u_s\xi_s\in Q\}} \\ &\times \sum_{1\le i<j\le |S|} \sum_{v\in\mathcal{M}\setminus\{0\}} \sum_{v'\in\mathcal{M}\setminus L(v)} \mathbf{1}_{\mathcal{E}_{ij}^{v,v'}(S,\xi;p,q)}. \end{aligned}$$

Equation (6.17) gives the exact identity

$$\mathbb{P}(r(\xi)=2,\mathcal{P}_T(S),D\in Q) = \mathbb{E}_{S\sim\mu} \sum_{\substack{p,q\in S\\ p\neq q}} F_2(S;p,q).$$

For $p,q\notin A$, the support and mark coefficient after deletion is

$$\begin{aligned} \mu(A\cup\{p,q\})9^{-(|A|+2)} &= \mu(A)\frac{1/(p+1)}{1-1/(p+1)}\frac{1/(q+1)}{1-1/(q+1)}9^{-(|A|+2)} \\ &= \frac{\mu(A)}{pq}\,9^{-(|A|+2)}. \end{aligned}$$

Thus the two-point case of Lemma 6.1, followed by the unique decomposition of the marks into (v,v',ζ) described above, gives the exact equality

$$\begin{aligned} &\mathbb{P}(r(\xi)=2,\ \mathcal{P}_T(S),\ D\in Q) \\ &= \sum_{A\subseteq\mathcal{B}(T)} \mu(A) \sum_{1\le i<j\le |A|+2} 9^{-(|A|+2)} \sum_{v\in\mathcal{M}\setminus\{0\}} \sum_{v'\in\mathcal{M}\setminus L(v)} \sum_{\zeta\in\mathcal{C}_{ij}(A;v)} \\ &\quad\times \sum_{(p,q)\in\mathcal{G}_{ij}(A)} \frac{1}{pq} \mathbf{1}_{\mathcal{P}_T(A\cup\{p,q\})} \mathbf{1}_{\{D_A(\zeta)+u_pv+u_qv'\in Q\}}. \end{aligned} \tag{6.18}$$

Here factorial insertion changes the support mass from $\mu(A\cup\{p,q\})$ to $\mu(A)/(pq)$; (6.14) restores the two deleted ranks; and (6.15) restores exactly the residual mark restrictions. Conversely, every tuple counted on the right reconstructs a rank-two marked support whose canonical data are (i,j,p,q,v,v'). Thus (6.18) has neither omission nor multiplicity.

Fix a term through the choice of A,i,j,v,v',ζ. On the profile event, Lemma 6.2 gives

$$u_p\ge \ell_i, \qquad u_q\ge \ell_j. \tag{6.19}$$

Let $M_{v,v'}$ be the matrix with columns v,v'. It has nonzero integer determinant. Every cofactor has absolute value at most one, while the determinant has absolute value at least one; hence every entry of $M_{v,v'}^{-1}$ has absolute value at most one. Define

$$\mathfrak{R} := M_{v,v'}^{-1}(Q-D_A(\zeta)), \qquad J_h := \pi_h(\mathfrak{R}) \quad (h=1,2).$$

Each J_h is an interval of length at most $2C_Q w$. Indeed, if $z, z' \in Q$ and ρ_h is row h of $M_{v,v'}^{-1}$, then

$$|\rho_h(z - z')| \le \|\rho_h\|_1 \max_{k=1,2} |z_k - z'_k| \le 2C_Q w.$$

Let H_i, H_j denote the two individual rank gaps in (6.14); when $j = i+1$, both equal (b_i, b_{i-1}), while for $j \ge i+2$ they are

$$H_i = (b_i, b_{i-1}), \qquad H_j = (b_{j-1}, b_{j-2}).$$

Taking closures only enlarges the sets, so define

$$I_i := \overline{J_1 \cap H_i \cap [\ell_i, 1] \cap [T^{\vartheta-1}, 1]}, \qquad I_j := \overline{J_2 \cap H_j \cap [\ell_j, 1] \cap [T^{\vartheta-1}, 1]}.$$

If either intersection is empty, the corresponding term in (6.18) is zero. Otherwise $I_i, I_j \subseteq [T^{\vartheta-1}, 1]$, both have length at most $2C_Q w$, and

$$\inf I_i \ge \ell_i, \qquad \inf I_j \ge \ell_j.$$

Every pair retained by (6.18) lies in the Cartesian product determined by I_i, I_j. Hence

$$\begin{aligned} &\sum_{(p,q)\in\mathcal{G}_{ij}(A)} \frac{1}{pq} \mathbf{1}_{\mathcal{P}_T(A\cup\{p,q\})} \mathbf{1}_{\{D_A(\zeta)+u_p v+u_q v'\in Q\}} \\ &\quad \le \Bigg(\sum_{\substack{p\in\mathcal{B}(T)\\ u_p\in I_i}} \frac{1}{p}\Bigg)\Bigg(\sum_{\substack{q\in\mathcal{B}(T)\\ u_q\in I_j}} \frac{1}{q}\Bigg) \ll_{C_Q} \frac{w^2}{\ell_i \ell_j}. \end{aligned} \tag{6.20}$$

The product on the right deliberately adds pairs with $p = q$, primes belonging to A, the wrong order $u_p \le u_q$ in the adjacent-rank case, and points of $J_1 \times J_2$ outside the parallelogram $\mathfrak{R}$. It also discards the profile restriction. All summands are nonnegative, so each enlargement has the required upper-bound direction.

The final bound in (6.20) is uniform in A, i, j, v, v', ζ. The mark-counting sublemma therefore allows the mark sums in (6.18) to be performed before the remaining restrictions are discarded:

$$\begin{aligned} &\mathbb{P}(r(\xi) = 2,\ \mathcal{P}_T(S),\ D \in Q) \\ &\quad \ll_{C_Q} w^2 \sum_{A\subseteq\mathcal{B}(T)} \mu(A) \sum_{1\le i<j\le|A|+2} \frac{16\, 3^{-(i+j)}}{\ell_i \ell_j} \\ &\quad \le 16Cw^2 \Bigg(\sum_{k\ge1} \frac{3^{-k}}{\ell_k}\Bigg)^2 \sum_{A\subseteq\mathcal{B}(T)} \mu(A) \ll_{C_Q} w^2, \end{aligned} \tag{6.21}$$

where (6.10) and $\sum_A \mu(A) = 1$ were used.

Rank one. Let p be the canonical first nonzero pivot and let i be its rank. After deleting p, put $A = S \setminus \{p\}$ and $K = |A| + 1$. Order A with weights $b_1 > \cdots > b_{K-1}$ and put $b_0 = +\infty$, $b_K = 0$. The pivot has rank i exactly when

$$u_p \in G_i(A) := (b_i, b_{i-1}).$$

For $v \ne 0$, let $\mathcal{C}_i^{(1)}(A; v)$ consist of the residual marks $\zeta = (\zeta_1, \dots, \zeta_{K-1})$ satisfying

$$\zeta_k = 0 \quad (k < i), \qquad \zeta_k \in L(v) \quad (k \ge i).$$

When $A = \varnothing$, this residual list is empty and the endpoint conventions still apply. Put

$$D_A(\zeta) := \sum_{k=1}^{K-1} b_k \zeta_k.$$

Let $t_1, \ldots, t_K$ denote the reconstructed decreasing-weight order. It is uniquely

$$(t_1, \ldots, t_K) = (r_1, \ldots, r_{i-1}, p, r_i, \ldots, r_{K-1}), \tag{6.22}$$

which is equivalent to $b_{i-1} > u_p > b_i$. For an ordered marked support (S, ξ), let $\mathcal{E}_i^v(S, \xi; p)$ denote the event

$$p = t_i, \qquad \xi_{t_k} = 0 \ (k < i), \qquad \xi_p = v \neq 0, \qquad \xi_{t_k} \in L(v) \ (k > i).$$

Then, pointwise,

$$\mathbf{1}_{\{r(\xi)=1\}} = \sum_{p \in S} \sum_{i=1}^{K} \sum_{v \in \mathcal{M} \setminus \{0\}} \mathbf{1}_{\mathcal{E}_i^v(S, \xi; p)}. \tag{6.23}$$

The unique nonzero summand uses the first nonzero marked prime. Deleting that prime from (6.22) makes residual indices $k < i$ the marks before the pivot and indices $k \geq i$ the marks after it. Thus the residual conditions are exactly those defining $\mathcal{C}_i^{(1)}(A; v)$, and in particular $|\mathcal{C}_i^{(1)}(A; v)| = 3^{K-i}$.

For $p \in S$, define

$$F_1(S; p) := 9^{-|S|} \sum_{\xi \in \mathcal{M}^S} \mathbf{1}_{\mathcal{P}_T(S)} \mathbf{1}_{\{\sum_{s \in S} u_s \xi_s \in Q\}} \sum_{i=1}^{|S|} \sum_{v \in \mathcal{M} \setminus \{0\}} \mathbf{1}_{\mathcal{E}_i^v(S, \xi; p)}.$$

Equation (6.23) gives

$$\mathbb{P}(r(\xi) = 1, \mathcal{P}_T(S), D \in Q) = \mathbb{E}_{S \sim \mu} \sum_{p \in S} F_1(S; p).$$

For $p \notin A$, the coefficient after deletion is exactly

$$\mu(A \cup \{p\}) 9^{-(|A|+1)} = \mu(A) \frac{1/(p+1)}{1 - 1/(p+1)} 9^{-(|A|+1)} = \frac{\mu(A)}{p} 9^{-(|A|+1)}.$$

The one-point case of Lemma 6.1, followed by the unique decomposition of the remaining marks as ζ, therefore gives the exact equality

$$\begin{aligned} &\mathbb{P}(r(\xi) = 1,\ \mathcal{P}_T(S),\ D \in Q) \\ &= \sum_{A \subseteq \mathcal{B}(T)} \mu(A) \sum_{i=1}^{|A|+1} 9^{-(|A|+1)} \sum_{v \in \mathcal{M} \setminus \{0\}} \sum_{\zeta \in \mathcal{C}_i^{(1)}(A; v)} \sum_{\substack{p \in \mathcal{B}(T) \setminus A \\ u_p \in G_i(A)}} \frac{1}{p} \mathbf{1}_{\mathcal{P}_T(A \cup \{p\})} \mathbf{1}_{\{D_A(\zeta) + u_p v \in Q\}}. \end{aligned} \tag{6.24}$$

Fix A, i, v, ζ and choose a coordinate h with $|v_h| = 1$. Write $Q = Q_1 \times Q_2$. The last indicator in (6.24) forces

$$u_p \in J_i^0 := v_h(Q_h - D_{A,h}(\zeta)),$$

an interval of length at most $C_Q w$. On the profile event, Lemma 6.2 also gives $u_p \geq \ell_i$. Intersect J_i^0 with the rank gap, the profile lower bound, and the prime band, and set

$$I_i^{(1)} := \overline{J_i^0 \cap G_i(A) \cap [\ell_i, 1] \cap [T^{\vartheta - 1}, 1]}.$$

If the intersection is empty, the term is zero. Otherwise $I_i^{(1)} \subseteq [T^{\vartheta-1}, 1]$ is an interval of length at most $C_Q w$ and $\inf I_i^{(1)} \geq \ell_i$. Therefore

$$\begin{aligned}
&\sum_{\substack{p\in\mathcal{B}(T)\setminus A\\ u_p\in G_i(A)}} \frac{1}{p}\mathbf{1}_{\mathcal{P}_T(A\cup\{p\})}\mathbf{1}_{\{D_A(\zeta)+u_p v\in Q\}} \\
&\qquad \leq \sum_{\substack{p\in\mathcal{B}(T)\\ u_p\in I_i^{(1)}}} \frac{1}{p} \ll_{C_Q} \frac{w}{\ell_i}.
\end{aligned} \tag{6.25}$$

The right side adds back primes in A and discards the profile and the unused coordinate condition from $D_A(\zeta) + u_p v \in Q$; nonnegativity gives the displayed upper-bound direction.

The profile event further implies $K \geq a\lambda_T$. We may now drop all of its other restrictions and sum the residual marks explicitly. The mark-counting sublemma gives

$$\begin{aligned}
&\mathbb{P}(r(\xi) = 1,\ \mathcal{P}_T(S),\ D \in Q) \\
&\leq Cw \sum_{A\subseteq\mathcal{B}(T)} \mu(A)\mathbf{1}_{\{|A|+1\geq a\lambda_T\}} \sum_{i=1}^{|A|+1} \frac{8\,3^{-(|A|+1+i)}}{\ell_i} \\
&\leq 8Cw\,3^{-a\lambda_T} \left(\sum_{i\geq 1}\frac{3^{-i}}{\ell_i}\right) \sum_{A\subseteq\mathcal{B}(T)} \mu(A) = o(w^2),
\end{aligned} \tag{6.26}$$

by (6.10), (6.9), and $\sum_A \mu(A) = 1$.

Rank zero. All K marks vanish, which has conditional probability 9^{-K}. The profile gives $K \geq a\lambda_T$, so

$$\mathbb{P}(r(\xi) = 0,\ \mathcal{P}_T(S),\ D \in Q) \leq 9^{-a\lambda_T} = o(w^2) \tag{6.27}$$

by (6.9). The three rank cases prove (6.11). □

The finite counts in (6.12) and (6.13), as well as the inverse-matrix bound for all 48 ordered noncollinear mark pairs, are independently enumerated by the companion verifier.

6.3 The missing-petal pivot

Lemma 6.4 (Canonical top missing-petal pivot)**.** *Fix a root support $S \subseteq \mathcal{B}(T)$. Independently select each $p \in \mathcal{B}(T) \setminus S$ with probability $q_p = 1/(3p)$, and let*

$$X = \sum_{\text{selected } p} u_p.$$

For every interval I of length at most $C_X w$,

$$\mathbb{P}\big(X \in I,\ N_x(R_0) \geq 1 \mid S\big) \ll_{C_X} w, \tag{6.28}$$

uniformly in S and I. For two conditionally independent completions and two target intervals fixed after the root data have been exposed, the corresponding joint probability is $O_{C_X}(w^2)$.

Proof. On the event in (6.28), choose canonically the selected prime p_* of largest weight among those of depth at most R_0, and delete it. Let ν be the product law with parameters q_p on $\mathcal{B}(T) \setminus S$. The one-point case of Lemma 6.1 gives the exact identity

$$\begin{aligned}
&\mathbb{P}\big(X \in I,\ N_x(R_0) \geq 1 \mid S\big) \\
&\quad = \sum_{A\subseteq\mathcal{B}(T)\setminus S} \nu(A) \sum_{p\in(\mathcal{B}(T)\setminus S)\setminus A} \frac{q_p}{1-q_p}\mathbf{1}_{\{p \text{ is canonical}\}}\mathbf{1}_{\{u_p+\sum_{r\in A} u_r\in I\}}.
\end{aligned} \tag{6.29}$$

There is exactly one canonical point on the event, so no multiplicity is present in this identity.

The canonical restriction gives $u_p \geq \mathrm{e}^{-R_0}$. For fixed A, the last indicator in (6.29) confines u_p to a translate of I. Moreover,

$$\frac{q_p}{1-q_p} = \frac{1}{3p-1} \leq \frac{1}{p}. \tag{6.30}$$

Discard the remaining canonical order restriction and apply Lemma 5.1 on $[\mathrm{e}^{-R_0}, 1]$. The inner sum in (6.29) is $O_{C_X}(w)$ uniformly in A, S. Summing $\nu(A)$ proves (6.28).

Given the root data, the two missing-x processes are independent, so the two one-copy estimates multiply. The product already includes the possibility that both copies select the same prime; no diagonal correction is needed. □

6.4 The second moment and rooted balance

Corollary 6.5 (Rooted second moment)**.** *For the function g_T in (5.15),*

$$\mathbb{E}_{S\sim\mu_{\mathcal{B}(T)}} g_T(S)^2 = \mathbb{P}(B_T^{(1)} \cap B_T^{(2)}) \ll w^4. \tag{6.31}$$

Proof. Let $B_T^{(1)}$ and $B_T^{(2)}$ be formed from two independent completions conditional on the same root S. Then, conditionally on S,

$$\mathbb{P}(B_T^{(1)} \cap B_T^{(2)} \mid S) = \mathbb{P}(B_T \mid S)^2 = g_T(S)^2.$$

Taking expectations proves the equality in (6.31).

Let

$$\mathcal{F}_{\mathrm{root}} := \sigma\big(S, (\xi_p)_{p\in S}\big),$$

where the two coordinates of ξ_p record the two complete colourings of the root support as in (6.3). Thus $\mathcal{F}_{\mathrm{root}}$ reveals the common root and both root colourings, but neither missing-x process. Define the root-good event

$$\mathcal{R}_T := \{N_S(s) \geq as \text{ for every } R_0 \leq s \leq \lambda_T\} \cap \{D \in [-w,w]^2\}.$$

It is $\mathcal{F}_{\mathrm{root}}$-measurable. If both completions satisfy B_T, then, in completion h, the root labels c, y, z partition S. Consequently, for $R_0 \leq s \leq \lambda_T$,

$$N_S(s) = N_{c,h}(s) + N_{y,h}(s) + N_{z,h}(s) \geq 3\lfloor \alpha s\rfloor \geq as$$

by (5.13). Moreover $D_h = Y_h - Z_h \in [-w,w]$. Hence $B_T^{(1)} \cap B_T^{(2)} \subseteq \mathcal{R}_T$, and Lemma 6.3, with $Q = [-w,w]^2$, gives

$$\mathbb{P}(\mathcal{R}_T) \leq C_0 w^2 \tag{6.32}$$

for an absolute constant C_0.

For $h = 1, 2$, let $\mathcal{X}_h \subseteq \mathcal{B}(T) \setminus S$ be the primes assigned to the missing x-petal in completion h. By (6.2), for every fixed realization of $\mathcal{F}_{\mathrm{root}}$ the indicators

$$\Big(\mathbf{1}_{\{p\in\mathcal{X}_h\}} : p \in \mathcal{B}(T)\setminus S,\ h\in\{1,2\}\Big)$$

are mutually independent and satisfy

$$\mathbb{P}(p \in \mathcal{X}_h \mid \mathcal{F}_{\mathrm{root}}) = \frac{1}{3p}.$$

Indeed, conditional on S, the two completions are independent, and within each completion the colouring variables on S and the missing-x variables on $\mathcal{B}(T)\setminus S$ are disjoint independent factors. Thus further conditioning on the two root colourings leaves the two missing processes independent and leaves each with exactly the product law used in Lemma 6.4.

The variables Y_h, Z_h are $\mathcal{F}_{\mathrm{root}}$-measurable. On $\mathcal{R}_T$, define the corresponding $\mathcal{F}_{\mathrm{root}}$-measurable interval

$$I_h := [\max(Y_h, Z_h) - w,\ \min(Y_h, Z_h) + w], \tag{6.33}$$

and define $I_h := \{0\}$ off $\mathcal{R}_T$. On $\mathcal{R}_T$ one has $|Y_h - Z_h| \le w$, so I_h is nonempty and

$$|I_h| = 2w - |Y_h - Z_h| \le 2w.$$

The range condition in $B_T^{(h)}$ forces $X_h \in I_h$, whereas its x-profile at R_0 forces

$$N_{x,h}(R_0) \ge \lfloor \alpha R_0 \rfloor = 24 \ge 1.$$

Put

$$E_h := \{X_h \in I_h,\ N_{x,h}(R_0) \ge 1\}.$$

We have the explicit event inclusion

$$B_T^{(1)} \cap B_T^{(2)} \subseteq \mathcal{R}_T \cap E_1 \cap E_2. \tag{6.34}$$

All the other window and profile requirements in $B_T^{(1)} \cap B_T^{(2)}$ have been discarded on the right; because an upper bound is sought, this only enlarges the event.

After conditioning on $\mathcal{F}_{\mathrm{root}}$, the root S and the two intervals I_h are fixed. On $\mathcal{R}_T$ their lengths are at most $2w$. The uniform one-copy estimate in Lemma 6.4 and the conditional independence of $\mathcal{X}_1, \mathcal{X}_2$ therefore give

$$\mathbb{P}(E_1 \cap E_2 \mid \mathcal{F}_{\mathrm{root}}) = \prod_{h=1}^{2} \mathbb{P}(E_h \mid \mathcal{F}_{\mathrm{root}}) \le C_1^2 w^2$$

on $\mathcal{R}_T$. Finally, (6.34), the tower property, and (6.32) yield

$$\begin{aligned} \mathbb{P}(B_T^{(1)} \cap B_T^{(2)}) &\le \mathbb{E}[\mathbf{1}_{\mathcal{R}_T} \mathbb{P}(E_1 \cap E_2 \mid \mathcal{F}_{\mathrm{root}})] \\ &\le C_1^2 w^2 \mathbb{P}(\mathcal{R}_T) \le C_0 C_1^2 w^4. \end{aligned}$$

This proves the upper bound in (6.31). □

Proposition 6.6 (Uniform rooted category balance)**.** *There is a constant $K < \infty$ with the following uniformity property. For every function $\eta(T)$ satisfying (5.5), and all sufficiently large T,*

$$\mathbb{E}_{S \sim \mu_{\mathcal{B}(T)}} \left(\frac{g_T(S)}{\beta_T} \right)^2 \le K. \tag{6.35}$$

The constant K is independent of T, of η, of the choice of a band at a later stage, and of the later integers H and M.

Proof. Corollary 6.5 gives a constant C_2 such that the numerator in

$$\mathbb{E} \left(\frac{g_T}{\beta_T} \right)^2 = \frac{\mathbb{E} g_T^2}{\beta_T^2}$$

is at most $C_2 w^4$. Proposition 5.5 gives a constant $c_1 > 0$, independent of T and η, such that $\beta_T \ge c_1 w^2$. Therefore

$$\mathbb{E} \left(\frac{g_T}{\beta_T} \right)^2 \le \frac{C_2}{c_1^2}.$$

Take $K = C_2/c_1^2$. All structural and buffer constants were fixed before T and η varied, so this K has the asserted uniformity. □

7 Flattening by alternative bands

The conditioned law on one band has a uniformly bounded L^2 density, but Proposition 4.2 needs every word marginal to be close to the ambient product law in L^1. We obtain that stronger conclusion by letting each cube coordinate choose its active band uniformly among many independent alternatives.

Proposition 7.1 (Alternative-band flattening). *Fix positive integers H, M. For each $1 \le i \le H$, let*

$$\mathcal{B}_{i,1}, \dots, \mathcal{B}_{i,M}$$

be mutually disjoint bands of the form $\mathcal{B}_{i,j} = \mathcal{B}(T_{i,j})$. Choose a balancing parameter $\eta_{i,j} > 0$ for each band, put

$$w_{i,j} := \frac{\eta_{i,j}}{T_{i,j}}, \tag{7.1}$$

and let $B_{i,j}$ denote the event (5.14) with (T, η, w) replaced by $(T_{i,j}, \eta_{i,j}, w_{i,j})$. Assume that every $T_{i,j}$ is in the valid range of Sections 5–6, that the same constant K in (6.35) works on all these bands, and that $\eta_{i,j} \le \eta_*$ for all i, j. Let E be any finite prime set containing all HM bands.

Then there is a probability law on dimension-H pair-product cubes in 2^E such that:

(i) every petal is nonempty;

(ii) for every word ω, if G_ω is the density of its marginal relative to μ_E, then

$$\mathbb{E}_{\mu_E} |G_\omega - 1| \le H \sqrt{\frac{K}{M}}; \tag{7.2}$$

(iii) every sampled cube satisfies

$$\left| \log d(S_\omega) - \frac{1}{3^H} \sum_{\tau \in \mathbb{F}_3^H} \log d(S_\tau) \right| \le H \eta_* \tag{7.3}$$

for every word ω.

Proof. For every group i, choose independently an active index J_i uniformly from $\{1, \dots, M\}$. On the active band $\mathcal{B}_{i,J_i}$, sample the five-state law conditioned on its event B_{i,J_i}. Put its c-labelled primes into the common support and use its x, y, z labels as the three petals of coordinate i. On every inactive band, select each prime with probability $1/(p+1)$ and put every selected prime into the common support. Do the same for primes of E outside all candidate bands. All choices on disjoint bands are independent.

The bands are disjoint, so petals from different coordinates are disjoint. Within an active band, the five-state law gives mutually disjoint petals. Finally, the profile at R_0 supplies at least $\lfloor \alpha R_0 \rfloor = 24$ points in each petal. Thus the construction is a genuine pair-product cube and proves (i).

We next compute a word marginal. Fix one group i and one of the three coordinate states selected by the word. For candidate band j, write $g_{i,j}(S) := \mathbb{P}(B_{i,j} \mid S^{(x)} = S)$ and $\beta_{i,j} := \mathbb{P}(B_{i,j})$, as in (5.15).

We verify here that this density is the same for all three states. For $r \in \{x, y, z\}$, put

$$g_{i,j}^{(r)}(S) := \mathbb{P}(B_{i,j} \mid S^{(r)} = S), \qquad S \subseteq \mathcal{B}_{i,j},$$

so $g_{i,j} = g_{i,j}^{(x)}$. A permutation of the labels x, y, z that sends x to r preserves both the five-state law and $B_{i,j}$ and sends $S^{(x)}$ to $S^{(r)}$. Hence, for every $S \subseteq \mathcal{B}_{i,j}$,

$$\mathbb{P}(B_{i,j} \cap \{S^{(r)} = S\}) = \mathbb{P}(B_{i,j} \cap \{S^{(x)} = S\}).$$

Every state has marginal law $\mu_{\mathcal{B}_{i,j}}$, so

$$\mathbb{P}(S^{(r)} = S) = \mathbb{P}(S^{(x)} = S) = \mu_{\mathcal{B}_{i,j}}(S) > 0.$$

Dividing the preceding joint-probability identity by this common marginal gives

$$g_{i,j}^{(r)}(S) = g_{i,j}^{(x)}(S) = g_{i,j}(S).$$

Bayes' formula therefore gives, for each of the three possible states,

$$\frac{\mathrm{d}\,\mathrm{Law}(S^{(r)} \mid B_{i,j})}{\mathrm{d}\mu_{\mathcal{B}_{i,j}}}(S) = \frac{g_{i,j}(S)}{\beta_{i,j}}. \tag{7.4}$$

Relative to the product support law on all M bands in group i, the word-selected state density is consequently exactly

$$G_i = \frac{1}{M}\sum_{j=1}^{M} \frac{g_{i,j}}{\beta_{i,j}}. \tag{7.5}$$

Indeed, conditional on $J_i = j$, every inactive band has density one, while the active state has density $g_{i,j}/\beta_{i,j}$ by (7.4); averaging over j gives (7.5).

Under the ambient product law, the M summands in (7.5) are independent, have mean one, and have second moment at most K. Cross terms vanish after centering, so

$$\mathbb{E}(G_i - 1)^2 = \frac{1}{M^2}\sum_{j=1}^{M} \mathbb{E}\left(\frac{g_{i,j}}{\beta_{i,j}} - 1\right)^2 = \frac{1}{M^2}\sum_{j=1}^{M}\left[\mathbb{E}\left(\frac{g_{i,j}}{\beta_{i,j}}\right)^2 - 1\right] \le \frac{K}{M}. \tag{7.6}$$

The groups are independent. Therefore the density of the full word is

$$G_\omega = \prod_{i=1}^{H} G_i.$$

The telescoping identity

$$\prod_{i=1}^{H} G_i - 1 = \sum_{i=1}^{H}(G_i - 1)\prod_{k<i} G_k$$

together with independence, $\mathbb{E}G_k = 1$, and Cauchy–Schwarz gives

$$\begin{aligned}\mathbb{E}|G_\omega - 1| &\le \sum_{i=1}^{H} \mathbb{E}|G_i - 1| \prod_{k<i} \mathbb{E}G_k \\ &\le \sum_{i=1}^{H} \sqrt{\mathbb{E}(G_i - 1)^2} \le H\sqrt{\frac{K}{M}}.\end{aligned}$$

This proves (ii).

It remains to check balance. On an active band (i,j) of scale $T_{i,j}$, after removing the common c-contribution, the three normalized logarithmic contributions are

$$Y + Z, \qquad X + Z, \qquad X + Y.$$

Their pairwise differences are $X - Y, X - Z, Y - Z$. On $B_{i,j}$ each has absolute value at most $w_{i,j}$, so the corresponding actual logarithms have range at most $T_{i,j}w_{i,j} = \eta_{i,j} \le \eta_*$. Two words differ in at most this amount in each of the H coordinates; hence all word logarithms have range at most $H\eta_*$. Their arithmetic mean is the logarithm of the geometric mean of the word products and lies between the minimum and maximum. This proves (7.3). □

8 Band placement and completion of the proof

We now place finitely many candidate bands inside the prime set used in the squarefree capacity. The order of limits is

$$H \text{ first}, \qquad M \text{ second}, \qquad y \to \infty \text{ last}.$$

For $y > \mathrm{e}^{\mathrm{e}}$, put

$$E_y := \{p : \log y < p \le y\}, \qquad L := \log y. \tag{8.1}$$

Fix H, M, and for $1 \le j \le HM$ define

$$b_j := \frac{1}{2}\vartheta^{2(j-1)} \tag{8.2}$$

and

$$\mathcal{B}_j := \left\{p : \mathrm{e}^{L^{\vartheta b_j}} < p \le \mathrm{e}^{L^{b_j}}\right\}. \tag{8.3}$$

Assign bands $(i-1)M+1, \ldots, iM$ to the ith coordinate group.

These bands are pairwise disjoint. Indeed,

$$b_{j+1} = \vartheta^2 b_j < \vartheta b_j,$$

so the upper endpoint of $\mathcal{B}_{j+1}$ is smaller than the lower endpoint of $\mathcal{B}_j$. They also lie in E_y for all sufficiently large y. Their upper endpoints are below y because $b_j \le 1/2 < 1$, while for the smallest fixed exponent $b_{HM} > 0$,

$$L^{\vartheta b_{HM}} > \log L$$

eventually. Thus the lower endpoint of every band exceeds $\mathrm{e}^{\log L} = L = \log y$.

Use the common balancing parameter

$$\eta_y := \frac{1}{\log\log y}. \tag{8.4}$$

For band j, its scale in (5.3) is $T_j = L^{b_j}$; define

$$w_j := \frac{\eta_y}{T_j}. \tag{8.5}$$

Thus, in the notation of Proposition 7.1,

$$\begin{aligned} \mathcal{B}_{i,j} &:= \mathcal{B}_{(i-1)M+j}, & T_{i,j} &:= T_{(i-1)M+j}, \\ w_{i,j} &:= w_{(i-1)M+j}, & \eta_{i,j} &:= \eta_y \end{aligned} \qquad (1 \le i \le H,\ 1 \le j \le M). \tag{8.6}$$

The slow-decay condition holds because

$$\eta_y = \frac{b_j}{\log T_j} \longrightarrow 0, \qquad \frac{\log(1/\eta_y)}{\log T_j} = \frac{\log\log\log y}{b_j \log\log y} \longrightarrow 0.$$

Since HM is fixed and $b_{HM} > 0$, one threshold $y_0(H, M)$ makes every estimate in Sections 5–6 valid simultaneously on all candidate bands; increase this threshold, if necessary, so that $H\eta_y \le 1$, as required by Proposition 4.2. Proposition 6.6 supplies the same constant K for every band. Moreover, on band j,

$$T_j w_j = \eta_y,$$

so its actual logarithmic balance error is at most η_y.

Apply Proposition 7.1 on the finite set E_y. Primes in inactive bands and outside all candidate bands are sampled into the common support with their product probabilities, exactly as in that proposition. Thus the ambient word law is μ_{E_y}. The resulting cube law has

$$\varepsilon_{H,M} = H\sqrt{\frac{K}{M}}, \qquad \zeta_y = H\eta_y$$

in Proposition 4.2. Hence

$$\frac{I^{\mathrm{sf}}_{E_y}}{Z_{E_y}} \leq 3\kappa^{H}_{\mathrm{cap}} + C\left(H\sqrt{\frac{K}{M}} + H\eta_y\right) \tag{8.7}$$

for an absolute constant C.

We now execute the stated order of limits. Given $\varepsilon > 0$, first choose H so that

$$3\kappa^{H}_{\mathrm{cap}} < \frac{\varepsilon}{3};$$

this is possible because $\kappa_{\mathrm{cap}} < 1$. Keeping H fixed, choose M so large that

$$CH\sqrt{\frac{K}{M}} < \frac{\varepsilon}{3}.$$

Finally keep H, M fixed and let $y \to \infty$ until $CH\eta_y < \varepsilon/3$. Equation (8.7) then gives, for every sufficiently large y,

$$0 \leq \frac{I^{\mathrm{sf}}_{E_y}}{Z_{E_y}} < \varepsilon.$$

Thus the limsup is at most ε. Since $\varepsilon > 0$ was arbitrary and the ratio is nonnegative, we have proved

$$\frac{I^{\mathrm{sf}}_{E_y}}{Z_{E_y}} \longrightarrow 0. \tag{8.8}$$

The least prime in E_y exceeds $\log y \to \infty$. Thus (8.8) verifies both hypotheses of Corollary 3.3. That corollary yields $f(N) = o(N)$ and completes the proof of Theorem 1.1.

9 Conclusion

The proof separates the problem into three reusable interfaces. The finite-prime envelope and unit-exponent deletion convert integer density into a normalized squarefree capacity. Joint-prefix transference converts flat balanced pair-product cubes into a cap-set saving. Finally, the prime-band construction provides those cubes: the Poisson anchor argument gives the first moment, the discrete canonical pivots give the matching second moment, and alternative bands flatten the conditioned marginals. The order H, then M, then y closes these interfaces without requiring any quantitative rate for $f(N)/N$.

A Numerical Verification

A companion numerical verifier is supplied with the paper as `numerical_verifier.py`. It can be run with

```
python3 numerical_verifier.py.
```

The verifier certifies the explicit finite numerical and combinatorial comparisons appearing in the proof. More precisely, it verifies:

- the squarefree product-law and unit-deletion normalizations;
- the substitution in the cap-set constant;
- the explicit structural, survival, buffer, and anchor constants;

- the five-state and conditional missing-petal probabilities;
- atom-by-atom finite tests of factorial insertion;
- all nine-mark rank counts and all noncollinear 2×2 mark matrices;
- the elementary balance identities and separation of the candidate bands.

The program uses only the Python standard library. Rational identities are checked exactly with `Fraction` arithmetic, and comparisons involving $\log 3$ use a rigorous rational interval obtained from the atanh series with an explicit remainder bound. No binary floating-point arithmetic or random simulation is used. The optional flag `-self-test` additionally exhausts small valuation models, squarefree set systems, and pair-product cubes as regression checks.

References

[1] H. L. Abbott and B. Gardner, An extremal problem in number theory, *Canadian Mathematical Bulletin* **10** (1967), 173–177. https://doi.org/10.4153/CMB-1967-015-8

[2] T. F. Bloom, *Erdős Problem #536*, Erdős Problems, accessed 13 July 2026. https://www.erdosproblems.com/536

[3] J. S. Ellenberg and D. Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Annals of Mathematics* **185** (2017), 339–343. https://doi.org/10.4007/annals.2017.185.1.8

[4] P. Erdős, On a problem in elementary number theory and a combinatorial problem, *Mathematics of Computation* **18** (1964), 644–646. https://doi.org/10.2307/2002950

[5] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[6] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[7] H. L. Montgomery and R. C. Vaughan, The large sieve, *Mathematika* **20** (1973), 119–134. https://doi.org/10.1112/S0025579300004708