A Tilted Residue-Class Construction for Long Prime-Free Intervals

GPT 5.6 Sol

25 August 2026

Abstract

We study the finite residue-class covering problem that produces long intervals containing no primes. A preliminary random sieve is biased toward the zero residue by a small Rankin-type tilt. For a squarefree rough composite, the tilt gives the exact survival law $Av^{-\tau}$; this permits an all-fiber block construction for composites, while a Maynard weight and a quantitative hypergraph covering theorem handle the surviving primes. The resulting bounds are

$$Y(X) \gg X\frac{\log X}{\log_3 X}, \qquad G(T) \gg \frac{\log T\,\log_2 T}{\log_4 T}.$$

This edition supplies the elementary rough-number count, all local probability factors, witness counts, normalizer caps, concentration calculations, assembly, and Chinese-remainder transfer in full detail. The only external inputs are classical prime-distribution theorems and two precisely stated results of Ford–Green–Konyagin–Maynard–Tao; every hypothesis of those results is verified explicitly.

Contents

1. Introduction and road map

Throughout, log means the natural logarithm. For $j \geq 1$, define the iterated logarithms recursively by

$$\log_1 u := \log u, \qquad \log_{j+1} u := \log(\log_j u).$$

All assertions involving $\log_3 u$ or $\log_4 u$ are made only for sufficiently large u, so that these quantities are defined and positive.

Let

$$2 = p_1 < p_2 < p_3 < \cdots$$

be the sequence of primes. For $T \geq 3$, define

$$G(T) := \max_{p_{n+1} \leq T} (p_{n+1} - p_n).$$

Thus $G(T)$ is the largest gap whose right-hand endpoint is a prime not exceeding T.

We also need a finite covering problem. For a real number $z \geq 2$, let $Y(z)$ be the largest nonnegative integer N for which one can choose one residue class $a_p \pmod p$ for every prime $p \leq z$ in such a way that

$$\{1, 2, \ldots, N\} \subseteq \bigcup_{p \leq z} (a_p + p\mathbb{Z}). \tag{1.1}$$

In plain language, every integer from 1 through N must belong to at least one of the selected residue classes.

This maximum really exists. Put

$$P(z) := \prod_{p \leq z} p.$$

There are only finitely many ways to choose the residues $(a_p)_{p \leq z}$. For any one such choice, choose a residue $b_p \pmod p$ different from $a_p \pmod p$ for every $p \leq z$. The Chinese remainder theorem gives an integer b satisfying $b \equiv b_p \pmod p$ simultaneously for all these primes. Every integer congruent to $b \pmod{P(z)}$ avoids all of the selected classes. Consequently, no $P(z)$ consecutive integers can all be covered. Thus $Y(z) \leq P(z) - 1$, and the maximum in (1.1) is finite.

The two target statements are as follows.

Covering theorem. There is an effective constant $c_0 > 0$ such that, for every sufficiently large X,

$$Y(X) \geq c_0 X \frac{\log X}{\log_3 X}. \tag{1.2}$$

Prime-gap corollary. There is an effective constant $c_1 > 0$ such that, for every sufficiently large T,

$$G(T) \geq c_1 \frac{\log T \, \log_2 T}{\log_4 T}. \tag{1.3}$$

The logic of the proof is best separated into three parts.

1. The earlier sections construct residue classes for primes up to x that cover all but $O(x/\log x)$ integers in a much longer interval $(x, y]$, where $y \asymp x \log x / \log_3 x$. This is the deep part: it contains the preliminary tilted sieve, the composite covering, the Maynard weight, and the hypergraph covering argument.
2. The assembly argument assigns each of the remaining $O(x/\log x)$ integers to its own new prime. It then translates the covered interval $(x, y]$ to an initial interval $[1, N]$. This proves (1.2).
3. A completely elementary Chinese-remainder argument turns (1.2) into a run of composite integers below T. This proves (1.3).

The claimed bound (1.3) is substantially stronger than the established Ford–Green–Konyagin–Maynard–Tao scale

$$\frac{\log T \, \log_2 T \, \log_4 T}{\log_3 T}.$$

Indeed, the ratio of (1.3) to that scale is

$$\frac{\log_3 T}{(\log_4 T)^2} \longrightarrow \infty.$$

This is not a contradiction with any known upper bound, but it makes the dependency statement important: the new unconditional result follows only if every input from the deep first part is proved with the stated uniformity. The accompanying assembly and transfer arguments do not repair, replace, or conceal any missing proof of those inputs.

Reader's guide and prerequisites

This chapter supplies every piece of notation and every elementary estimate used later. A reader who has taken undergraduate courses in number theory, real analysis, and probability should be able to begin here without having studied sieve theory.

There are, however, four established results whose full proofs would each take us far outside the argument being audited:

1. the prime number theorem and Mertens' product theorem;
2. the Bombieri–Vinogradov distribution theorem for primes, in an effective form with one possible exceptional prime removed;
3. the parameter-uniform Maynard weight theorem derived from that distribution theorem;
4. the quantitative hypergraph covering theorem proved by Ford, Green, Konyagin, Maynard, and Tao.

The first two are classical theorems of analytic number theory. The last two are stated in full, in the precise special forms used here, in the prime-layer chapter. Every hypothesis is then verified line by line. Thus no *sieve-theory vocabulary or technique* is presumed of the reader, and there is no appeal to an unnamed "standard sieve estimate." As in most research papers, the proofs of the explicitly identified input theorems are cited rather than reproduced.

This distinction matters. The supplied manuscript's new work is the tilted preliminary law, its correlation estimates, and the way the composite and prime layers are selected simultaneously. Those parts are proved in full below.

Basic notation

All logarithms are natural. For a sufficiently large real number x, put

$$L = \log x, \qquad L_2 = \log\log x, \qquad L_3 = \log\log\log x,$$

and define $L_j = \log_j x$ recursively for $j \geq 4$. Every statement containing these symbols is asserted only when x is so large that the indicated iterated logarithms are positive.

For two positive quantities $f = f(x)$ and $g = g(x)$:

- $f = O(g)$, or $f \ll g$, means that $|f| \leq Cg$ for an absolute constant C once x is large;
- $f \asymp g$ means both $f \ll g$ and $g \ll f$;
- $f = o(g)$ means $f/g \to 0$;
- $x^{o(1)}$ denotes a positive quantity whose logarithm is $o(\log x)$.

The letter $\mathbb{P}$ denotes the set of primes as well as probability; context always distinguishes the two. We use $\mathbf{1}_E$ for the indicator of an event E. If $n \geq 1$, then

- $\omega(n)$ is the number of distinct prime factors of n;
- $P^-(n)$ is the least prime factor of n, with $P^-(1) = +\infty$;
- n is *squarefree* if no prime square divides n;
- n is *w-rough* if $P^-(n) > w$.

Five elementary probability facts

We record the exact forms needed later.

Union bound. For events $E_1, \dots, E_m$,

$$\mathbb{P}\left(\bigcup_{j=1}^{m} E_j\right) \leq \sum_{j=1}^{m} \mathbb{P}(E_j).$$

Indeed, pointwise $\mathbf{1}_{\cup E_j} \leq \sum_j \mathbf{1}_{E_j}$; take expectations.

Markov's inequality. If $X \geq 0$ and $a > 0$, then

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}X}{a}.$$

This follows from $X \geq a\mathbf{1}_{X \geq a}$.

Chebyshev's inequality. If X has finite variance and $a > 0$, then

$$\mathbb{P}(|X - \mathbb{E}X| \geq a) \leq \frac{\operatorname{Var} X}{a^2}.$$

Apply Markov to the nonnegative random variable $(X - \mathbb{E}X)^2$.

Conditional second-moment form. If E is an event of positive probability and

$$\mathbb{E}\big((X-1)^2 \mid E\big) \leq \varepsilon,$$

The conclusion is

$$\mathbb{P}(E \cap \{X < 1/2\}) \leq 4\varepsilon\, \mathbb{P}(E).$$

On $X < 1/2$ one has $(X-1)^2 > 1/4$, so conditional Markov gives the result.

Weighted Bernstein inequality. Let $\xi_1, \dots, \xi_m$ be independent Bernoulli random variables of mean θ, and let $0 \leq a_j \leq M$. Put $S = \sum_j a_j$. Then, for every $u > 0$,

$$\mathbb{P}\left(\left|\sum_j (\xi_j - \theta)a_j\right| \geq u\right) \leq 2\exp\left(-\frac{u^2}{2\theta MS + 2Mu/3}\right).$$

For completeness, here is the standard moment-generating-function proof. If $0 < \lambda < 3/M$, the inequality $e^v - 1 - v \le v^2/[2(1 - v/3)]$, valid for $0 \le v < 3$, gives

$$\begin{aligned}\mathbb{E}e^{\lambda(\xi_j-\theta)a_j} &= e^{-\lambda\theta a_j}\left(1-\theta+\theta e^{\lambda a_j}\right)\\ &\le \exp\left(\frac{\theta\lambda^2 a_j^2}{2(1-\lambda M/3)}\right).\end{aligned}$$

Multiplying over j, using $\sum a_j^2 \le M\sum a_j = MS$, and applying Markov to the exponential gives

$$\mathbb{P}\left(\sum_j(\xi_j-\theta)a_j \ge u\right) \le \exp\left(-\lambda u + \frac{\theta\lambda^2 MS}{2(1-\lambda M/3)}\right).$$

Choosing $\lambda = u/(\theta MS + Mu/3)$ proves the upper-tail bound. Apply the same argument to $-\sum_j(\xi_j-\theta)a_j$ and add the two probabilities. The absolute value here is important: it corrects the one-sided display in the supplied manuscript.

Elementary analytic inequalities

We repeatedly use the following estimates. Their constants are absolute.

Logarithm estimate. If $|z| \le 1/2$, then

$$\log(1+z) = z + O(z^2), \qquad e^z = 1 + z + O(z^2).$$

To prove the first formula, write

$$\log(1+z) - z = -\int_0^z \frac{t}{1+t}\,dt;$$

the integrand has absolute value at most $2|t|$. The exponential estimate follows from Taylor's theorem with remainder.

A divisor-reciprocal estimate. If $d \ne 0$ and every prime divisor under consideration exceeds $w \ge 3$, then

$$\sum_{\substack{s\mid d\\ s>w}} \frac{1}{s} \le \frac{\log|d|}{w\log w}.$$

There are at most $\log|d|/\log w$ distinct primes $s > w$ dividing d, and each contributes at most $1/w$.

A prime-square tail. For $w \ge 2$,

$$\sum_{\substack{s>w\\ s\ \text{prime}}} \frac{1}{s^2} \le \sum_{n>w}\frac{1}{n^2} \le \frac{1}{w-1}.$$

The final inequality follows by comparison with $\int_{w-1}^{\infty} t^{-2}\,dt$.

Divisor expansion for a squarefree integer. If g is squarefree and $\tau \ge 0$, then

$$g^\tau = \prod_{s\mid g} s^\tau = \sum_{d\mid g}\prod_{s\mid d}(s^\tau - 1).$$

Expand the product $\prod_{s\mid g}[1 + (s^\tau - 1)]$. Each choice of a subset of the prime factors corresponds to one divisor $d \mid g$.

Rankin's trick. If $X \ge 0$, $z > 0$, and $a > 0$, then

$$\mathbf{1}_{X>z} \le (X/z)^a.$$

This is immediate: the right side is at least 1 when $X > z$. We shall use the same idea inside sums, replacing a hard lower bound on a divisor by a small negative power of that divisor.

Classical facts about primes

The proof uses the following named theorems in familiar forms.

Prime number theorem. As $z \to \infty$,

$$\pi(z) \sim \frac{z}{\log z}.$$

Uniformly when z_2/z_1 is bounded below by a fixed number greater than 1, this implies

$$\pi(z_2) - \pi(z_1) = (1 + o(1)) \int_{z_1}^{z_2} \frac{dt}{\log t}.$$

In all applications below, $\log z_1 \sim \log z_2$, so the integral equals $(1 + o(1))(z_2 - z_1)/\log z_1$.

Mertens' product theorem. As $z \to \infty$,

$$\prod_{\substack{s \le z \\ s \text{ prime}}} \left(1 - \frac{1}{s}\right) = \frac{e^{-\gamma} + o(1)}{\log z},$$

where γ is Euler's constant.

Mertens' reciprocal-prime estimate. As $z \to \infty$,

$$\sum_{\substack{s \le z \\ s \text{ prime}}} \frac{1}{s} = \log\log z + O(1).$$

Bertrand's postulate. For every integer $n > 1$, there is a prime strictly between n and $2n$.

Only the statements above are used. In particular, no unmentioned information about the spacing of primes is smuggled into the argument.

Constructing the admissible tuple

A finite set of distinct integers $\mathcal{H} = \{h_1, \dots, h_r\}$ is called *admissible* if, for every prime s, its residues do not fill all of $\mathbb{Z}/s\mathbb{Z}$. Equivalently, for each prime s there is some residue class modulo s avoided by every h_i.

The supplied manuscript asked the reader simply to "fix" an admissible r-tuple in $[1, 2r^2]$. Here is the missing construction.

Lemma (short admissible tuple). For all sufficiently large r, there are distinct integers

$$r < h_1 < \cdots < h_r \le 2r^2$$

forming an admissible r-tuple.

Proof. For all sufficiently large r, the prime number theorem gives

$$\pi(2r^2) - \pi(r) \sim \frac{2r^2}{\log(2r^2)} > r.$$

Choose $h_1, \dots, h_r$ to be any r primes in $(r, 2r^2]$. If $s \le r$ is prime, none of the h_i is congruent to 0 (mod s), so the residue 0 is missing. If $s > r$, the set has only $r < s$ elements and therefore cannot occupy all s residue classes. Thus the tuple is admissible. □

Why a residue covering produces a prime gap

For $X \ge 2$, define $Y(X)$ to be the largest $N \ge 0$ for which there are residues $a_p \pmod p$, one for every prime $p \le X$, such that every $1 \le h \le N$ lies in at least one of the chosen residue classes.

If $m \equiv -a_p \pmod p$ for every $p \le X$, then any covered h satisfies $p \mid m + h$. Choosing $m > X$ makes $m + h > p$, so $m + h$ is composite. The final chapter makes this Chinese-remainder argument quantitative. The entire probabilistic construction in the middle of the paper has one purpose: prove a large lower bound for $Y(X)$.

2. Parameters and the preliminary target sets

Throughout, $\mathbb{P}$ denotes the set of primes, all logarithms are natural, and

$$L := \log x, \qquad L_j := \log_j x \quad (j \geq 2). \tag{2.1}$$

Thus $L_2 = \log\log x$, $L_3 = \log\log\log x$, and so on. We use the prime number theorem and the following two standard consequences of it:

$$\prod_{p \leq z}\left(1 - \frac{1}{p}\right) = (e^{-\gamma} + o(1))\frac{1}{\log z}, \qquad \sum_{p \leq z}\frac{1}{p} = \log\log z + O(1), \qquad \sum_{p > z}\frac{1}{p^2} \ll \frac{1}{z\log z}. \tag{2.2}$$

The first formula is Mertens' product theorem. These are the only previously established analytic facts needed in Sections 2–5; in particular, we will not invoke a sieve theorem.

Fix

$$B_0 = 100, \qquad D > 2,$$

and later choose a sufficiently small fixed constant $c > 0$. Let $t > 1$ be the larger solution of

$$t - \log t = DL_3, \tag{2.3}$$

and define

$$\tau := \frac{t}{L}, \qquad H := \frac{cL}{t}, \qquad y := xH, \qquad w := L^{B_0}. \tag{2.4}$$

The function $u \mapsto u - \log u$ is strictly increasing for $u > 1$, so the stated solution is unique. Equation (2.3) gives

$$te^{-t} = e^{-DL_3} = L_2^{-D}, \quad \text{and hence} \quad e^{-t} = \frac{1}{tL_2^D}. \tag{2.5}$$

Also $t = DL_3 + \log t$. First this shows $t \asymp L_3$, and substituting this back into $\log t$ gives

$$t = DL_3 + L_4 + O(1), \qquad H \asymp \frac{L}{L_3}. \tag{2.6}$$

We record several consequences that will be used repeatedly:

$$\begin{aligned} &\log y = L + O(L_2), \qquad 2H < w, \\ &\tau\log w = \frac{B_0 tL_2}{L} = o(1), \qquad \tau\log H = O\!\left(\frac{tL_2}{L}\right) = o(1). \end{aligned} \tag{2.7}$$

All four statements follow directly from $t \asymp L_3$ and the fact that every fixed power of $\log x$ is smaller than x^{ε}.

For the later prime layer, put

$$r := \left\lfloor (\log(x/4))^{1/5} \right\rfloor. \tag{2.8}$$

We need an admissible r-tuple $\mathcal{H} = (h_1, \ldots, h_r) \subseteq [1, 2r^2]$. Recall that admissible means that, for every prime ℓ, the residues $h_1, \ldots, h_r \pmod{\ell}$ do not occupy every residue class. Here is an explicit existence proof. By the prime number theorem, for all large r there are at least r primes in $(r, 2r^2]$; choose any r of them as the h_i. If $\ell \leq r$, then none of the h_i is $0 \pmod{\ell}$, so the zero class is missing. If $\ell > r$, then r integers cannot occupy all ℓ classes. Thus this tuple is admissible.

Later we will split the primes in $(x/2, x]$ into disjoint sets $\mathcal{P}_C$ and $\mathcal{P}_Q$ satisfying

$$|\mathcal{P}_C| \asymp \frac{x}{L}, \qquad |\mathcal{P}_Q| \asymp \frac{x}{L}. \tag{2.9}$$

Only these size bounds, and the fact that every member of either set exceeds $x/2$, are used in Sections 4–5.

3. The tilted residue law

For every prime $s \le w$, choose the residue $a_s = 0$. For each prime $w < s \le x/2$, independently define

$$\beta_s := \frac{(s-1)s^{-\tau}}{s-1+s^{-\tau}} \tag{3.1}$$

and choose

$$a_s = \begin{cases} 0, & \text{with probability } 1-\beta_s, \\ b, & \text{with probability } \beta_s/(s-1) \quad \text{for each } b \in (\mathbb{Z}/s\mathbb{Z})^\times. \end{cases} \tag{3.2}$$

Since $0 < s^{-\tau} < 1$, one has $0 < \beta_s < 1$, and the probabilities in (3.2) sum to one. Put

$$B_s := 1 - \frac{\beta_s}{s-1} = \frac{s-1}{s-1+s^{-\tau}}, \qquad A := \prod_{w<s\le x/2} B_s. \tag{3.3}$$

For a realization $a = (a_s)$, define its surviving set by

$$S(a) := \{n \in \mathbb{Z} : n \not\equiv a_s \pmod{s} \text{ for every prime } s \le x/2\}. \tag{3.4}$$

When no confusion is possible, we write simply S.

3.1. The exact one-point probabilities

Lemma 3.1 (local likelihood ratio). If $w < s \le x/2$ is prime, then

$$\mathbb{P}(n \not\equiv a_s \pmod{s}) = \begin{cases} \beta_s, & s \mid n, \\ B_s, & s \nmid n. \end{cases} \tag{3.5}$$

Moreover,

$$\frac{\beta_s}{B_s} = s^{-\tau}. \tag{3.6}$$

Proof. If $s \mid n$, then n survives exactly when $a_s \neq 0$, an event of probability β_s. If $s \nmid n$, the only bad choice is the particular nonzero residue $n \pmod{s}$, which has probability $\beta_s/(s-1)$; hence the survival probability is B_s. Finally, substituting (3.1) and (3.3) gives

$$\frac{\beta_s}{B_s} = \frac{(s-1)s^{-\tau}}{s-1+s^{-\tau}} \frac{s-1+s^{-\tau}}{s-1} = s^{-\tau}.$$

This proves the lemma. □

Let C be the set of squarefree, w-rough composite integers in $(x, y]$. Here w-rough means that every prime factor is larger than w.

Lemma 3.2 (exact one-point law). For all sufficiently large x, every prime factor of every $v \in C$ is at most $x/2$.

Moreover,

$$q_v := \mathbb{P}(v \in S) = Av^{-\tau}. \tag{3.7}$$

Every prime $q \in (x, y]$ survives with probability exactly A.

Proof. Suppose that a prime $s > x/2$ divides $v \in C$. Since v is composite, v/s is an integer greater than one. Because v is w-rough, this integer is at least w. On the other hand,

$$\frac{v}{s} < \frac{2y}{x} = 2H < w$$

by (2.7), a contradiction. Thus all prime factors of v occur among the random coordinates $w < s \le x/2$.

At a coordinate not dividing v, the local survival probability is B_s; at a coordinate dividing v, it is $\beta_s = B_s s^{-\tau}$. Independence, (3.6), and squarefreeness therefore give

$$q_v = \prod_{w<s\le x/2} B_s \prod_{s\mid v} s^{-\tau} = Av^{-\tau}.$$

A prime $q > x$ is divisible by none of the coordinate primes, so every local factor is B_s, and its survival probability is A. □

3.2. Analytic estimates, with an elementary rough-number count

We first estimate A. From (3.3),

$$-\log A = \sum_{w<s\le x/2} \log\left(1+\frac{s^{-\tau}}{s-1}\right) = \sum_{w<s\le x/2} s^{-1-\tau} + O\left(\sum_{s>w} s^{-2}\right). \tag{3.8}$$

The error is $o(1)$ by (2.2). Partial summation in the prime number theorem gives, uniformly for the present $\tau = t/L$,

$$\sum_{w<s\le x/2} s^{-1-\tau} = \int_w^{x/2} \frac{du}{u^{1+\tau}\log u} + o(1) = E_1(\tau\log w) - E_1(\tau\log(x/2)) + o(1), \tag{3.9}$$

where

$$E_1(z) := \int_z^\infty \frac{e^{-u}}{u}\,du.$$

Indeed, the last equality follows from the substitution $u = e^{v/\tau}$, or equivalently $v = \tau\log u$. Now $\tau\log w = o(1)$, $\tau\log(x/2) = t + o(1)$, and

$$E_1(z) = -\gamma - \log z + O(z) \quad (z\downarrow 0), \qquad E_1(t) \le \frac{e^{-t}}{t} = o(1).$$

Consequently

$$A = (e^\gamma + o(1))\tau\log w = (e^\gamma B_0 + o(1))\frac{tL_2}{L}. \tag{3.10}$$

We next count the w-rough integers in $(x, y]$ without using the fundamental lemma of sieve theory.

Lemma 3.3 (elementary rough-number count). If

$$R(x,y;w) := \#\{x < n \le y : (n, \prod_{p\le w} p) = 1\},$$

In this notation,

$$R(x,y;w) = (e^{-\gamma} + o(1))\frac{y}{\log w}. \tag{3.11}$$

Proof. Put $N = y - x$, so $N = (1+o(1))y = x^{1+o(1)}$. For a squarefree divisor d of $P(w) := \prod_{p\le w} p$, let

$$A_d := \#\{x < n \le y : d \mid n\}.$$

The multiples of d are spaced exactly d apart, so

$$A_d = \frac{N}{d} + O(1) \tag{3.12}$$

with an absolute error at most one.

We now use only finite inclusion-exclusion. If an integer n has exactly k prime divisors at most w, then for every $J \geq 0$

$$\sum_{j=0}^{J} (-1)^j \binom{k}{j} = \begin{cases} 1, & k = 0, \\ (-1)^J \binom{k-1}{J}, & k \geq 1. \end{cases} \tag{3.13}$$

The second identity follows from Pascal's identity by induction on J. Thus the sum in (3.13) is an upper bound for $1_{k=0}$ when J is even and a lower bound when J is odd. Summing over $x < n \leq y$ gives the Bonferroni bounds

$$\sum_{\substack{d \mid P(w) \\ \omega(d) \leq 2m+1}} \mu(d) A_d \leq R(x, y; w) \leq \sum_{\substack{d \mid P(w) \\ \omega(d) \leq 2m}} \mu(d) A_d. \tag{3.14}$$

Choose

$$m := \left\lfloor \frac{L}{1000 \log w} \right\rfloor.$$

This tends to infinity. Replacing every A_d in (3.14) by N/d creates a total endpoint error at most

$$\sum_{j \leq 2m+1} \binom{\pi(w)}{j} \leq (2m+2) w^{2m+1} = \exp\{(0.002 + o(1))L\} = x^{0.002+o(1)}. \tag{3.15}$$

This is $o(N/\log w)$.

It remains to show that the truncated main sums approximate the full Euler product. Put

$$R_w := \sum_{p \leq w} \frac{1}{p} = O(\log \log w).$$

For each j, the sum of $1/d$ over squarefree $d \mid P(w)$ with $\omega(d) = j$ is at most $R_w^j / j!$: expanding R_w^j counts every product of j distinct primes at least $j!$ times. Therefore, for $J = 2m$ or $2m+1$,

$$\begin{aligned} \left| \sum_{\substack{d \mid P(w) \\ \omega(d) \leq J}} \frac{\mu(d)}{d} - \prod_{p \leq w} \left(1 - \frac{1}{p}\right) \right| &\leq \sum_{j > J} \frac{R_w^j}{j!} \\ &\leq 2 \left(\frac{e R_w}{J+1} \right)^{J+1} = o\left(\frac{1}{\log w} \right). \end{aligned} \tag{3.16}$$

For the second inequality, note that $R_w/(J+2) = o(1)$, so the tail is at most twice its first term, and use $(J+1)! \geq ((J+1)/e)^{J+1}$. In fact the last expression in (3.16) decays like x^{-c_0} for a fixed $c_0 > 0$.

Combining (3.12)–(3.16) shows

$$R(x, y; w) = N \prod_{p \leq w} \left(1 - \frac{1}{p}\right) + o\left(\frac{N}{\log w} \right).$$

Mertens' formula (2.2), together with $N \sim y$, proves (3.11). $\square$

The primes among these rough integers contribute only

$$O\left(\frac{y}{L}\right) = o\left(\frac{y}{\log w} \right). \tag{3.17}$$

The number that are not squarefree is at most

$$\sum_{w < s \leq \sqrt{y}} \left(\frac{y}{s^2} + 1 \right) \ll \frac{y}{w \log w} + \frac{\sqrt{y}}{\log y} \ll \frac{y}{w \log w} = o\left(\frac{x}{L} \right). \tag{3.18}$$

Indeed, a nonsquarefree w-rough integer is divisible by s^2 for some prime $s > w$; the first term uses (2.2), and the second uses the prime number theorem. Since $w = x^{o(1)}$ and $y = x^{1+o(1)}$, the second term is smaller than the first for all sufficiently large x. It follows from (3.11), (3.17), and (3.18) that

$$|C| = (e^{-\gamma} + o(1))\frac{y}{\log w}. \tag{3.19}$$

For clarity, define the exact comparison quantity

$$q_C^* := Ae^{-t}. \tag{3.20}$$

For $x < v \le y$,

$$v^{-\tau} = x^{-\tau}\exp\{-\tau\log(v/x)\} = e^{-t}(1 + o(1)) \tag{3.21}$$

uniformly, because $0 \le \log(v/x) \le \log H$ and $\tau \log H = o(1)$. Hence, uniformly for $v \in C$,

$$q_v = (1 + o(1))q_C^*, \qquad q_C^* = (e^{\gamma}B_0 + o(1))\frac{1}{LL_2^{D-1}}. \tag{3.22}$$

Combining (3.19), (3.22), and $y = cxL/t$ gives

$$N_C := \sum_{v\in C} q_v = (c + o(1))\frac{x}{tL_2^D}. \tag{3.23}$$

Finally, the prime number theorem gives $\pi(y) - \pi(x) = (1 + o(1))y/L$, and therefore

$$A(\pi(y) - \pi(x)) = (ce^{\gamma}B_0 + o(1))\frac{xL_2}{L}. \tag{3.24}$$

3.3. Concentration of the number of surviving primes

Let $Q(a)$ be the number of primes in $(x, y]$ that survive. Its mean is the quantity in (3.24). We prove that $Q(a)$ is concentrated around this mean.

Fix distinct primes $q, q' \in (x, y]$, and put $\rho_s = \beta_s/(s - 1)$, so $B_s = 1 - \rho_s$. Neither prime is zero modulo any coordinate $s \le x/2$. If $q \not\equiv q' \pmod s$, their joint local survival probability is $1 - 2\rho_s$; if $q \equiv q' \pmod s$, it is $1 - \rho_s = B_s$. Since $\rho_s \ll 1/s$, the elementary inequality

$$|\log(1 - u) + u| \le 2u^2 \qquad (0 \le u \le 1/2) \tag{3.25}$$

It follows that

$$\log\frac{\mathbb{P}(q, q' \in S)}{A^2} = O\left(\sum_{s>w}\frac{1}{s^2}\right) + O\left(\sum_{\substack{s>w\\ s|q-q'}}\frac{1}{s}\right). \tag{3.26}$$

If $d \ne 0$ and all primes in a set divide d and exceed w, then

$$\sum_{\substack{s>w\\ s|d}}\frac{1}{s} \le \frac{\omega(d)}{w} \le \frac{\log|d|}{w\log w}. \tag{3.27}$$

Here $|q - q'| \le y = x^{1+o(1)}$. Thus (2.2), (3.26), and (3.27) give, uniformly in distinct q, q',

$$\mathbb{P}(q, q' \in S) = (1 + o(1))A^2. \tag{3.28}$$

Writing $I_q = 1_{q\in S}$ and $M = A(\pi(y) - \pi(x))$, we obtain

$$\begin{aligned}\operatorname{Var} Q &= \sum_q \operatorname{Var} I_q + \sum_{q\neq q'} \operatorname{Cov}(I_q, I_{q'}) \\ &\le M + o(1)M^2.\end{aligned} \tag{3.29}$$

By (3.24), $M \to \infty$. Chebyshev's inequality now gives

$$Q(a) = (1+o(1))M$$

with probability $1 - o(1)$.

4. All-fiber blocks and unrooted correlations

Fix a sufficiently small constant $\kappa > 0$, independent of x, and put

$$K := \left\lfloor \kappa \frac{|C|}{x} \right\rfloor. \tag{4.1}$$

By (2.4) and (3.19),

$$K \asymp \frac{L}{tL_2}. \tag{4.2}$$

In particular,

$$K \longrightarrow \infty, \qquad K = o(w), \qquad \tau K \log y = O\left(\frac{L}{L_2}\right) = o(L). \tag{4.3}$$

Let $p \in \mathcal{P}_C$. For each residue $b \pmod p$, define the fiber

$$C_{p,b} := \{v \in C : v \equiv b \pmod p\}.$$

Order each nonempty fiber in any fixed way, divide it into consecutive sets of K elements, and keep the last set even if it has fewer than K elements. Let Ω_p be the collection of all resulting blocks, and put $B_p := |\Omega_p|$.

Lemma 4.1 (all-fiber decomposition). Every $v \in C$ lies in exactly one block of Ω_p, and

$$\frac{|C|}{K} \le B_p \le \frac{|C|}{K} + p = (1+O(\kappa))\frac{|C|}{K} \asymp x. \tag{4.4}$$

Distinct members of one block are coprime. Consequently, the product of the members of a block is squarefree.

Proof. If a fiber contains m_b elements, it contributes $\lceil m_b/K \rceil$ blocks. Thus

$$\frac{m_b}{K} \le \left\lceil \frac{m_b}{K} \right\rceil \le \frac{m_b}{K} + 1.$$

Summing over at most p nonempty fibers gives the first two inequalities in (4.4). Since $K = (1+o(1))\kappa|C|/x$ and $p \le x$, the additive term p is at most $(\kappa + o(1))|C|/K$, proving the rest of (4.4).

Now let $n \neq m$ belong to one fiber. Then $m - n = hp$ for a nonzero integer h, and

$$0 < |h| < \frac{y-x}{p} < 2H.$$

If a prime s divided both n and m, then $s > w$, because both numbers are w-rough. Also $s \neq p$, because Lemma 3.2 says that every prime factor of a member of C is at most $x/2$, whereas $p > x/2$. Hence $s \mid h$. But $0 < |h| < 2H < w < s$, which is impossible. Thus $(n, m) = 1$. Each member of C is squarefree, so a product of pairwise coprime members is squarefree. □

For $E \in \Omega_p$, set

$$\pi(E) := \mathbb{P}(E \subseteq S), \qquad X_E := \frac{1_{E\subseteq S}}{\pi(E)}, \qquad Z_p := \frac{1}{B_p}\sum_{E\in\Omega_p} X_E. \tag{4.5}$$

Every local factor in $\pi(E)$ is positive because $|E| \le K = o(w)$. Also $\mathbb{E}X_E = 1$, and therefore

$$\mathbb{E}Z_p = 1 \tag{4.6}$$

exactly. In this section, $\mathbb{E}_{E\neq F}$ and $\mathbb{P}_{E\neq F}$ mean expectation and probability for a uniformly chosen ordered pair of distinct blocks of Ω_p.

Lemma 4.2 (local block factor). Let $E \in \Omega_p$, let $k = |E|$, and let $w < s \le x/2$ be prime. The s-coordinate factor in $\pi(E)$ is

$$\begin{cases} 1 - k\beta_s/(s-1), & s \nmid \prod_{n\in E} n, \\ \beta_s\big(1 - (k-1)/(s-1)\big), & s \mid \prod_{n\in E} n. \end{cases} \tag{4.7}$$

Proof. The members of E occupy distinct residues modulo s. Indeed, two congruent members would have difference hp with $0 < |h| < 2H < s$ and $s \neq p$, which is impossible. If no member is zero modulo s, the random residue must avoid k specified nonzero classes. Their total probability is $k\beta_s/(s-1)$. If one member is zero, the residue must first be nonzero, an event of probability β_s, and then avoid the other $k-1$ classes; conditionally on being nonzero, all $s-1$ classes are equally likely. This gives (4.7). □

For distinct blocks $E, F \in \Omega_p$, define

$$G(E,F) := \gcd\left(\prod_{n\in E} n, \prod_{m\in F} m\right). \tag{4.8}$$

The following explicit error term is convenient:

$$\varepsilon_x := L^{-90} + x^{-1/8}. \tag{4.9}$$

It tends to zero. The exponent -90 has no special meaning; it merely leaves ample room because $B_0 = 100$.

Proposition 4.3 (unrooted block correlations). Uniformly for $p \in \mathcal{P}_C$, the following statements hold for all sufficiently large x.

1. Every block satisfies
$$\pi(E)^{-1} = x^{o(1)}. \tag{4.10}$$
2. For distinct blocks,
$$\frac{\mathbb{P}(E \cup F \subseteq S)}{\pi(E)\pi(F)} \le G(E,F)^{\tau} \exp\{O(L^{3-B_0+o(1)})\}. \tag{4.11}$$
3. If $d \le x$ is squarefree, every prime factor of d exceeds w, and $j = \omega(d)$, then
$$\mathbb{P}_{E\neq F}(d \mid G(E,F)) \ll \frac{(C_b H^2)^j}{d^2} \tag{4.12}$$
for an absolute constant C_b.
4. The large-gcd tail satisfies
$$\mathbb{E}_{E\neq F}\big(G(E,F)^{\tau};\, G(E,F) > x\big) \ll x^{-1/8}. \tag{4.13}$$

Proof of (1). Let E have k members and put $N_E = \prod_{n\in E} n$. We compare the exact local factor (4.7) with the product of the one-point factors of the members of E. Write $D_s = s-1$, $\rho_s = \beta_s/D_s$, and $B_s = 1 - \rho_s$. By (4.3), $k/D_s = o(1)$, uniformly in every coordinate.

If $s \nmid N_E$, the exact factor is $1 - k\rho_s$, whereas the product of the k one-point factors is B_s^k. Using (3.25) once with $u = k\rho_s$ and once with $u = \rho_s$, we obtain

$$\log \frac{1 - k\rho_s}{B_s^k} \ge -C\frac{k^2}{s^2}. \tag{4.14}$$

If $s \mid N_E$, exactly one member is zero modulo s. The exact factor is $\beta_s(1 - (k-1)/D_s)$, while the product of the one-point factors is

$$\beta_s B_s^{k-1} = B_s^k s^{-\tau}.$$

Again using (3.25),

$$\log \frac{1-(k-1)/D_s}{B_s^{k-1}} \geq -C\left(\frac{k}{s}+\frac{k^2}{s^2}\right). \tag{4.15}$$

Multiplying (4.14) and (4.15) over all coordinates gives

$$\pi(E) \geq A^k N_E^{-\tau} \exp\left\{-C\left(k\sum_{s\mid N_E}\frac{1}{s}+k^2\sum_{s>w}\frac{1}{s^2}\right)\right\}. \tag{4.16}$$

Every prime in the first sum exceeds w, so (3.27) and $N_E \leq y^k$ give

$$\sum_{s\mid N_E}\frac{1}{s} \leq \frac{\log N_E}{w\log w} \leq \frac{k\log y}{w\log w}. \tag{4.17}$$

Taking logarithms in (4.16) and using (3.10), (4.2), (4.3), and (2.2), we find

$$\log \pi(E)^{-1} \leq K\log A^{-1} + \tau K\log y + O\left(\frac{K^2\log y}{w\log w}\right) = o(L). \tag{4.18}$$

Exponentiating proves (4.10).

Proof of (2). Fix a coordinate s. Let $z_E, z_F \in \{0,1\}$ indicate whether E, F, respectively, contain a number divisible by s. Put

$$a = |E| - z_E, \qquad b = |F| - z_F,$$

Let c_s be the number of nonzero residue classes modulo s occupied by both blocks. Within each block the occupied classes are distinct. With $D_s = s-1$, direct division of the joint local probability by the two marginal local probabilities gives

$$\begin{aligned}
R_{00} &= \frac{1-\beta_s(a+b-c_s)/D_s}{(1-\beta_s a/D_s)(1-\beta_s b/D_s)},\\
R_{10} &= \frac{1-(a+b-c_s)/D_s}{(1-a/D_s)(1-\beta_s b/D_s)},\\
R_{01} &= \frac{1-(a+b-c_s)/D_s}{(1-\beta_s a/D_s)(1-b/D_s)},\\
R_{11} &= \beta_s^{-1}\frac{1-(a+b-c_s)/D_s}{(1-a/D_s)(1-b/D_s)}.
\end{aligned} \tag{4.19}$$

For example, in the 10 case the first block contains zero, so its marginal is $\beta_s(1-a/D_s)$; the joint event requires a nonzero residue avoiding $a+b-c_s$ classes and has probability $\beta_s(1-(a+b-c_s)/D_s)$. Cancelling β_s gives the second line. The other lines follow in the same way.

Apply (3.25) to every logarithm in (4.19). The linear terms in $\log R_{00}$ add to $\beta_s c_s/D_s$. Those in $\log R_{10}$ add to

$$\frac{c_s-(1-\beta_s)b}{D_s} \leq \frac{c_s}{D_s},$$

and the 01 case is symmetric. Every discarded quadratic term is $O(K^2/s^2)$. Hence

$$R_{00}, R_{10}, R_{01} \leq \exp\left\{C\left(\frac{c_s}{s}+\frac{K^2}{s^2}\right)\right\}. \tag{4.20}$$

Since $\beta_s = B_s s^{-\tau}$,

$$\beta_s^{-1} = s^\tau B_s^{-1} = s^\tau \exp\{O(1/s)\}.$$

The rational factor in R_{11} is treated as above, so

$$R_{11} \le s^{\tau} \exp\left\{ C\left(\frac{1}{s} + \frac{c_s}{s} + \frac{K^2}{s^2} \right) \right\}. \tag{4.21}$$

The 11 case occurs exactly when $s \mid G(E,F)$.

We now sum the small exponential errors. Blocks in one partition are disjoint, so if $n \in E$, $m \in F$, then $n \neq m$. A common nonzero residue modulo s gives one pair (n,m) with $s \mid n-m$. Therefore (3.27) yields

$$\begin{aligned} \sum_{s>w} \frac{c_s}{s} &\le \sum_{n\in E} \sum_{m\in F} \sum_{\substack{s>w \\ s\mid n-m}} \frac{1}{s} \ll \frac{K^2 L}{w \log w}, \\ K^2 \sum_{s>w} \frac{1}{s^2} &\ll \frac{K^2}{w \log w}, \\ \sum_{s \mid G(E,F)} \frac{1}{s} &\le \frac{\log G(E,F)}{w \log w} \le \frac{K \log y}{w \log w}. \end{aligned} \tag{4.22}$$

All three right sides are $O(L^{3-B_0+o(1)})$. Multiplying (4.20)–(4.21) over s proves (4.11).

Proof of (3). Fix p. For every residue $b \pmod p$, let r_b be its unique representative in $(x, x+p]$. Every member of the fiber $C_{p,b}$ is of the form

$$r_b + hp,$$

The number of possible integer offsets h is at most

$$U := 2H + 2. \tag{4.23}$$

Suppose $d \mid G(E,F)$, and write $j = \omega(d)$. Because the members of a block are pairwise coprime, for every $s \mid d$ there is a unique member of E divisible by s, and a unique member of F divisible by s. Record their two offsets. There are at most U^{2j} possible offset arrays.

Fix one array. The congruence $s \mid r_b + h_s p$ fixes $r_b \pmod s$. The Chinese remainder theorem therefore fixes $r_b \pmod d$, and similarly fixes $r_{b'} \pmod d$. An interval of length $p \le x$ contains at most $p/d + 1 \le 2x/d$ numbers in a prescribed class modulo d. Hence there are at most

$$\left(\frac{p}{d} + 1\right)^2 \le \frac{4x^2}{d^2}$$

possible ordered pairs $(r_b, r_{b'})$.

Once a representative and its recorded offsets are fixed, the resulting prescribed elements lie in at most one block of the fixed partition: any one prescribed element already determines its unique block. Thus this encoding is injective on valid ordered block pairs. The number of such pairs is at most $4x^2 U^{2j}/d^2$. Since (4.4) gives $B_p(B_p-1) \gg x^2$, division by the number of all ordered distinct block pairs proves (4.12), after enlarging the absolute constant C_b.

Proof of (4). The integer $G = G(E,F)$ is squarefree. Suppose first that $G > x$ has a prime factor $s > x^{1/3}$. A union bound and (4.12), applied with $d = s$, give

$$\mathbb{P}_{E\neq F}(\text{such an } s) \ll H^2 \sum_{s > x^{1/3}} \frac{1}{s^2} = x^{-1/3+o(1)}. \tag{4.24}$$

If every prime factor of G is at most $x^{1/3}$, multiply its prime factors one at a time until the product first exceeds $x^{1/3}$. The resulting divisor $d \mid G$ lies in $(x^{1/3}, x^{2/3}]$. Therefore

$$\mathbb{P}_{E\neq F}(G>x) \ll x^{-1/3+o(1)} + \sum_{\substack{x^{1/3}<d\le x^{2/3}\\ d \text{ squarefree}\\ P^-(d)>w}} \frac{(C_bH^2)^{\omega(d)}}{d^2} \tag{4.25}$$
$$\le x^{-1/3+o(1)} + x^{-1/6}\prod_{s>w}\left(1+\frac{C_bH^2}{s^{3/2}}\right).$$

Here $d > x^{1/3}$ implies $d^{-2} \le x^{-1/6}d^{-3/2}$. Moreover,

$$H^2\sum_{s>w} s^{-3/2} \ll H^2 w^{-1/2} = o(1),$$

so the product in (4.25) is $1+o(1)$. Thus

$$\mathbb{P}_{E\neq F}(G>x) \le x^{-1/6+o(1)}. \tag{4.26}$$

Finally,

$$G^\tau \le (y^K)^\tau = \exp\{\tau K\log y\} = x^{O(1/L_2)}. \tag{4.27}$$

Multiplying (4.26) by (4.27) gives (4.13) for all sufficiently large x. This finishes the proof of Proposition 4.3.□
We next prove that the random normalizer Z_p is usually close to one.

Auxiliary prime-sum estimate. For the present $\tau = t/L$,

$$\sum_{s>w}\frac{s^\tau-1}{s^2} \ll \frac{\tau}{w}. \tag{4.28}$$

Proof. Since $e^u - 1 \le ue^u$ for $u \ge 0$,

$$s^\tau - 1 \le \tau(\log s)s^\tau.$$

Consequently

$$\sum_{s>w}\frac{s^\tau-1}{s^2} \le \tau\sum_{s>w}\frac{\log s}{s^{2-\tau}} \le \tau\sum_{n>w}\frac{\Lambda(n)}{n^{2-\tau}}.$$

The elementary Chebyshev bound $\sum_{n\le u}\Lambda(n) \ll u$, followed by partial summation, makes the last sum $O(w^{-1+\tau})$. By (2.7), $w^\tau = e^{\tau\log w} = 1+o(1)$, so this is $O(1/w)$. □

Proposition 4.4 (normalizer variance). Uniformly in $p \in \mathcal{P}_C$,

$$\mathrm{Var}(Z_p) \ll \varepsilon_x. \tag{4.29}$$

Proof. Since G is squarefree,

$$G^\tau = \prod_{s\mid G} s^\tau = \sum_{d\mid G}\prod_{s\mid d}(s^\tau-1). \tag{4.30}$$

All summands are nonnegative. On the event $G \le x$, every divisor $d \mid G$ also satisfies $d \le x$, so Tonelli's theorem and (4.12) give

$$\begin{aligned}\mathbb{E}_{E\neq F}(G^\tau-1;G\le x) &\le \sum_{\substack{d>1\\ d\ \text{squarefree}\\ P^-(d)>w}} \frac{(C_bH^2)^{\omega(d)}}{d^2}\prod_{s\mid d}(s^\tau-1)\\ &\le \prod_{s>w}\left(1+C_bH^2\frac{s^\tau-1}{s^2}\right)-1\\ &\le \exp\left(C_bH^2\sum_{s>w}\frac{s^\tau-1}{s^2}\right)-1\\ &\ll \frac{H^2\tau}{w}=L^{1-B_0+o(1)}.\end{aligned} \tag{4.31}$$

In the last line we used the auxiliary prime-sum estimate (4.28) and the fact that $H^2\tau/w=o(1)$.

Now average the correlation bound (4.11). On $G\le x$, (4.31) shows that the average of G^τ is at most $1+L^{1-B_0+o(1)}$; on $G>x$, (4.13) applies. Hence

$$\frac{1}{B_p(B_p-1)}\sum_{E\neq F}\frac{\mathbb{P}(E\cup F\subseteq S)}{\pi(E)\pi(F)}\le 1+O(\varepsilon_x). \tag{4.32}$$

For the diagonal, the definition of X_E gives the exact identity

$$\mathbb{E}X_E^2=\frac{1}{\pi(E)}.$$

Thus (4.10) and $B_p\asymp x$ imply

$$\frac{1}{B_p^2}\sum_E\mathbb{E}X_E^2\le\frac{x^{o(1)}}{B_p}=x^{-1+o(1)}\ll x^{-1/8}. \tag{4.33}$$

Expanding $\mathbb{E}Z_p^2$, using (4.32) for the off-diagonal terms and (4.33) for the diagonal terms, gives $\mathbb{E}Z_p^2\le 1+O(\varepsilon_x)$. Since $\mathbb{E}Z_p=1$, this is (4.29). □

5. Rooted correlations and the composite layer

Fix $v\in C$. For each $p\in\mathcal{P}_C$, let $E_{p,v}$ be the unique block of Ω_p containing v. Let $E_{p,v}^\circ$ be its *companion set*, obtained by deleting v. Write

$$N_{p,v}:=\prod_{n\in E_{p,v}^\circ}n.$$

If $p\neq p'$, define

$$G_v(p,p'):=\gcd(N_{p,v},N_{p',v}). \tag{5.1}$$

The two rooted blocks $E_{p,v}$ and $E_{p',v}$ have no common member other than v. Indeed, a second common member would have the forms

$$v+hp=v+h'p'$$

with nonzero integers $|h|,|h'|<2H$. Then $hp=h'p'$. Since p,p' are distinct primes and $2H<\min(p,p')$, the divisibility $p\mid h'$ is impossible.

5.1. Conditioning exactly on a surviving root

The event $v \in S$ is the intersection, over the coordinate primes s, of an event depending only on a_s. Since the coordinates were independent before conditioning, they remain independent after conditioning on $v \in S$: for events A_s depending on individual coordinates,

$$\mathbb{P}\left(\bigcap_s A_s \;\middle|\; v \in S\right) = \prod_s \mathbb{P}(A_s \mid v \not\equiv a_s \pmod{s}). \tag{5.2}$$

Thus, at a coordinate s, the conditional law is obtained by deleting the single forbidden outcome $a_s \equiv v \pmod{s}$ and renormalizing.

Lemma 5.1 (exact rooted expectations). For every $p \in \mathcal{P}_C$,

$$\mathbb{E}(X_{E_{p,v}} \mid v \in S) = \frac{1}{q_v}, \tag{5.3}$$

and

$$q_v^2 \mathbb{E}(X_{E_{p,v}}^2 \mid v \in S) = \frac{q_v}{\pi(E_{p,v})}. \tag{5.4}$$

Proof. Because $v \in E_{p,v}$, the event $E_{p,v} \subseteq S$ is contained in the event $v \in S$. Therefore

$$\mathbb{P}(E_{p,v} \subseteq S \mid v \in S) = \frac{\pi(E_{p,v})}{q_v}.$$

On this event $X_{E_{p,v}} = \pi(E_{p,v})^{-1}$, and off it $X_{E_{p,v}} = 0$. Multiplication by $\pi(E_{p,v})^{-1}$ and by $\pi(E_{p,v})^{-2}$, respectively, gives (5.3) and (5.4). □

Put

$$M_0 := \sum_{p \in \mathcal{P}_C} \frac{1}{B_p}, \qquad T_v := \sum_{p \in \mathcal{P}_C} \frac{X_{E_{p,v}}}{B_p}. \tag{5.5}$$

From (2.9) and (4.4),

$$M_0 \asymp \frac{\kappa}{L}. \tag{5.6}$$

Lemma 5.1 gives the exact conditional mean

$$\mathbb{E}(T_v \mid v \in S) = \frac{M_0}{q_v}. \tag{5.7}$$

We now bound its conditional variance.

Proposition 5.2 (rooted block correlations). Uniformly for $v \in C$,

$$\mathbb{E}\left[\left(\frac{q_v T_v}{M_0} - 1\right)^2 \;\middle|\; v \in S\right] \ll \varepsilon_x. \tag{5.8}$$

More precisely, for distinct $p, p' \in \mathcal{P}_C$,

$$q_v^2 \mathbb{E}(X_{E_{p,v}} X_{E_{p',v}} \mid v \in S) \le G_v(p,p')^\tau \exp\{O(L^{3-B_0+o(1)})\}, \tag{5.9}$$

and, if $d \le x$ is squarefree and $P^-(d) > w$,

$$\mathbb{P}_{p \ne p'}(d \mid G_v(p,p')) \ll \frac{L^2 (C_b H^2)^{\omega(d)}}{d^2}. \tag{5.10}$$

The probability in (5.10) may be uniform on ordered distinct pairs (p,p'), or may use weights proportional to $(B_p B_{p'})^{-1}$.

Proof: exact local ratios. Write each rooted block as its root v together with its companion set. Fix a coordinate s. Let $z_E, z_F \in \{0,1\}$ indicate whether the two companion sets contain zero modulo s; let a, b count their nonzero residue classes; and let c_s count the nonzero classes common to the two companion sets. Within each companion set all classes are distinct, and no companion occupies the root class.

First suppose $v \equiv 0 \pmod{s}$. Then no companion can be zero: otherwise s would divide both v and $v + hp$, hence would divide h, contrary to $0 < |h| < 2H < s$. Conditioning on the root forces $a_s \neq 0$, and (3.2) is then uniform on the $D_s = s - 1$ nonzero classes. The conditional probabilities that the first companion set, the second companion set, and their union survive are respectively

$$1 - \frac{a}{D_s}, \qquad 1 - \frac{b}{D_s}, \qquad 1 - \frac{a + b - c_s}{D_s}.$$

Thus their joint-to-product ratio is

$$\frac{1 - (a + b - c_s)/D_s}{(1 - a/D_s)(1 - b/D_s)}. \tag{5.11}$$

Now suppose $v \not\equiv 0 \pmod{s}$. The root survives this coordinate with probability $B_s = 1 - \beta_s/D_s$. Under the conditional law,

$$\mathbb{P}(a_s = 0 \mid v \text{ survives}) = \frac{1 - \beta_s}{B_s}, \tag{5.12}$$

and every nonzero class different from $v \pmod{s}$ has conditional probability

$$\frac{\beta_s}{D_s B_s}. \tag{5.13}$$

Using (5.12)–(5.13), and listing the cases $(z_E, z_F) = (0,0), (1,0), (0,1), (1,1)$, direct division gives the four ratios

$$\begin{aligned}
R_{00}^{(v)} &= B_s \frac{1 - \beta_s(1 + a + b - c_s)/D_s}{(1 - \beta_s(1 + a)/D_s)(1 - \beta_s(1 + b)/D_s)}, \\
R_{10}^{(v)} &= B_s \frac{1 - (1 + a + b - c_s)/D_s}{(1 - (1 + a)/D_s)(1 - \beta_s(1 + b)/D_s)}, \\
R_{01}^{(v)} &= B_s \frac{1 - (1 + a + b - c_s)/D_s}{(1 - \beta_s(1 + a)/D_s)(1 - (1 + b)/D_s)}, \\
R_{11}^{(v)} &= \frac{B_s}{\beta_s} \frac{1 - (1 + a + b - c_s)/D_s}{(1 - (1 + a)/D_s)(1 - (1 + b)/D_s)}.
\end{aligned} \tag{5.14}$$

For instance, in the 10 case, after conditioning on the root the joint event has probability

$$\frac{\beta_s}{B_s}\left(1 - \frac{1 + a + b - c_s}{D_s}\right),$$

The first marginal probability is

$$\frac{\beta_s}{B_s}\left(1 - \frac{1 + a}{D_s}\right),$$

and the second marginal has probability

$$\frac{1}{B_s}\left(1 - \frac{\beta_s(1 + b)}{D_s}\right).$$

Their quotient is the second line of (5.14); the other cases are identical calculations.

Applying (3.25) to (5.11) and (5.14), exactly as in (4.20), shows that (5.11) and the first three ratios in (5.14) are at most

$$\exp\left\{C\left(\frac{c_s}{s}+\frac{K^2}{s^2}\right)\right\}. \tag{5.15}$$

In the 11 case, $B_s/\beta_s = s^\tau$ exactly. The linear term from the remaining rational factor is $(1+c_s)/D_s$, and all higher terms are $O(K^2/s^2)$. Hence

$$R_{11}^{(v)} \le s^\tau \exp\left\{C\left(\frac{1}{s}+\frac{c_s}{s}+\frac{K^2}{s^2}\right)\right\}. \tag{5.16}$$

The 11 case occurs exactly when $s \mid G_v(p,p')$. The root itself has disappeared from the ratio: a large factor s^τ occurs only when both companion sets contain a multiple of s.

The two companion sets have no common integer. Therefore the same difference count as in (4.22) gives

$$\sum_{s>w}\frac{c_s}{s} \ll \frac{K^2 L}{w\log w}, \qquad K^2\sum_{s>w}s^{-2} \ll \frac{K^2}{w\log w}. \tag{5.17}$$

Also

$$\sum_{s\mid G_v(p,p')}\frac{1}{s} \le \frac{\log G_v(p,p')}{w\log w} \le \frac{K\log y}{w\log w}. \tag{5.18}$$

All these errors are $O(L^{3-B_0+o(1)})$. Multiplication over the coordinates proves (5.9). To see that its left side really is the conditional joint-to-product ratio just calculated, observe that

$$\begin{aligned} q_v^2\mathbb{E}(X_E X_F \mid v\in S) &= q_v^2\frac{\mathbb{P}(E\cup F\subseteq S \mid v\in S)}{\pi(E)\pi(F)} \\ &= q_v\frac{\mathbb{P}(E\cup F\subseteq S)}{\pi(E)\pi(F)} \\ &= \frac{\mathbb{P}(E\cup F\subseteq S \mid v\in S)}{\mathbb{P}(E\subseteq S\mid v\in S)\mathbb{P}(F\subseteq S\mid v\in S)}. \end{aligned} \tag{5.19}$$

Here $E = E_{p,v}$ and $F = E_{p',v}$.

Proof of the rooted divisor count. Every companion in the color-p block has the form

$$v+hp, \qquad 0<|h|<2H. \tag{5.20}$$

If a coordinate prime s divides this companion, then $s \nmid v$. Otherwise $s \mid hp$; since $s\neq p$, this would give $s\mid h$, impossible because $|h|<s$. Thus h is invertible modulo s, and

$$p \equiv -vh^{-1} \pmod{s}. \tag{5.21}$$

Suppose $d \mid G_v(p,p')$. For each $s\mid d$, record the unique companion offset in each rooted block that witnesses divisibility by s. There are at most $(CH^2)^{\omega(d)}$ arrays. Once an array is fixed, the Chinese remainder theorem and (5.21) fix both p and p' modulo d. For an upper bound, we may drop the condition that they are prime.

The number of possible pairs is at most

$$\left(\frac{x}{d}+1\right)^2 \le \frac{4x^2}{d^2}$$

pairs in $(x/2,x]^2$. Since (2.9) gives

$$|\mathcal{P}_C|(|\mathcal{P}_C|-1) \asymp \frac{x^2}{L^2},$$

division proves (5.10) for the uniform pair measure. Because $B_p \asymp x$ uniformly, changing to weights proportional to $(B_pB_{p'})^{-1}$ changes probabilities by at most a fixed multiplicative constant.

Completion of the second-moment proof. Repeating (4.24)–(4.27), now with the outside factor L^2 in (5.10), gives for either of the pair measures just described

$$\mathbb{E}_{p\neq p'}\big(G_v(p,p')^\tau; G_v(p,p')>x\big) \ll x^{-1/8}. \tag{5.22}$$

On $G_v \le x$, (4.30), (5.10), and the auxiliary estimate (4.28) give

$$\begin{aligned}\mathbb{E}_{p\neq p'}(G_v^\tau - 1; G_v \le x) &\ll \exp\left\{CL^2H^2\sum_{s>w}\frac{s^\tau-1}{s^2}\right\}-1\\ &\ll \frac{L^2H^2\tau}{w} = L^{3-B_0+o(1)}.\end{aligned} \tag{5.23}$$

The insertion of L^2 into each Euler factor is a harmless upper bound: for every nonempty squarefree d, the single outside factor L^2 in (5.10) is at most $L^{2\omega(d)}$.

Define

$$\lambda_p := \frac{1}{B_pM_0}. \tag{5.24}$$

Then

$$\sum_p \lambda_p = 1, \qquad \sum_p \lambda_p^2 \ll \frac{L}{x}. \tag{5.25}$$

The normalized random variable in (5.8) is

$$\frac{q_vT_v}{M_0} = \sum_p \lambda_p q_v X_{E_{p,v}}.$$

Its conditional mean is one by (5.7). Equations (5.9), (5.22), and (5.23), averaged with the weights $\lambda_p\lambda_{p'}$, show that its off-diagonal second moment is at most $1+O(\varepsilon_x)$. For the diagonal, Lemma 5.1, (4.10), (3.22), and (5.25) give

$$\begin{aligned}\sum_p \lambda_p^2 q_v^2 \mathbb{E}(X_{E_{p,v}}^2 \mid v\in S) &= \sum_p \lambda_p^2 \frac{q_v}{\pi(E_{p,v})}\\ &\le q_v x^{o(1)}\sum_p \lambda_p^2 = x^{-1+o(1)} \ll \varepsilon_x.\end{aligned} \tag{5.26}$$

Subtracting the square of the mean proves (5.8). □

5.2. Capping the color law

The importance weights X_E have the correct expectation, but for a particular preliminary realization their sum need not be exactly B_p. The normalizer Z_p measures this discrepancy. We now turn the weights into a genuine probability law without dividing by the random quantity Z_p.

For each $p \in \mathcal{P}_C$:

- If $Z_p \le 2$, choose an internal label $E \in \Omega_p$ with probability

$$\frac{X_E}{2B_p}, \tag{5.27}$$

  and put the unused probability on a dummy label.
- If $Z_p > 2$, choose only the dummy label.

This is a probability distribution because, when $Z_p \le 2$,

$$\sum_{E\in\Omega_p}\frac{X_E}{2B_p} = \frac{Z_p}{2} \le 1. \tag{5.28}$$

If the label E is chosen, use modulo p the residue class of the fiber containing E. Every member of E is then covered. Distinct labels from one fiber may correspond to the same residue; this causes no difficulty, because the experiment first selects a label and then uses its fiber residue. Use residue zero for the dummy label. No member of C is divisible by $p > x/2$, so the dummy choice covers no intended vertex. Conditional on the preliminary realization, make these choices independently for different p.

For $v \in C$, define the incidence discarded by the cap:

$$D_v := \sum_{p \in \mathcal{P}_C} \frac{X_{E_{p,v}}}{B_p} 1_{Z_p > 2}. \tag{5.29}$$

Lemma 5.3 (capped-tail identity). For every $p \in \mathcal{P}_C$,

$$\frac{1}{B_p} \sum_{E \in \Omega_p} \mathbb{P}(Z_p > 2 \mid E \subseteq S) = \mathbb{E}(Z_p 1_{Z_p > 2}) \le 2 \operatorname{Var}(Z_p). \tag{5.30}$$

Proof. Since $\pi(E) > 0$,

$$\begin{aligned} \frac{1}{B_p} \sum_E \mathbb{P}(Z_p > 2 \mid E \subseteq S) &= \frac{1}{B_p} \sum_E \frac{\mathbb{P}(Z_p > 2,\ E \subseteq S)}{\pi(E)} \\ &= \mathbb{E}\left[1_{Z_p > 2} \frac{1}{B_p} \sum_E \frac{1_{E \subseteq S}}{\pi(E)} \right] \\ &= \mathbb{E}(Z_p 1_{Z_p > 2}). \end{aligned}$$

For every $z \ge 0$,

$$z 1_{z > 2} \le 2(z-1)^2.$$

Indeed, for $z > 2$, the difference between the right and left sides is $(2z-1)(z-2) \ge 0$, and for $z \le 2$ the left side is zero. Taking expectations and using $\mathbb{E} Z_p = 1$ proves (5.30). □

Proposition 5.4 (covering the squarefree rough composites). For a preliminary realization a, let $R_C(a)$ denote the conditional expected number of elements of $C \cap S(a)$ that remain uncovered after the independent capped choices (5.27). Then

$$\mathbb{E}_a R_C(a) = o(x/L). \tag{5.31}$$

Proof. First we bound the surviving vertices whose total uncapped incidence T_v is unusually small. From (5.8) and conditional Chebyshev,

$$\mathbb{P}\left(T_v < \frac{M_0}{2q_v} \;\middle|\; v \in S \right) \ll \varepsilon_x.$$

Multiplying by $q_v = \mathbb{P}(v \in S)$ and summing over $v \in C$ gives

$$\sum_{v \in C} \mathbb{P}\left(v \in S,\ T_v < \frac{M_0}{2q_v} \right) \ll \varepsilon_x N_C. \tag{5.32}$$

Next we bound the vertices for which too much incidence is discarded by the caps. Because $E_{p,v}$ contains v, every term in D_v vanishes unless $v \in S$; hence $D_v = 0$ when $v \notin S$. Put $q_{\max} = \max_{v \in C} q_v$. Markov's inequality gives

$$\sum_{v \in C} \mathbb{P}\left(v \in S,\ D_v > \frac{M_0}{4q_v} \right) \le \frac{4q_{\max}}{M_0} \sum_{v \in C} \mathbb{E} D_v. \tag{5.33}$$

To estimate the last sum, interchange the sums over vertices and blocks. For a fixed p, every block E contributes the same value X_E/B_p to each of its $|E| \le K$ members. Therefore

$$\begin{aligned}\sum_{v\in C} D_v &= \sum_{p\in\mathcal{P}_C} \frac{1_{Z_p>2}}{B_p} \sum_{E\in\Omega_p} |E| X_E \\ &\le K \sum_{p\in\mathcal{P}_C} Z_p 1_{Z_p>2}.\end{aligned}$$

Taking expectations and using Lemma 5.3 and Proposition 4.4 yields

$$\sum_{v\in C} \mathbb{E}D_v \ll K|\mathcal{P}_C|\varepsilon_x. \tag{5.34}$$

By (3.22), $q_{\max} \asymp q_C^*$. Equations (4.1), (4.4), (2.9), and (5.6) give

$$\frac{K|\mathcal{P}_C|}{M_0} \ll |C|.$$

Consequently

$$\frac{q_{\max} K|\mathcal{P}_C|}{M_0} \ll q_C^*|C| \asymp N_C. \tag{5.35}$$

Combining (5.33)–(5.35), the expected number of vertices with excessive discarded incidence is $O(\varepsilon_x N_C)$.

Call a surviving vertex good if neither bad event in (5.32) or (5.33) occurs. For such a vertex,

$$\frac{1}{2}(T_v - D_v) \ge \frac{M_0}{8q_v}. \tag{5.36}$$

The left side is exactly the sum, over colors p, of the probabilities that the particular label $E_{p,v}$ is selected: the uncapped probability is $X_{E_{p,v}}/(2B_p)$, and the terms with $Z_p > 2$ are removed. Selecting another block in the same fiber could also cover v, so (5.36) is a lower bound for the full covering intensity.

By (3.22) and (5.6), uniformly in $v \in C$,

$$\frac{M_0}{8q_v} \gg L_2^{D-1}. \tag{5.37}$$

Conditional on the preliminary sieve, different colors are selected independently. If their individual probabilities of covering a fixed vertex are θ_p, then

$$\mathbb{P}(v \text{ is uncovered} \mid a) = \prod_p (1-\theta_p) \le \exp\left(-\sum_p \theta_p\right). \tag{5.38}$$

Thus every good surviving vertex has conditional uncovered probability at most $\exp\{-c_1 L_2^{D-1}\}$.

The expected total number of uncovered vertices is therefore at most

$$O(\varepsilon_x N_C) + N_C \exp\{-c_1 L_2^{D-1}\}. \tag{5.39}$$

The first term is $o(x/L)$, because by (3.23)

$$\frac{\varepsilon_x N_C}{x/L} \ll \frac{L^{-89}}{tL_2^D} + \frac{Lx^{-1/8}}{tL_2^D} \longrightarrow 0.$$

For the second term,

$$\log\left(\frac{N_C e^{-c_1 L_2^{D-1}}}{x/L}\right) \le L_2 - c_1 L_2^{D-1} + O(\log L_2) \longrightarrow -\infty,$$

because $D > 2$. This proves (5.31). $\square$

Proposition 5.4 concerns only the squarefree w-rough composite set C. Integers with a prime factor at most w were already covered by the preliminary choices $a_s = 0$, while (3.18) shows that the remaining nonsquarefree w-rough integers number only $o(x/L)$. Those exceptional integers, together with the $o(x/L)$ expected leftovers in (5.31), can be covered one at a time later using unused primes.

6. Covering the surviving primes

This section covers the primes that survive the preliminary random sieve. The argument has two deep inputs, both due to Ford, Green, Konyagin, Maynard, and Tao (abbreviated **FGKMT**):

1. a multidimensional sieve weight that favours residue classes containing several primes; and
2. a hypergraph covering theorem that chooses one such residue class for each final prime without wasting too much coverage through overlaps.

We state precisely the versions of those two established theorems that we use. We do **not** reprove them here. Reproving the first would require the Bombieri–Vinogradov theorem and the Maynard multidimensional sieve, while reproving the second would require the full Rödl-nibble argument. Everything else in this section is proved in detail from those two inputs and the results of the preceding sections.

> **Dependency note.** This is standard proof-by-citation: once the two quoted FGKMT theorems are accepted, the proof below is complete. The cited source is K. Ford, B. Green, S. Konyagin, J. Maynard, and T. Tao, *Long gaps between primes*, Journal of the American Mathematical Society **31** (2018), 65–105.

6.1. Notation and facts already proved

Let x tend to infinity and put

$$L := \log x, \qquad L_2 := \log\log x, \qquad L_3 := \log\log\log x. \tag{6.1}$$

The symbol $O_{\le}(\eta)$ denotes a quantity whose absolute value is at most η. Thus $X = Y + O_{\le}(\eta)$ means that the distance between X and Y is at most η.

The parameters from the preliminary sieve satisfy

$$y = xH, \qquad H = \frac{cL}{t}, \qquad t \asymp L_3, \qquad w = L^{B_0}, \qquad B_0 = 100, \tag{6.2}$$

where $c > 0$ is a sufficiently small fixed constant. In particular,

$$y = x^{1+o(1)}, \qquad xr^2 = o(y), \qquad x \le y \le \frac{xL^2}{10} \tag{6.3}$$

for the value of r defined below.

For each prime $s \le w$, the preliminary sieve chooses the residue 0 modulo s. For each prime $w < s \le x/2$, independently, it chooses a random residue $a_s \pmod s$ according to

$$\beta_s := \frac{(s-1)s^{-t/L}}{s-1+s^{-t/L}}, \tag{6.4}$$

$$\mathbb{P}(a_s = 0) = 1 - \beta_s, \qquad \mathbb{P}(a_s = b) = \frac{\beta_s}{s-1} \quad (b \not\equiv 0 \pmod s). \tag{6.5}$$

For a realization $a = (a_s)$, let

$$S(a) := \{n \in \mathbb{Z} : n \not\equiv a_s \pmod s \text{ for every prime } s \le x/2\}. \tag{6.6}$$

We usually write simply S when the random realization is understood. Let

$$\mathcal{Q} := \{q : q \text{ is prime and } x < q \le y\}. \tag{6.7}$$

Because every $q \in \mathcal{Q}$ is coprime to every coordinate prime $s \le x/2$, its survival probability is

$$A := \prod_{w < s \le x/2} \left(1 - \frac{\beta_s}{s-1}\right). \tag{6.8}$$

The preliminary estimates proved earlier give

$$A = (e^{\gamma}B_0 + o(1))\frac{tL_2}{L}, \qquad AH = (ce^{\gamma}B_0 + o(1))L_2, \tag{6.9}$$

and, with probability $1 - o(1)$,

$$|\mathcal{Q} \cap S| = (1 + o(1))A|\mathcal{Q}| = (ce^{\gamma}B_0 + o(1))\frac{xL_2}{L}. \tag{6.10}$$

For the prime layer set

$$r := \left\lfloor (\log(x/4))^{1/5} \right\rfloor. \tag{6.11}$$

Fix distinct positive integers

$$\mathcal{H} = \{h_1, \dots, h_r\} \subseteq [1, 2r^2] \tag{6.12}$$

that form an *admissible tuple*. This means that, for every prime ℓ, the residues $h_1, \dots, h_r \pmod{\ell}$ do not occupy all ℓ residue classes. Equivalently, for every prime ℓ, there is an integer n for which none of $n + h_1, \dots, n + h_r$ is divisible by ℓ.

For a finite set $U \subseteq \mathcal{Q}$, define

$$\nu_s(U) := |U \pmod{s}|, \tag{6.13}$$

the number of distinct residue classes modulo s occupied by U. Its exact preliminary survival probability is

$$\Pi(U) := \mathbb{P}(U \subseteq S) = \prod_{w < s \le x/2} \left(1 - \frac{\beta_s \nu_s(U)}{s - 1}\right). \tag{6.14}$$

Indeed, all elements of U are nonzero modulo s, and the only forbidden outcomes for a_s are the $\nu_s(U)$ nonzero residue classes occupied by U. In particular, $\Pi(\{q\}) = A$ for every $q \in \mathcal{Q}$.

Lemma 6.1 (survival of a small set)
Uniformly for every set $U \subseteq \mathcal{Q}$ with $|U| \le 2r$,

$$\Pi(U) = \big(1 + O(L^{-18})\big)A^{|U|}. \tag{6.15}$$

Consequently,

$$A^{-r} = x^{o(1)}, \qquad \log A^{-r} = O(rL_2) = o(L). \tag{6.16}$$

Proof. Put $k := |U|$ and

$$c_s := \frac{\beta_s}{s - 1}.$$

Since $0 < c_s \le 1/(s-1)$, we have $c_s \ll 1/s$. From (6.8) and (6.14),

$$\log \frac{\Pi(U)}{A^k} = \sum_{w < s \le x/2} \left(\log(1 - c_s\nu_s(U)) - k\log(1 - c_s)\right). \tag{6.17}$$

First suppose that the elements of U occupy distinct residues modulo s. Then $\nu_s(U) = k$. Taylor's formula $\log(1 - z) = -z + O(z^2)$, valid here because $kc_s = o(1)$, gives

$$\log(1 - kc_s) - k\log(1 - c_s) = O(k^2c_s^2) = O(k^2s^{-2}). \tag{6.18}$$

In general, the same Taylor expansion gives

$$\log(1 - c_s\nu_s(U)) - k\log(1 - c_s) = O\left(\frac{k - \nu_s(U)}{s} + \frac{k^2}{s^2}\right). \tag{6.19}$$

Whenever $k - \nu_s(U) > 0$, at least two elements of U are congruent modulo s. More quantitatively,

$$k - \nu_s(U) \le \#\{\{u, v\} \subseteq U : u \ne v,\ s \mid u - v\}. \tag{6.20}$$

For any nonzero integer d,

$$\sum_{\substack{s > w \\ s \mid d}} \frac{1}{s} \le \frac{1}{w \log w} \sum_{s \mid d} \log s \le \frac{\log |d|}{w \log w}. \tag{6.21}$$

Every difference $u - v$ has size at most $y = x^{1+o(1)}$, so $\log |u - v| = O(L)$. There are $O(k^2)$ pairs. Summing (6.19), and using (6.20)–(6.21), gives

$$\left|\log \frac{\Pi(U)}{A^k}\right| \ll k^2 \sum_{s > w} \frac{1}{s^2} + \frac{k^2 L}{w \log w}. \tag{6.22}$$

Now $k \le 2r \ll L^{1/5}$ and $w = L^{100}$. Both terms on the right are $O(L^{-18})$. Exponentiating proves (6.15).

Finally, (6.9) implies $\log(1/A) = O(L_2)$. Since $r = O(L^{1/5})$,

$$\log A^{-r} = r \log(1/A) = O(rL_2) = o(L),$$

which is equivalent to (6.16). □

6.2. The two external FGKMT inputs

The following statements are quoted results. They are included with all the hypotheses needed below so that every application can be checked explicitly.

FGKMT Input I (uniform multidimensional sieve estimate)

This is the specialization to the integers and to level $\vartheta = 1/3$ of FGKMT, Lemma 7.2 and Theorem 6.

Let X be sufficiently large. There is an integer $B \le X$, equal to 1 or to a prime, with the following property. The same B works for every base T with

$$X \le T \le X(\log X)^2. \tag{6.23}$$

Let k be an integer satisfying

$$k_0 \le k \le (\log X)^{1/5}, \tag{6.24}$$

where k_0 is an absolute constant. Let

$$\mathcal{L} = (L_1, \dots, L_k), \qquad L_j(n) = a_j n + b_j, \tag{6.25}$$

be an admissible family of distinct integer linear forms satisfying

$$1 \le |a_j| \le \log X, \qquad |b_j| \le X(\log X)^2. \tag{6.26}$$

In FGKMT Theorem 6 one also writes a coefficient-size hypothesis $|a_j|, |b_j| \le T^\alpha$. Here $\alpha = 2$; (6.23) and (6.26) imply that hypothesis for all sufficiently large X.

Let $\mathcal{L}' \subseteq \mathcal{L}$ be a distinguished subfamily such that every $L_j \in \mathcal{L}'$ is nonnegative on $[T, 2T]$ and $(a_j, B) = 1$. Finally choose

$$T^{1/30} \le R \le T^{1/9}. \tag{6.27}$$

For a prime $\ell \nmid B$, let

$$\omega_{\mathcal{L}}(\ell) := \#\{n \pmod{\ell} : \ell \mid L_1(n)\cdots L_k(n)\}. \tag{6.28}$$

Define the B-deleted singular series

$$\mathfrak{S}_B(\mathcal{L}) := \prod_{\ell \nmid B} \left(1 - \frac{\omega_{\mathcal{L}}(\ell)}{\ell}\right)\left(1 - \frac{1}{\ell}\right)^{-k}. \tag{6.29}$$

Then FGKMT construct a nonnegative weight

$$w_{k,\mathcal{L},B,R}(n) = \left(\sum_{d_1 \mid L_1(n),\dots,d_k \mid L_k(n)} \lambda_{d_1,\dots,d_k}(\mathcal{L})\right)^2. \tag{6.30}$$

There is one smooth function F_k, depending only on k, from which all the coefficients λ are constructed. Associated with F_k are positive quantities I_k, J_k satisfying

$$I_k \gg (2k\log k)^{-k}, \qquad J_k \asymp \frac{\log k}{k} I_k. \tag{6.31}$$

The following estimates hold uniformly for every family and every base satisfying the hypotheses above:

$$\sum_{T \le n \le 2T} w_{k,\mathcal{L},B,R}(n) = \Big(1 + O((\log T)^{-1/10})\Big)\left(\frac{B}{\varphi(B)}\right)^k \mathfrak{S}_B(\mathcal{L}) T (\log R)^k I_k, \tag{6.32}$$

where replacing T by the exact number of integers in $[T, 2T]$ changes only an $O(T^{-1})$ relative error. If $L_*(n) = a_* n + b_*$ belongs to $\mathcal{L}'$, is larger than R throughout $[T, 2T]$, and

$$N_*(T) := \#\{n \in [T, 2T] \cap \mathbb{Z} : L_*(n) \text{ is prime}\},$$

then

$$\begin{aligned}
&\sum_{T \le n \le 2T} 1_{\{L_*(n) \text{ prime}\}} w_{k,\mathcal{L},B,R}(n) \\
&\quad = \Big(1 + O((\log T)^{-1/10})\Big)\frac{\varphi(|a_*|)}{|a_*|}\left(\frac{B}{\varphi(B)}\right)^{k-1} \mathfrak{S}_B(\mathcal{L}) N_*(T) (\log R)^{k+1} J_k \\
&\qquad + O\left(\left(\frac{B}{\varphi(B)}\right)^k \mathfrak{S}_B(\mathcal{L}) T (\log R)^{k-1} I_k\right).
\end{aligned} \tag{6.33}$$

Finally,

$$w_{k,\mathcal{L},B,R}(n) \ll T^{2/9 + o(1)}. \tag{6.34}$$

All implied constants above are absolute in this specialization.

We shall also use one explicit feature of the coefficient construction. Put

$$W_k := \prod_{\substack{\ell \le 2k^2 \\ \ell \nmid B}} \ell \tag{6.35}$$

and define

$$\mathfrak{S}_{W_k B}(\mathcal{L}) := \prod_{\ell \nmid W_k B} \left(1 - \frac{\omega_{\mathcal{L}}(\ell)}{\ell}\right)\left(1 - \frac{1}{\ell}\right)^{-k}. \tag{6.36}$$

Every prime divisor of a supported product $d_1 \cdots d_k$ in (6.30) is at most R, is coprime to $W_k B$, and divides at most one of the d_j. Apart from factors depending only on k, B, R, F_k, the coefficient array depends on $\mathcal{L}$ only through $\mathfrak{S}_{W_k B}(\mathcal{L})$, the root counts $\omega_{\mathcal{L}}(\ell)$, and the coordinate labels of those roots. Consequently, if two families have the same supported divisor set, the same root counts and coordinate labels at every supported prime, and their deleted singular series differ by a scalar ρ, then their entire coefficient arrays differ by that same scalar ρ. This is immediate from FGKMT's coefficient definition and is the feature used in (6.57) and (6.67) below.

FGKMT Input II (quantitative hypergraph covering)

This is FGKMT, Corollary 4, together with the essential-support conclusion of their Theorem 3(a), which is used in the proof of that corollary.

Let $\mathcal{P}'$ and $\mathcal{Q}'$ be finite sets such that

$$|\mathcal{P}'| \le x, \qquad |\mathcal{Q}'| > L_2^3. \tag{6.37}$$

For each $p \in \mathcal{P}'$, let $\mathbf{e}_p$ be a random subset of $\mathcal{Q}'$. The random sets for different p need not be independent. Suppose that:

1. **Rank bound.** Almost surely,
$$|\mathbf{e}_p| \le r_*, \qquad r_* = O\left(\frac{LL_3}{L_2^2}\right). \tag{6.38}$$
2. **Point sparsity.** For every $p \in \mathcal{P}'$ and $q \in \mathcal{Q}'$,
$$\mathbb{P}(q \in \mathbf{e}_p) \le x^{-3/5}. \tag{6.39}$$
3. **Almost-uniform degrees.** There is a number C, independent of q, with
$$\frac{5}{4}\log 5 \le C \ll 1, \tag{6.40}$$
such that, for all but at most $|\mathcal{Q}'|/L_2^2$ vertices q,
$$\sum_{p \in \mathcal{P}'} \mathbb{P}(q \in \mathbf{e}_p) = C + O_{\le}(L_2^{-2}). \tag{6.41}$$
4. **Small codegrees.** For every two distinct vertices $q_1, q_2 \in \mathcal{Q}'$,
$$\sum_{p \in \mathcal{P}'} \mathbb{P}(q_1, q_2 \in \mathbf{e}_p) \le x^{-1/20}. \tag{6.42}$$

Then, for every positive integer

$$m \le \frac{L_3}{\log 5}, \tag{6.43}$$

there are new random sets $\mathbf{e}'_p \subseteq \mathcal{Q}'$ such that, with probability $1 - o(1)$,

$$\#\{q \in \mathcal{Q}' : q \notin \mathbf{e}'_p \text{ for every } p \in \mathcal{P}'\} \sim 5^{-m}|\mathcal{Q}'|. \tag{6.44}$$

Moreover, for every p, every value in the essential range of $\mathbf{e}'_p$ is either the empty set or a value in the essential range of $\mathbf{e}_p$. In plain language, the covering theorem may discard a color, but it never invents an edge that the original random law could not select.

6.3. The Maynard weight in the present parameters

For a prime $p \in (x/2, x]$ and an integer n, put

$$E_{p,n} := \{n + h_i p : 1 \le i \le r\} \cap \mathcal{Q}. \tag{6.45}$$

Thus $E_{p,n}$ is the set of target primes captured by the residue class $n \pmod p$ through the chosen shifts h_i.

Proposition 6.2 (uniform Maynard weight)

Assume

$$xr^2 = o(y), \qquad x \le y \le \frac{xL^2}{10}, \qquad \log y \sim L. \tag{6.46}$$

There are nonnegative weights $w_M(p,n)$, supported on primes $x/2 < p \le x$ and integers $-y \le n \le y$, a quantity $u \asymp \log r$, and a normalization $\alpha_M \ge x^{-o(1)}$ such that, uniformly in p, $q \in \mathcal{Q}$, and $1 \le i \le r$,

$$\sum_n w_M(p,n) = \big(1 + O(L_2^{-10})\big)\,\alpha_M \frac{y}{L^r}, \tag{6.47}$$

$$\sum_{\substack{x/2<p\le x\\ p \text{ prime}}} w_M(p, q - h_i p) = \big(1 + O(L_2^{-10})\big)\,\alpha_M \frac{u}{r}\frac{x}{2L^r}, \tag{6.48}$$

and

$$w_M(p,n) \le x^{1/3+o(1)}. \tag{6.49}$$

The exceptional prime B, the smooth function F_r, the sieve level R, and the reference coefficient array are the same for all p, q, i and throughout the range (6.46).

Proof. We apply FGKMT Input I with the outer parameter

$$X_0 := x/2, \qquad R := (x/4)^{1/9}. \tag{6.50}$$

We now verify every hypothesis.

The two bases. We shall use the bases $T = X_0$ and $T = 2y$. The lower bound $T \ge X_0$ is clear. Also,

$$2y \le \frac{xL^2}{5} < \frac{x}{2}(\log(x/2))^2 = X_0(\log X_0)^2$$

for all sufficiently large x. Hence both bases lie in (6.23), and the same exceptional prime $B \le X_0$ works for both.

The number of forms. From (6.11),

$$r \le (\log(x/4))^{1/5} < (\log X_0)^{1/5}.$$

Also $r \to \infty$, so $r \ge k_0$ eventually.

The sieve level. For $T = X_0$ and $T = 2y$,

$$T^{1/30} \le R \le T^{1/9}.$$

The upper inequality is immediate because $T \ge x/2 > x/4$. For the lower inequality, take logarithms. Since $\log(2y) = L + O(L_2)$,

$$\frac{1}{9}\log(x/4) - \frac{1}{30}\log(2y) = \left(\frac{1}{9} - \frac{1}{30}\right)L + O(L_2) > 0$$

for large x.

We next make the coefficient system common to all the families. For a prime ℓ, let

$$\nu_\ell := |\{h_1, \dots, h_r\} \pmod{\ell}|.$$

Define two reference singular series

$$\mathfrak{S}_0 := \prod_{\ell \nmid B} \left(1 - \frac{\nu_\ell}{\ell}\right) \left(1 - \frac{1}{\ell}\right)^{-r}, \tag{6.51}$$

$$\mathfrak{S}_{W_r,0} := \prod_{\ell \nmid W_r B} \left(1 - \frac{\nu_\ell}{\ell}\right) \left(1 - \frac{1}{\ell}\right)^{-r}, \tag{6.52}$$

Here

$$W_r := \prod_{\substack{\ell \le 2r^2 \\ \ell \nmid B}} \ell.$$

The tuple $\mathcal{H}$ is admissible, so $\nu_\ell < \ell$. For $\ell \le 2r^2$, every factor in (6.51) is positive and their product is at least $\exp(-O(r^2 \log r))$. For $\ell > 2r^2$, the h_i are distinct modulo ℓ, so $\nu_\ell = r$, and Taylor expansion gives

$$\sum_{\ell > 2r^2} \left| \log \left[\left(1 - \frac{r}{\ell}\right) \left(1 - \frac{1}{\ell}\right)^{-r} \right] \right| \ll r^2 \sum_{\ell > 2r^2} \ell^{-2} = O(1).$$

Therefore

$$\mathfrak{S}_0 \ge \exp(-O(r^2 \log r)) = x^{-o(1)}. \tag{6.53}$$

For a final prime $p \in (x/2, x]$, consider the ordered family

$$\mathcal{L}_p = (n + h_1 p, \dots, n + h_r p). \tag{6.54}$$

If $\ell \ne p$, multiplication by p permutes the residue classes modulo ℓ, so this family has ν_ℓ roots modulo ℓ. At $\ell = p$, every form is congruent to n, so there is exactly one root. Since $p > 2r^2$, the reference family has r roots at p. Put

$$\rho(z) := \frac{1 - 1/z}{1 - r/z}. \tag{6.55}$$

Comparing the single Euler factor at p gives the exact identities

$$\mathfrak{S}_B(\mathcal{L}_p) = \rho(p)\mathfrak{S}_0, \qquad \mathfrak{S}_{W_r B}(\mathcal{L}_p) = \rho(p)\mathfrak{S}_{W_r,0}. \tag{6.56}$$

Every supported prime in the coefficient array is at most $R < p$, exceeds $2r^2$, and is coprime to $W_r B$. At such a prime all r roots are distinct and all r coordinate labels occur. Hence the supported divisor set and every nonsingular part of the coefficient construction are common. If $\lambda^{(0)}_{d_1,\dots,d_r}$ is the reference array formed using $\mathfrak{S}_{W_r,0}$, then

$$\lambda_{d_1,\dots,d_r}(\mathcal{L}_p) = \rho(p)\lambda^{(0)}_{d_1,\dots,d_r}. \tag{6.57}$$

Define

$$w_M(p, n) := 1_{[-y,y]}(n)\, w_{r,\mathcal{L}_p,B,R}(n). \tag{6.58}$$

Let I_r, J_r be the quantities from FGKMT Input I and set

$$\alpha_M := 2 \left(\frac{B}{\varphi(B)}\right)^r \mathfrak{S}_0 (\log R)^r L^r I_r, \tag{6.59}$$

$$u := \frac{\varphi(B)}{B} \frac{\log R}{L} \frac{r J_r}{2 I_r}. \tag{6.60}$$

Since B is 1 or a prime, $\varphi(B)/B \asymp 1$. Also $\log R \asymp L$, and (6.31) gives

$$u \asymp \log r. \tag{6.61}$$

Equations (6.31) and (6.53) imply

$$\alpha_M \geq x^{-o(1)}. \tag{6.62}$$

We prove (6.47). Choose an integer a with $|a - 3y| \leq 1$, put $m = n + a$, and define

$$L_{p,a,j}(m) := m + h_j p - a. \tag{6.63}$$

The integer intervals $[a - y, a + y]$ and $[2y, 4y]$ differ in only $O(1)$ points. Translation does not change root counts, singular series, or coefficient arrays. The new slopes all equal 1, while

$$|h_j p - a| \leq 3y + 2xr^2 + 1 < X_0(\log X_0)^2 \tag{6.64}$$

for large x, by (6.46). Thus all coefficient hypotheses hold at the base $T = 2y$. We use (6.32) with no distinguished prime-detecting form. Equations (6.56) and (6.59) give

$$\sum_n w_M(p, n) = \Big(1 + O((\log(2y))^{-1/10}) + O(r/x)\Big)\, \alpha_M \frac{y}{L^r} \tag{6.65}$$

apart from the $O(1)$ endpoint values. By (6.34), those values contribute at most $x^{2/9+o(1)}$, whereas

$$\alpha_M \frac{y}{L^r} = x^{1-o(1)}.$$

They are therefore negligible. Since $(\log(2y))^{-1/10} = o(L_2^{-10})$ and $r/x = o(L_2^{-10})$, this proves (6.47).

We next prove (6.48). Fix $q \in \mathcal{Q}$ and $i \in \{1, \dots, r\}$. Consider the transformed family

$$\widetilde{L}_i(n) := n, \qquad \widetilde{L}_j(n) := q + (h_j - h_i)n \quad (j \neq i). \tag{6.66}$$

This family is admissible. To see this without skipping the modular check, fix a prime $\ell \neq q$. The roots are 0, together with

$$n \equiv -q(h_j - h_i)^{-1} \pmod{\ell}$$

for those j for which $h_j \not\equiv h_i \pmod{\ell}$. The map $z \mapsto -q/(z - h_i)$ is injective on residue classes $z \neq h_i$, so the number of roots is exactly ν_ℓ. At $\ell = q$, every form has the single root 0, because $q > 2r^2 > |h_j - h_i|$. Thus

$$\mathfrak{S}_B(\widetilde{\mathcal{L}}) = \rho(q)\mathfrak{S}_0, \qquad \lambda_d(\widetilde{\mathcal{L}}) = \rho(q)\lambda_d^{(0)}. \tag{6.67}$$

At $n = q - h_i p$,

$$n + h_j p = q + (h_j - h_i)p = \widetilde{L}_j(p). \tag{6.68}$$

In coordinate i, any supported divisor must divide q on the left and p on the right. Since $p, q > R$, that divisor is 1. Using (6.57) and (6.67) term by term in the square (6.30) gives the exact comparison

$$w_{r,\mathcal{L}_p,B,R}(q - h_i p) = \left(\frac{\rho(p)}{\rho(q)}\right)^2 w_{r,\widetilde{\mathcal{L}},B,R}(p). \tag{6.69}$$

This is the point at which both scalar factors must be retained. The prime-detecting estimate for the transformed family contains the singular series factor $\rho(q)$; after (6.69), the net factor is $\rho(p)^2/\rho(q)$, not $\rho(q)$. Uniformly for $p, q > x/2$,

$$\frac{\rho(p)^2}{\rho(q)} = 1 + O(r/x). \tag{6.70}$$

The support condition also holds:

$$q - h_i p \leq y, \qquad q - h_i p \geq -2xr^2 > -y$$

for large x, by $xr^2 = o(y)$.

We apply (6.33) at the base $T = X_0 = x/2$, with the distinguished form $\widetilde{L}_i(n) = n$. Its slope is 1, it is positive and larger than R on $[x/2, x]$, and it is coprime to B. All transformed slopes have absolute value at most $2r^2 < \log X_0$, and all intercepts have absolute value at most $y < X_0(\log X_0)^2$. The prime number theorem gives

$$\#\{p : x/2 < p \le x,\ p \text{ prime}\} = \big(1 + o(L_2^{-10})\big)\,\frac{x}{2L}. \tag{6.71}$$

The main term in (6.33), after using (6.70), is therefore

$$\big(1 + O(L_2^{-10})\big)\left(\frac{B}{\varphi(B)}\right)^{r-1} \mathfrak{S}_0 \frac{x}{2L}(\log R)^{r+1} J_r. \tag{6.72}$$

Using (6.59)–(6.60), this is exactly

$$\big(1 + O(L_2^{-10})\big)\,\alpha_M \frac{u}{r}\frac{x}{2L^r}. \tag{6.73}$$

The secondary term in (6.33), divided by this main term, is

$$O\left(\frac{r}{u\log R}\right) = O\left(\frac{L^{1/5}}{L\log r}\right) = o(L_2^{-10}). \tag{6.74}$$

This proves (6.48). Finally, (6.34), at either of our two bases, gives

$$w_M(p,n) \ll x^{2/9+o(1)} \le x^{1/3+o(1)},$$

which proves (6.49). All choices $B, F_r, R, \lambda^{(0)}$ were common, so the asserted uniformity has also been proved. □

6.4. Splitting the final primes into two colors

Normalize the Maynard weights separately for each final prime p:

$$W_p := \sum_n w_M(p,n), \qquad \mu_p(n) := \frac{w_M(p,n)}{W_p}. \tag{6.75}$$

Thus μ_p is a probability distribution on the integers $n \in [-y, y]$.

Lemma 6.3 (deterministic color split)

There is a partition

$$\{p : x/2 < p \le x,\ p \text{ prime}\} = \mathcal{P}_Q \sqcup \mathcal{P}_C \tag{6.76}$$

such that both sets have cardinality $\asymp x/L$,

$$M := \max_{p,n} \mu_p(n) \le x^{-2/3+o(1)}, \tag{6.77}$$

and, uniformly for $q \in \mathcal{Q}$,

$$\sum_{p\in\mathcal{P}_Q}\sum_{i=1}^{r} \mu_p(q - h_i p) = \big(1 + O(L_2^{-10})\big)\,C_0, \tag{6.78}$$

where, on putting

$$\delta := L^{-4}, \tag{6.79}$$

we have chosen

$$C_0 := 3A(1+\delta). \tag{6.80}$$

The partition is fixed before the preliminary residues are sampled.

Proof. From (6.47), (6.49), and $\alpha_M \geq x^{-o(1)}$,

$$W_p = x^{1-o(1)}, \qquad \mu_p(n) \leq \frac{x^{1/3+o(1)}}{x^{1-o(1)}} \leq x^{-2/3+o(1)},$$

which proves (6.77).

Divide (6.48) by (6.47) and sum over i. Uniformly for q,

$$\Sigma_q^{\mathrm{all}} := \sum_{\substack{x/2<p\leq x \\ p \text{ prime}}} \sum_{i=1}^{r} \mu_p(q-h_i p) = \left(1+O(L_2^{-10})\right)\Sigma_0, \tag{6.81}$$

where

$$\Sigma_0 := \frac{ux}{2y} = \frac{u}{2H}. \tag{6.82}$$

Choose

$$\theta_x := \frac{C_0}{\Sigma_0} = \frac{6Ay(1+\delta)}{ux}. \tag{6.83}$$

By (6.9), $AH \asymp cL_2$, while $u \asymp \log r \asymp L_2$. Thus $\theta_x \asymp c$. Choosing the fixed constant $c > 0$ small enough puts θ_x in a fixed closed subinterval of $(0,1)$.

Independently retain each final prime p with probability θ_x. Let ξ_p be its 0-1 indicator and set

$$a_{p,q} := \sum_{i=1}^{r} \mu_p(q-h_i p).$$

By (6.77),

$$0 \leq a_{p,q} \leq rM = x^{-2/3+o(1)}. \tag{6.84}$$

We use the following standard two-sided form of Bernstein's inequality. If X_j are independent, mean-zero random variables with $|X_j| \leq b$ and $\sum_j \mathbb{E}X_j^2 \leq V$, then

$$\mathbb{P}\left(\left|\sum_j X_j\right| \geq z\right) \leq 2\exp\left(-\frac{z^2}{2(V+bz/3)}\right). \tag{6.85}$$

Apply this with $X_p = (\xi_p - \theta_x)a_{p,q}$. Here

$$b \leq rM, \qquad V \leq \theta_x(rM)\sum_p a_{p,q}.$$

Take $z = L_2^{-10}\theta_x\Sigma_0$. Since $\theta_x\Sigma_0 = C_0 \asymp A = x^{-o(1)}$, Bernstein's exponent is $x^{2/3-o(1)}$. Hence

$$\mathbb{P}\left(\left|\sum_p (\xi_p-\theta_x)a_{p,q}\right| > L_2^{-10}\theta_x\Sigma_0\right) \leq 2e^{-x^{2/3-o(1)}}. \tag{6.86}$$

There are at most $y = x^{1+o(1)}$ possible q. A union bound shows that, with probability $1-o(1)$, the estimate in (6.86) holds simultaneously for all $q \in \mathcal{Q}$. The ordinary Chernoff inequality also shows that the number of retained primes is $\asymp x/L$, and so is the number not retained. Fix one outcome for which all these conclusions hold. Let $\mathcal{P}_Q$ be the retained set and $\mathcal{P}_C$ its complement. Combining (6.81), (6.83), and (6.86) proves (6.78). □

6.5. Turning the weights into genuine probability laws

The distribution $\mu_p(n)$ was constructed before the preliminary sieve. After the preliminary residues are known, we want to favour only those choices for which the entire candidate set $E_{p,n}$ survived. Define the exact importance factor

$$G_{p,n} := \frac{1_{\{E_{p,n}\subseteq S\}}}{\Pi(E_{p,n})} \tag{6.87}$$

Define its normalizing sum by

$$Y_p := \sum_n \mu_p(n) G_{p,n}. \tag{6.88}$$

For every fixed p, n,

$$\mathbb{E}G_{p,n} = \frac{\mathbb{P}(E_{p,n}\subseteq S)}{\Pi(E_{p,n})} = 1.$$

Since $\sum_n \mu_p(n) = 1$,

$$\mathbb{E}Y_p = 1 \tag{6.89}$$

exactly.

Proposition 6.4 (concentration of the prime-color normalizer)
Uniformly for $p \in \mathcal{P}_Q$,

$$\mathrm{Var}(Y_p) \ll L^{-18} + x^{-2/3+o(1)}. \tag{6.90}$$

Proof. For brevity write $E_n := E_{p,n}$ and $\mu(n) := \mu_p(n)$. Expanding the square and using (6.14),

$$\mathbb{E}Y_p^2 = \sum_{n,m} \mu(n)\mu(m) \frac{\Pi(E_n \cup E_m)}{\Pi(E_n)\Pi(E_m)}. \tag{6.91}$$

If $E_n \cap E_m = \varnothing$, Lemma 6.1 applied to E_n, E_m, and $E_n \cup E_m$ gives

$$\frac{\Pi(E_n \cup E_m)}{\Pi(E_n)\Pi(E_m)} = 1 + O(L^{-18}). \tag{6.92}$$

If the two sets overlap, then for some i, j,

$$n + h_i p = m + h_j p, \qquad m = n + (h_i - h_j)p. \tag{6.93}$$

For fixed i, j, n, equation (6.93) determines m. Therefore

$$\begin{aligned} \sum_{n,m:E_n\cap E_m\neq\varnothing} \mu(n)\mu(m) &\le \sum_{i,j}\sum_n \mu(n)\mu(n+(h_i-h_j)p) \\ &\le r^2 M \sum_n \mu(n) = r^2 M. \end{aligned} \tag{6.94}$$

Lemma 6.1 also gives, for an overlapping pair,

$$\begin{aligned} \frac{\Pi(E_n \cup E_m)}{\Pi(E_n)\Pi(E_m)} &\le (1 + O(L^{-18}))A^{|E_n\cup E_m|-|E_n|-|E_m|} \\ &= (1 + O(L^{-18}))A^{-|E_n\cap E_m|} \le 2A^{-r} \end{aligned} \tag{6.95}$$

for large x. Inserting (6.92), (6.94), and (6.95) into (6.91), and using $\mathbb{E}Y_p = 1$, gives

$$\mathrm{Var}(Y_p) \ll L^{-18} + A^{-r} r^2 M.$$

By (6.16), $A^{-r} = x^{o(1)}$; also $r^2 = x^{o(1)}$, and (6.77) gives $M \le x^{-2/3+o(1)}$. This proves (6.90). □

We now define the actual random edge of color p, after conditioning on a fixed preliminary realization. If $Y_p \le 1+\delta$, choose the integer n with probability

$$\frac{\mu_p(n)G_{p,n}}{1+\delta}. \tag{6.96}$$

The total mass assigned in (6.96) is $Y_p/(1+\delta) \le 1$; put the remaining mass on a dummy outcome whose edge is empty. If $Y_p > 1+\delta$, use only the dummy outcome. Thus this is a genuine probability law in every preliminary realization. No division by a random normalizer is being hidden.

6.6. Conditional degrees of the surviving target primes

For $q \in \mathcal{Q}$, define the uncapped weighted degree

$$D_q := \sum_{p\in\mathcal{P}_Q}\sum_{i=1}^{r}\mu_p(q-h_ip)G_{p,q-h_ip}. \tag{6.97}$$

Also put

$$S_q := \sum_{p\in\mathcal{P}_Q}\sum_{i=1}^{r}\mu_p(q-h_ip). \tag{6.98}$$

Every candidate set appearing in (6.97) contains q. Therefore

$$\begin{aligned}\mathbb{E}(G_{p,q-h_ip} \mid q\in S) &= \frac{\mathbb{P}(E_{p,q-h_ip}\subseteq S)}{\Pi(E_{p,q-h_ip})\mathbb{P}(q\in S)}\\ &= \frac{1}{A}.\end{aligned}$$

Consequently,

$$\mathbb{E}(D_q \mid q\in S) = \frac{S_q}{A}. \tag{6.99}$$

By Lemma 6.3,

$$S_q = \big(1+O(L_2^{-10})\big)\,C_0 \tag{6.100}$$

uniformly in q.

Proposition 6.5 (conditional prime-degree concentration)
With probability $1-o(1)$ over the preliminary sieve, all but

$$o\left(\frac{x}{LL_2}\right)$$

surviving primes $q \in \mathcal{Q}\cap S$ have capped marginal degree

$$3 + O_{\le}(L_2^{-2}). \tag{6.101}$$

Proof. Fix q. Write

$$a_{p,i} := \mu_p(q-h_ip), \qquad b_p := \sum_{i=1}^{r} a_{p,i}. \tag{6.102}$$

Since every $a_{p,i} \le M$,

$$b_p \le rM, \qquad \sum_p b_p^2 \le rM\sum_p b_p = rMS_q. \tag{6.103}$$

We need the second moment of D_q. Consider first two distinct colors $p \neq p'$. The rooted candidate sets

$$E_{p,q-h_i p}, \qquad E_{p',q-h_k p'}$$

intersect exactly in q. They certainly both contain q. If they had another common element, then for suitable indices $j \neq i$ and $\ell \neq k$,

$$(h_j - h_i)p = (h_\ell - h_k)p'. \tag{6.104}$$

Both nonzero coefficients have absolute value below $2r^2 < \min(p, p')$. Because p, p' are distinct primes, (6.104) would force p to divide the nonzero integer $h_\ell - h_k$, which is impossible. Thus the intersection is exactly $\{q\}$. Lemma 6.1 now gives

$$\mathbb{E}\,(G_{p,q-h_i p} G_{p',q-h_k p'}) = \frac{1 + O(L^{-18})}{A}. \tag{6.105}$$

Summing (6.105) over i, k and $p \neq p'$, the cross-color contribution is

$$\frac{1 + O(L^{-18})}{A} \sum_{p \neq p'} b_p b_{p'}. \tag{6.106}$$

For equal colors we do not need to determine all possible intersections. The crude bound (6.95) gives a contribution at most

$$O\left(A^{-r} \sum_p b_p^2\right) = O(A^{-r} r M S_q). \tag{6.107}$$

Since

$$\sum_{p \neq p'} b_p b_{p'} = S_q^2 - \sum_p b_p^2,$$

and the omitted diagonal term is also absorbed by (6.107), equations (6.103), (6.106), and (6.107) give

$$\mathbb{E} D_q^2 = \frac{S_q^2}{A} + O\left(\frac{L^{-18} S_q^2}{A} + A^{-r} r M S_q\right). \tag{6.108}$$

Unconditionally, $\mathbb{E} D_q = S_q$, and $D_q = 0$ whenever $q \notin S$. Hence $D_q 1_{\{q \in S\}} = D_q$. Expanding the square gives the exact cancellation

$$\begin{aligned} \mathbb{E}\left(D_q - \frac{C_0}{A} 1_{\{q \in S\}}\right)^2 &= \mathbb{E} D_q^2 - 2\frac{C_0}{A}\mathbb{E} D_q + \frac{C_0^2}{A^2}\mathbb{P}(q \in S) \\ &= \mathbb{E} D_q^2 - \frac{S_q^2}{A} + \frac{(S_q - C_0)^2}{A}. \end{aligned} \tag{6.109}$$

By (6.100), $S_q - C_0 = O(L_2^{-10}) C_0$, and $S_q \asymp C_0 \asymp A$. The terms involving M are smaller than any fixed negative power of L_2, by (6.16) and (6.77). Summing (6.109) over $q \in \mathcal{Q}$, we obtain

$$\sum_{q \in \mathcal{Q}} \mathbb{E}\left(D_q - \frac{C_0}{A} 1_{\{q \in S\}}\right)^2 \ll L_2^{-20} |\mathcal{Q}| \frac{C_0^2}{A}. \tag{6.110}$$

Let

$$\rho_* := \frac{1}{10 L_2^2}. \tag{6.111}$$

Since $C_0 \asymp A$ and $A|\mathcal{Q}| \asymp xL_2/L$, the right side of (6.110) is $O(x/(LL_2^{19}))$. If N surviving targets satisfy

$$\left| D_q - \frac{C_0}{A} \right| > \rho_*,$$

then the random sum in (6.110) is at least $N\rho_*^2$. Markov's inequality therefore shows that, with probability $1 - o(1)$,

$$N = o\left(\frac{x}{LL_2}\right). \tag{6.112}$$

It remains to account for colors discarded by the cap $Y_p > 1 + \delta$. Put

$$\sigma_p^2 := \mathrm{Var}(Y_p).$$

Chebyshev's inequality gives

$$\mathbb{P}(Y_p > 1 + \delta) \le \frac{\sigma_p^2}{\delta^2}.$$

Also, by Cauchy–Schwarz,

$$\begin{aligned} \mathbb{E}[(Y_p - 1)1_{\{Y_p > 1+\delta\}}] &\le \sigma_p \mathbb{P}(Y_p > 1 + \delta)^{1/2} \\ &\le \frac{\sigma_p^2}{\delta}. \end{aligned}$$

Adding the contribution of 1,

$$\mathbb{E}[Y_p 1_{\{Y_p > 1+\delta\}}] \le \frac{\sigma_p^2}{\delta^2} + \frac{\sigma_p^2}{\delta}. \tag{6.113}$$

Let L_q^{cap} be the part of D_q coming from discarded colors. For a fixed discarded p, interchange the sums over q and n. A candidate edge $E_{p,n}$ contains at most r targets, so, deterministically,

$$\begin{aligned} \sum_{q \in \mathcal{Q} \cap S} L_q^{\mathrm{cap}} &\le r \sum_{p : Y_p > 1+\delta} \sum_n \mu_p(n) G_{p,n} \\ &= r \sum_{p : Y_p > 1+\delta} Y_p. \end{aligned} \tag{6.114}$$

By Proposition 6.4, $\sigma_p^2 \ll L^{-18} + x^{-2/3+o(1)}$. Since $\delta = L^{-4}$, $r = L^{1/5+o(1)}$, and $|\mathcal{P}_Q| \ll x/L$, equations (6.113)–(6.114) give

$$\mathbb{E} \sum_{q \in \mathcal{Q} \cap S} L_q^{\mathrm{cap}} \ll xL^{-10.8+o(1)}. \tag{6.115}$$

This is much smaller than $\rho_* x/(LL_2)$. A final application of Markov's inequality shows that, with probability $1 - o(1)$, only

$$o\left(\frac{x}{LL_2}\right)$$

surviving targets have $L_q^{\mathrm{cap}} > \rho_*$.

At every target outside the two exceptional sets, its marginal degree under the capped law (6.96) is exactly

$$\frac{D_q - L_q^{\mathrm{cap}}}{1 + \delta}. \tag{6.116}$$

Now $C_0/[A(1 + \delta)] = 3$, both additive errors are at most ρ_*, and $\delta = L^{-4} = o(L_2^{-2})$. Therefore (6.116) equals

$$3 + O_{\le}(L_2^{-2})$$

for all sufficiently large x. This proves the proposition. □

6.7. Applying the hypergraph covering theorem

Fix a preliminary realization that satisfies both Proposition 6.5 and the prime-survivor estimate (6.10). Let

$$\mathcal{B}(a) \subseteq \mathcal{Q} \cap S(a)$$

be the exceptional set from Proposition 6.5, so

$$|\mathcal{B}(a)| = o\left(\frac{x}{LL_2}\right),$$

and put

$$\mathcal{Q}' := \{q \in \mathcal{Q} \cap S(a) : q \notin \mathcal{B}(a)\}. \tag{6.117}$$

For a non-dummy outcome n of color p, define

$$\mathbf{e}_p := E_{p,n} \cap \mathcal{Q}'. \tag{6.118}$$

For the dummy outcome, set $\mathbf{e}_p = \varnothing$. These are random sets only with respect to the capped color laws; the preliminary realization is now fixed.

Proposition 6.6 (quantitative hypergraph application)
The random edges (6.118) satisfy every hypothesis of FGKMT Input II. Hence there is one supported residue choice for each $p \in \mathcal{P}_Q$ that leaves only $O(x/L)$ vertices of $\mathcal{Q}'$ uncovered.

Proof. We check the four hypotheses one at a time.

Sizes of the index and vertex sets. Clearly

$$|\mathcal{P}_Q| \le x.$$

By (6.10) and the smaller deletion in (6.117),

$$|\mathcal{Q}'| \asymp \frac{xL_2}{L} > L_2^3. \tag{6.119}$$

Rank. Every edge has at most r vertices. Moreover,

$$r = L^{1/5+o(1)} = o\left(\frac{LL_3}{L_2^2}\right), \tag{6.120}$$

so (6.38) holds.

Point sparsity. Fix $p \in \mathcal{P}_Q$ and $q \in \mathcal{Q}'$. If $q \in E_{p,n}$, then $n = q - h_i p$ for some i. Under the capped law,

$$\begin{aligned}\mathbb{P}(q \in \mathbf{e}_p) &\le \frac{1}{1+\delta} \sum_{i=1}^{r} \mu_p(q - h_i p)\Pi(E_{p,q-h_i p})^{-1} \\ &\le rM \max_{|U|\le r} \Pi(U)^{-1}.\end{aligned} \tag{6.121}$$

Lemma 6.1 and (6.16) imply

$$\max_{|U|\le r} \Pi(U)^{-1} \le 2A^{-r} = x^{o(1)}.$$

Together with (6.77),

$$\mathbb{P}(q \in \mathbf{e}_p) \le x^{-2/3+o(1)} \le x^{-3/5}. \tag{6.122}$$

This is (6.39).

Uniform degrees. No vertex of $\mathcal{Q}'$ is exceptional in Proposition 6.5. Therefore, for every $q \in \mathcal{Q}'$,

$$\sum_{p \in \mathcal{P}_Q} \mathbb{P}(q \in \mathbf{e}_p) = 3 + O_{\leq}(L_2^{-2}). \tag{6.123}$$

Thus (6.41) holds with $C = 3$, and

$$3 > \frac{5}{4} \log 5.$$

Codegrees. Let $q_1 \neq q_2$ be vertices of $\mathcal{Q}'$. If one edge of color p contains both, then all points in that edge are congruent modulo p, so

$$p \mid q_1 - q_2. \tag{6.124}$$

At most one prime $p \in (x/2, x]$ can divide $q_1 - q_2$. Indeed, if two distinct such primes p, p' divided the difference, then

$$pp' > x^2/4 > |q_1 - q_2|,$$

because $|q_1 - q_2| < y = o(x^2)$, which is impossible for a nonzero multiple of pp'. Therefore (6.122) gives

$$\sum_{p \in \mathcal{P}_Q} \mathbb{P}(q_1, q_2 \in \mathbf{e}_p) \leq x^{-3/5} < x^{-1/20}. \tag{6.125}$$

This proves the codegree hypothesis.

We may now apply FGKMT Input II with

$$m := \left\lfloor \frac{L_3}{\log 5} \right\rfloor. \tag{6.126}$$

It supplies output edges whose uncovered set has cardinality

$$(1 + o(1))5^{-m}|\mathcal{Q}'|. \tag{6.127}$$

Since

$$5^{-m} \asymp e^{-L_3} = L_2^{-1},$$

equation (6.119) turns (6.127) into

$$O(x/L). \tag{6.128}$$

Finally, the essential-support conclusion is important. By FGKMT Theorem 3(a), as inherited by their Corollary 4, every nonempty output edge is an edge that the capped input law could actually attain. Hence it has the form $E_{p,n} \cap \mathcal{Q}'$ for some supported integer n, and the residue class $n \pmod p$ covers it. Intersecting with $\mathcal{Q}'$ can only have removed vertices; the actual residue class may cover additional targets, which is harmless. Choosing one outcome for which (6.127) holds proves the proposition. □

6.8. What has been obtained

For one preliminary realization in the high-probability event, the primes in $\mathcal{P}_Q$ can be assigned residue classes that leave only

$$O(x/L)$$

nonexceptional surviving target primes uncovered. The exceptional set $\mathcal{B}(a)$ has size $o(x/(LL_2))$, which is smaller. The colors in $\mathcal{P}_C$ were never used in this section and remain available for the composite layer.

The two non-elementary facts in the argument are exactly FGKMT Inputs I and II. All parameter checks, changes of variables, probability normalizations, correlation calculations, cap-loss estimates, and the verification of the hypergraph hypotheses have been supplied above.

7. Assembly of the covering theorem

Throughout, log is the natural logarithm, $\log_1 u = \log u$, and $\log_{j+1} u = \log(\log_j u)$. For a real number $Z \geq 2$, let $Y(Z)$ be the largest nonnegative integer N for which one can select one residue class modulo every prime $p \leq Z$ so that all of $1, 2, \ldots, N$ are covered.

7.1. Parameters and exact inputs

We now state precisely what the assembly needs from the preceding sections. This list is also a dependency checklist for the main theorem.

Let

$$L := \log x, \qquad L_j := \log_j x \quad (j \geq 2).$$

Fix constants $D > 2$ and $c > 0$. Let $t > 1$ be the larger solution of

$$t - \log t = DL_3, \tag{7.1}$$

and set

$$H := \frac{cL}{t}, \qquad y := xH, \qquad w := L^{100}. \tag{7.2}$$

For sufficiently large x, the earlier sections must supply the following facts.

Input 1: the prime split. The primes in $(x/2, x]$ are partitioned into two disjoint sets $\mathcal{P}_C$ and $\mathcal{P}_Q$. The first set is reserved for the composite layer and the second for the prime layer.

Input 2: the preliminary residues. A random preliminary choice assigns one residue to every prime $p \leq x/2$. For every prime $p \leq w$, this residue is $0 \pmod p$. Let $\mathcal{S}(a)$ denote the integers not covered by a realization a of these preliminary residues.

Input 3: the good event and prime layer. There is a good event $\mathcal{G}$, depending only on the preliminary choice, with

$$\mathbb{P}(\mathcal{G}) = 1 - o(1). \tag{7.3}$$

For every $a \in \mathcal{G}$, the prime-layer result supplies one residue for each $p \in \mathcal{P}_Q$ and leaves only $O(x/L)$ prime targets in $(x, y]$ uncovered. This count includes the exceptional prime targets deleted before the hypergraph argument.

Input 4: the composite layer. Let $\mathcal{C}$ be the set of squarefree, w-rough composite integers in $(x, y]$, where w-rough means that no prime $p \leq w$ divides the integer. For a preliminary realization a, let $R_C(a)$ be the conditional expected number of elements of $\mathcal{C} \cap \mathcal{S}(a)$ left uncovered after the composite-layer residues for $p \in \mathcal{P}_C$ are chosen. The composite-layer result must give

$$\mathbb{E}_a R_C(a) = o(x/L). \tag{7.4}$$

Input 5: the remaining elementary count. The preliminary-stage estimate gives

$$\#\{n \in (x, y] : n \text{ is } w\text{-rough and not squarefree}\} = o(x/L). \tag{7.5}$$

No other property of the sieve weights or hypergraphs is used in the assembly. Accordingly, if any item in this checklist has not been proved, the conclusion of the next subsection is conditional on that item.

7.2. Proof of the covering theorem

We first record the size of the interval $(x, y]$. Put $A := DL_3$. For large x, the function $u \mapsto u - \log u$ is strictly increasing for $u > 1$. Moreover,

$$A - \log A < A$$

For sufficiently large A, we also have

$$2A - \log(2A) > A.$$

Therefore the solution $t > 1$ of (7.1) lies between A and $2A$. From (7.1),

$$t = A + \log t.$$

Since $A < t < 2A$, we have $\log t = \log A + O(1)$. Also

$$\log A = \log(DL_3) = \log D + \log L_3 = L_4 + O(1).$$

It follows that

$$t = DL_3 + L_4 + O(1) = (D + o(1))L_3. \tag{7.6}$$

Consequently,

$$H = \frac{cL}{t} = \Big(\frac{c}{D} + o(1)\Big)\,\frac{L}{L_3} \longrightarrow \infty. \tag{7.7}$$

Using $y = xH$, we obtain

$$y - x = x(H-1) = \Big(\frac{c}{D} + o(1)\Big)\,x\frac{L}{L_3}. \tag{7.8}$$

We now choose one preliminary realization that is good for both final layers. Since $R_C(a) \geq 0$, equations (7.3) and (7.4) imply

$$\begin{aligned} \mathbb{E}\big(R_C(a) \mid a \in \mathcal{G}\big) &= \frac{\mathbb{E}\big(R_C(a)\mathbf{1}_{a\in\mathcal{G}}\big)}{\mathbb{P}(\mathcal{G})} \\ &\leq \frac{\mathbb{E}R_C(a)}{\mathbb{P}(\mathcal{G})} = o(x/L). \end{aligned} \tag{7.9}$$

At least one $a \in \mathcal{G}$ therefore satisfies

$$R_C(a) = o(x/L); \tag{7.10}$$

otherwise every value in the conditional average in (7.9) would be larger than that average. Fix such an a.

For this fixed preliminary realization, $R_C(a)$ is the average, over the composite-layer choices, of the actual number of uncovered composite targets. Hence at least one deterministic set of composite-layer residues leaves at most $R_C(a) = o(x/L)$ of those targets uncovered. Fix such a choice. Because $a \in \mathcal{G}$, fix also the prime-layer residues whose existence is asserted in the third input above. These two selections are compatible: they use the disjoint prime sets $\mathcal{P}_C$ and $\mathcal{P}_Q$.

We next check that all kinds of surviving integers have been counted. Let $n \in (x, y]$ survive the preliminary residues. If a prime $p \leq w$ divided n, then $n \equiv 0 \pmod p$, and the preliminary residue $0 \pmod p$ would have covered n. Thus every preliminary survivor is w-rough. Such an integer is exactly one of the following:

1. a prime;
2. a squarefree composite, and hence an element of $\mathcal{C}$;
3. a nonsquarefree w-rough integer.

The prime layer leaves $O(x/L)$ integers of the first kind. The chosen composite-layer residues leave $o(x/L)$ integers of the second kind. By (7.5), there are $o(x/L)$ integers of the third kind even before the final two layers act. It follows that there is a fixed constant $M > 0$, independent of x, such that for every sufficiently large x the number of integers in $(x, y]$ still uncovered is at most

$$M\frac{x}{L}. \tag{7.11}$$

We cover these last integers one at a time. By the prime number theorem, for every fixed $C > 1$,

$$\pi(Cx) - \pi(x) = (C - 1 + o(1)) \frac{x}{\log x}. \tag{7.12}$$

Choose a fixed $C_{\mathrm{cl}} > 1$ so large that $C_{\mathrm{cl}} - 1 > M + 1$. Equation (7.12) then gives

$$\pi(C_{\mathrm{cl}}x) - \pi(x) > M\frac{x}{L} \tag{7.13}$$

for every sufficiently large x. Thus there are more primes in $(x, C_{\mathrm{cl}}x]$ than there are uncovered integers in $(x, y]$. Assign a different prime $\ell \in (x, C_{\mathrm{cl}}x]$ to each remaining integer n, and choose

$$a_\ell \equiv n \pmod{\ell}. \tag{7.14}$$

This residue covers n. Give every as-yet unused prime at most $C_{\mathrm{cl}}x$ the residue 0. We have now assigned exactly one residue to every prime $p \le C_{\mathrm{cl}}x$, and these residues cover every integer in $(x, y]$. Adding new residue classes cannot undo any coverage already obtained.

It remains to translate this interval to the initial interval in the definition of Y. Put

$$b := \lfloor x \rfloor, \qquad N := \lfloor y \rfloor - \lfloor x \rfloor. \tag{7.15}$$

If $1 \le h \le N$, then

$$x < b + h \le \lfloor y \rfloor \le y,$$

so some selected residue $a_p \pmod p$ covers $b + h$. Define the translated residue

$$\widetilde{a}_p \equiv a_p - b \pmod{p}. \tag{7.16}$$

From $b + h \equiv a_p \pmod p$ we get $h \equiv \widetilde{a}_p \pmod p$. Hence the translated residues cover every integer $h \in [1, N]$. Therefore

$$Y(C_{\mathrm{cl}}x) \ge N. \tag{7.17}$$

Since $N = y - x + O(1)$, equation (7.8) yields a constant $c_2 > 0$ such that

$$Y(C_{\mathrm{cl}}x) \ge c_2 x \frac{\log x}{\log_3 x} \tag{7.18}$$

for every sufficiently large x.

Finally, let X be an arbitrary sufficiently large real number and take

$$x := \left\lfloor \frac{X}{C_{\mathrm{cl}}} \right\rfloor.$$

Then $C_{\mathrm{cl}}x \le X$. The function Y is nondecreasing: a covering using primes up to a smaller bound remains a covering when arbitrary residues are assigned to the additional primes allowed by a larger bound. Consequently,

$$Y(X) \ge Y(C_{\mathrm{cl}}x).$$

Moreover, $x \asymp X$, $\log x = \log X + O(1)$, and $\log_3 x = \log_3 X + o(1)$. After decreasing the constant in (7.18), if necessary, we conclude that

$$Y(X) \ge c_0 X \frac{\log X}{\log_3 X}.$$

This proves the covering theorem, conditional only on the inputs listed above. If those inputs have effective constants and effective thresholds, then c_0 and the threshold here are effective as well.

8. From residue-class coverings to prime gaps

Throughout, log is the natural logarithm, $\log_1 u = \log u$, and $\log_{j+1} u = \log(\log_j u)$. Write

$$P(z) := \prod_{p \le z} p.$$

Let $Y(z)$ be the largest nonnegative integer N for which one residue class modulo each prime $p \le z$ covers $1, \dots, N$. If $2 = p_1 < p_2 < \cdots$ are the primes, let

$$G(T) := \max_{p_{n+1} \le T} (p_{n+1} - p_n).$$

8.1. An elementary exponential bound for the primorial

For the transfer to prime gaps, a bound of the form $P(z) \le C^z$ with one fixed C is enough. We prove a slightly weaker but especially transparent version of the usual Chebyshev bound.

Lemma 8.1. For every real $z \ge 1$,

$$P(z) = \prod_{p \le z} p \le 16^z. \tag{8.1}$$

Proof. First let n be a positive integer. Every prime $p \in (n, 2n]$ divides the central binomial coefficient

$$\binom{2n}{n} = \frac{(n+1)(n+2)\cdots(2n)}{1 \cdot 2 \cdots n}.$$

Indeed, the numerator contains the factor p, while the denominator has no factor divisible by p, because $p > n$. The primes in $(n, 2n]$ are distinct and therefore pairwise coprime, so their product divides $\binom{2n}{n}$. Also,

$$\binom{2n}{n} < \sum_{j=0}^{2n} \binom{2n}{j} = (1+1)^{2n} = 4^n. \tag{8.2}$$

Apply this with $n = 2^{j-1}$, for $j = 1, 2, \dots, k$. The intervals $(2^{j-1}, 2^j]$ partition the primes at most 2^k, and (8.2) gives

$$\begin{aligned} P(2^k) &= \prod_{j=1}^{k} \prod_{2^{j-1} < p \le 2^j} p \\ &< \prod_{j=1}^{k} 4^{2^{j-1}} = 4^{1+2+\cdots+2^{k-1}} = 4^{2^k - 1}. \end{aligned} \tag{8.3}$$

For $1 \le z \le 2$, (8.1) is immediate. Otherwise choose k so that $2^{k-1} < z \le 2^k$. The function P is nondecreasing, and $2^k < 2z$. Equation (8.3) therefore gives

$$P(z) \le P(2^k) < 4^{2^k} < 4^{2z} = 16^z.$$

This proves the lemma. □

The constant 16 is not important. A sharper elementary argument gives 4^z, but (8.1) leads to exactly the same iterated-logarithmic order in the prime-gap theorem.

8.2. The bounded CRT realization and covering-to-gap transfer

Lemma 8.2 (bounded Chinese-remainder realization). Suppose that $z \geq 2$ is an integer, that one residue $a_p \pmod p$ has been chosen for every prime $p \leq z$, and that these residues cover $[1, N]$. Then there is an integer m satisfying

$$z < m \leq z + P(z), \qquad m \equiv -a_p \pmod p \quad (p \leq z), \tag{8.4}$$

such that all of

$$m+1, m+2, \ldots, m+N$$

are composite.

Proof. The moduli in (8.4) are distinct primes and hence pairwise coprime. The Chinese remainder theorem therefore gives one solution class modulo $P(z)$. Any interval containing exactly $P(z)$ consecutive integers contains exactly one representative of that class, so there is a unique representative m in the half-open interval $(z, z + P(z)]$.

Let $1 \leq h \leq N$. Since the selected residues cover $[1, N]$, some prime $p \leq z$ satisfies $h \equiv a_p \pmod p$. Equation (8.4) then gives

$$m + h \equiv -a_p + a_p \equiv 0 \pmod p.$$

Thus $p \mid m + h$. But $m + h > z \geq p$, so $m + h$ is a positive multiple of p strictly larger than p. It is therefore composite. This holds for every $h \in [1, N]$. □

We can now prove the transfer theorem without any sieve theory.

Proposition 8.3 (covering-to-gap transfer). Suppose that constants $c_0 > 0$ and z_0 satisfy

$$Y(z) \geq c_0 z \frac{\log z}{\log_3 z} \tag{8.5}$$

for every $z \geq z_0$. Then there is a constant $c_1 > 0$ such that

$$G(T) \geq c_1 \frac{\log T \, \log_2 T}{\log_4 T} \tag{8.6}$$

for every sufficiently large T.

Proof. For an integer $z \geq z_0$, define

$$\ell(z) := \left\lfloor c_0 z \frac{\log z}{\log_3 z} \right\rfloor. \tag{8.7}$$

By (8.5), there are residues modulo the primes $p \leq z$ that cover at least the first $\ell(z)$ positive integers. Lemma 8.2 supplies an integer $m \leq z + P(z)$ for which

$$m+1, m+2, \ldots, m+\ell(z)$$

are all composite.

Bertrand's postulate already guarantees that there is a prime larger than $m + \ell(z)$. Let q be the first such prime, and let q_- be the prime immediately preceding q. (The predecessor exists because $q > m > z \geq 2$.) Since there is no prime in the displayed composite run, $q_- \leq m$. Hence

$$q - q_- > \ell(z). \tag{8.8}$$

We must also bound q from above. Bertrand's postulate says that for every integer $n > 1$ there is a prime strictly between n and $2n$. Apply it to $n = m + \ell(z)$. Since q is the first prime after n,

$$q < 2(m + \ell(z)) \leq 2\big(P(z) + z + \ell(z)\big). \tag{8.9}$$

Now let T be large and choose

$$z := \left\lfloor \frac{\log T}{2 \log 16} \right\rfloor. \tag{8.10}$$

Then $z \to \infty$ with T, so $z \geq z_0$ for all sufficiently large T. Lemma 8.1 and (8.10) give

$$P(z) \leq 16^z \leq T^{1/2}. \tag{8.11}$$

Also $z = O(\log T)$ and, from (8.7), $\ell(z) = O(\log T \, \log_2 T)$. Thus

$$z + \ell(z) = T^{o(1)} = o(T^{1/2}). \tag{8.12}$$

Substituting (8.11) and (8.12) into (8.9) gives

$$q \leq 2T^{1/2} + o(T^{1/2}) < T \tag{8.13}$$

for all sufficiently large T. Therefore the gap in (8.8) is one of the gaps included in the definition of $G(T)$. It remains to express $\ell(z)$ in terms of T, keeping track of every iteration of the logarithm. If

$$\alpha := \frac{1}{2 \log 16},$$

then (8.10) says

$$z = \alpha \log T + O(1). \tag{8.14}$$

Taking one logarithm gives

$$\log z = \log(\alpha \log T + O(1)) = \log_2 T + \log \alpha + o(1) = \log_2 T + O(1), \tag{8.15}$$

and therefore $\log z \sim \log_2 T$. Taking a second logarithm gives

$$\log_2 z = \log_3 T + o(1), \tag{8.16}$$

and taking a third gives

$$\log_3 z = \log_4 T + o(1). \tag{8.17}$$

The quantity inside the floor in (8.7) tends to infinity, so the floor changes it by a relative $o(1)$. Combining (8.7) and (8.14)–(8.17),

$$\ell(z) = (c_0 \alpha + o(1)) \, \frac{\log T \, \log_2 T}{\log_4 T}. \tag{8.18}$$

Equations (8.8), (8.13), and (8.18) prove (8.6), for example with any fixed $c_1 < c_0/(2 \log 16)$ after increasing the threshold for T. $\square$

Taking (8.5) from the covering theorem proves the prime-gap corollary. Notice that the proof of Proposition 8.3 used only the Chinese remainder theorem, the elementary primorial estimate in Lemma 8.1, and Bertrand's postulate.

Appendix A. Exact root-conditioned local factors

This appendix derives the one-coordinate conditional probabilities that are useful when checking the rooted second moment.

Fix a coordinate prime $s \in (w, x/2]$, a root v, and another target n. Let $\mathbb{P}_s$ denote probability only over the random residue a_s in the tilted law. By the definition of conditional probability,

$$\mathbb{P}_s(n \text{ survives} \mid v \text{ survives}) = \frac{\mathbb{P}_s(n \text{ and } v \text{ survive})}{\mathbb{P}_s(v \text{ survives})}.$$

If $n \equiv v \pmod{s}$, the two survival events are identical and the answer is 1. Suppose their residues differ, and write $D = s - 1$ and $B_s = 1 - \beta_s/D$.

Neither n nor v is divisible by s. Conditioning on survival of v removes the one possible outcome $a_s \equiv v \pmod{s}$. The conditional probability of the different forbidden outcome $a_s \equiv n \pmod{s}$ is

$$\frac{\beta_s/D}{B_s} = \frac{\beta_s}{D - \beta_s}.$$

Therefore

$$\mathbb{P}_s(n \text{ survives} \mid v \text{ survives}) = 1 - \frac{\beta_s}{s - 1 - \beta_s}.$$

Only n is divisible by s. Here n survives precisely when $a_s \neq 0$. Under the additional condition $a_s \neq v$, the allowed nonzero outcomes have total original probability $\beta_s(D-1)/D$. Hence

$$\mathbb{P}_s(n \text{ survives} \mid v \text{ survives}) = \frac{\beta_s(D-1)/D}{B_s} = s^{-\tau}\frac{s-2}{s-1},$$

where the final equality uses $\beta_s/B_s = s^{-\tau}$.

Only v is divisible by s. Conditioning on survival of v is the same as conditioning on $a_s \neq 0$. The residue a_s is then uniform over the $s-1$ nonzero classes, exactly one of which is forbidden by n. Thus

$$\mathbb{P}_s(n \text{ survives} \mid v \text{ survives}) = \frac{s-2}{s-1}.$$

The case in which both n and v are divisible by s cannot occur when their residues differ.

These are one-target conditional probabilities, not the two-block joint-to-product ratios used in the rooted correlation proposition. For example, when the root is nonzero and both rooted blocks have a companion in the zero residue, that ratio is

$$s^{\tau}\frac{1 - (1 + a + b - c_s)/(s-1)}{[1 - (1+a)/(s-1)][1 - (1+b)/(s-1)]}.$$

The rational factor is

$$\exp\left(O\left(\frac{1 + c_s}{s} + \frac{K^2}{s^2}\right)\right),$$

and it need not be at most one. The important structural point is that the shared root has already been conditioned away: the potentially large factor s^{τ} occurs exactly when the two *companion* sets both contain a multiple of s.

References

1. W. D. Banks, K. Ford, and T. Tao, "Large prime gaps and probabilistic models," *Inventiones Mathematicae* **233** (2023), 1471–1518. doi:10.1007/s00222-023-01199-0.
2. H. Davenport, *Multiplicative Number Theory*, 3rd ed., revised by H. L. Montgomery, Graduate Texts in Mathematics 74, Springer, 2000.

3. K. Ford, B. Green, S. Konyagin, J. Maynard, and T. Tao, “Long gaps between primes,” *Journal of the American Mathematical Society* **31** (2018), 65–105. doi:10.1090/jams/876; arXiv:1412.5029v3.
4. J. Friedlander and H. Iwaniec, *Opera de Cribro*, American Mathematical Society Colloquium Publications 57, 2010.
5. R. A. Rankin, “The difference between consecutive prime numbers,” *Journal of the London Mathematical Society* **13** (1938), 242–247.