hexagon:2611.01204v1

# Joint cofactor signs in prime-factor windows on consecutive integers

Brice Pouly*

**Abstract**

Choose an integer $V < v \le 2V$ with prescribed large prime factors at $v$ and $v-1$. We count these integers and show that all four pairs of cofactor Liouville signs occur equally often, with a power-of-logarithm saving uniform over separate bounded marks on the two prime labels. One site may retain its entire rough part while the other reveals a finite partition of its prime. Revealing both prime labels reconstructs the integer. For the number of prime factors of the short cofactor of $v$, we determine what each label contributes. The opposite label is independent of this count up to every inverse power of $\log V$. The first label produces an explicit discrepancy of order $(\log\log V)^{-1/2}$, whose leading term is already captured by one bit. We also compute the leading error in a common rounded Gaussian approximation and the transition for residues modulo $k$. The argument for the count uses dispersion alone. For the signs, we extend two-point graph estimates to rough-part observations and transfer the classical nine-state Möbius law to these populations. The asymmetric four-prime window counts tuples. For four balanced primes the signed bounds hold at every large scale, and the mass asymptotic holds on a set of scales of logarithmic density one.

## 1 What do the prescribed prime factors reveal?

Write $\Omega(n)$ for the number of prime factors with multiplicity, $\lambda(n) = (-1)^{\Omega(n)}$, and $\mu(n)$ for the Möbius function. Consider the two equations

$$v = wt, \qquad v - 1 = qm, \qquad V < v \le 2V.$$

At the second site $q$ is a prime a little larger than $\sqrt{V}$. At the first, $w$ is a prime close to $V$ on a logarithmic scale, or a prescribed product of two primes. The remaining factors $t$ and $m$ have very different sizes. The problem is to predict their signs from the observed primes.

A small instance of the congruence is $130 = 13 \cdot 10$ and $129 = 43 \cdot 3$. The same first label occurs at $143 = 13 \cdot 11$, with $142 = 71 \cdot 2$. The signs of the first cofactors are opposite. In $100 < v \le 200$, observing both 13 and 43 specifies $v$ uniquely, because two solutions must differ by a multiple of $13 \cdot 43$. The exact populations below guarantee this uniqueness for every pair of labels. The substantive question is what survives when the two labels are observed separately.

All logarithms are natural, and limits are taken as $V \to \infty$. Put $L = \log V$, $T = \log L$. In the following table $p$ belongs to a window if $\log p/L$ is in its indicated interval. All letters in the table denote primes, and a shared boundary belongs to just one interval.

| window | range of exponents | role |
|---|---|---|
| $\mathcal{P}$ | $[1-2/T, 1-1/T]$ | $w = p$ |
| $\mathcal{R}$ | $[1/2-2/T, 1/2-3/(2T))$ | $w = rs$ |
| $\mathcal{S}$ | $[1/2-3/(2T), 1/2-1/T]$ | $w = rs$ |
| $\mathcal{Q}$ | $[0.5053, 0.5057]$ | $q \mid v-1$ |

*Research and writing with AI assistance.

Let $E_1$ be the population with $w = p \in \mathcal{P}$, and $E_2$ the one with $(r, s) \in \mathcal{R} \times \mathcal{S}$. The windows determine $w, q$ uniquely for all sufficiently large $V$. Indeed two primes from $\mathcal{P}$ or from $\mathcal{Q}$ exceed $2V$ in product, as do three factors from $\mathcal{R} \cup \mathcal{S}$, counting multiplicity. We use $d = 1, 2$ to indicate these populations, and put $c_Q = \log(0.5057/0.5053)$. Bounded marks have modulus at most one and may be complex and depend on $V$.

**Theorem 1.1** (Separate observations)**.** *For $d = 1, 2$ one has $|E_d| \sim c_Q V/T^d$. There is an absolute $c > 0$ such that for any bounded functions $A(w), B(q)$ and each $\epsilon, \eta \in \{1, -1\}$,*

$$\sum_{v \in E_d} A(w)B(q) \left(\mathbf{1}_{\lambda(t)=\epsilon,\ \lambda(m)=\eta} - \tfrac{1}{4}\right) \ll VL^{-c}. \tag{1}$$

*The same estimate with constant depending on $b$ holds in each residue class $v \bmod b$ for fixed $b \geq 1$. The mark $A(rs)$ can be an arbitrary matrix of the two first-site primes.*

The error is absolute and is uniform over the marks. With $A = B = 1$, each sign pair has probability $1/4 + O(T^d L^{-c})$ under uniform sampling from $E_d$. With complex marks (1) is an estimate for a signed sum. No lower bound on its weighted mass is implicit.

**Theorem 1.2** (Partial disclosure)**.** *In the uniform population $E_d$ let $\Sigma = (\lambda(t), \lambda(m))$. Observe $O = (w, \pi(q))$, where the map $\pi$ takes at most $K$ values. With $\mathsf{U}_4$ uniform on the four signs,*

$$d_{\mathrm{TV}}(\mathcal{L}(\Sigma, O), \mathsf{U}_4 \otimes \mathcal{L}(O)) \ll \sqrt{K}\, T^d L^{-c}.$$

*The bound also holds for $O = (q, \pi(w))$. For the observation $O = (w, q)$ the distance equals $3/4$ for every sufficiently large $V$.*

Here total variation between probabilities on a countable space is half their $\ell^1$ distance. In particular one entire label, even the pair $(r, s)$, gives an asymptotically fair sign pair in mean over the observed label. Both exact labels give the two signs by a Chinese remainder calculation. The partition bound quantifies an intermediate experiment.

## The analytic mechanism and its antecedents

The proof separates three signed moments. Fouvry and Radziwiłł's dispersion method [1] counts the physical integers and controls the sign on the short side. Their prime-modulus theorem already permits arbitrary short coefficients. Section 2 retains the modulus support in their variance calculation to cover the composite moduli needed here, provided every prime factor of a modulus exceeds the short variable. For the two moments containing the other sign, Section 3 adapts Part I of *Ordinary two-point correlations of multiplicative functions* [3]. An arbitrary observation of the large prime factors is unchanged by the small prime multipliers used in that argument. This local covariance is the substitute for global multiplicativity.

The graph method builds on Helfgott and Radziwiłł's *Expansion, divisibility and parity* [24] and Pilatte's quantitative development [4]. Wang and Yi already obtain a quantitative transfer for the phases $\exp(\mathrm{i}\theta\Omega)$, uniformly in the angles, and choose the endpoint providing the Fourier estimate [18, Proposition 3.3 and Lemmas 3.4–3.5]. The additional freedom here is an arbitrary bounded function of an entire rough part, including the identities and multiplicities of its primes. Our transfer is uniform as these functions and the prime windows vary with the scale. The proof isolates the unit scalar attached to each numerical path before taking the absolute value of its residue expectation, and recovers increasing edges from two Hermitian matrices. The required short-interval Fourier estimate is supplied by the corrected text of Matomäki, Radziwiłł and Tao [2, Theorem 1.7 and (1.12)]. Part II of [3] also allows arbitrary bounded graph tests, but fixes the graph parameters before the limit in the averaging length; the quantitative imports used here are those of Part I.

In *Reconstruction and sign prediction from sampled prime factors* [5], the integer is sampled harmonically and a prime is disclosed by a random rule. Here the prime windows select a population of vanishing density under ordinary counting. An absolute $o(V)$ error can exceed its mass $V/T^d$. The observation in *Parity information in a rough cofactor and exact reconstruction thresholds* [30] is instead the rough cofactor together with a residue of the origin. In the present experiment, either prime label may be kept in full, and knowing both determines the origin.

The nine Möbius constants are classical in logarithmic averages [19, Corollary 7], with squarefree support given by Mirsky [7]. Section 5 transfers that table to the cofactor populations. Section 6 computes the dependence of the short cofactor's prime count on either exact label. For $w$ the length of the remaining interval accounts for the entire first-order dependence. For $q'$ the dispersion main term factors.

### Guide to the proofs

The following dependencies distinguish the quantities available by dispersion alone from those requiring the two-point estimate.

| Quantity | Main input | Location |
|---|---|---|
| Masses and the short-side sign | Rough-modulus dispersion and a one-variable mean value | Section 2 |
| Two cofactor signs with observations | Marked graph transfer, followed by the three-moment expansion | Sections 3–4 |
| Short-cofactor count with one exact label | Dispersion and a quantitative one-variable approximation | Section 6 |
| Balanced four-prime mass | Quantitative correlations on good scales | Section 7 |

Section 5 uses tuple counting for the asymmetric four-prime population. Section 7 uses integer counting and obtains the mass asymptotic on a set of scales of logarithmic density one. The positive sieve applications are collected in Appendix A, and Appendix B separates a further proved observation from the open question with two arbitrary rough parts. Two different mechanisms produce the scale $\sqrt{T}$: periodic Gaussian concentration in Section 6, and the accumulation of prime-band density corrections in Proposition 4.2.

## 2 Counting the integers and their short cofactors

We will use a third population in the proof. Keep $(r, s) \in \mathcal{R} \times \mathcal{S}$ at the first site and write $v - 1 = abu$ at the second. Set $\gamma = 0.25275$ and take

$$\mathcal{A} = [\gamma(1 - 2/T), \gamma(1 - 1/T)], \qquad \mathcal{B} = [\gamma(1 + 1/T), \gamma(1 + 2/T)]$$

in exponent notation. Let $E_3$ consist of the *tuples* $(r, s, t, a, b, u)$ so obtained. For large $V$ the three large second-site primes that might occur are all $> V^{1/4}$ when they belong to these bands. Thus an integer gives at most two cross-band pairs $(a, b)$, including the case of repeated primes, and each product $ab$ specifies its pair. In all three families put $w = p$ or $rs$, and let $q' = q$ or $ab$ be the second-site product. Let $m' = (v - 1)/q'$.

**Proposition 2.1** (Physical dispersion)**.** *For any $J > 0$ and bounded $A(w), B(q'), \beta(t)$, in any of the three populations,*

$$\sum_{E_d} A(w)B(q')\beta(t) = \sum_{q'} \frac{B(q')}{\varphi(q')} \sum_{w} A(w) \sum_{V/w < t \le 2V/w} \beta(t) + O_J(VL^{-J}). \tag{2}$$

*Furthermore, if one prescribes $v \equiv x \bmod b_0$ with $b_0 \le L^H$, the same assertion holds with the restriction $wt \equiv x \bmod b_0$ in the last sum. The implied constant may depend on the fixed $H$. Both estimates are on tuples for $d = 3$.*

The first step adapts the published dispersion estimate to moduli whose least prime exceeds the short variable. Fouvry and Radziwiłł's prime-modulus theorem already allows arbitrary short coefficients [1, Theorem 1.2]. We retain the same freedom for the rough composite moduli used in $E_3$. The support of the moduli, rather than a Siegel–Walfisz condition on the short sequence, controls the principal variance term.

**Lemma 2.2** (Dispersion for rough moduli)**.** *Fix $0 < a < 1$, $B > 0$, and an integer $a_0 \neq 0$. Put $X = MN$ and assume*

$$\exp((\log X)^a) \le N \le \exp\Big(\frac{B \log X}{\sqrt{\log\log X}}\Big), \qquad X^{0.50} \le Q \le X^{0.506}.$$

*Write $m \sim M$ for $M < m \le 2M$, and similarly for $n \sim N$ and $q \sim Q$. For coefficients $|\alpha_m|, |\beta_n| \le 1$ and any set $\mathcal{D}$ of moduli $q \sim Q$ satisfying $(q, a_0) = 1$ and $P^-(q) > 2N$, one has*

$$\sum_{q \in \mathcal{D}} \Bigg| \sum_{\substack{m \sim M,\ n \sim N \\ mn \equiv a_0 \bmod q}} \alpha_m \beta_n - \frac{1}{\varphi(q)} \sum_{\substack{m \sim M,\ n \sim N \\ (mn,q)=1}} \alpha_m \beta_n \Bigg| \ll_{a,B,a_0} X N^{-c_0}$$

*for an absolute $c_0 > 0$. The estimate is uniform over all the displayed lengths, coefficients and modulus sets once $X$ exceeds a threshold depending only on $a, B, a_0$. Here $P^-(q)$ is the least prime factor of $q$.*

*Proof.* Put $L_X = \log(2X)$. Use the dispersion identity and estimates of Fouvry and Radziwiłł [1, proof of Theorems 1.1 and 1.2, especially (49)–(52) in the linked full text]. Choose unit phases to attain the absolute discrepancy and set the modulus coefficient $c_q = 0$ outside $\mathcal{D}$. Their Cauchy–Schwarz step bounds the result by $\|\alpha\|_2$ times the square root of a variance. The only main term requiring information on the short sequence has the form, up to an absolute bounded factor,

$$M \sum_{\delta} \frac{1}{\delta} \sum_{\substack{k_1, k_2 \asymp Q/\delta \\ (k_1,k_2)=1}} \frac{c_{\delta k_1} \overline{c_{\delta k_2}}}{k_1 k_2} \sideset{}{^*}\sum_{e \bmod \delta} E^*(e; k_1) \overline{E^*(e; k_2)}.$$

Here $E^*(e; k)$ is the discrepancy of $\beta_n \mathbf{1}_{(n,k)=1}$ in the unit class $e$ mod $\delta$ from its average on the unit classes. At $\delta = 1$ it vanishes. In every remaining term with nonzero coefficient $\delta > 2N$, since $\delta$ divides a retained modulus. Each class contains at most one $n \sim N$. Subtracting the mean is an orthogonal projection, so $\sum_e^* |E^*(e; k)|^2 \le \sum_{n \sim N} |\beta_n|^2 \ll N$. Cauchy–Schwarz in $e$ bounds each inner product by $O(N)$. Since $\sum_{k \sim Q/\delta} 1/k \ll 1$ for $\delta \le 2Q$, this entire contribution is $O(MN \log(2Q))$.

All remaining terms in the proof of their Theorem 1.1 use only $|c_q| \le 1$. With $D = N^{1/20}$ their combined variance bound is

$$\ll L_X^{C_1}(N^2 Q + D^{-1/2} M N^2) + D^{C_2} X^{\varepsilon}(M^{3/20} N^{59/20} Q^{33/20} + N^{23/8} Q^{15/8}).$$

Indeed $M > Q X^{\varepsilon}$ and $M > N > D^{10}$ for small fixed $\varepsilon > 0$. After division by $MN^2$, the two final monomials, without the common factor $D^{C_2} X^{\varepsilon}$, are

$$X^{-17/20} N^{9/5} Q^{33/20}, \qquad X^{-1} N^{15/8} Q^{15/8}.$$

Their fixed exponents of $X$ at $Q = X^{0.506}$ are $-0.0151$ and $-0.05125$. The upper bound on $N$ makes every remaining fixed power of $N$ equal to $X^{o(1)}$, uniformly in this lemma. The lower

bound gives $L_X^{C_1} = N^{o(1)}$, again uniformly. Choose a fixed sufficiently small $\varepsilon > 0$, followed by the threshold for $X$. The normalised variance is then $O(N^{-1/40+o(1)} + X^{-c_2})$ for an absolute $c_2 > 0$, including the $O(\log(2Q)/N)$ main-term contribution. Since $\|\alpha\|_2 \ll \sqrt{M}$, taking the square root proves the claim for a sufficiently small absolute $c_0 > 0$. The modulus support has been retained in the main term before its phases are discarded. □

*Proof of Proposition 2.1.* Separate $w, t, q'$ into dyadic intervals. In either first-site window, $\log(V/w)$ lies between $(1 + o(1))L/T$ and $(7/2 + o(1))L/T$. The second product satisfies the modulus bounds of Lemma 2.2, with room at the endpoints. Every prime entering $q'$ exceeds $2N$. The ranges of the first-site primes and the second-site primes are disjoint for large $V$. Hence $(wt, q') = 1$ in the density term on all these rectangles.

To impose $V < wt \le 2V$, subdivide the $t$-range with successive ratio $1 + L^{-K}$ and freeze the two $w$-endpoints in each interval. Every actual boundary error comes from an origin within $O(VL^{-K} + 1)$ of $V$ or $2V$, and each has at most two representations also in the enlarged range. In the density term the matching tails cost

$$\ll \sum_{q'} \frac{1}{\varphi(q')} \left( VL^{-K} \sum_{w} \frac{1}{w} + \#\{w\} \right).$$

The first sum is bounded and $w \ll V^{1-1/T}$. The polynomially many rectangles and dyads cost only powers of $L$ in Lemma 2.2. Choose $K$ after $J$. For the progression variant, all $w$ are units modulo $b_0$ since their prime factors exceed $L^H$. Split $w$ by its residue and restrict $t$ to $x\bar{w} \bmod b_0$. This costs at most $L^H$ applications of the lemma, with the boundary error estimated only once on the restricted population. □

The prime number theorem and partial summation give $\sum_{\mathcal{P}} 1/p \sim 1/T$, $\sum_{\mathcal{R}} 1/r \sim \sum_{\mathcal{S}} 1/s \sim 1/T$, and $\sum_{\mathcal{Q}} 1/(q-1) \sim c_Q$. At each of $\mathcal{A}, \mathcal{B}$ the reciprocal prime sum is again $(1 + o(1))/T$. Errors in replacing floors by interval lengths are $O(\#\{w\})$ before the bounded second-site sum. Thus

$$|E_1| \sim c_Q V/T, \qquad |E_2| \sim c_Q V/T^2, \qquad |E_3| \sim V/T^4.$$

In $E_3$ the sum with the four genuine prime weights $\log r \log s \log a \log b$ is $(\gamma^2/4 + o(1))VL^4/T^4$. Their ratios to $L$ converge uniformly to $1/2, 1/2, \gamma, \gamma$. In all three families the moment with $\beta(t) = \lambda(t)$ in (2) is $O_J(VL^{-J})$ for arbitrary $A, B$. This follows from the prime number theorem for Liouville on each dyad of $t$, since $\log(V/w) \gg L/T$. In a fixed residue class first extract its fixed gcd, then apply the prime number theorem for the twists of Liouville in the reduced classes. Thus the mass and the first sign already have proofs independent of the binary correlation argument.

With a logarithm on each selected prime, the masses in $E_1, E_2$ are respectively

$$(\Delta_Q + o(1))VL^2/T, \qquad (\Delta_Q/4 + o(1))VL^3/T^2, \qquad \Delta_Q = 0.5057 - 0.5053.$$

Here $\sum_{q \in \mathcal{Q}} \log q/((q-1)L) \to \Delta_Q$ by partial summation. The distinction between $\Delta_Q$ and $c_Q$ is needed because the second window has fixed exponent width. Normalizing each logarithm by $L + 1$ absorbs it into the separate marks. These insertion rules also apply if the anchor slots range over all integers with their von Mangoldt weights $\Lambda$. In $E_1, E_2, E_3$ and the balanced population of Section 7, every slot exceeds $V^{1/5}$ eventually. The count of origins having a divisor $p^j \ge V^{1/5}$, $j \ge 2$, at either site is $O(V^{9/10}L + \sqrt{V}L)$. Indeed on each dyad $[H, 2H]$ there are $O(\sqrt{H}L)$ proper prime powers, and each gives $O(V/H + 1)$ multiples. The number of weighted tuples over such an origin is $O(L^8)$ since each of the four slots has $O(L)$ prime-power divisors and weight $O(L)$. Thus the difference costs $O_J(VL^{-J})$ even for the unnormalised logarithms and inserted bounded tests.

A related experiment starts from the divisor weights in *Critical divisor weights and the localisation of parity defects* [31]. In *Visible prime factors under neighbouring critical divisor weights* [32], reweighting at $n+2$ changes the observed prime and the origin by different amounts in total variation. Here the windows at $v$ and $v-1$ prescribe a single sampling law. Formula (2) compares observations within that law.

# 3 A marked two-point estimate at a fixed shift

The mass estimate allows a mark to be any bounded function of the long variable. Once the sign of $m$ is inserted, complete multiplicativity moves that sign onto $v-1$. We need cancellation there even when a function of the exact large factors of $v$ is retained. We prove the required extension at the level of the graph argument, since a qualitative estimate for each fixed multiplicative function would give insufficient uniformity after normalisation.

## 3.1 A transferable graph estimate

The uniform phase argument of Wang and Yi [18, Proposition 3.3] treats phases determined by $\Omega$. We formulate the small-prime covariance and the one-endpoint Fourier input separately, so that the other endpoint may retain a nonmultiplicative observation of its large prime factors.

**Lemma 3.1** (Numerical phases in a path expansion)**.** *Fix an integer $h \geq 1$. Suppose that a family of real symmetric matrices is built from numerical steps $n \mapsto n \pm hqd$, where the disjoint prime supplies determine $d, q$ uniquely from $qd$. Its entries may contain residue tests, vertex weights, and site deletions. Assign a unit scalar $\zeta(d,q)$ to each increasing step, independently of its initial site and of the residue coordinates, and assign $\overline{\zeta(d,q)}$ to the reverse step.*

*For an integer $k \geq 1$ and orientations $\varepsilon_a \in \{-1,1\}$, fix a numerical closed word $\omega = ((\varepsilon_a, d_a, q_a))_{1\leq a\leq 2k}$ arising by joining the two paths in a Hilbert–Schmidt square, let $W_\omega$ and $W_\omega^\zeta$ be its unmarked and marked contributions, including all vertex restrictions. Then pointwise, and hence under either the integer-origin law or the independent-residue law,*

$$W_\omega^\zeta = \zeta_\omega W_\omega, \qquad \zeta_\omega = \prod_{a=1}^{2k} \zeta(d_a, q_a)^{\varepsilon_a}, \qquad |\zeta_\omega| = 1. \tag{3}$$

*Consequently a finite-law comparison based on the total absolute coefficient mass has the same error budget after marking. A moment bound obtained by summing absolute values of individual product-law word expectations is also unchanged.*

*Proof.* Hermitian reversal of the second path in the Hilbert–Schmidt square changes every multiplier to its complex conjugate. For a fixed numerical word all orientations and pairs $(d_a, q_a)$ have already been specified. Their product is therefore the scalar in (3); it is independent of the origin, all residue draws, and all decisions to retain or delete a site. This proves the identity, even when one of the vertex restrictions makes the word contribution zero.

Write $\mathbb{E}_I$ for an integer-origin average and $\mathbb{E}_*$ for the independent-residue expectation. The identity gives

$$\left|(\mathbb{E}_I - \mathbb{E}_*)W_\omega^\zeta\right| = \left|(\mathbb{E}_I - \mathbb{E}_*)W_\omega\right|, \qquad \left|\mathbb{E}_* W_\omega^\zeta\right| = \left|\mathbb{E}_* W_\omega\right|.$$

More explicitly, in any expansion $W_\omega = \sum_\alpha c_{\omega,\alpha}\mathbf{1}_{\mathcal{E}_{\omega,\alpha}}$, marking replaces $c_{\omega,\alpha}$ by $\zeta_\omega c_{\omega,\alpha}$. The events and the sum of absolute coefficient values are unchanged. Apply the finite-law comparison to these signed expansions first, and only then sum the absolute values of the product-law word expectations. This order preserves every cancellation among residue choices belonging to the same word and proves both assertions. □

We next state the two analytic hypotheses. Write $e(t) = e^{2\pi i t}$. Let $x$ be a prefix length, $A_*$ a sufficiently large absolute constant, and $U \to \infty$ with $U^{A_*} \le \log x$. Fix integers $h, l \ge 1$ and a residue $b \bmod l$. Suppose $|f|, |g| \le 1$ on the positive integers. For completely multiplicative unit phases $\kappa_f, \kappa_g$ assume

$$f(kn) = \kappa_f(k)f(n), \qquad g(kn) = \kappa_g(k)g(n) \quad \text{whenever every prime of } k \text{ is } \le e^{102U}. \tag{G1}$$

Only the divisors actually used by the graph are needed. Set $\gamma = \kappa_f \kappa_g$.

For each integer $D$ with $\frac{1}{2}\exp(U^{0.995}) \le D \le \exp(2U)$ and each $Y$ with $|\log(Y/x)| \le 4U$, assume that at least one of the following holds, uniformly in real $\theta$:

$$\frac{1}{YH}\sum_{v \le Y}\left|\sum_{1 \le m \le H} F(v+m)e(\theta m)\right| \ll_h U^{-0.8}, \qquad (F,H) = (f,D) \text{ or } (g,(2h+1)D). \tag{G2}$$

The endpoint satisfying this estimate can depend on $D, Y$ and on the functions. All constants in the conclusion are uniform for a common bound in (G2).

**Lemma 3.2** (Marked graph transfer)**.** *Under (G1)–(G2) there is an absolute $\delta_* > 0$ such that*

$$\left|\frac{1}{x}\sum_{\substack{n \le x \\ n \equiv b \bmod l}} f(n)g(n+h)\right| \ll_{h,l} U^{-\delta_*}.$$

*Proof.* We use the prime supplies, finite-residue comparison, deletion estimates and nonbacktracking trace bound proved in Part I of *Ordinary two-point correlations of multiplicative functions* [3].

The imported estimates in Part I of [3] can be located as follows.

| Source file | Results | Use here |
|---|---|---|
| `02-finite-law.tex` | Lemma 3.1; Corollaries 3.2–3.3 | Finite-origin and weight comparisons |
| `03-analytic.tex` | Lemmas 4.2 and 4.6 | Positive Fourier bounds and deletion cost |
| `04-paths-trace.tex` | Theorem 5.4, §5.2 | Signed word comparison before absolute values |
| `06-forest.tex` | Completion of Theorem 5.4, §7.4 | Positive majorants for the word expectations |
| `07-spectral.tex` | Lemmas 8.1–8.2; Propositions 8.3–8.4 | Hermitian transfer and testing on blocks |

The centered sum, its Fourier estimate and its spectral test are the three steps affected by the phase. In the notation of that proof take graph size parameter $U$, its absolute constant $W$ large and fixed, $J = \lfloor \log U/(1200W) \rfloor$, $\eta = e^{-J}$, $K = e^{4J}$, and $k = \lfloor U \rfloor$. Each centered divisor $d$ uses one prime from each of $J$ disjoint supplies. Each padding divisor $q$ is squarefree and uses a separate supply. All these primes are coprime to $l$. The weights are $u(q) = 4^{\omega(q)}$, with $\omega$ counting distinct primes. In a bin $j$, $j\eta \le \log(dq) < (j+1)\eta$, the prefix is $T_j = xe^{j\eta}$ and $dq < e^{101U}$. The harmonic mass $S_0 = \sum_{j,d,q} u(q)/(dq)$ is between $SV_*/2$ and $SV_*$, with $S$ the unrestricted padding mass and $V_*$ the product of the $J$ centered reciprocal masses. Each such mass is at least $W$.

*Centering.* In the centered sum of Part I, §2, multiply an increasing step $n \mapsto n + hqd$ by $\overline{\gamma(qd)}$ and use $f(n)g(n+hqd)$ at its endpoints. Its other factors include $\mathbf{1}_{q|n}\mathbf{1}_{n \equiv bqd \bmod l}\prod_{p|d}(\mathbf{1}_{p|n} - 1/p)$. In the term with every divisibility retained, put $n = qdm$. The phase cancels (G1), even if $m$ shares primes with $qd$. Boundedness permits replacing $T_j/(qd)$ by $x$ in the normalised prefix at cost $O(\eta + x^{-1})$. The full-divisibility terms thus recover the desired correlation with positive mass $S_0$.

In a partial term put $d = d_0 z$ where $z$ uses the nonempty set of omitted supplies. Division by $qd_0$ leaves $\overline{\gamma(z)}$, in addition to the constant sign from the omitted factors. With $Y = T_j/(qd_0)$ the summand becomes $f(v)g(v+hz)$ on $v \equiv bz \bmod l$ with weight $\overline{\gamma(z)}/z$. The value of $\log(Y/x)$ lies within $4U$. Split $z$ by its class modulo $l$ and into the at most two dyadic intervals of its bin. On each resulting set $\mathcal{Z} \subset [D, 2D)$ the polynomial $B(\theta) = \sum_{z\in\mathcal{Z}} \overline{\gamma(z)} e(hz\theta)/z$ satisfies

$$\|B\|_\infty \ll U^{-0.99}, \qquad \int_0^1 |B(\theta)|^4 \,\mathrm{d}\theta \ll D^{-1}U^{-3.96}.$$

Indeed Part I, Lemma 4.2 bounds the reciprocal count and the positive weighted number of solutions $z_1 + z_2 = z_3 + z_4$ on these rough sets. Triangle inequality after the fourth-moment expansion proves the same bounds with any deterministic coefficients of modulus at most one.

For completeness, average over initial translations $1 \le m \le D$. Orthogonality leaves the integral of $BF_vG_v$ with normalisation $1/(YD)$, where $F_v$ is the length-$D$ sum of $f(v+m)e(\theta m)$ in a fixed residue class and $G_v$ is the length-$(2h+1)D$ sum of $g(v+a)e(-\theta a)$. The endpoint error is $O_h(D/Y)$. A fixed class costs no factor in (G2), by its additive character expansion of coefficient norm one. Where $|B| \le U^{-1.05}$, the two Parseval bounds give $O_h(U^{-1.05})$. The complement has measure $O(D^{-1}U^{0.24})$. Use (G2) there for the designated polynomial and the trivial length bound for the other. Its contribution is $O_h(U^{0.24-0.99-0.8})$ in either orientation. Summing over the $O(U/\eta)$ bins and at most $2^J$ omissions gives the normalised error $O_{h,l}(2^J\eta^{-1}U^{-1/20})$ after division by $SV_*$. This is the centering saving needed in Part I, §4. No estimate for a free endpoint's own Fourier sum has entered.

*The spectral estimate with phases.* Use the same deleted sites, padding cutoffs, blocks and degree projection $E$ as in Part I, §§4–8. Deletion was bounded by the sum of absolute weights, so its cost still applies. In a bin write $A_d$ for the original symmetric matrix of steps with centered divisor $d$. For a unit scalar $\xi$ form $A_d^\xi$ by multiplying the increasing entry with increment $hqd$ by $\xi\overline{\gamma(qd)}$, and giving the opposite entry its conjugate multiplier. The matrix is Hermitian. We claim the compressed norm bound of Part I, Proposition 8.3 (the proposition just before testing on blocks), with the same exceptional fraction and constants for each fixed $\xi$.

The weighted absolute row bounds and the invariant level sets of $\prod_{p|d} |\mathbf{1}_{p|n} - 1/p|$ are unaffected by the phases. The proof of [3, Part I, Lemma 8.2] therefore gives

$$\|A_d^\xi\| \le 2K, \qquad \sum_d \|A_d^\xi v\|^2 \le 4K^2(8W)^J\|v\|^2 \quad (v = Ev).$$

Form $H_{d,d'}^\xi = \mathbf{1}_{d\ne d'} A_{d'}^\xi$. Apply Lemma 3.1 to its path expansion with $\zeta(d,q) = \xi\overline{\gamma(qd)}$. In the source proof, scalar finite-law comparison is Corollary 3.2; its application to the signed word expansion precedes absolute values in the proof of Theorem 5.4, §5.2. The positive bounds are completed in §7.4. Thus the proof of [3, Part I, Theorem 5.4] gives

$$\mathbb{E}_I \|(H^\xi)^k\|_{\mathrm{HS}}^2 \le [K(C_2\sqrt{W})^J]^{2k}$$

with the same absolute constant, after its original harmless enlargement. Lemma 8.1 of that source applies to self-adjoint operators on a complex Hilbert space and requires no commutation between distinct edge operators. It gives Proposition 8.3 for the Hermitian family $A_d^\xi$, with exceptional fraction at most $e^{-2k}$. All these bounds are uniform over the numerical phases.

Let $C$ denote the operator containing only increasing entries with multiplier $\overline{\gamma(qd)}$, summed over $d$ and compressed by $E$. On the block,

$$C = \tfrac{1}{2}\left(E\sum_d A_d^1 E - \mathrm{i}E\sum_d A_d^{\mathrm{i}} E\right).$$

The two Hermitian estimates therefore bound $\|C\|$ by $K(C_3\sqrt{W})^J$ outside a fraction at most $2e^{-2k}$ of block origins. Test $C$ on two vectors with entries $\sqrt{w(n)}\,\overline{f(n)}$ and $\sqrt{w(n)}\,g(n)$,

inserting the original padding-degree cutoff. Here $w(n)=\sum_{q\mid n}u(q)$ on the padding supply. Each squared norm has the averaged upper bound $2MS$ on blocks of length $M=\lceil e^{103U}\rceil$, by the truncated weight comparison. Cauchy–Schwarz gives the same bound for the mean product of their norms. On the exceptional blocks the original absolute-row estimate applies to both vectors. Counting the occurrences of each increasing edge in overlapping blocks gives the same endpoint error as in Part I, §8.3. This yields the original bound $O(SK(C_3\sqrt{W})^J/U+e^{-U})$ per bin.

The finite-origin comparisons require prefix lengths at least $\exp(U^{A_*}/2)$, which follows from the chosen range of $U$. Their degree and coefficient bounds use $U$ alone. The final estimate, after division by $S_0$, is bounded by a constant times

$$\eta+2^J\eta^{-1}U^{-1/20}+2^J(K^{-1}+U^{-100}+e^{-2WJ})+K\eta^{-1}(C_3/\sqrt{W})^J+e^{-U^{0.9}}+x^{-1}.$$

Choose $W$ so that $e^5C_3/\sqrt{W}\le e^{-1}$. Every term is $O(e^{-J})$. This proves the lemma with an absolute positive $\delta_*$. □

### 3.2 One arbitrary observation of large factors

For $y>1$ write $\mathcal{R}_y(n)=\prod_{p^j\parallel n,\ p>y}p^j$. The primes in the next theorem can have arbitrary multiplicities.

**Theorem 3.3.** *Fix $h,l\ge1$, let $y=\exp((\log X)^{0.9})$, and let $s$ be completely multiplicative, $|s|\le1$, with $s(p)=1$ for $p\le y$. Let $|H|\le1$ be any function of a positive integer, and $i=0$ or $1$. Uniformly in these data and $b \bmod l$,*

$$\left|\sum_{\substack{X<n\le2X\\ n\equiv b \bmod l}}\lambda(n)s(n)\lambda(n+h)^iH(\mathcal{R}_y(n+h))\right|\ll_{h,l}X(\log X)^{-c}$$

*for an absolute $c>0$. The estimate also holds after interchanging the two endpoint functions.*

*Proof.* Take prefixes $x=X,2X$ with the same marks and $U=(\log x)^{1/A}$ for a fixed sufficiently large $A\ge\max(5000,A_*)$. For large $X$, $e^{102U}<y$. At the first endpoint the small-prime multiplier is $\lambda$, and at the second it is $\lambda^i$ since $\mathcal{R}_y(kn)=\mathcal{R}_y(n)$ for such $k$. For (G2) use the short-interval estimate of Matomäki, Radziwiłł and Tao [2, Theorem 1.7 and (1.12), corrected text] at the $\lambda s$ endpoint. Write

$$\mathbb{D}_Z(f,g)^2=\sum_{p\le Z}\frac{1-\Re(f(p)\overline{g(p)})}{p},\qquad M(f;Z,Q)=\inf_{\substack{q\le Q,\,\psi \bmod q\\ |\tau|\le Z}}\mathbb{D}_Z(f,\psi(n)n^{\mathrm{i}\tau})^2.$$

For $Z=\lfloor Y\rfloor+1$ and $Q\le(\log Z)^{1/125}$, (1.12) gives $M(\lambda;Z,Q)\ge(1/3-o(1))\log\log Z$. The loss in replacing $\lambda$ by $\lambda s$ is at most $2\sum_{y<p\le Z}1/p=(1/5+o(1))\log\log Z$. Thus $M(\lambda s;Z,Q)\ge\frac{1}{10}\log\log Z$. Theorem 1.7 bounds the normalised mean in (G2) by a constant times

$$e^{-M(\lambda s;Z,Q_H)/20}+\frac{\log\log H_0}{\log H_0}+(\log Z)^{-1/700},\qquad Q_H=\min((\log Z)^{1/125},(\log H_0)^5),$$

where $H_0$ is the short interval length $D$ or $(2h+1)D$. Each term is $O_h(U^{-0.8})$. The integral over real starting points gives the integer estimate by integrating over $(v,v+1)$ at integer length. Lemma 3.2 proves the assertion on each prefix, hence on the dyad. □

# 4 Which observations can be retained?

*Proof of Theorem 1.1.* Take $n = v-1$. At the second site let $s_B, s_0$ be completely multiplicative, equal to one on primes off $\mathcal{Q}$, and equal respectively to $B(q), 0$ at $q \in \mathcal{Q}$. On the dyad the difference $s_B(n) - s_0(n)$ is zero unless $q \mid n$ for a unique $q \in \mathcal{Q}$, and then equals $B(q)$. Indeed two such prime factors, counted with multiplicity, have product $> 2V$. At the first site all primes in the windows exceed $y = \exp(L^{0.9})$. Thus a single bounded function of $\mathcal{R}_y(v)$ tests the required first-site pattern and returns $A(w)$ on it. This includes arbitrary joint marks of $r, s$.

Now $\lambda(t) = (-1)^d \lambda(v)$ and $\lambda(m) = -\lambda(v-1)$. Apply Theorem 3.3 to the two masks separately. It bounds both moments containing $\lambda(m)$ by $O_b(VL^{-c})$. Shifting the dyad by one changes at most two terms and $v \equiv x \bmod b$ becomes $n \equiv x-1 \bmod b$. Proposition 2.1 and the subsequent prime number estimate give the short-side moment. Expand $(1+\epsilon\lambda(t))(1+\eta\lambda(m))/4$ to obtain (1). □

*Proof of Theorem 1.2.* Fix one sign pair. Let $e_{w,j}$ be its count on $(w, \pi(q)) = (w, j)$ minus one quarter of the total count in that cell. For any choices $|a_w|, |b_j| \le 1$, (1) gives $|\sum_{w,j} a_w b_j e_{w,j}| \ll VL^{-c}$. Let $\varepsilon_j$ be independent uniform signs. For $Z = \sum b'_j \varepsilon_j$ with fixed real coefficients, $\mathbb{E}Z^2 = \sum |b'_j|^2$ and $\mathbb{E}Z^4 \le 3(\mathbb{E}Z^2)^2$. Hölder gives $\mathbb{E}|Z| \ge (\sum |b'_j|^2/3)^{1/2}$. Choose $a_w$ as the sign of the corresponding row sum after the $\varepsilon_j$ have been drawn. It follows that

$$\sum_{w,j} |e_{w,j}| \le \sqrt{K} \sum_{w} \Big( \sum_{j} |e_{w,j}|^2 \Big)^{1/2} \le \sqrt{3K}\, \mathbb{E} \sum_{w} \left| \sum_{j} e_{w,j} \varepsilon_j \right| \ll \sqrt{K}\, VL^{-c}.$$

Sum over the four pairs and divide by $2|E_d|$. Transposing the rows and columns proves the reversed assertion. Finally $v \equiv 0 \bmod w$ and $v \equiv 1 \bmod q$ specify a single class modulo $wq > V$. In an interval of length $V$ each nonempty cell therefore contains a single origin. The distance between its point mass on signs and $\mathsf{U}_4$ is $3/4$. Averaging this equality over the actual label law proves the last assertion. □

The preceding theorem also handles partitions chosen after fixing $V$ and before sampling $v$. For example it allows $K = L^\theta$ for any fixed $\theta < 2c$. The signed progression estimates concern fixed moduli. A diagonal argument gives some unspecified $b_*(V) \to \infty$ for which the four asymptotics are relative to the class mass uniformly at $b \le b_*(V)$. Indeed take $b_*$ to grow slowly enough that $\max_{j \le b_*} jC_j T^d L^{-c} \to 0$ for the fixed-modulus constants $C_j$, including their thresholds. The explicit polylogarithmic modulus range of Proposition 2.1 concerns the dispersion estimate.

## 4.1 Observing every prime factor above a cutoff

An exact prime label is a small part of the information in one integer. We next allow the entire list of its primes above a threshold, with their multiplicities. This also covers observations combining the primes in the prescribed bands with the other large primes of the short cofactor.

**Theorem 4.1.** *Fix $0 < \alpha < 1$ and $z \ge \exp(L^\alpha)$. In $E_d$, $d = 1, 2$, let $O_z = (w, \mathcal{R}_z(v))$. Uniformly over all bounded $H(O_z), B(q)$, equation (1) holds with $A(w)$ replaced by $H(O_z)$ and $c$ replaced by some $c_\alpha > 0$. Consequently observing $(O_z, \pi(q))$ with $|\operatorname{im} \pi| \le K$ gives total variation discrepancy $\ll_\alpha \sqrt{K} T^d L^{-c_\alpha}$ from the uniform sign law times the observation law.*

*Proof.* It suffices to use $z_* = \exp(L^\alpha)$. The rough part at $z_*$ contains all anchor primes for large $V$ and determines $O_z$ for every larger cutoff. Write $z = z_*$. For moments containing $\lambda(m)$ apply Lemma 3.2 to the opposite prime mask and the function $\lambda(v)^i H(O_z)$ at the first site, including the anchor existence test there. Choose $A > 2/\alpha$ sufficiently large, so every graph prime is $< z$. The controlled endpoint $\lambda s_B$ is altered only in $\mathcal{Q}$, of bounded reciprocal

prime mass, and hence satisfies (G2) just as in Theorem 3.3. This proves the two required signed estimates.

For the remaining moment write $t = uh$ where $P^-(u) > z$ and $P^+(h) \le z$, with the value one allowed for either empty product. Put $D = \exp(L^{\alpha/2})$ and discard $h < D$ or $h > z^T$. We bound these losses even if the condition on $q$ is removed. In $E_1$ the number of first-site anchors for each $u, h$ is $\ll V/(uhL)$ by counting primes in the required dyad. In $E_2$ fix $r$ before counting $s$ and sum $1/r \ll 1/T$ over its band. The resulting bound in both cases is $\ll V/(uhLT^{d-1})$. All feasible remaining prime intervals have logarithm comparable with $L$. Since $t \le 2V^{4/T}$, Mertens' theorem gives

$$\sum_{\substack{u \le 2V^{4/T} \\ P^-(u) > z}} 1/u \le \prod_{z < p \le 2V^{4/T}} (1 - 1/p)^{-1} \ll L/(T \log z).$$

The $h < D$ contribution is thus $\ll (V/T^d)L^{-\alpha/2}$. For the upper tail, keep $w$ fixed and take $\sigma = 1/\log z$. The function $t \mapsto (t_{\le z})^\sigma$, with $t_{\le z}$ its smooth part, is the convolution of 1 with a nonnegative function. Therefore its average up to any $x \ge 1$ is at most

$$\prod_{p \le z} \frac{1 - 1/p}{1 - p^{\sigma - 1}} \ll 1.$$

The last bound follows by taking logarithms and using $p^\sigma - 1 \ll \sigma \log p$. Markov's inequality bounds the contribution with $h > z^T$ by $\ll V e^{-T} \sum_w 1/w \ll (V/T^d)e^{-T}$.

On the remaining range use $W = wu = \mathcal{R}_z(v)$ as the long dispersion variable and $h$ as the short one. The coefficient at $W$ is $H(w, W)\lambda(u)$ on the allowed support. Since $u \le 2V^{4/T}$, $W$ determines the distinguished prime or pair uniquely. Its prime factors all exceed $z$. Define the completely multiplicative function

$$g_z(h) = \lambda(h)\mathbf{1}_{P^+(h) \le z} \qquad (h \ge 1).$$

The short coefficient in dispersion is $\beta(h) = g_z(h)\mathbf{1}_{D \le h \le z^T}$; it need not be multiplicative. Each prime modulus $q$ exceeds $u, h$, lies outside the first-site bands, and is coprime to $Wh$. Apply Lemma 2.2 with $a = \alpha/3$ and, for example, $B = 5$. For all sufficiently large $V$ its range contains

$$D/2 \le N \le 2V^{4/T}.$$

since $X \asymp V$ on the rectangles. The constants and the threshold may depend on $\alpha$. Thus the dispersion error remains smaller than every fixed negative power of $L$ after the polynomially many subdivisions. The same mesh proof of Proposition 2.1 applies because every origin on the boundary has unique smooth and rough parts and the density floor error costs $O(V/D)$ per total bounded reciprocal modulus mass. One can use a mesh on $W$, with the cutoffs on $h$ left on the short side. The discrepancy for this replacement is $O_{J,\alpha}(VL^{-J})$.

In the resulting main term keep the sign of the smooth factor. Uniformly for $D/2 \le x \le 5V^{4/T}$, Halász's mean-value theorem [9] gives

$$\left| \sum_{h \le x} g_z(h) \right| \ll_\alpha x L^{-c'\alpha}$$

for some absolute $c' > 0$. In fact, at primes up to $Y = \min(x, z)$ the values of $g_z$ agree with $\lambda$. For frequencies $|\tau| \le L^2 < Y$, (1.12) of [2] gives squared distance at least $(1/3 - o(1)) \log\log Y \ge \alpha T/7$. The remaining summands in the distance are nonnegative. Halász's bound is $\ll x((1 + M)e^{-M} + L^{-1})$ with this minimum $M$. For fixed $W$ the retained sum is a difference of two prefix sums of $g_z$, with endpoints between $D/2$ and $5V^{4/T}$ whenever the

interval is nonempty. Halász's theorem is applied to $g_z$ itself, and subtraction of those prefixes supplies the bound for $\beta$.

Finally for each fixed $w$ the possible $u$ lie in $[V/(wz^T), 2V/(wD)]$. The reciprocal sum of the $z$-rough integers in this interval is $O(1+T)$. Indeed the term 1 costs one and partial summation of $\#\{n \le x : P^-(n) > z\} \ll x/\log z$ for $x \ge z$ bounds the rest by $O(1 + \log(2z^T/D)/\log z)$. The latter rough-number estimate follows from the dimension-one upper sieve, or from *Numerically explicit estimates for the distribution of rough numbers* [8]. Hence $\sum_W 1/W \ll (1+T)/T^d$. The density sum over $q$ is bounded. Combining this with the two discarded tails proves the third signed estimate with a positive power of $L$ saved. Walsh inversion and the row argument of Theorem 1.2 complete the proof. □

In particular for $K = 1$, if $\varepsilon_V = T^d L^{-c_\alpha}$, all four conditional probabilities differ from $1/4$ by $O(\sqrt{\varepsilon_V})$ outside fibres of total population probability $O(\sqrt{\varepsilon_V})$. This averages by the number of integers in each fibre.

### 4.2 A growing number of prescribed prime factors

In the two-prime model the first mark can test a whole ordered list. This remains true for a growing list, with a correction to the leading mass near length $\sqrt{T}$.

**Proposition 4.2.** *Fix $0 < \rho_0 < \sqrt{2/3}$, let $1 \le d \le \rho_0\sqrt{T}$ be integral, and split $[1/d - 2/T, 1/d - 1/T]$ into $d$ disjoint consecutive bands of equal width, assigning their shared endpoints once. Require $v = t\prod_{i=1}^d p_i$ with one prime from each band and $v - 1 = qm$ with $q \in \mathcal{Q}$. On the resulting integer population $E(d)$,*

$$|E(d)| = (c_Q + o(1))\frac{V}{T^d}\exp\left(\frac{3d^2}{2T}\right)$$

*uniformly for this range of $d$. The four-sign estimate (1) holds with any bounded joint mark of $(p_1, \dots, p_d)$ and a separate mark on $q$, with error $O_{\rho_0}(VL^{-c})$. In particular the single-label total variation conclusion holds with error $O_{\rho_0}(T^d L^{-c}) = L^{-c+o(1)}$.*

*Proof.* Write $a_i = 1/d - 2/T + (i-1)/(dT)$ for the left endpoints. Their sum is $1-(3d+1)/(2T)$, and the sum of the right endpoints is $1-(3d-1)/(2T)$. Thus on the enlarged dyad the cofactor is at most $3V^{(3d+1)/(2T)}$, strictly below the smallest band prime $V^{1/d-2/T}$ for all sufficiently large $V$. Indeed the exponent gap is $1/d - (3d+5)/(2T) \gg_{\rho_0} 1/d$. Any origin having a prime in every band therefore has exactly one in each, including multiplicity. The product $w$ is a single bounded-coefficient variable in the dispersion argument. The dyadic short lengths satisfy $\log N \asymp dL/T$, uniformly for $1 \le d \le \rho_0\sqrt{T}$. They therefore fall in Lemma 2.2 with, for example, $a = 1/2$ and $B = 5$, for all sufficiently large $V$. The threshold may also depend on $\rho_0$. The mesh and its boundary losses are uniform since $w \ll V^{1-1/T}$ and each origin admits at most one tuple. This gives the mass formula with error $O_{J,\rho_0}(VL^{-J})$ before evaluating the reciprocal sums, and gives the same bound for the single short sign by the prime number theorem for $\lambda$. All band primes exceed $\exp(L^{0.9})$, so Theorem 3.3 gives the other signed moments with a single arbitrary function of the rough part, independently of $d$. The identity at the first site is $\lambda(t) = (-1)^d\lambda(v)$.

For the $i$th band put $s_i = 2d - i + 1$. Its reciprocal prime mass is $\log(1 + 1/(T - s_i)) + O_J(L^{-J})$, uniformly by the prime number theorem with its classical remainder since $a_i L \gg L/\sqrt{T}$. If this mass is $m_i$, Taylor expansion yields

$$\log(Tm_i) = (s_i - 1/2)/T + O((d^2+1)/T^2).$$

Summing gives $3d^2/(2T) + O((d^3+d)/T^2)$, and $\sum_{q\in\mathcal{Q}} 1/(q-1) \to c_Q$. All absolute log-power errors are negligible relative to $T^{-d} = L^{-o(1)}$. This proves the claims, including the factor $e^{3\rho^2/2}$ in the relative mass if $d/\sqrt{T} \to \rho$. □

### 4.3 A phase on the quotient of the two cofactors

Let $\chi$ be a unitary completely multiplicative function, extended to the positive rationals. Hereafter in this section let $P$ denote the mark $A(p)B(q)$ in $E_1$ or $A(r)C(s)B(q)$ in $E_2$. All three prime marks are bounded. Factoring the mark inside the first-site pair allows us to choose which endpoint supplies cancellation.

**Proposition 4.3.** *Let $P = A(p)B(q)$ on $E_1$ and $P = A(r)C(s)B(q)$ on $E_2$, with each prime mark bounded by one. Uniformly over $\chi$ and these marks, for $u = t$ and for $u = m$ one has $\left|\sum_{E_d} P\lambda(u)\chi(t/m)\right| \ll VL^{-c}$ for an absolute $c > 0$. In particular the first-site mark on $E_2$ is factorised between $r$ and $s$. For a homomorphism $\pi : \mathbb{Q}_{>0}^{\times} \to K$ to a finite abelian group and $|H| \le 1$,*

$$\left|\sum_{E_d} P\lambda(u)H(\pi(t/m))\right| \ll \sqrt{|K|}\, VL^{-c}.$$

*For a single prescribed image the factor $\sqrt{|K|}$ can be omitted.*

*Proof.* Use $\chi(t/m) = \chi(v)\overline{\chi(v-1)}\chi(q)\overline{\chi(w)}$ and absorb the last two factors into the marks. At a single-prime site use the two masks from the proof of Theorem 1.1. At a two-prime site let a multiplicative mask have prime values $zA(r)$ and $z'C(s)$ in the bands and one elsewhere. On each integer in the dyad it gives a monomial of total degree at most two. Extract degree $(1,1)$ by averaging at the three cube roots in each variable, with multiplier $\overline{zz'}$. Thus it suffices to consider two bounded multiplicative functions $f, g$, with anchor masks $s_1, s_2$ differing from one only on the displayed prime windows. Their unmasked factors are $\lambda\chi, \bar{\chi}$ or $\chi, \lambda\bar{\chi}$ in the order of the sites. They satisfy (G1) of Section 3. Each endpoint is available for (G2).

Choose $U = (\log x)^{1/A}$, enlarging the absolute $A$ used in Theorem 3.3 if necessary. For the two candidate intervals of (G2) let $Q_* = (C_h(1+U))^5$ dominate both conductor bounds in Theorem 1.7 of [2]. In the notation of Section 3, take $M_1 = M(f; Z, Q_*)$ and $M_2 = M(g; Z, Q_*)$. The product triangle inequality for the pretentious distance gives, with minimising twists,

$$\sqrt{M_1} + \sqrt{M_2} \ge \mathbb{D}_Z(\lambda s_1 s_2, \psi_1\psi_2(n)n^{\mathrm{i}(\tau_1+\tau_2)}).$$

Here $\psi_1\psi_2$ is viewed modulo the least common multiple. One may extend from $Z$ to $2Z$ at cost $O(1)$ in squared distance. The conductor is at most $Q_*^2 \le (\log(2Z))^{1/125}$ and the frequency at most $2Z$. Removing the anchor masks costs at most $2\sum_{p \text{ in the bands}} 1/p = O(1)$. By (1.12) of [2], $\max M_i \ge (1/12 - o(1)) \log\log Z$. The endpoint attaining the maximum meets (G2) by the same calculation as in Theorem 3.3. Lemma 3.2 now applies to both prefixes. For the group observation, normalized Fourier inversion on $K$ has coefficient norm at most $\sqrt{|K|}$ by Parseval. An indicator of one group element has coefficient norm one. □

For instance a prescribed vector $(v_p(t) - v_p(m))_{p\in I}$ for any finite prime set $I$ can be tested at the same absolute cost $O(VL^{-c})$ by integrating the product of coordinate phases against Haar measure on $(\mathbb{R}/\mathbb{Z})^I$. No factor depends on $|I|$. The bound concerns either individual sign. The observable $\lambda(t/m)$ already fixes their product. In fact $t, m$ are coprime, so reading the complete rational number $t/m$ as a reduced fraction gives both cofactors. The Fourier bound for one prescribed value does not sum uniformly over all such values.

### 4.4 Splitting the prime multiplicities

For a set $S$ of primes, possibly depending on $V$, let $\Omega_S(t) = \sum_{p\in S} v_p(t)$. Divide the prime factors from the two cofactors into complementary counts

$$C_S = \Omega_S(t) + \Omega_{S^c}(m), \qquad D_S = \Omega_{S^c}(t) + \Omega_S(m).$$

For $k = 2$ the two separate parities are instances of Proposition 4.3, and their product is $\lambda(t)\lambda(m)$. The marginal equidistribution of $\Omega$ modulo a fixed integer is classical; see Coons and Dahmen [26]. Here the prime partition, the cofactor population and the separate prime marks are retained simultaneously. The power of two in the following extension enters through an elementary fact about roots of unity.

**Theorem 4.4.** *With the factorised marks $P = A(p)B(q)$ on $E_1$ and $P = A(r)C(s)B(q)$ on $E_2$, for $k = 2^j \geq 2$ satisfying $k^2 \log(2T) = o(T)$ and each $a_0, b_0 \bmod k$,*

$$\sum_{E_d} P\left(\mathbf{1}_{(C_S, D_S) \equiv (a_0, b_0) \bmod k} - k^{-2}\right) \ll V \exp(-c_1 T/k^2)$$

*with absolute $c_1 > 0$, uniformly in $S$, the cells, and the separate prime marks. For $P = 1$ the error is $o(|E_d|/k^2)$ throughout this range.*

*Proof.* In a Fourier mode $(a, b)$ with $\zeta = e(1/k)$, the two multiplicative functions before anchor selection have prime values $\zeta^a, \zeta^b$ at the first site and $\zeta^b, \zeta^a$ at the other, according to membership in $S$. Factors at the anchors are absorbed into $P$ and projected out as above. If $a + b \neq 0 \bmod k$, choose $1 \leq r \leq k/2$ with $r(a + b) \equiv k/2 \bmod k$. Such an $r$ exists by writing $a + b = 2^\nu u$ with $u$ odd and taking $r = k/2^{\nu+1}$. Thus the $r$th power of the product before masking is $\lambda$. If $a + b = 0$ and the mode is nonzero, choose $r$ with $ra = k/2 \bmod k$. The first-site values are now inverses, and its $r$th power is $\lambda$ at every prime.

For bounded multiplicative functions and unit comparison phases (including Dirichlet characters), $\mathbb{D}(f^r, g^r) \leq r\mathbb{D}(f, g)$. This follows prime by prime from $1 - \Re z^r \leq r^2(1 - \Re z)$ for $|z| \leq 1$, by the unit-circle inequality and harmonic extension. At a zero of a character the inequality is immediate. Apply this and the product triangle inequality to the minimisers used above. Their powers have modulus of definition at most $Q_*^2$ and frequencies at most $2rZ$. Moving the cutoff from $Z$ to $2rZ$ costs $O(1)$ in squared distance since $r \leq k = O(\sqrt{T})$. Removing the already powered anchor masks costs $O(1)$ as well. Formula (1.12) of [2] at $2rZ$ now implies $\max(M_1, M_2) \gg T/k^2$.

Set $U = \exp(\varepsilon T/k^2)$ with sufficiently small fixed $\varepsilon > 0$. This tends to infinity, $U^{A_*} \leq \log x$ and the bound on $Q_*^2$ above holds. The short-interval error $e^{-\max(M_1, M_2)/20}$ is $O(U^{-0.8})$ by the choice of $\varepsilon$. The other two terms meet (G2) as before. The small-prime phases in (G1) are the two unmasked root-of-unity functions. Lemma 3.2 gives $O(V \exp(-c_1 T/k^2))$ per nonzero mode. The inverse Fourier coefficients of one cell have total absolute mass one, independent of $k$. Finally $k^2 T^d \exp(-c_1 T/k^2) = o(1)$ by the stated range. □

# 5 Four prime factors and the nine Möbius states

For $E_3$, recall $v = rst$, $v - 1 = abu$, with $r, s$ close to the square-root scale and $a, b$ in the two bands about $V^\gamma$. Choosing a tuple uniformly weights an origin by its number of representations. We also use the probability proportional to $\log r \log s \log a \log b$ on the same tuples. Section 2 computed the two normalising constants.

The nine constants in the next table are classical. Mirsky [7] supplies the squarefree support, while Tao [19, Corollary 7] obtains the full signed table for consecutive integers in logarithmic averages. Matomäki, Radziwiłł and Tao [20, Theorem 1.9] prove positive lower natural density for each of the nine patterns. Here we transport the table to cofactor populations under ordinary counting. This requires square-divisibility tails small relative to the shrinking population, and the measure on $E_3$ remains a measure on tuples.

**Theorem 5.1.** *In $E_3$ each pair $(\lambda(t), \lambda(u))$ has limiting probability $1/4$ under either law. More precisely, the absolute discrepancy in (1) is $O_{b_0}(VL^{-c})$ for a bounded mark $A(rs)B(a)C(b)$ on*

*any fixed class modulo $b_0$. Under either law, $(\mu(t), \mu(u))$ has the following limiting table. The same table holds for $(\mu(t), \mu(m))$ on $E_1, E_2$ with uniform counting. Put*

$$D_1 = 1/\zeta(2), \qquad D_2 = \prod_p (1 - 2/p^2).$$

| | $+1$ | $-1$ | $0$ |
|---|---|---|---|
| $+1$ | $D_2/4$ | $D_2/4$ | $(D_1 - D_2)/2$ |
| $-1$ | $D_2/4$ | $D_2/4$ | $(D_1 - D_2)/2$ |
| $0$ | $(D_1 - D_2)/2$ | $(D_1 - D_2)/2$ | $1 - 2D_1 + D_2$ |

*Proof.* At the first site $\mathcal{R}_y(v)$ reads the entire prime pair and hence its mark. At the second site an integer may admit two distinct choices. We describe exactly how to count both. Apart from integers divisible by the square of a band prime, the site weight is

$$\left( \sum_{a \in \mathcal{A},\ a \mid v-1} B(a) \right) \left( \sum_{b \in \mathcal{B},\ b \mid v-1} C(b) \right).$$

For $n = v - 1$ let $d_A, d_B$ count the two kinds of prime factors with multiplicity, so $d_A + d_B \le 3$. Give the primes of the bands respectively the bounded multiplicative values $(1 + zB(a))/2$ and $(1 + z'C(b))/2$, and every other prime the value one. On $n$ extract bidegree $(1, 1)$ by Fourier averaging over fourth roots in both variables. The coefficient is the displayed product times $2^{-d_A - d_B}$ when the band factors are simple. Restore $2^{d_A + d_B}$ by the four-term discrete Fourier expansion of the function $j \mapsto 2^j$, $0 \le j \le 3$. Its character factors multiply the prime masks by a unit phase on the two bands. Thus the weight is a linear combination of admissible multiplicative masks with absolute coefficient sum bounded by an absolute constant. The number of omitted origins is

$$\ll \sum_{V^{1/4} < p \le \sqrt{2V}} (V/p^2 + 1) \ll V^{3/4} + \sqrt{V}.$$

Their multiplicity is bounded. Both anchors have even parity, so $\lambda(t) = \lambda(v)$ and $\lambda(u) = \lambda(v-1)$. Apply Theorem 3.3 and Proposition 2.1 as in the proof of Theorem 1.1. To insert the prime logarithms divide each one by $L+1$ before applying the bounded-mark estimates. The uniform limits of their ratios to $L$ also show that the two probability measures have total variation distance $o(1)$.

For the zero states fix $Y \ge 2$ and test divisibility by $p^2$ at both sites for $p \le Y$. Eventually all selected primes exceed $Y$. The finite test is a union of fixed residue classes modulo $\prod_{p \le Y} p^2$. For the unsigned mass each such class has asymptotic relative density equal to its natural density, by Proposition 2.1 and direct counting of $t$ in a progression. In each fixed class the three signed moments have absolute error $O_Y(VL^{-c})$. The two forbidden residues are $0, 1 \bmod p^2$, which gives the finite products for $D_1, D_2$, the classical squarefree-pattern densities of Mirsky [7].

We justify removing the cutoff in the rare population. Put $L_0 = \log V$. For $Y < p \le L_0$, apply Proposition 2.1 with $b_0 = p^2$. With unit marks the count for either class is the unrestricted density main term divided by $p^2$, up to $O_J(VL^{-J})$. The floor errors are $O(\#\{w\})$ times a bounded reciprocal sum. After summation and normalisation the tail is $O(1/Y) + o(1)$. Above $L_0$ the unrestricted bound is $O(V/L_0 + \sqrt{V})$, also negligible at mass $V/T^4$. A square in a cofactor implies a square in its site. For any fixed prime the converse holds once the anchors exceed it. Consequently the truncation argument applies to the cofactors themselves. First take $V \to \infty$, then $Y \to \infty$. On the event that both cofactors are squarefree the two signed coordinates are each equal to their Liouville signs, giving $D_2/4$ per cell. On the event that only the first is squarefree its sign is fair, giving $(D_1 - D_2)/2$ per sign. Interchange the sites for the other two cells. The last entry follows by subtraction. □

A fixed congruence can also be retained in this argument. For example, conditional on a fixed sign pair, the residues $v \bmod M$ are asymptotically uniform for every fixed $M$. Thus finite lists of valuations of fixed integer linear forms in $v$ have the same limiting law as for a uniform residue at every prime. To see this, truncate the valuations at $j$, take $M$ to be the product of the corresponding $p^j$, then let $j$ increase. For a nonconstant form the Haar probability of a root modulo $p^j$ tends to zero. Primes dividing a coefficient are handled by these same congruences. Conditioning on $v \equiv x \bmod M$ instead can change the constants in the table, with local factors obtained by counting the allowed residues in that class.

## 6 The number of prime factors of the short cofactor

Write $K = \Omega(t)$. Knowing $w$ fixes the interval $(V/w, 2V/w]$ of the short cofactor. Its length varies with $w$. The opposite label $q'$ enters the dispersion main term through a separate factor. We quantify the effect of that difference on the count $K$.

Throughout this section the probability is uniform on $E_1, E_2$ or on the tuples of $E_3$. All assertions also hold on those tuples with probability proportional to $(\log r)(\log s)(\log a)(\log b)$. Marginals and conditional expectations refer to the chosen probability. Put

$$m_0 = T - \log T, \qquad u(w) = \frac{T \log(V/w)}{L}, \qquad \ell(w) = \log u(w).$$

Let $G_T$ be the law of $\lfloor m_0 + \sqrt{T}\, Z + 1/2 \rfloor$ for a standard normal $Z$. For an observation $O$ in a measurable space, put

$$\Delta(O) = d_{\mathrm{TV}}(\mathcal{L}(O, K), \mathcal{L}(O) \otimes \mathcal{L}(K)).$$

Since $K$ is integer-valued, this distance is half the expectation of $\sum_j |\mathbb{P}(K = j \mid O) - \mathbb{P}(K = j)|$. A comparison with $G_T$ also measures the discrepancy of the marginal law of $K$ from a Gaussian. The two distances will have different first-order behaviour.

Wang, Wei, Yan and Yi [17, Theorem 1.1] prove a Gaussian limit for $\Omega(n)$ when the largest prime factor belongs to a fixed set of primes of positive natural density. In $E_1$ we observe the value of that largest prime in a shrinking exponent window and impose a second condition at $v - 1$. For consecutive integers, Charamaras and Richter [22] study the two counts in logarithmic averages. Wang and Yi [18] treat ordinary averages, and Andreevski [23] gives a joint local limit on scales of logarithmic density one. The population here is selected by prime windows at both sites. The dependence in $\Delta(O)$ compares the count with the actual label on that population.

**Theorem 6.1** (Gaussian approximation and residues)**.** *For $O = w$ and for $O = q'$,*

$$d_{\mathrm{TV}}(\mathcal{L}(O, K), \mathcal{L}(O) \otimes G_T) \ll T^{-1/2}.$$

*For $k \geq 2$, let $d_k$ be the distance of $\mathcal{L}(O, K \bmod k)$ from $\mathcal{L}(O)$ times the uniform law modulo $k$. If $k/\sqrt{T}$ tends to $0$, to $\rho \in (0, \infty)$, or to infinity, then $d_k$ tends respectively to $0$, to*

$$\Psi(\rho) = \frac{1}{2} \int_0^1 \left| \sum_{h \in \mathbb{Z}} e^{-2\pi^2 h^2/\rho^2} e(hx) - 1 \right| \mathrm{d}x,$$

*or to $1$. On any compact subinterval of $0 < k/\sqrt{T} < \infty$ one has $d_k = \Psi(k/\sqrt{T}) + O(T^{-1/2})$, with the constant depending on that subinterval.*

**Theorem 6.2** (Dependence on the observed label)**.** *For every fixed $A > 0$ one has $\Delta(q') \ll_A L^{-A}$. Given $w$, draw a message $H$ from any probability kernel at $w$, conditionally independently of $K$. Uniformly over these kernels,*

$$\Delta(H) = \frac{\mathbb{E}|\mathbb{E}(\ell(w) \mid H) - \mathbb{E}\ell(w)|}{\sqrt{2\pi T}} + O(T^{-1}).$$

*In particular $H = w$ gives the first-order dependence on the entire first label. The single bit $\mathbf{1}_{\ell(w)>\mathbb{E}\ell(w)}$ attains the same coefficient.*

**Corollary 6.3** (Constants for the prime windows)**.** *Let $U_1$ be uniform on $[1,2]$. Let $U_2$ and $U_3$ each have the law of the sum of independent uniforms on $[1,3/2]$ and $[3/2,2]$. With $\gamma_{\mathrm{E}}$ denoting Euler's constant, set*

$$b_\Omega = \gamma_{\mathrm{E}} + \sum_p \left( \log(1-1/p) + \frac{1}{p-1} \right).$$

*In population $E_d$ the window gives the two limits*

$$\begin{aligned} \sqrt{T}\,\Delta(w) &\longrightarrow \frac{\mathbb{E}|\log U_d - \mathbb{E}\log U_d|}{\sqrt{2\pi}}, \\ \sqrt{T}\,d_{\mathrm{TV}}(\mathcal{L}(O,K), \mathcal{L}(O)\otimes G_T) &\longrightarrow \frac{b_\Omega + \mathbb{E}\log U_d - 1/6}{\sqrt{2\pi}} \qquad (O = w,\ q'). \end{aligned}$$

*Both constants are strictly positive. In particular $T^{-1/2}$ is the exact order for the rounded normal specified in Theorem 6.1.*

For $E_1$ the two limits are about 0.06800 and 0.50039. For $E_2$ and $E_3$ they are about 0.02224 and 0.78363. The second number includes the error in replacing the count itself by a Gaussian.

## 6.1 Transferring a one-variable law with the label

In the model supplied by Proposition 2.1, choose $w$ with weight $N_w = \lfloor 2V/w \rfloor - \lfloor V/w \rfloor$, choose $t$ uniformly from that dyad, and choose $q'$ independently of $(w,t)$ with weight $1/\varphi(q')$. In the weighted population multiply the first and last weights respectively by $\log r \log s$ and $\log a \log b$. These are functions of their respective products because the two bands at each site are disjoint. The conditional law of $t$ given $w$ is still uniform.

For fixed $j$, take $\beta(t) = \mathbf{1}_{\Omega(t)=j}$ in (2). Taking the sign of the discrepancy at each value of one label as its mark sums the absolute discrepancies for that label. In the weighted case use the two stated weights divided by $(L+1)^2$ each, and multiply the chosen mark by the sign. The resulting marks are bounded and the mass after this scaling is $\gg V/T^4$ in all cases. There are $O(L)$ values of $j$. The elementary normalisation inequality for two positive measures therefore gives, for either label and every fixed $A$,

$$d_{\mathrm{TV}}(\mathcal{L}(O,K), \mathcal{L}_{\mathrm{mod}}(O,K)) \ll_A L^{-A}.$$

Indeed the bound before choosing the dispersion exponent $J$ is $O_J(T^4 L^{1-J})$. Contraction gives the same estimate for each marginal. In the model, $q'$ and $K$ are independent. The product of their actual marginals differs from the model product by at most the sum of the two marginal errors. This proves the assertion about $q'$ in Theorem 6.2, including with the four logarithms.

Write $\pi_m$ for the Poisson law of mean $m$. For a uniform integer $U_x$ up to $x$, put $m = \log\log x$. Barbour, Kowalski and Nikeghbali [21, Section 8.3, Theorem 8.2, case $\Omega$] give the order-one approximation

$$\big\| \mathcal{L}(\Omega(U_x)) - \pi_m - b_\Omega\, \dot\pi_m \big\|_1 \ll m^{-1} + m/\log x, \qquad \dot\pi_m(j) = \pi_m(j)(j/m - 1). \tag{4}$$

The coefficient is the derivative at $z=1$ of $\Gamma(z)^{-1}\prod_p (1-1/p)^z (1-z/p)^{-1}$. Differentiating the locally convergent product gives the displayed $b_\Omega$. Each prime contributes a positive $O(p^{-2})$ term, so $b_\Omega > \gamma_{\mathrm{E}} > 1/2$.

The second derivative in the Poisson mean satisfies

$$\ddot\pi_m(j) = \pi_m(j)\frac{(j-m)^2 - j}{m^2}, \qquad \|\ddot\pi_m\|_1 \le 2/m.$$

Thus Taylor's formula replaces $\pi_m + b_\Omega \dot{\pi}_m$ by the probability $\pi_{m+b_\Omega}$ at cost $O(1/m)$ in $\ell^1$. Subtract the two prefix counts for $(Y, 2Y]$ and divide by $\lfloor 2Y \rfloor - \lfloor Y \rfloor$. The two coefficients are bounded and sum to one with their signs. The Poisson parameters at $Y$ and $2Y$ differ by $O(1/\log Y)$, which also bounds their distance by the independent-increment coupling. Replacing real endpoints by their integer parts costs $O(1/(Y \log Y))$ in the parameters. For $t_Y$ uniform on that dyad we obtain

$$d_{\mathrm{TV}}(\mathcal{L}(\Omega(t_Y)), \pi_{\log\log Y + b_\Omega}) \ll 1/\log\log Y. \tag{5}$$

Mean adjustment for Poisson approximation is classical [27, 21]. Chhaibi, Delbaen, Méliot and Nikeghbali [28, Section 4.5] obtain precise distance asymptotics for the count of distinct prime divisors. Here (5) is used at the scale selected by an observed label.

### 6.2 How much of the count a message retains

For $Y = V/w$ the identity $\log\log Y = m_0 + \ell(w)$ is exact. In our windows $1 \le u(w) \le 7/2$. By (5) and the preceding transfer, the actual joint law of $(w, K)$ is within $O(1/T)$ of the law obtained by taking the actual $w$ and drawing, conditionally on it, a Poisson variable of mean $m_0 + b_\Omega + \ell(w)$. This estimate averages over the label.

Put $b = m_0 + b_\Omega$, $D = \ell(w)$ and $\bar{D} = \mathbb{E}D$. Apply the same message kernel in both laws. Conditionally on the information $\sigma(H)$, the Poisson mixture has expansion

$$\pi_b + \mathbb{E}(D \mid H)\dot{\pi}_b + O_{\ell^1}(T^{-1}) \qquad \text{almost surely.}$$

Here and below $O_{\ell^1}$ bounds the sum of absolute masses. The remainder is uniform because $D$ is bounded and $\|\ddot{\pi}_{b+s}\|_1 \ll 1/T$ for bounded $s$. The marginal expansion has $\bar{D}$ in place of the conditional mean. Taking the expected $\ell^1$ norm of the difference gives

$$\Delta(H) = \tfrac{1}{2}\|\dot{\pi}_b\|_1 \, \mathbb{E}|\mathbb{E}(D \mid H) - \bar{D}| + O(T^{-1}).$$

For $P \sim \pi_b$, summing $(j-b)\pi_b(j)$ over $j > b$ telescopes and gives $\mathbb{E}|P - b| = 2b\,\pi_b(\lfloor b \rfloor)$. Stirling's formula yields $\frac{1}{2}\|\dot{\pi}_b\|_1 = (2\pi T)^{-1/2} + O(\log T/T^{3/2})$. Within either value of $\mathbf{1}_{D>\bar{D}}$ the variable $D - \bar{D}$ has constant weak sign. This proves the equality of its coefficient with the one for $H = w$ and completes Theorem 6.2. The same Taylor bound gives $d_{\mathrm{TV}}(\mathcal{L}(K), \pi_{b+\bar{D}}) \ll 1/T$.

To evaluate the coefficient, floors in $N_w$ may be replaced by $V/w$ with relative total error $o(1)$, by the window bounds in Section 2. On a prime band write $\log p/L = \beta - x/T$, with $\beta = 1$ or $1/2$. Partial summation of the prime number theorem gives reciprocal-prime mass asymptotic to $\int dx/[T(\beta - x/T)]$ on each fixed nondegenerate subinterval of the $x$-band. Thus the normalised $x$ tends to uniform measure there. For $w = rs$, the weight $1/(rs)$ factors and the two coordinates become independent. The weights $\log r \log s/L^2$ converge uniformly to $1/4$. Consequently $u(w)$ tends in law to $U_d$ under the actual measure in each case. Boundedness and continuity of log on these intervals prove the first limit of Corollary 6.3.

### 6.3 The Gaussian error and its reduction modulo an integer

For bounded $a$ let $m = m_0$, $z_j = (j-m)/\sqrt{m}$ and $\phi(z) = e^{-z^2/2}/\sqrt{2\pi}$. Ouimet [29, Lemma 2.1] gives the local expansion at the Poisson mean. Expanding instead around $m$ gives

$$\pi_{m+a}(j) = \frac{\phi(z_j)}{\sqrt{m}} + \frac{\phi(z_j)}{m}\left(a z_j + \frac{z_j^3 - 3z_j}{6}\right) + r_j, \qquad \sum_{j\in\mathbb{Z}} |r_j| \ll m^{-1}.$$

To see the uniform remainder, expand $-m - a + j\log(m+a) - \log(j!)$ for $|z_j| \le m^{1/12}$. The linear correction to the log-density is $(az_j + (z_j^3 - 3z_j)/6)/\sqrt{m}$, and the remainder is $O((1 + |z_j|^4)/m)$. Exponentiating and summing the Gaussian times $(1 + |z_j|^6)/m^{3/2}$ costs

$O(1/m)$. The complement costs $O(e^{-cm^{1/6}})$ by the exponential Markov bound for a Poisson variable and the Gaussian tail.

For a normal density $f$ of mean and variance $m$, summing the midpoint integration errors over the unit cells costs $O(\|f''\|_1) = O(1/m)$, since the linear term integrates to zero. Changing the variance to $T$ costs $O(\log T/T)$ by integrating the absolute variance derivative. Hence uniformly for bounded $a$,

$$d_{\mathrm{TV}}(\pi_{m+a}, G_T) = \frac{1}{2\sqrt{T}} \int_{\mathbb{R}} \phi(z) \left| az + \frac{z^3 - 3z}{6} \right| \mathrm{d}z + O(\log T/T). \tag{6}$$

The corresponding Riemann sums differ from the integral by $O(1/\sqrt{m})$ before the outside scaling, since the integrands have uniformly bounded variation. In our applications $a = b_\Omega + \ell(w)$ or $a = b_\Omega + \bar{D}$, and $a > 1/2$. The polynomial has the sign of $z$. Integrating $z\phi(z)$ and $z^3\phi(z)$ on the positive half-line reduces the integral including its factor $1/2$ to $(a - 1/6)/\sqrt{2\pi}$.

For $O = w$, integrate (6) against the actual first label and use the $O(1/T)$ joint Poisson comparison. For $O = q'$, its dependence on $K$ is $O_A(L^{-A})$ and the marginal count is within $O(1/T)$ of $\pi_{b+\bar{D}}$. Both calculations give the same leading coefficient. Letting $u(w)$ tend to its window law proves the second limit in Corollary 6.3 and the Gaussian bound of Theorem 6.1.

For fixed modulus, Coons and Dahmen [26] study $\Omega$ on unrestricted integers. Here the modulus may vary with the cofactor scale. Projection modulo $k$ contracts total variation. Wrapping the continuous normal around a circle of circumference $k$ and rescaling to circumference one gives, up to translation, the density $\sum_{h\in\mathbb{Z}} e^{-2\pi^2 h^2 T/k^2} e(hx)$. Rounding and reducing modulo $k$ integrate this density on $k$ consecutive arcs of length $1/k$. When $k/\sqrt{T}$ lies in a compact subinterval of $(0, \infty)$ its derivative is uniformly bounded. The sum of the absolute arc discrepancies therefore differs from the integral of the absolute deviation by $O(1/k)$, uniformly in the translation. This proves the moving-profile estimate and, by continuity of $\Psi$, the critical limit. When $k/\sqrt{T} \to 0$, the nonzero Fourier coefficients have vanishing absolute sum. When $k/\sqrt{T} \to \infty$, $O(H\sqrt{T}) = o(k)$ consecutive residues carry probability tending to one for a sufficiently slowly growing $H \to \infty$. Their uniform mass tends to zero. This proves the remaining limits.

The one-variable local limit also follows from (5) and Stirling's formula. In particular, with $m_Y = \log\log Y$,

$$\sup_j \left| \sqrt{m_Y}\, \mathbb{P}(\Omega(t_Y) = j) - \phi((j - m_Y)/\sqrt{m_Y}) \right| \longrightarrow 0.$$

The bound on the error in (5) is uniform as the minimum of $Y$ tends to infinity. The bounded shift $b_\Omega$ disappears in this local limit and enters the constant for the Gaussian comparison.

## 7 When both prime pairs are near the square root

Now take $r, a \in \mathcal{R}$ and $s, b \in \mathcal{S}$, so the same two bands occur at both sites. Let $E_4$ count integers $V < v \le 2V$ with $v = rst$, $v - 1 = abu$. For large $V$ both pairs are unique. In this setting neither site supplies the small modulus product in Lemma 2.2. Separating the sign estimates from the size of the population matters here.

**Proposition 7.1.** *For bounded separate prime marks $A(r)B(s)C(a)D(b)$, each of the three nonconstant Liouville moments of the cofactors is $O_M(VL^{-c})$ on every dyad, also in a fixed residue class modulo $M$. There is a set $\mathcal{G}$ of positive integer scales and an absolute $\delta > 0$ with*

$$\sum_{V \le x,\ V \notin \mathcal{G}} \frac{1}{V} \ll (\log(2x))^{1-\delta}.$$

*Along $\mathcal{G}$ one has $|E_4| \sim V/T^4$, four Liouville probabilities $1/4 + o(1)$ and the limiting Möbius table of Theorem 5.1. The mass with the four positive prime logarithms is $(1/16 + o(1))VL^4/T^4$ and has the same limiting probabilities.*

*Proof.* At a site bearing a Liouville sign, select its two bands by the cubic-root coefficient extraction in the proof of Proposition 4.3. This gives a finite linear combination of $\lambda s$ with $s$ equal to one on small primes. The opposite site is an arbitrary bounded function of its rough part, with any required opposite Liouville sign retained. Theorem 3.3 gives the signed estimates on every scale. This proof permits the sign-bearing site to be either $v$ or $v-1$.

For the mass use Tao and Teräväinen's *Quantitative correlations and some problems on prime factors of consecutive integers*, Theorem 3.1(i) [10]. We explain the two uniformities concerning residue classes and moving windows. Freeze a pair of bands at a scale $N_0 \in [\sqrt{X}, X]$ for a moment. For a set $I$ of band primes put $s_I(n) = \mathbf{1}_{p\nmid n\ \forall p\in I}$. The selector for the presence of a prime in each band is $(1-s_{\mathcal{R}})(1-s_{\mathcal{S}})$, a linear combination of four real completely multiplicative functions $s_I$ (including the constant function). All their modified primes exceed $X^{0.23}$ for large $X$. For $X^{0.4} \le N \le X$ put $\delta_N = N^{-1}\sum_{N<n\le 2N} s_I(n)$. For $q \le (\log X)^{1/3}$ inclusion–exclusion over the divisors composed of distinct modified primes gives, uniformly in $a$,

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \bmod q}} s_I(n) = N\delta_N/q + O(1 + N/\log X).$$

Indeed each divisor $d \le 2N$ is a unit modulo $q$ and contributes $O(1)$ floor error. There are at most four prime factors. Fixing all but the last gives $O(N/(d'\log X))$ choices for that last prime by Chebyshev, since its upper bound, if feasible, exceeds $X^{0.23}$. The sum of $1/d'$ is bounded by Mertens' theorem on $[X^{0.23}, 2X]$. For larger $q$ the trivial $O(N/q+1)$ suffices for (3.1) of [10] with $\mathcal{L} = (\log X)^{1/3}$. All masks equal one in the prime range required by (3.2).

Apply the cited theorem to each pair of masks, with shifts $0, -1$. Its exceptional logarithmic measure on $[\sqrt{X}, X]$ is $O((\log X)^{1-\delta_0})$ for some absolute $\delta_0 > 0$, simultaneously for every modulus $M \le (\log X)^{\delta_0}$ and its classes. On a good scale its conclusion replaces the first mask by its dyadic mean in the correlation. The progression calculation just given evaluates the remaining one-site mean. The shift of one contributes $O(1)$ to a count. Consequently the main term for the two selectors in a class is $N/M$ times the square of the one-site density, up to $O(N(\log X)^{-\delta_0})$ after decreasing $\delta_0$.

Now put the grid for $\log N/\log X$ at mesh $(\log X)^{-\delta_0/4}$ and choose $N_0$ in each cell. The union of the exceptional sets still has logarithmic measure $O((\log X)^{1-3\delta_0/4})$. In its own cell, each frozen band endpoint differs from its moving endpoint by $O((\log X)^{-\delta_0/4})$ in units of $\log X$, since the four endpoint functions have bounded derivatives there. The proportion of integers with a prime factor in a changed slice of width $\xi$ is $O(\xi + 1/\log X)$, by partial summation for primes and the floor term $O(\pi(N^{0.6})/N)$ near the actual cell. This error bounds the change of either one-site selector and hence of the two-site count, independently of $M$ and its class.

For the actual bands at $N$, uniqueness gives one-site density $s(N) = \sum_{r,s} 1/(rs)$ up to $O(N^{-2/\log\log N})$, since each product is $\le N^{1-5/(2\log\log N)}$. Thus for a sufficiently small fixed $\kappa > 0$ the class count equals $(N/M)s(N)^2 + O(N(\log X)^{-\kappa})$ whenever $M \le (\log X)^{\kappa}$ on the good scales. To pass to integers, assign each $V$ an interval $[V, V+1)$. For the frozen masks, the dyadic counts vary by $O(1)$ on it. The frozen main terms have the same variation: if $C_I(N) = \sum_{N<n\le 2N} s_I(n)$, then $C_I(N) = O(N)$ and $C_I(N') - C_I(N) = O(1)$ for $|N'-N| \le 1$. Hence

$$\frac{C_I(N')C_J(N')}{MN'} - \frac{C_I(N)C_J(N)}{MN} = O(1)$$

uniformly for $M \ge 1$. Thus the estimate at a good real point transfers to the integer endpoint before adding the moving-window slice error. Discard intervals meeting a grid boundary and those wholly exceptional up to measure zero. The former cost $O((\log X)^{\delta_0/4}/\sqrt{X})$ in harmonic mass, and the latter are bounded by the exceptional logarithmic measure. Telescoping over $X = \exp(2^j)$ gives the asserted global bound, after a possible reduction of the exponent.

Finally $s(V) \sim T^{-2}$. The relative square tail for both sites is $O(1/Y) + o(1)$ on $\mathcal{G}$. For $Y < p \le L^{\kappa/3}$ sum the class estimate at $M = p^2$, and above that cutoff use the unrestricted $O(VL^{-\kappa/3} + \sqrt{V})$ estimate. Fixed square tests retain the all-scale signed moments just proved. The truncation in Theorem 5.1 now gives the Möbius probabilities. Each prime logarithm is $(1/2 + o(1))L$ uniformly, proving the last statement. □

The asymptotic mass of $E_4$ on every scale remains open in this argument. An absolute estimate for its signed moments can be much larger than its positive mass on an exceptional scale.

# A Counting almost primes and prescribing the sign of an origin

## A.1 Two factorizations at distance two

The following examples use positive weights. They give actual counts on each dyad while leaving the parity of the number of factors of the shifted integer unspecified. We write $P^-(n)$ for its least prime factor.

**Proposition A.1.** *Fix $0 < \alpha < 1$ and put $z = \exp(L^\alpha)$. For some $\eta > 0$ there are $\asymp_\alpha V/L^2$ integers $V < v \le 2V$ of the form $v = pt$, with $p, t$ prime, $z < t \le z^2$, such that $v - 2$ is squarefree, $P^-(v-2) > V^\eta$ and $\Omega(v-2) \le 2$. In the box $r/\sqrt{V} \in [1, 1.1]$, $s/\sqrt{V} \in [1.3, 1.5]$, both prime, there are $\asymp V/L^3$ integers $v = rs$ for which $v-2$ is squarefree, $P^-(v-2) \ge V^{0.1225}$ and $\Omega(v-2) \le 3$.*

*Proof.* The unshifted first population has cardinality $Y \sim (\log 2)V/L$, by the prime number theorem for $p$ on each dyad and $\sum_{z<t\le z^2} 1/t \sim \log 2$ over primes. The two factors are distinguishable by size. At level $D = V^{0.514}$, the congruence $pt \equiv 2 \bmod d$ for odd $d$ has density $g(d) = 1/\varphi(d)$ with total absolute error $O_J(VL^{-J})$. Corollary 1.1(i) of Fouvry and Radziwiłł applies since $0.514 < 17/33$ and the shorter prime variable exceeds $\exp(L^{\alpha/2})$ and has size $V^{o(1)}$ [1]. Its prime indicator satisfies Siegel–Walfisz on intervals by subtraction. Excluding primes dividing an auxiliary modulus $k$ costs at most $\omega(k)$, absorbed by its allowed divisor factor. Endpoints of the product strip can be frozen on a mesh $1 + L^{-H}$. When summing over all $d$, an origin in the boundary strips contributes at most $\tau(v-2)$ times its bounded multiplicity. Cauchy–Schwarz with $\sum_{n\le 3V} \tau(n)^2 \ll VL^3$ gives $O(VL^{(3-H)/2})$. The averaged main-term boundary is just as small, with the harmonic modulus sum $\ll L$. The restricted-density correction with $t \mid d$ is $O(YL/z)$ on summing $d$, since $\sum_{t|d,\ d\le D} 1/\varphi(d) \ll L/(t-1)$ for each prime $t$ in this range. The larger prime exceeds $D$. Even moduli have density zero and give zero actual count.

For squarefree $d$ write $R_d$ for the error in $Yg(d)$. The preceding bound and $|R_d| \ll V/d + 1 + Y/\varphi(d)$ give $\sum_{d\le D} \tau(d)^j |R_d| \ll_{j,J} VL^{-J}$ by Cauchy–Schwarz and divisor moments. Its dimension-one hypotheses hold since $g(2) = 0$, $g(p) = 1/(p-1)$ for odd $p$, and $\sum_{w\le p<y} g(p) \log p = \log(y/w) + O(1)$.

We apply Greaves's weighted linear sieve [11, Theorem 1, pp. 73–75, and Section 5, p. 95] with sieve level $y = D$, factor bound $R = 2$, and degree $g_G = 1.95$ (denoted $g$ there). Write $U_G, V_G$ for his fixed weight parameters, to distinguish them from the size parameter $V$ here. The $R = 2$ evaluation in that source supplies an admissible profile $W_G$ with

$$\tfrac{1}{2} < U_G < 1, \qquad V_G < \tfrac{1}{4}, \qquad V_G + 2U_G \ge 1.95, \qquad \mathcal{M}(W_G) > 0.$$

The profile is chosen as in Section 5 from the bounds (1.8); in particular $W_G(1) = U_G - V_G$ and

$$0 \le W_G(t) \ll \min(1, t^2).$$

The numerical input being imported is $\delta_2 = 0.044560$ in (1.4): $1.95 < 2 - \delta_2$ leaves a strict margin. Also $2V < D^{1.95}$ for large $V$, because $0.514 \cdot 1.95 = 1.0023 > 1$. The theorem gives a

lower bound $\gg Y/L$ for the sum of $W_G(\log P^-(n)/\log D)$ over the members $n = v-2$ whose hybrid factor count $\nu_{D,U_G}(n)$ is at most two. This hybrid count counts small prime divisors once and primes at least $D^{U_G}$ with multiplicity; it always dominates the number of distinct prime factors. The divisor-weighted remainder estimate above supplies the error required by that theorem. We next remove small least primes and repeated factors.

To enforce the least-prime restriction, apply the upper-bound sieve to multiples of $p$ with no prime below $p$ dividing $v-2$. It gives $O(Yg(p)/\log(2p))$ plus $O(\sum_{d\le p^4,(d,p)=1}\tau(d)^2|R_{pd}|)$, with squarefree $d$. For $p \le V^\eta$, take $5\eta < 0.514$. The total weighted remainders are negligible by the preceding divisor estimate. The small-prime contribution to the lower bound is therefore

$$\ll (Y/L^2)\sum_{p\le V^\eta}\log(2p)/p \ll \eta Y/L + Y/L^2.$$

Choose $\eta$ small and discard repeated larger primes with the unrestricted bound $O(V^{1-\eta}+\sqrt{V})$. Since the least-prime weight is bounded, the required lower count follows. The upper-bound sieve at $V^\eta$ gives the matching order.

For the second box the mass is $Y \asymp V/L^2$. At level $D = V^{0.49}$ both primes exceed every modulus. To get the same divisor-weighted remainder estimate, use character orthogonality in the two prime variables. At primitive conductor $F < f \le 2F$ the contribution is $\ll L\sqrt{V}(\sqrt{V}+F^2)/F$ by Cauchy–Schwarz and the multiplicative large sieve. Each primitive character enters with total weight $\ll L/\varphi(f)$. Here the residue affects only a unit phase, a fact we shall use below. For $f \le L^K$ use Siegel–Walfisz on one prime sum. For the remaining dyads the bound sums to $L^{O(1)}(V/L^K+\sqrt{V}D)$. Choose $K$ after the desired precision.

For the square condition in [12, Lemma 11.1, pp. 37–38 of the linked text], put $A_{p^2} = \#\{v \text{ in the box} : p^2 \mid v-2\}$. The elementary bound

$$\sum_{p\ge D^{1/4}} A_{p^2} \ll VD^{-1/4}+\sqrt{V} = O(V^{1-0.1225}+\sqrt{V}) = o(Y/L^3)$$

is stronger than the square-sum hypothesis there, since $Y \asymp V/L^2$. The same bound permits removal of all remaining squareful values after the almost-prime count. Lemma 11.1 in *Infinitude of palindromic almost-prime numbers* [12] applies because $2V < D^g$ for some $g$ strictly between $1/0.49$ and $3+\log(3(1+3^{-3})/4)/\log 3$. It retains least prime at least $D^{1/4} = V^{0.1225}$ and at most three prime factors. The upper sieve gives the upper count. □

## A.2 Two signs of the origin and three neighbouring valuations

We record an application where the sign itself is specified by prime factorization. Put $M = \sqrt{V}$, $I = [M, 1.1M]$, $J = [1.3M, 1.5M]$. In both populations $n = rs$, $r \in I$ and $s \in J$ is prime. In $\mathcal{N}_+$ take $r$ prime, and in $\mathcal{N}_-$ require $r = qt$ with $q, t$ prime and $V^a < q \le V^b$, where $a = 1/(3e+2)$, $b = 1/5$. These are distinct integers in each population since the bands distinguish the factors. Their respective signs are $\lambda(n) = +1, -1$. On the class $n \equiv 1 \bmod 24$ define

$$U_1 = (n+1)/2, \quad U_2 = (3n+1)/4, \quad U_3 = 3n+2.$$

For $n = 24k+1$ these forms are $U_1 = 12k+1$, $U_2 = 18k+1$, and $U_3 = 72k+5$. The relations $3U_1 - 2U_2 = 1$, $U_3 - 6U_1 = -1$, and $U_3 - 4U_2 = 1$ show that the three integers are pairwise coprime; each is also coprime to 6. At a prime $p \ge 5$ at most one site has positive valuation.

**Proposition A.2.** *In either population there are $\asymp V/L^5$ origins $n \equiv 1 \bmod 24$ for which $U_1, U_2, U_3$ are squarefree and have least prime $> V^{0.075}$. Each then has at most thirteen prime factors. For uniform $n$ in either population restricted only by $n \equiv 1 \bmod 24$, observe the assigned site and its exact valuation for each $5 \le p \le z$, writing 0 for absence. For*

$z = \exp(L^{\alpha})$, $0 < \alpha < 1$, *their joint law has total variation distance* $O_{J_0,\alpha}(L^{-J_0})$ *for each* $J_0 > 0$ *from independent coordinates with*

$$\mathbb{P}(Y_p = 0) = 1 - 3/(p-1), \qquad \mathbb{P}(Y_p = (i,j)) = p^{-j} \quad (1 \le i \le 3,\ j \ge 1).$$

*More generally for* $5 \le z \le V$ *the distance to that model tends to zero exactly when* $\log z = o(\log V)$.

*Proof.* Write $W_{\pm} = |\mathcal{N}_{\pm}|$. The prime number theorem gives $W_+ \sim .08V/L^2$. The ratio $W_-/W_+$ tends to $\frac{1}{2}\int_a^b [u(1/2-u)]^{-1}\,\mathrm{d}u = 1$, by counting $t$ for each $q$. The endpoints give $\log((b/(1/2-b))/(a/(1/2-a))) = 1$. For either population let $E(m)$ be the largest discrepancy of a unit class modulo $m$ from $W_{\pm}/\varphi(m)$. For all fixed $B, j$,

$$\sum_{m \le 24V^{.499}} \tau(m)^j E(m) \ll_{B,j} VL^{-B}. \tag{7}$$

To see this, repeat the large sieve calculation of Proposition A.1 at $Q = 24V^{.499}$. For modulus $m = fk$ an imprimitive character requires $(r,k) = 1$ on the $r$-sequence. Keep that restriction on its coefficients while summing primitive conductors for each fixed $k$. The $s$-sequence needs no restriction, as $s > Q$. With coefficient $1/\varphi(k)$ outside, the same bound $\ll M(M+F^2)/F$ at each conductor dyad follows from the large sieve with both coefficient norms $O(\sqrt{M})$. Sum $1/\varphi(k) \ll L$ in total. For small conductors Siegel–Walfisz applies on $s$. The principal correction is $\ll W_{\pm}LV^{-.09}$ in total by summing $1/\varphi(m)$ over multiples of any prime dividing $n$, since each origin has at most three factors, all $> V^{.09}$. The unrestricted dyadic remainder is $\ll L^{O(1)}(V/L^K + MQ)$ and divisor weights are absorbed as above.

For $p \ge 5$ the three roots of the neighbour product are the distinct units $-1, -1/3, -2/3$ mod $p$. In the fixed class modulo 24 the density for squarefree $d$ is $\frac{1}{8}\prod_{p \mid d} 3/(p-1)$. Thus the sieve has dimension three with level $V^{.499}$ by (7), also when remainders carry $3^{\omega(d)}$. The dimension-three lower sifting limit is at most 6.6409 [13, Section 3, Table 1], and $.499/.075 > 6.6409$. The lower and upper sieves give order $(W_{\pm}/8)\prod_{5 \le p \le z}(1 - 3/(p-1))$ at $z = V^{.075}$. Squares beyond $z$ cost $O(V/z + \sqrt{V})$ by summing the three progressions modulo $p^2$. Finally $U_i < 7V$ and $14 \cdot .075 > 1$.

For the joint law we use the Kubilius-model comparison for sequences with a positive level of distribution, as developed by Gorodetsky [15]. Ford's comparison for shifted primes [25] is a direct antecedent for local valuation laws. We retain the assignment to the three sites and verify the comparison in each population of prescribed origin sign. Given an observation with product $m = \prod_{Y_p = (i,j)} p^j \le V^{.05}$, expand exact powers at the assigned sites over $e \mid \mathrm{rad}(m)$ by inclusion–exclusion. Before this expansion multiply the positive exact-power indicator by the upper and lower fundamental-lemma weights at level $D' = V^{.39}$ excluding all three roots at the remaining primes $5 \le p \le z$. For each $e$ and each sieve divisor $d$ the term involves $3^{\omega(d)}$ unit classes modulo $24med$. The dimension bound is uniform after removing primes dividing $m$. On the assigned part the density factor for $p^j$ is $1/\varphi(p^j) - 1/\varphi(p^{j+1}) = p^{-j}$. The sieving main term on the complementary part gives the remaining product factors with relative error $\ll \exp(-cs\log s)$ for $s = \log D'/\log z$ sufficiently large, by the fundamental lemma [14, Chapters 6 and 12]. The exact-power factor is positive and factors out of the main term before this estimate. All progression errors for all these observations are bounded by

$$\sum_{m \le V^{.05}} 3^{\omega(m)} \sum_{e \mid \mathrm{rad}(m)} \sum_{d \le D'} 3^{\omega(d)} E(24med) \ll_B VL^{-B}.$$

Indeed $med \le V^{.49}$ and all multiplicities at a fixed product are bounded by a fixed power of its divisor function. Apply (7), also at 24 for the normalisation.

Under the independent model, $\mathbb{E}(\log m)^j \ll_j (\log z)^j$ for every fixed positive integer $j$ by expanding the moment using independence, $\sum_{k \ge 1} k^j/p^k \ll_j 1/p$ and $\sum_{p \le z}(\log p)^j/p \ll_j$

$(\log z)^j$. Thus $m > V^{.05}$ has model mass $\ll_j (\log z/L)^j$. Equality of the two total masses and the preceding summed discrepancy on $m \le V^{.05}$ bound the arithmetic tail as well. This proves the asserted log-power estimate and sufficiency of $\log z = o(L)$. Conversely if $z \ge V^\theta$ on a subsequence, the independent model assigns to the first site on $(V^{\theta/2}, V^\theta]$ a sum of independent Bernoulli indicators of mean tending to $\log 2$ and maximum success probability tending to zero. Its count converges to Poisson($\log 2$) by its generating function. More than $2/\theta$ such primes have positive limiting probability and would make their product exceed $7V$, impossible for the physical first site. This proves necessity. □

At $z = \exp(\sqrt{L})$ the fundamental lemma also gives the asymptotic $(W_\pm/8)\prod_{5\le p\le z}(1 - 3/(p-1))$ for three squarefree $z$-rough neighbours. Its ratio between the two origin signs tends to one. The full signs of the neighbours still depend on their prime factors above the observed cutoff.

These three neighbours occur in the four forms of *An explicit bound below one half for a quartic Liouville correlation* [6]. On the class used here their full product of signs, including the origin, is $-\lambda(n)\lambda(U_1)\lambda(U_2)\lambda(U_3)$ by complete multiplicativity. The counts above specify the origin sign and the three local factorizations. The parity of the remaining three cofactors is the further information needed for that product.

# B Two observations above slowly growing prime cutoffs

Theorem 3.3 admits one arbitrary function of an entire rough part when the other site meets a short-interval estimate. The first site can in fact keep its whole rough part while the second discloses a count of its largest factors.

**Corollary B.1.** *For $E_d$, $d = 1, 2$, put $Z \ge V^{1/T}$ and $N_Z(m) = \sum_{p>Z} v_p(m)$. Uniformly over $|H|, |B| \le 1$ and $i = 0, 1$ one has*

$$\left|\sum_{E_d} B(q)\lambda(m)\lambda(t)^i H(\mathcal{R}_y(v), N_Z(m))\right| \ll TVL^{-c}, \qquad y = \exp(L^{0.9}).$$

*Proof.* One has $0 \le N_Z(m) \le 2T$ for large $V$. Extend $H$ in its second argument to a function on $\mathbb{Z}/M\mathbb{Z}$, $M = \lceil 3T \rceil$, setting unused values to zero. Its Fourier coefficients $H_j(\mathcal{R}_y(v))$ are each bounded by one. In mode $j$ the opposite-site phase is $e(jN_Z(v-1)/M)e(-j\mathbf{1}_{q>Z}/M)$. Absorb the second factor into $B(q)$. The first is completely multiplicative and equals one on primes $\le y$, because $Z \ge V^{1/T} > y$ eventually. Select $q$ by the two masks of Section 4, and the first-site pattern and $H_j$ are read from $\mathcal{R}_y(v)$. Theorem 3.3 then applies per mode and the sum over $M$ modes gives the result. □

For two unbounded observations there is a separate question. It is qualitative and concerns the full dyad.

**Question B.2.** Fix $b \ge 1$, let $y_1, y_2 \to \infty$ at arbitrary rates, and let $i + j = 1$ with $i, j \in \{0, 1\}$. Does

$$\sup_{G,H,\chi}\left|\frac{1}{X}\sum_{\substack{X<n\le 2X\\ b\mid n}} \lambda(n)^i\lambda(n+1)^j\chi\left(\frac{n}{n+1}\right)G(\mathcal{R}_{y_1}(n))H(\mathcal{R}_{y_2}(n+1))\right| \longrightarrow 0$$

hold for separate bounded $G, H$ and all unitary completely multiplicative $\chi$?

The single-sign restriction is necessary with this range of twists. For two signs choose $G = H = 1$ and $\chi = \lambda$. The summand becomes identically one and the mean tends to $1/b$. In Question B.2 both sites may carry a rough-part observation, so neither one need satisfy (G2). Also an absolute $o(1)$ on this full dyad gives no rate relative to a shrinking prime-window population.

### Where a signed estimate enters

The amplification in *The joint Dickman law for consecutive integers* [16] suggests an approach. Hold its auxiliary parameter $B$ fixed while $X \to \infty$, so the dilation primes $(B^{1000}, \exp(4B)]$ eventually lie below both cutoffs. Choose $B$ large enough that they are units modulo $b$. For $a' - b' = hc$, subtracting the two products yields the pair

$$u = b'(a'n + 1)/c, \qquad u + h = a'(b'n + 1)/c.$$

In the squared one-sign correlation, the paired character ratios at the two sites cancel the dilation factors, while the sign contributes $\lambda(a'b')$. For example the ratios $(a'n/(a'n+1))/(b'n/(b'n+1))$ and $(u + h)/u$ are equal, so no additional character of an anchor is left. With two signs the analogous two dilation signs multiply to one.

The channel in [16, §7.1] records the size and a residue of a fairly retained prime product. We specify its signed counterpart. Put $P_0 = B^{1000}$, $R = B/\log P_0$, and $\mathcal{P}_B = \{p : P_0 < p \le e^{4B}\}$. Let $\Omega_B$ be the set of subsets $S \subset \mathcal{P}_B$ with independent inclusion probabilities $1/p$. Given $S$, retain each of its primes independently with probability $1/2$ and call their product $b_S$. Write $\ell_B(a) = \log a/B$ and $I_0 = [1/2, 3]$. For $1 \le d \le B$ let $\mathcal{R}_d = (\mathbb{Z}/d\mathbb{Z})^\times$ carry uniform probability (the one-point group when $d = 1$). Partition $I_0$ into equal fine cells of length $\Delta \asymp B^{-11/10}$, with consistent endpoint conventions. The output measure is

$$d\mathfrak{m}_d(x, r) = \mathrm{d}x\, d\mathrm{unif}_{\mathcal{R}_d}(r); \qquad \mathfrak{m}_d(J \times \{r\}) = \frac{\Delta}{\varphi(d)} =: \mu_d.$$

For $g \in L^2(\Omega_B)$ define the step function

$$(U_d^\lambda g)(x, r) = \frac{\varphi(d)}{\Delta}\mathbb{E}\Big[g(S)\lambda(b_S)\mathbf{1}_{\ell_B(b_S)\in J_x,\ b_S\equiv r \bmod d}\Big], \qquad x \in J_x \subset I_0. \tag{8}$$

The expectation includes the random site and its fair split. The domain has the site probability measure and the range is $L^2(I_0 \times \mathcal{R}_d, \mathfrak{m}_d)$. The desired bound is on the full operator, including its residue-constant output modes.

Here is the still unproved local estimate that would give such a bound. In the independent comparison for two fair splits, the products are $CE_1, CE_2$, with $C, E_1, E_2$ independent and each including $p \in \mathcal{P}_B$ with probability $1/(4p)$. For $\theta = B^{-1/10}$ one would need, uniformly for $d \le B$, $r \in \mathcal{R}_d$, and every interval $J \subset [\theta, 3]$ of length at most $\Delta$,

$$\Big|\mathbb{E}\Big[\lambda(E_i)\mathbf{1}_{\ell_B(E_i)\in J}\mathbf{1}_{E_i\equiv r \bmod d}\Big]\Big| \ll \frac{\Delta}{\varphi(d)}R^{-1/2}\theta^{-5/4} + B^{-80}. \tag{9}$$

This must include the partial cells obtained by translation and truncation at $\theta$. The exponent $-5/4$ is suggested by the signed parameter $-1/4$ for squarefree products; an Euler product by itself does not establish this uniform local estimate.

We record precisely its operator consequence. For an output cell $A = J \times \{r\}$ put $p_A^\lambda(S) = \mathbb{E}_{\mathrm{split}}[\lambda(b_S)\mathbf{1}_{(\ell_B(b_S), b_S \bmod d)\in A} \mid S]$. The step kernel of $U_d^\lambda(U_d^\lambda)^*$ is

$$K_d^\lambda(A, A') = \frac{\mathbb{E}[p_A^\lambda(S)p_{A'}^\lambda(S)]}{\mu_d^2}.$$

The coupling in [16, Lemma 7.2] changes this kernel by $O(B^5/P_0)$ in operator norm: on a successful coupling the products and their signs agree, and on failure their absolute values are at most one. In the independent model, $\lambda(CE_1)\lambda(CE_2) = \lambda(E_1)\lambda(E_2)$. For fixed $C = c$, apply (9) to $(J - \ell_B(c)) \cap (\theta, \infty)$ and residue $rc^{-1}$, and similarly for $A'$. Conditional independence of $E_1, E_2$ then bounds the retained kernel, pointwise in absolute value, by

$$O(R^{-1}\theta^{-5/2}) + O(B^{-70}).$$

The total output measure is $|I_0|$, independent of $d$, so the same bound controls its operator norm by Schur's inequality. The part where one exclusive product has logarithm at most $\theta$ is dominated entrywise by the unsigned positive kernel. The two-sided Schur estimate of [16, Lemma 7.3] applies to this majorant. Consequently (9) would imply

$$\|U_d^\lambda\|^2 \ll R^{-1}\theta^{-5/2} + (\theta + 1/R)^{1/4} + B^{-70} + B^5/P_0, \qquad 1 \le d \le B. \tag{10}$$

For the specified $P_0$ and $\theta$ this yields $\sup_{d \le B} \|U_d^\lambda\| \ll B^{-1/80}$. This conditional implication neither proves (9) nor verifies its insertion into the full amplification. Both steps remain open here. The ordinary largest-prime law and Theorem 1.1 concern different estimates.

## Acknowledgment of AI assistance

OpenAI ChatGPT assisted with the mathematical development and English exposition, the comparison with the cited dispersion and graph proofs, and the verification programs. The accompanying finite checks cover coefficient extraction, elementary counting, operator identities and the probability calculations in Section 6. The asymptotic estimates are justified in the proofs and by the analytic inputs cited there.

## References

[1] É. Fouvry and M. Radziwiłł, *Level of distribution of unbalanced convolutions*, Ann. Sci. Éc. Norm. Supér. (4) 55 (2022), 537–568. 10.24033/asens.2501. arXiv:1811.08672.

[2] K. Matomäki, M. Radziwiłł and T. Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory 9 (2015), 2167–2196. 10.2140/ant.2015.9.2167. Corrected full text, 1 March 2022.

[3] OpenAI, *Ordinary two-point correlations of multiplicative functions*, 24 September 2026. PDF and sources. Part I, Sections 2–8, and Part II.

[4] C. Pilatte, *Improved bounds for the two-point logarithmic Chowla conjecture*, J. Amer. Math. Soc., published electronically 10 September 2026. 10.1090/jams/1084. arXiv:2310.19357.

[5] B. Pouly, *Reconstruction and sign prediction from sampled prime factors*, Zenodo, 2026. 10.5281/zenodo.24443266.

[6] B. Pouly, *An explicit bound below one half for a quartic Liouville correlation*, Zenodo, 2026. 10.5281/zenodo.24443184.

[7] L. Mirsky, *Arithmetical pattern problems relating to divisibility by $r$th powers*, Proc. London Math. Soc. (2) 50 (1948), 497–508. 10.1112/plms/s2-50.7.497.

[8] K. Fan, *Numerically explicit estimates for the distribution of rough numbers*, J. Number Theory 260 (2024), 120–150. 10.1016/j.jnt.2024.01.008.

[9] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, third edition, Graduate Studies in Mathematics 163, American Mathematical Society, 2015. 10.1090/gsm/163.

[10] T. Tao and J. Teräväinen, *Quantitative correlations and some problems on prime factors of consecutive integers*, 2026. 10.48550/arXiv.2512.01739.

[11] G. Greaves, *The weighted linear sieve and Selberg's $\lambda^2$-method*, Acta Arith. 47 (1986), 71–96. 10.4064/aa-47-1-71-96. Article text.

[12] A. Tuxanidy and D. Panario, *Infinitude of palindromic almost-prime numbers*, Int. Math. Res. Not. IMRN 2024, no. 18, 12466–12503. 10.1093/imrn/rnae174. Full text, Lemma 11.1.

[13] K. Ford, *Sieve methods lecture notes*, Spring 2023, Section 3, Table 1. Author's text.

[14] J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications 57, 2010. 10.1090/coll/057.

[15] O. Gorodetsky, *A Kubilius model for sieve-theoretic sequences*, Analysis Mathematica 52 (2026), 999–1022. 10.1007/s10476-026-00178-w. Full text.

[16] OpenAI, *The joint Dickman law for consecutive integers*, 24 September 2026. PDF and sources. Section 7, especially `channels.tex` and `kernel.tex`.

[17] B. Wang, Z. Wei, P. Yan and S. Yi, *Generalizations of the Erdős–Kac theorem and the prime number theorem*, Commun. Math. Stat. 13 (2025), 1177–1197. 10.1007/s40304-023-00354-6.

[18] B. Wang and S. Yi, *Dynamical generalizations of the two-point Chowla conjecture*, 2026. 10.48550/arXiv.2610.16394.

[19] T. Tao, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, Forum Math. Pi 4 (2016), e8. 10.1017/fmp.2016.6. Corollary 7.

[20] K. Matomäki, M. Radziwiłł and T. Tao, *Sign patterns of the Liouville and Möbius functions*, Forum Math. Sigma 4 (2016), e14. 10.1017/fms.2016.6. Full text, Theorem 1.9.

[21] A. D. Barbour, E. Kowalski and A. Nikeghbali, *Mod-discrete expansions*, Probab. Theory Related Fields 158 (2014), 859–893. 10.1007/s00440-013-0498-8. Full text, Section 8.3, Theorem 8.2, pp. 24–25.

[22] D. Charamaras and F. K. Richter, *Asymptotic independence of $\Omega(n)$ and $\Omega(n+1)$ along logarithmic averages*, 2026. 10.48550/arXiv.2412.17583.

[23] J. Andreevski, *A joint local limit theorem for prime divisor counting functions*, 2026. 10.48550/arXiv.2610.28328.

[24] H. A. Helfgott and M. Radziwiłł, *Expansion, divisibility and parity*, 2021. 10.48550/arXiv.2103.06853.

[25] K. Ford, *Poisson approximation of prime divisors of shifted primes*, Int. Math. Res. Not. IMRN 2025, no. 7, rnaf079. 10.1093/imrn/rnaf079. Author's text.

[26] M. Coons and S. R. Dahmen, *On the residue class distribution of the number of prime divisors of an integer*, Nagoya Math. J. 202 (2011), 15–22. 10.1215/00277630-1260423. arXiv:0906.1029.

[27] H.-K. Hwang, *Asymptotics of Poisson approximation to random discrete distributions: an analytic approach*, Adv. Appl. Probab. 31 (1999), 448–491. 10.1239/aap/1029955143.

[28] R. Chhaibi, F. Delbaen, P.-L. Méliot and A. Nikeghbali, *Mod-$\phi$ convergence: approximation of discrete measures and harmonic analysis on the torus*, Ann. Inst. Fourier (Grenoble) 70 (2020), 1115–1197. 10.5802/aif.3332.

[29] F. Ouimet, *On the Le Cam distance between Poisson and Gaussian experiments and the asymptotic properties of Szász estimators*, J. Math. Anal. Appl. 499 (2021), no. 1, 125033. 10.1016/j.jmaa.2021.125033.

[30] B. Pouly, *Parity information in a rough cofactor and exact reconstruction thresholds*, Zenodo, 2026. 10.5281/zenodo.24578849.

[31] B. Pouly, *Critical divisor weights and the localisation of parity defects*, Zenodo, 2026. 10.5281/zenodo.24578863.

[32] B. Pouly, *Visible prime factors under neighbouring critical divisor weights*, Zenodo, 2026. 10.5281/zenodo.24578867.