Averaged Vaughan decompositions for a quartic Liouville correlation

Brice Pouly

Research and writing with AI assistance

hexagon:2611.01200v1

Abstract

A smoothed Vaughan decomposition of the Liouville function contains a short-divisor term whose mean square stays positive at a single cutoff. We compute the mixed second moment of its actual coefficient, with the product of the two short divisors required to divide the integer. The estimate is uniform in two cutoff scales on an ordinary dyadic interval and on fixed affine images. For a cutoff derivative supported on $[1/2, 1]$, the limiting kernel vanishes when the logarithms of the cutoffs differ by a factor of at least two. With this profile fixed, the kernel determines the optimal first-order error among all probability averages on an interval of scales. On an interval of length H in $\log\log U$, the uniform average attains the leading one-site mean square κ_w/H. Its excess over the optimum has exact order H^{-2}. We compute the gain from assigning additional mass to the two endpoint cutoffs and bound the mass of each subinterval for any nearly optimal distribution. On a fixed finite grid both factors of the remaining convolution have power size. Uniform moments allow simultaneous replacement at $n, n+1, 3n+1, 3n+2$. The remaining bilinear sums retain three Liouville signs on the same product, whose correlation is open.

1 One cutoff, several cutoffs, and the four signs

Let $\lambda(n) = (-1)^{\Omega(n)}$, where prime factors are counted with multiplicity. Write

$$C(n) = \lambda(n)\lambda(n+1)\lambda(3n+1)\lambda(3n+2), \qquad \mathbb{E}_X h = \frac{1}{X}\sum_{X<n\le 2X} h(n), \qquad L = \log X.$$

Every limiting logarithmic mean c of this product satisfies $|c| \le 196567331/402653184$, as proved in *An explicit bound below one half for a quartic Liouville correlation* [14]. We study the ordinary dyadic mean. Its cancellation on every sufficiently large dyad would imply logarithmic cancellation by partial summation. Even if only $\lambda(n)$ is replaced by a convolution at $n = mr$, the other three signs are tested at $mr+1, 3mr+1, 3mr+2$. They stay on the same integer. The first question is how accurately the single sign can be replaced before those three signs have to be estimated.

At one cutoff the short-divisor error has a positive limiting mean square. Averaging cutoffs can reduce this error only through their mixed moments at the same integer. We compute those moments for the coefficient produced by the exact convolution identity, and then choose how to distribute the cutoff scales. This separates the accuracy of the replacement from the cancellation subsequently required of the bilinear sum.

Fix $0 < \zeta < 1$, $z = \exp(L^{\zeta})$, $W = X^{1/100}$, and $0 < \theta_0 < 1/4$. We use the least and largest prime factors P^-, P^+, with $P^-(1) = \infty$ and $P^+(1) = 1$. Define

$$R(n) = \mathbf{1}_{P^-(n)>z}, \qquad \nu(n) = \mu(n)R(n), \qquad f(n) = \lambda(n)\mathbf{1}_{P^+(n)\le z}\mathbf{1}_{n\le W}.$$

Choose a nonnegative $w \in C_c^\infty(\mathbb{R})$ supported on $[1/2, 1]$ with $\int w = 1$. On $u \ge 0$ let $\chi(u) = \int_u^\infty w(v)\,dv$ and take a smooth compactly supported extension to the negative axis. For

$z \le U \le X^{\theta_0}$ set

$$a_U(d) = \nu(d)\chi\Big(\frac{\log d}{\log U}\Big), \qquad b_U = \nu - a_U, \qquad \mathcal{B}_U = b_U * R * f, \qquad V_U = b_U * \mathcal{B}_U.$$

Here $*$ is Dirichlet convolution. We call V_U the type II term because on its support both factors in mr exceed $U^{1/2}$. The type I term and its smooth-part sign are identified exactly in Section 2.

For a fixed affine form $F(n) = an + b$ with positive integer slope and integer constant, $\|h(F)\|_{p,X} = (\mathbb{E}_X|h(F(n))|^p)^{1/p}$. The statements below include every such form, primitive or otherwise. Implicit constants may depend on the fixed forms and profiles. There is no uniformity in their coefficients.

Theorem 1.1 (Separated cutoffs)**.** *Take $J \ge 1$ fixed, $\theta_j = \theta_0 2^{-j}$, $U_j = X^{\theta_j}$ for $0 \le j < J$, and $\overline{V}_J = J^{-1}\sum_j V_{U_j}$. At each fixed affine form F,*

$$\lim_{X\to\infty} \|\lambda(F) - \overline{V}_J(F)\|_{2,X}^2 = K_w/J, \quad K_w = \int_{\mathbb{R}_+^4} \prod_{i=1}^{2} w(x_{i1}+x_{i2}) \prod_{j=1}^{2} w(x_{1j}+x_{2j})\, d\mathbf{x} > 0.$$

For any fixed complex weights c_j of sum one the corresponding limit is $K_w\sum_j |c_j|^2$. The same statements hold for any fixed cutoffs X^{θ_j}, $0 < \theta_j \le \theta_0$, with logarithms pairwise separated by a factor of at least two.

For two scales θ and $\theta/2$, the two row sums of the matrix (x_{ij}) lie in $[\theta/2, \theta]$ and the two column sums lie in $[\theta/4, \theta/2]$. The total of all four entries would have to be θ, a set of four-dimensional measure zero. This simple constraint accounts for the endpoint factor two and for the exact cost of assigning unequal weights to separated cutoffs.

Overlapping scales yield a different constant. Let Y_1, Y_2 be independent real random variables with density w, and set

$$\kappa_w = \int_0^{1/2} \Pr\Big(\frac{\min(Y_1,Y_2)}{Y_1+Y_2} \ge u\Big)^2\, du, \qquad \frac{1}{3} \le \kappa_w \le \frac{1}{2}. \tag{1}$$

Theorem 1.2 (Optimal averaging over an interval of scales)**.** *Fix $\zeta < \zeta' < 1$. Put $A_- = L^{\zeta'}$, $A_+ = \theta_0 L$ and $H = \log(A_+/A_-)$, for X large. For a probability measure σ on $[0, H]$ write*

$$U(s) = \exp(A_- e^s), \qquad V_\sigma = \int V_{U(s)}\, d\sigma(s).$$

Let σ_0 be uniform Lebesgue probability on $[0, H]$. At each fixed F,

$$\|\lambda(F) - V_{\sigma_0}(F)\|_{2,X}^2 = \kappa_w/H + O(H^{-2}), \tag{2}$$

$$\inf_\sigma \|\lambda(F) - V_\sigma(F)\|_{2,X}^2 = \kappa_w/H + O(H^{-2}). \tag{3}$$

The infimum includes atomic measures and measures depending on X. The estimates keep w fixed. Proposition 4.2 gives the uniform average to second order.

The endpoints explain why the uniform law can be improved. Each endpoint interacts with half the kernel seen by an interior cutoff. Giving the two endpoints a total mass of order $1/H$ lowers the mean square by $\kappa_w^2/(2K_wH^2) + O(H^{-3})$. Proposition 4.4 computes the best weights in this family. Proposition 4.5 also shows that every distribution attaining the leading constant becomes uniform on rescaling the interval to $[0, 1]$.

Corollary 1.3 (Four simultaneous replacements)**.** *Put $(F_1, F_2, F_3, F_4) = (n, n+1, 3n+1, 3n+2)$. For the fixed grid,*

$$\limsup_{X\to\infty} \mathbb{E}_X \Big| C(n) - \prod_{i=1}^{4} \overline{V}_J(F_i(n)) \Big| \ll J^{-1/2}.$$

The implicit constant is independent of J. Every factor in every bilinear term exceeds $X^{\theta_{J-1}/2}$. For the uniform average of Theorem 1.2, the four-site L^1 error is $O((\log\log X)^{-1/2})$ and every bilinear factor exceeds $\exp(L^{\zeta'}/2)$.

Only the fixed-grid statement provides a fixed positive exponent for the length of the factors. The second statement and its one-site optimality concern the accuracy of the approximation. They give no decay estimate for the quartic remainder itself.

Vaughan's identity separates short and long convolution factors [22]. Weighted and smoothed forms are developed by Sedunova [18, Lemma 1.5] and Srivastav [20, Lemma 2.1]. The latter assembles sieve coefficients by least common multiples. Lichtman's rough Möbius identity supplies the algebra used here [12, Proposition 5.2]. His bilinear square estimates, imported in Section 5, are separate inputs from [12, Propositions 8.2 and 9.1 and the proof of Proposition 7.1].

The moment calculation uses prime sums, Euler products and Fourier inversion. These methods occur in the divisor correlations of Goldston and Yıldırım [6], including their general k-correlation theorem [7, Theorem 1.1], and in the moments of smooth Möbius divisor weights studied by Granville, Koukoulopoulos and Maynard [9, Theorems 1.3 and 1.7]. Graham evaluates the mean square for the two-parameter Barban–Vehov weights [8]; Carneiro, Chirre, Helfgott and Mejía-Cordero optimize the test function in that quadratic sieve [3, Theorem 1.2 and Corollary 1.3].

Here the short coefficient has the product condition $d_1d_2 \mid n$, and the same cutoff is retained in its two factors before averaging. We determine its two-scale covariance, transfer it to the actual error $\lambda - V_U$, and optimize the distribution of scales to first order for a fixed profile w. The two rough divisors in each copy lead to the four allocation variables in Theorem 1.1. The uniform moments then yield the four-site replacement. Theorems 1.1 and 1.2 use no binary correlation estimate.

A related arithmetic question is treated in *Local Compensation of Vaughan Remainders with Prescribed Divisors* [16, Theorem 1.1 and Corollary 6.2]. There one cutoff for Λ is chosen at each X, the divisor pair is prescribed, and the other cofactor carries a genuine prime weight. Here the mean square is taken at one affine site while the cutoff varies. Section 7 explains which multipliers a local cancellation can retain.

Sections 2–4 establish the reduction. Sections 5 and 6 examine the bilinear fibres and the prescribed slope three. An independent harmonic comparison and the localization of signed divisor expansions are proved in Appendices A–B. Their supports and weights are recorded together in Section 7.

2 The short-divisor coefficient in the exact identity

Let $s_z(n)$ be the full part of n on primes $\le z$ and let $r_z(n) = n/s_z(n)$. On z-rough integers μ is the convolution inverse of the constant function one. In the full Dirichlet convolution algebra this says $\nu * R = \mathbf{1}_{\{1\}}$. Consequently

$$\nu * f = 2a_U * f - a_U * a_U * R * f + b_U * (b_U * R * f). \tag{4}$$

Indeed expanding $b_U = \nu - a_U$ on the right and using $\nu * R = \mathbf{1}_{\{1\}}$ proves the equality at every integer. This is the algebra in [12, (5.8)–(5.9)], with a smooth cutoff on each short variable and the Liouville sign in f. On any $F((X, 2X])$ under consideration, $2a_U * f$ vanishes for large X since $UW = o(X)$.

The type I term in (4) is exactly $f(s_z(n))T_U(n)$, where

$$T_U(n) = \sum_{d_1d_2\mid n} a_U(d_1)a_U(d_2). \tag{5}$$

Equivalently, $T_U(n) = (a_U * a_U * R)(r_z(n))$. The condition on the product of the divisors is essential. For a prime $p > z$, put $c = \chi(\log p/\log U)$. Then

$$T_U(p) = 1 - 2c, \qquad \left(\sum_{d|p} a_U(d)\right)^2 = (1-c)^2.$$

Thus independent divisibility conditions already change the coefficient at a prime. In contrast, imposing $(d_1, d_2) = 1$ in (5) changes it only when a square of a prime exceeding z divides n. This is the restriction used in the mixed-moment calculation below.

The factor $f(s_z(n))$ is common to all cutoffs. Averaging the identity keeps the same U in the two occurrences of a_U; it does not average those occurrences independently. On the long side,

$$\mathcal{B}_U(r) = f(s_z(r)) \sum_{d|r_z(r)} b_U(d), \qquad |\mathcal{B}_U(r)| \le \tau(r), \qquad b_U(d) = 0 \quad (d \le U^{1/2}).$$

In particular $\mathcal{B}_U(r)$ is also zero for $r \le U^{1/2}$.

Finally $(\nu * f)(n) = \mu(r_z(n))f(s_z(n))$. This agrees with $\lambda(n)$ unless $s_z(n) > W$ or $p^2 \mid n$ for some prime $p > z$. Among the integers up to CX the latter set has size $O_C(X/z)$. For the former use the nonnegative sum of $1/d$ over smooth divisors $d > W$, with Rankin's exponent $1/\log z$. Division by X gives at most

$$O_C\left(W^{-1/\log z} \prod_{p \le z} (1 - p^{-1+1/\log z})^{-1}\right) \ll_C (\log z)\exp\{-L/(100\log z)\}.$$

Mertens' product and $\sum_{p\le z}(\log p)/p \ll \log z$ give the last inequality. Counting all integers up to $C_F X$ bounds the count on a fixed affine image. Thus each exception has relative size $O_{A,F}(L^{-A})$ for every fixed A.

On the dyad the actual error is

$$E_U(F) := \lambda(F) - V_U(F) = (\lambda - \nu * f)(F) - f(s_z(F))T_U(F). \tag{6}$$

Away from the exceptional set the first term vanishes and $|f(s_z(F))| = 1$. Since $|T_U| \le \tau^2$, any fixed product moment involving these exceptions is $O_A(L^{-A})$, uniformly in U. Here and below one can use $\tau(n)^k \le \tau_{2^k}(n)$ and $\sum_{n\le x} \tau_m(n) \ll_m x(1+\log x)^{m-1}$, followed by Cauchy–Schwarz. This observation transfers the *full mixed second moment*, including its lower bound, from T_U to the genuine errors E_U.

3 The mixed second moment at two scales

Write $Z = \log z$, $A = \log U$, $B = \log V$, $a = \min(A,B)$, $b = \max(A,B)$, and $w_A(u) = A^{-1}w(u/A)$. For $A, B > 0$ define

$$K(A,B) = \int_{[0,\infty)^4} \prod_{i=1}^{2} w_A(x_{i1} + x_{i2}) \prod_{j=1}^{2} w_B(x_{1j} + x_{2j})\, d\mathbf{x}. \tag{7}$$

A common dilation of A, B leaves this integral invariant. Transposing the matrix exchanges them.

Proposition 3.1 (Uniform mixed moment)**.** *If $Z \le a \le b \le \theta_0 L$, then for any fixed $D > 0$,*

$$\mathbb{E}_X E_U(F)E_V(F) = K(A,B) + O_w\left(\frac{a^2}{b^2}\left(\frac{Z}{a} + \frac{1}{Z}\right)\right) + O_{D,F,w,\theta_0,\zeta}(L^{-D}). \tag{8}$$

The same estimate holds for $T_U(F)T_V(F)$. In particular $K(A,B) = 0$ for $b \ge 2a$, and $K(A,A) = K_w > 0$.

Proof. We start with T_U, T_V. In each copy of (5) impose $(d_1, d_2) = 1$. A discrepancy requires a rough square dividing F and has negligible mean by Section 2. For a term of the two coprime sums let $h = [d_1 d_2, e_1 e_2]$. Every prime dividing h exceeds z, hence is prime to the fixed slope of F for X large. Exactly one residue class of n mod h makes $h \mid F(n)$. Its normalized count is $1/h + O(1/X)$. There are at most U^2V^2 quadruples, so the total error is $O(X^{4\theta_0 - 1})$. The main term is precisely the expectation obtained by selecting primes $p > z$ independently with probability $1/p$ and allowing each selected prime at most one of the two positions in each copy.

For $\Re s = 1$ put $\Phi(s) = \int_{\mathbb{R}} \chi(v) e^{sv}\, dv$. Then $\chi(u) = \int_{\mathbb{R}} \Phi(1 + ih) e^{-(1+ih)u}\, dh/(2\pi)$ and $\Phi(1 + ih)$ decreases faster than every inverse power of $1 + |h|$. Use four such integrals with s_1, s_2, t_1, t_2 on the line $\Re s = 1$, and set $\alpha_i = s_i/A$, $\beta_j = t_j/B$. For a selected prime the first copy contributes $g_A(p) = 1 - \sum_i p^{-\alpha_i}$ and the second $g_B(p) = 1 - \sum_j p^{-\beta_j}$. The local expectation is $1 + H(\log p)/p$, where

$$H(v) = -\sum_i e^{-\alpha_i v} - \sum_j e^{-\beta_j v} + \sum_{ij} e^{-(\alpha_i + \beta_j)v}.$$

For justification take a finite prime range $z < p \le T$ with $T \ge \max(U, V)$. Fourier inversion gives the same compactly truncated divisor sums for every such T. For fixed contours the product converges absolutely as $T \to \infty$.

Here are bounds allowing this limit and a uniform comparison. Put $D_0 = 2 + \sum |s_i| + \sum |t_j|$, and let $g_A^0(p) = -1$ for $p \le e^A$, $+1$ otherwise. Mertens' estimates give

$$\sum_{p > z} \frac{|g_A(p) - g_A^0(p)|}{p} \ll 1 + \log D_0. \tag{9}$$

Indeed up to e^A use $|1 - e^{-s_i v/A}| \le \min(2, |s_i| v/A)$, split at $v = A/(1 + |s_i|)$, and use $\sum_{p \le y} \log p/p \ll \log(2y)$. Above e^A the sum of $p^{-1-1/A}$ is bounded by partial summation. The same bound holds for B. Since the factors g are bounded, replacing $g_A g_B$ by $g_A^0 g_B^0$ changes the sum of real parts in the logarithm by $O(1 + \log D_0)$. For X large the logarithm is its convergent series at 1, with total quadratic error $O(\sum_{p > z} p^{-2})$. The reference product has logarithmic main term $-2 \sum_{e^a < p \le e^b} 1/p$. Thus for the finite or infinite product P,

$$|P| \ll D_0^C (a/b)^2. \tag{10}$$

This majorant is integrable against the four transforms.

To compute P, use $H(0) = 0$, $|H(v)| \ll D_0 v/a$ and $\int_0^\infty |H'(v)|\, dv \ll D_0$. By Mertens' formula $\sum_{p \le e^v} 1/p = \log v + c + O(1/v)$ for $v \ge 2$ [21, Chapter I.1]. Stieltjes integration from Z to infinity gives

$$\log P = \int_0^\infty H(v) \frac{dv}{v} + O(D_0(Z/a + 1/Z)).$$

The omitted interval at zero costs $O(D_0 Z/a)$, and the boundary and derivative errors in partial summation cost $O(D_0/Z)$. Frullani's formula in the right half-plane gives the exponential of the integral as

$$P_0 = \frac{\alpha_1 \alpha_2 \beta_1 \beta_2}{\prod_{i,j} (\alpha_i + \beta_j)}, \qquad |P_0| \le D_0^4 (a/b)^2.$$

For the bound use $|\alpha_i + \beta_j| \ge 1/a$. Along the line segment joining the two logarithms, the modulus of the exponential never exceeds the larger endpoint modulus. With (10), integration against the four transforms therefore replaces P by P_0 at the error in (8).

Represent $(\alpha_i + \beta_j)^{-1}$ as $\int_0^\infty e^{-(\alpha_i + \beta_j) x_{ij}}\, dx_{ij}$. The four numerators differentiate the four cutoff transforms, giving $-\frac{d}{du}\chi(u/A) = w_A(u)$ for $u \ge 0$, and likewise for B. Absolute convergence follows before Fourier integration from the positive real parts and the rapid decrease of Φ. The resulting expression is (7). If $b \ge 2a$, the supports force incompatible total row and column sums except on a null boundary. For $A = B = 1$, choose r in an open interval

where $w > 0$ and all x_{ij} near $r/2$. This shows strict positivity. Equation (6) transfers the calculation to the actual errors. □

On a fixed separated grid the mixed second moment tends to K_w on the diagonal and to zero elsewhere. Expanding the squared norm now proves Theorem 1.1, including its assertion for complex coefficients. In particular equal weights uniquely minimize the limiting error on that grid.

4 Averaging overlapping cutoffs

The common total of the row sums and the column sums in (7) provides a coordinate in which changing the cutoff becomes translation. Put $d_0 = \log 2$ and $k(t) = K(1, e^t)$. This even, nonnegative, continuous function vanishes for $|t| \ge d_0$. For $0 \le v \le S/2$ let

$$h_A(S, v) = \int_v^{S-v} w_A(r) w_A(S - r)\, dr, \qquad P(\ell, u) = e^{\ell} h_1(e^{\ell}, e^{\ell} u) \quad (0 \le u \le 1/2).$$

Extend h_A and P by zero outside their stated domains. In the first variable P is supported on $[0, d_0]$.

Lemma 4.1. *For $s, t \in \mathbb{R}$,*

$$K(e^s, e^t) = \int_0^{1/2} \int_{\mathbb{R}} P(\ell - s, u) P(\ell - t, u)\, d\ell\, du, \qquad \int_{\mathbb{R}} k(t)\, dt = \kappa_w.$$

Proof. Use as coordinates the total S, the first row sum r, the first column sum c and $x = x_{11}$. The other entries are $r - x, c - x, S - r - c + x$ and the absolute Jacobian is one. The length of the allowed interval for x is

$$\min(r, c) - \max(0, r + c - S) = \min(r, S - r, c, S - c)$$

for $0 \le r, c \le S$. Write this minimum as the integral of the indicators that all four arguments exceed v. The r, c integrations separate, giving $K(A, B) = \iint h_A(S, v) h_B(S, v)\, dv\, dS$. Since $h_A(S, v) = A^{-1} h_1(S/A, v/A)$, the substitutions $S = e^{\ell}$, $v = Su$ give the first equality. Finally

$$\int_{\mathbb{R}} P(\ell, u)\, d\ell = \int_0^{\infty} \int_{uS}^{(1-u)S} w(r) w(S - r)\, dr\, dS = \Pr\{\min(Y_1, Y_2) \ge u(Y_1 + Y_2)\}.$$

Fubini's theorem gives the second equality. The last probability equals one for $u \le 1/3$ and is at most one throughout, giving the bounds in (1). Also $K_w > 0$ since w is strictly positive on an open interval and all four sums can lie there simultaneously. □

Proposition 4.2 (Finite-interval energy)**.** *For a probability measure σ supported in $[0, H]$, $H > 0$, define*

$$\mathcal{E}(\sigma) = \iint k(s - t)\, d\sigma(s)\, d\sigma(t).$$

Then $\mathcal{E}(\sigma) \ge \kappa_w/(H + d_0)$. For $H \ge d_0$, the uniform law σ_0 has the exact energy

$$\mathcal{E}(\sigma_0) = \frac{\kappa_w}{H} - \frac{m_w}{H^2}, \qquad m_w = \int_{-d_0}^{d_0} |t| k(t)\, dt.$$

On the interval of Theorem 1.2, $\|\lambda(F) - V_\sigma(F)\|_{2,X}^2 = \mathcal{E}(\sigma) + O_F(\epsilon_X)$ uniformly in σ, where $\epsilon_X = L^{-\min(\zeta, \zeta' - \zeta)}$.

Proof. At fixed u, convolve $P(\cdot, u)$ with σ. The result is supported on $[0, H + d_0]$ and its integral is $\int P(\ell, u)\, d\ell$. Its squared L^2 norm is at least the square of this integral divided by $H + d_0$. Integrate in u and apply Lemma 4.1. For the uniform law the difference of two samples has density $(H - |t|)_+/H^2$. Integration against k gives the stated formula. The last claim follows by integrating (8), since $Z/a + 1/Z \le 2\epsilon_X$ and the common translation $\log A_-$ disappears. This proves Theorem 1.2, as $\epsilon_X = o(H^{-j})$ for every fixed j. □

A proportion of order $1/H$ of the interval lies within one kernel width of an endpoint. The missing interactions there produce the correction $-m_w/H^2$ to the uniform energy. Section 4.2 uses the endpoint interaction to reduce the second term and compares the resulting law with the unrestricted minimum.

The same first-order optimum holds if complex measures of total mass one and total variation at most a fixed $T \ge 1$ are allowed. Indeed, in the energy one conjugates the second copy of the measure. The convolution in the proof still has the same integral, so Cauchy–Schwarz gives the identical lower bound; the integrated arithmetic error is now $O_F(T^2\epsilon_X)$. The uniform probability is available in every such class.

Two comparisons make the scale of this result more concrete. First, on the moving factor-two grid $A_j = \theta_0 L 2^{-j} \ge L^{\zeta'}$ there are $J_X = H/d_0 + O(1)$ points. The maximal row sum of the errors in (8) is $O(\epsilon_X \sum_{h\in\mathbb{Z}} 2^{-2|h|} + J_X L^{-D})$. Thus the equal-weight one-site mean square is $(K_w + o(1))/J_X$. Its leading constant per interval length is $d_0 K_w > \kappa_w$. Indeed Cauchy–Schwarz on $[0, d_0]$ gives the weak inequality, and for every $u < 1/3$ the continuous section $P(\cdot, u)$ has integral one and vanishes at its endpoints. The inequality is strict. Second, if all cutoffs must satisfy $X^{\theta_1} \le U \le X^{\theta_0}$ for a fixed $0 < \theta_1 < \theta_0$, the length is only $H = \log(\theta_0/\theta_1)$. The limiting inferior of the one-site mean square of *every* probability average, including X-dependent ones, is at least $\kappa_w/(H + d_0)$. Thus a uniform power floor for the factors has an explicit one-site cost.

As a geometric example, take the uniform density on $[1/2, 1]$ in (1). For $1/3 \le u \le 1/2$, set $r = u/(1 - u)$. Subtracting the two corner triangles from the square $[1/2, 1]^2$ gives probability $5 - 4r - 1/r$. Hence

$$\kappa = \tfrac{1}{3} + \int_{1/2}^{1} \frac{(5 - 4r - 1/r)^2}{(1 + r)^2}\, dr = 26 + 60\log 3 - 132\log 2 = 0.42130948\ldots.$$

For this example the kernel itself has the following values at rational ratios by polynomial integration.

B/A	1	5/4	3/2	7/4	2
$K(A,B)$	5/6	2189/4800	25/216	19/1568	0.

Indeed $h_A(S, v) = 4A^{-2}[S - 2\max(v, A/2, S - A)]_+$ for $A \le S \le 2A$, and is zero otherwise. In particular the factor-two grid has leading constant $5\log 2/6 = 0.57762\ldots$ in this limiting profile. Smooth densities supported in the same interval approximate κ by L^1 convergence, and bounded smooth approximants also recover the kernel values by dominated convergence. The arithmetic estimates use a smooth w.

4.1 Uniform moments and the quartic product

The one-site L^2 error can be used inside a product once the other factors have uniform higher moments. The following bound provides these moments without a further restriction on θ_0.

Lemma 4.3 (Uniform moments)**.** *For every fixed $p \ge 1$ and fixed affine form F, there is a constant $C_{p,F,w,\theta_0,\zeta}$ such that, for all sufficiently large X,*

$$\sup_{z \le U \le X^{\theta_0}} (\|T_U(F)\|_{p,X} + \|V_U(F)\|_{p,X}) \le C_{p,F,w,\theta_0,\zeta}.$$

The same bound for V_U holds for every probability average of these cutoffs. The averaging measure may depend on X but is independent of n. In particular the constant is independent of the number J of points in a finite grid and of their locations.

Proof. For the coprime divisor sum the inverse transforms in Proposition 3.1 integrate $\prod_{\ell>z,\ \ell\mid F} g_A(\ell)$, with two frequencies. Set $D_\ell = 1 + |g_A(\ell) - g_A^0(\ell)|$. Its range is $[1,5]$, and $D_\ell^p - 1 \ll_p D_\ell - 1$. Expand the nonnegative product and count multiples of its squarefree divisors $d \le C_F X$. The proportion for each is $\ll_F 1/d$, so

$$\mathbb{E}_X \prod_{\ell>z,\ \ell\mid F} D_\ell^p \ll_F \prod_{z<\ell\le C_F X} \left(1 + \frac{D_\ell^p - 1}{\ell}\right) \ll_{F,p} (2 + |s_1| + |s_2|)^{O_p(1)}$$

by (9). Minkowski's integral inequality proves the bound for the coprime sum. The exceptional sets and (6) give the bounds for T_U and V_U. Convex averaging preserves them. □

To prove Corollary 1.3, telescope the difference of the fourfold products. In each of the four terms put the one-site error in L^2 and the three remaining factors in L^6, since $1/2 + 3/6 = 1$. Lemma 4.3 bounds these three norms independently of J. Hölder's inequality therefore gives

$$\mathbb{E}_X \left| C(n) - \prod_{i=1}^{4} V_\sigma(F_i(n)) \right| \ll \sum_{i=1}^{4} \|\lambda(F_i) - V_\sigma(F_i)\|_{2,X},$$

for any of the probability averages under consideration. Theorems 1.1 and 1.2 prove the two assertions, with the original integer n common to the four sites. For the fixed grid the mean of the approximating product is exactly

$$J^{-4} \sum_{0\le j_1,\dots,j_4<J} \mathbb{E}_X \sum_{m_i r_i = F_i(n)\ (1\le i\le 4)} \prod_{i=1}^{4} b_{U_{j_i}}(m_i)\mathcal{B}_{U_{j_i}}(r_i).$$

Already at one site, a sufficient estimate for cancellation of C would be

$$\sum_{\substack{X<mr\le 2X \\ m,r>X^\delta}} \mu(m)R(m)\alpha_m\beta_r\lambda(mr+1)\lambda(3mr+1)\lambda(3mr+2) = o_{\delta,D}(X) \tag{11}$$

uniformly for $|\alpha_m| \le 1$, $|\beta_r| \le L^D$, with fixed $\delta, D > 0$. Only the actual coefficients are needed. They have $\alpha_m = 1 - \chi(\log m/\log U)$ and $\beta_r = \mathcal{B}_U(r)$. Truncating $|\beta_r| > L^D$ costs at most $2XL^{-D}\sum_{r\le 2X}\tau(r)^2/r \ll XL^{4-D}$, by counting the possible m and using $\tau(r)^2 \le \tau_4(r)$ as in [12, Lemma 5.3]. Take $D > 4$ and $\delta < \theta_{J-1}/2$, let X tend to infinity, then let J grow. Section 5 keeps the two signs at $3mr+1$ and $3mr+2$ in the corresponding fibre calculation.

4.2 Endpoint weights and the distribution of nearly optimal cutoffs

A cutoff at an endpoint has neighbours on only one side. Its interaction with the uniform law is therefore half the interaction of a cutoff well inside the interval. This gives an explicit way to improve the uniform average. We first compute the improvement, then ask how much any nearly optimal law can depart from uniformity on a long subinterval. Endpoint masses together with an interior density also occur in optimal design for correlated observations, for instance for the triangular covariance models of Dette, Pepelyshev and Zhigljavsky [4, Theorem 2.1 and (2.12), preprint version]. Here the weights are determined directly from the compactly supported arithmetic kernel k.

Proposition 4.4 (Adding mass to the endpoint cutoffs)**.** *Write $K = K_w = k(0)$, $\kappa = \kappa_w$, $m = m_w$, and suppose $H \ge d_0 = \log 2$. Among probability measures with constant density on $(0,H)$ and possible atoms at $0, H$, the unique minimizer of $\mathcal{E}$ is*

$$\sigma_* = (1-\tau_H)\sigma_0 + \frac{\tau_H}{2}(\delta_0 + \delta_H), \qquad \tau_H = \frac{A_H}{B_H}, \quad A_H = \frac{\kappa}{2H} - \frac{m}{H^2}, \quad B_H = \frac{K}{2} - \frac{m}{H^2}.$$

The gain over σ_0 is exactly A_H^2/B_H. On the growing interval of Theorem 1.2, at each fixed affine form F,

$$\|\lambda(F) - V_{\sigma_*}(F)\|_{2,X}^2 = \frac{\kappa}{H} - \frac{m + \kappa^2/(2K)}{H^2} + O_{F,w}(H^{-3}),$$

with the parameters of that theorem fixed. For the unrestricted minimum $I_H = \min_\sigma \mathcal{E}(\sigma)$ over probabilities on $[0, H]$,

$$\frac{\kappa^2}{2K} \;\le\; \liminf_{H\to\infty} H^2(\mathcal{E}(\sigma_0) - I_H) \;\le\; \limsup_{H\to\infty} H^2(\mathcal{E}(\sigma_0) - I_H) \;\le\; \kappa d_0 - m.$$

Proof. The minimum over all probabilities exists by weak compactness and continuity of k. Give the endpoints masses p, q and let $\tau = p + q$. The interaction of either endpoint with σ_0 is $\kappa/(2H)$, since k is even and supported on $[-d_0, d_0]$. The interaction between endpoints vanishes, including at $H = d_0$. Their self-interaction is $K(p^2 + q^2)$. Proposition 4.2 now gives the entire quadratic

$$\mathcal{E}((1-\tau)\sigma_0 + p\delta_0 + q\delta_H) = \mathcal{E}(\sigma_0) - 2A_H\tau + B_H\tau^2 + \tfrac{K}{2}(p-q)^2.$$

We verify $0 < A_H < B_H$ even at the shortest allowed interval. Put $g(u) = \int P(\ell, u)\, d\ell$. When $g(u) > 0$, take independent S, T with density $P(\cdot, u)/g(u)$ on $[0, d_0]$ and distribution function G_u. The identity

$$\mathbb{E}|S - T| = 2\int_0^{d_0} G_u(x)(1 - G_u(x))\, dx < d_0/2$$

follows by counting the points between S and T. Strictness follows from continuity of G_u and $G_u(0) = 0$. By Lemma 4.1, the integral of $g(u)^2\mathbb{E}|S - T|$ is m, whereas $\int g(u)^2\, du = \kappa$. Thus $2m < \kappa d_0$. Also $\kappa < d_0 K$ by the strict Cauchy–Schwarz comparison above. These inequalities give $A_H > 0$ and $B_H - A_H = (K - \kappa/H)/2 > 0$. Completing the square proves the minimization formula. Expanding A_H^2/B_H gives $\kappa^2/(2KH^2) + O_w(H^{-3})$. The uniform arithmetic error in Proposition 4.2 is $o(H^{-3})$. Finally $\kappa/(H + d_0) \le I_H \le \mathcal{E}(\sigma_*)$ gives both bounds. □

For the uniform density on $[1/2, 1]$ used in the geometric example, the improvement coefficient $3\kappa^2/5$ is approximately 0.10650101. This example evaluates the limiting kernel. The arithmetic conclusion uses a fixed smooth profile.

For $H > 2d_0$ one has $I_H < \mathcal{E}(\sigma_*)$. Suppose σ_* were a global minimizer. Moving a fraction of its mass to x would give the first-variation inequality $v(x) \ge \mathcal{E}(\sigma_*)$ for its potential $v(x) = \int k(x - s)\, d\sigma_*(s)$. Its integral against σ_* equals $\mathcal{E}(\sigma_*)$, so continuity makes it constant on the support. Here σ_* has positive density throughout the interval. On $[d_0, H - d_0]$ the potential is $(1 - \tau_H)\kappa/H$. Matching this value on $[0, d_0]$ would require

$$\tfrac{\tau_H}{2}k(x) = \frac{1 - \tau_H}{H}\int_x^{d_0} k(r)\, dr.$$

The integral on the right would solve a homogeneous first-order differential equation with zero terminal value. It would vanish identically, contrary to $k(0) > 0$. Determining the second-order coefficient of the unrestricted optimum remains open here.

Proposition 4.5 (Where nearly optimal cutoffs lie)**.** *For a probability σ on $[0, H]$, put $\Delta = \mathcal{E}(\sigma) - \kappa_w/(H + d_0)$. For each closed subinterval $J \subset [0, H]$ of length r, including $r = 0$,*

$$\left|\sigma(J) - \frac{r}{H + d_0}\right| \le \frac{d_0}{H + d_0} + \sqrt{3(r + d_0)\Delta}.$$

In particular, if $H\mathcal{E}(\sigma) \to \kappa_w$ as $H \to \infty$, the distribution function of s/H under σ tends uniformly to that of the uniform law on $[0, 1]$. If $\mathcal{E}(\sigma) \le \kappa_w/H + C/H^2$ for fixed $C \ge 0$, then for every such J

$$|H\sigma(J) - r| \ll_{w,C} 1 + \sqrt{r}.$$

Proof. Let $P_u = P(\cdot, u)$. Subtracting the mean on $[0, H + d_0]$ in the proof of Proposition 4.2 gives

$$\Delta = \int_0^{1/2} \int_0^{H+d_0} \left| (P_u * \sigma)(x) - \frac{g(u)}{H + d_0} \right|^2 dx\, du.$$

For $0 \le u \le 1/3$ the two arguments of w in h_1 both lie in $[1/2, 1]$ whenever their product contributes, so each is at least one third of their sum. Thus $P_u = P_0$ and $g(u) = 1$ throughout this range. Consequently $f = P_0 * \sigma$ is a probability density and its squared L^2 distance on $[0, H + d_0]$ from $1/(H + d_0)$ is at most 3Δ.

Write $J = [a, b]$. Starting at a point of J, adding a displacement with density P_0 lands in $[a, b + d_0]$. Conversely, landing in $[a + d_0, b]$ requires starting in J. The latter interval is taken as empty if $r < d_0$. Hence $\sigma(J)$ lies between the integrals of f over those two intervals. Their lengths are $(r - d_0)_+$ and $r + d_0$. Comparing each integral with its length divided by $H + d_0$ costs at most $\sqrt{3(r + d_0)\Delta}$ by Cauchy–Schwarz. This proves the displayed estimate, also for atoms. At first-order optimality $\Delta = o(1/H)$, so one may take $J = [0, tH]$ uniformly for $0 \le t \le 1$. Under the second hypothesis $\Delta \le (C + \kappa_w d_0)/H^2$. The replacement of $Hr/(H + d_0)$ by r costs at most d_0. Proposition 4.2 transfers both arguments to the arithmetic error. □

There is also a unique unrestricted minimizing probability for each $H > 0$, and it is invariant under $s \mapsto H - s$. Extending the quadratic form to signed measures, Lemma 4.1 gives $\mathcal{E}(\nu) \ge \|P_0 * \nu\|_2^2/3$ for every real signed difference ν of two probabilities. If this vanishes then $\widehat{\nu}$ vanishes near zero, since $\widehat{P_0}(0) = 1$. Compact support makes $\widehat{\nu}$ entire, so $\nu = 0$. The energy is therefore strictly convex, and reflection preserves it.

5 What remains on a bilinear fibre

Write the product variable as mn in this section. When $mn + 1 = aq$, with a a product of k selected primes, the adjacent sign is already determined by q. The other two signs become $\lambda(3aq - 2)\lambda(3aq - 1)$. A second endpoint b over the same q therefore leaves the four factors

$$\lambda(3aq - 2)\lambda(3aq - 1)\lambda(3bq - 2)\lambda(3bq - 1). \tag{12}$$

We give the precise set of functions of (m, n) that can be removed from this calculation.

5.1 The binary input and the fibre identity

Use a rectangular block $m \sim M$, $n \sim N$, where $u \sim U$ denotes $(U, 2U]$. In this section $X = MN$, $x/4 \le X \le 3x$, $M, N \ge x^\delta$ for fixed $\delta > 0$, and $L = \log x$, $z = e^{L^\zeta}$. Let $0 < \alpha < \beta < 1/2$, $\beta < \zeta < 1$ and choose $j_0 > 0$, $\beta + j_0 < \varrho < 1 - \beta - j_0$. Choose k prime intervals

$$\mathcal{P}_i = (e^{h_i}, e^{2h_i}] \cap \mathbb{P}, \qquad L^\alpha \le h_1 < \cdots < h_k \le L^\beta, \qquad h_{i+1}/h_i \ge r_0 > 19/3.$$

Let $\mathcal{A}$ consist of products of one prime from each interval. Write $\eta_Y(a) = \eta(a/Y)$ for a fixed nonnegative smooth dyadic partition of unity supported on $Y \le a \le 4Y$. The weights are

$$\alpha_m = \mu(m)R(m)m^{iv}w_0(m), \qquad |v| \le x^{C_0}, \quad \|w_0\|_\infty + \mathrm{Var}(w_0) \le C_w, \quad |\beta_n| \le L^C.$$

For a multiplier F on this rectangle set

$$(T_Y F)(q) = \sum_{\substack{a \in \mathcal{A},\ m \sim M,\ n \sim N \\ mn+1 = aq}} \eta_Y(a)\alpha_m\beta_n F(m, n), \qquad Q_Y(F) = \sum_q |(T_Y F)(q)|^2.$$

The unnormalized square $Q_Y(1)$ is the quantity S_Y in [12, (6.11)]. In passing to the major and minor arcs the normalization in that paper involves the product $\prod_i \sum_{p \in \mathcal{P}_i} 1/p \asymp_k 1$, and removing common selected primes costs $O_{D,k}(XYL^{-D})$.

We use the following external estimate in this subsection and in Proposition 5.2. Given a fixed precision D, choose k sufficiently large before x tends to infinity. Then

$$\|T_Y(m^{it}b(n))\|_2 \ll_D \sqrt{XY}L^{-D}, \qquad |t| \le x^{C_1}, \quad |b(n)| \le 1. \tag{13}$$

This is the square estimate underlying [12, Propositions 8.2 and 9.1, proof of Proposition 7.1], under the parameters just listed. One absorbs t into v and $b(n)$ into β_n and increases the fixed frequency exponent. The central Gram calculation in Proposition 3.1 does not use this input.

Put $G(m,n) = \lambda(mn+1)$ and $H(m,n) = \lambda(3mn+1)\lambda(3mn+2)$. Complete multiplicativity gives identically

$$T_Y(GF)(q) = (-1)^k \lambda(q) T_Y F(q). \tag{14}$$

It holds even when $(a,q) > 1$. Set $e_j(m) = \exp(2\pi i j \log(m/M)/\log 2)$. For $J \le L^K$ and $|b_{e,j}(n)| \le L^{C'}$, the functions

$$\sum_{e=0}^{1} \sum_{|j|\le J} G(m,n)^e e_j(m) b_{e,j}(n) \tag{15}$$

have image of norm $O_B(\sqrt{XY}L^{-B})$ by (14), (13) and the triangle inequality, provided the input precision exceeds $B + K + C' + 2$. This statement applies to balanced blocks as well. A construction extracting such functions from the actual H uses a shorter n variable.

5.2 A joint projection on the actual rough integers

Let $\pi_M(m) = \mu^2(m)R(m)/(mZ_M)$, where $Z_M = \sum_{m\sim M} \mu^2(m)R(m)/m$. This is a probability measure on squarefree rough m, independent of n.

Lemma 5.1. *Suppose $N \le X^{1/2-\varepsilon}$ with $\varepsilon > 0$ fixed. For fixed $K, A, D > 0$ and $J \le L^K$, outside $O_D(NL^{-D})$ columns n, the Gram matrix of $(G(\cdot,n)^e e_j)_{e=0,1,\ |j|\le J}$ in $L^2(\pi_M)$ differs from the identity by $O_A(L^{-A})$ in operator norm.*

Proof. For any fixed B use the truncated inclusion–exclusion for $R(m)$ through depth $r = \lceil c_B \log L \rceil$. Its pointwise error is bounded by the number of squarefree $d \mid m$ with $P^+(d) \le z$ and $\omega(d) = r+1$. Summing over $m \sim M$ costs

$$\ll M \frac{(\sum_{p\le z} 1/p)^{r+1}}{(r+1)!} + z^{r+1} \ll_B ML^{-B}.$$

The remaining rough squares cost $O(M/z)$. The same reciprocal estimate bounds the tail error of the truncated Euler product by the Bonferroni inequalities. For $|h| \le 2J$ the integral comparison on $d \mid m$ has error $O((1+J)/M)$ per d for the weight $e_h(m)/m$. Increasing c_B as needed gives

$$\sum_{m\sim M} \mu^2(m)R(m)e_h(m)/m = \prod_{p\le z}(1-1/p) \int_M^{2M} e_h(u)\, du/u + O_B(L^{-B}).$$

The integral is $\log 2$ for $h = 0$ and zero otherwise. Thus $Z_M \asymp L^{-\zeta}$. This treats both diagonal blocks since $G^2 = 1$.

For the cross block put $E_f(U;r) = \max_{y\le U,(a,r)=1} |\sum_{u\le y,\ u\equiv a\ (r)} f(u)|$. For $Q \le U^{1/2-\varepsilon'}$ and each fixed B,

$$\sum_{r\le Q} \tau(r) E_\lambda(U;r) \ll_{B,\varepsilon'} U(\log U)^{-B}. \tag{16}$$

Indeed $\lambda(u) = \sum_{d^2\mid u} \mu(u/d^2)$. Truncate $d \le (\log U)^H$. For these d, Bombieri–Vinogradov for μ applies to the reduced classes $a\bar{d}^2 \pmod r$, uniformly at the smaller scale U/d^2 [12, Lemma 5.4].

Non-coprime d, r are impossible in a reduced class. The tail costs $O(U(\log U)^{1-H} + Q\sqrt{U})$ by counting at most $U/(rd^2) + 1$ points per class and d. This proves the unweighted inequality to arbitrary precision. Cauchy–Schwarz using $E_\lambda \le U/r + 1$ and $\sum_{r\le Q} \tau(r)^2/r \ll (\log U)^4$ gives (16).

After the same rough sieve, each $d \mid m$ in a cross entry imposes $mn + 1 \equiv 1 \pmod{nd}$. Partial summation bounds its normalized contribution by $O((1 + J)E_\lambda(4X + 1; nd)/(MZ_M))$. Each modulus occurs at most $\tau(nd)$ times in the sum over n, d, and $nd \le 2Nz^r \le X^{1/2-\varepsilon/2}$. By (16) the mean over n of the largest cross entry is $O_B(L^{-B})$ to arbitrary precision. Markov's inequality and the dimension bound for the operator norm prove the lemma. □

Proposition 5.2. *Under the hypotheses of Lemma 5.1, and using* (13) *to sufficiently high fixed precision, there is a function H_* with $|H_*| \ll \sqrt{1 + J}$ such that for every column $H_*(\cdot, n)$ is orthogonal in $L^2(\pi_M)$ to all the functions e_j and Ge_j, $|j| \le J$. For every admissible Y,*

$$|\sqrt{Q_Y(H)} - \sqrt{Q_Y(H_*)}| \ll_A \sqrt{XY}L^{-A}.$$

The same choice works for all the admissible coefficients with the fixed bounds above.

Proof. On each good column orthogonally project H onto the stated span. Since the least Gram eigenvalue is at least $1/2$, the projected function has a coefficient vector of squared norm at most two. Its absolute value is $O(\sqrt{1 + J})$. Subtract it to obtain H_*. On the exceptional columns $\mathcal{D}$ set the projection and H_* to zero. The projection is of the form (15). A bounded function supported on $n \in \mathcal{D}$ satisfies

$$Q_Y(F) \ll YL^{O(1)} \sum_{m\sim M,\ n\in\mathcal{D}} \tau(mn)\tau(mn + 1) \ll XYL^{O(1)}(|\mathcal{D}|/N)^{1/2}.$$

For the first step apply Cauchy–Schwarz over $a \le 4Y$ and the factorizations $mn = aq - 1$. For the second use Cauchy–Schwarz on pairs, then group by mn and use the fixed divisor moments of u and $u + 1$. Take the exceptional-column precision large enough. The triangle inequality for the images proves the assertion. The final norm still uses the fibre weights from T_Y. □

5.3 A rank-one obstruction

On a full rectangle take $H_0(m, n) = \lambda(m)\lambda(n)$, $\alpha_m = \mu(m)R(m)$ and $\beta_n = \lambda(n)$. Then $\alpha_m\beta_n H_0(m, n) = \mu^2(m)R(m) \ge 0$. Write $s_Y = \sum_{a\in\mathcal{A}} \eta_Y(a)/a$. Each admissible a is $x^{o(1)}$ and is coprime to every rough m. For each such m the number of n in the required class modulo a is $N/a + O(1) \gg N/a$. The sieve for Z_M above gives $\gg M/L^\zeta$ squarefree rough m on every such dyad, whether or not N is short. Thus $\sum_q T_Y H_0(q) \gg Xs_Y/L^\zeta$. There are $O_k(h_k)$ dyadic Y since $\log a \le 2kh_k$, while $\sum_Y s_Y = \prod_i \sum_{p\in\mathcal{P}_i} 1/p \asymp_k 1$. For some Y, $s_Y \gg_k 1/h_k$. As $q \le (4X + 1)/Y$, Cauchy–Schwarz yields

$$Q_Y(H_0) = Q_Y(GH_0) \gg_k \frac{XY}{L^{2\zeta}h_k^2} \gg_k XYL^{-2(\zeta+\beta)}.$$

The equality is (14). Even the adjacent pair $GH_0 = \lambda(mn)\lambda(mn + 1)$ can therefore have a large weighted square. This obstruction uses different sites from H. An estimate for (12) would have to use their actual affine positions or other arithmetic structure in those signs.

6 The prescribed slope three

The same quartic has another representation, $C(n) = D(n)D(3n + 1)$ with $D(n) = \lambda(n)\lambda(n + 1)$. It raises a precise localization question. Put $A(t) = \mathbb{E}_X D(n)n^{it}$ and let $B(y) = D(\lfloor y\rfloor + 1)$ on

$[3\lfloor X\rfloor, 7\lceil X\rceil)$, zero elsewhere. Write $b(u) = B(e^u)$ and take a smooth probability density ϕ supported in $(0, 1/20)$. With $\widehat{f}(t) = \int f(u)e^{-itu}\,du$,

$$\mathbb{E}_X C = \frac{1}{2\pi}\int_{\mathbb{R}} A(t)\widehat{b}(t)3^{it}\widehat{\phi}(-t/X)\,dt. \tag{17}$$

For v in the support of $X\phi(Xv)$, $\lfloor 3e^v n\rfloor = 3n$ throughout the dyad for large X. The smooth average of $b(\log(3n) + v)$ therefore equals $D(3n + 1)$ exactly. Fourier inversion of this smooth average gives (17); it retains both coefficients and their relative phase.

The binary input used in this section is

$$\sum_{m\le y}\lambda(rm + 1)\lambda(sm + 1) = o_{r,s}(y) \qquad (r, s \in \mathbb{N},\ r \ne s \text{ fixed}). \tag{18}$$

This is the fixed-affine Liouville case of [13, Theorem 1.1]: the two slopes are positive and the determinant is $r - s \ne 0$. Only finitely many such pairs are used at a time below. No uniformity for slopes growing with y is required.

Proposition 6.1. *Using* (18)*, one has*

$$\eta_X = \sup_{\chi}|\mathbb{E}_X D(n)\chi(n)| = \sup_{t\in\mathbb{R}}|A(t)| = o(1), \quad \eta_X \ge \frac{\pi(2X) - \pi(X)}{X},$$

where χ ranges over completely multiplicative functions of modulus one. For fixed $0 < a < b$ and measurable B_X supported on $[aX, 2bX]$,

$$\int_a^b |\mathbb{E}_X D(n)B_X(sn)|^2\frac{ds}{s} \le \eta_X^2\int_{aX}^{2bX}|B_X(y)|^2\frac{dy}{y}.$$

Proof. For distinct fixed primes p, q, use (18) with $(r, s) = (p, q)$. We recall Kátai's argument [11] to justify the uniform supremum; a finite quantitative formulation is given by Bourgain, Sarnak and Ziegler [2, Theorem 2]. For a fixed finite set of primes $\mathcal{P}$, put $h = \sum_{p\in\mathcal{P}} 1/p$ and $d(n) = \sum_{p\in\mathcal{P},p|n} 1$. By counting multiples, the mean square of $d - h$ up to y is at most $h + o_{\mathcal{P}}(1)$. In $h\sum_{n\le y} g(n)\lambda(n + 1)$, $g = \lambda\chi$, replace h by $d(n)$ with error at most $y(\sqrt{h} + o_{\mathcal{P}}(1))$. The resulting expression is

$$\sum_m g(m)\sum_{p\in\mathcal{P},\ pm\le y} g(p)\lambda(pm + 1).$$

Cauchy–Schwarz in m bounds its square by $y^2(h + o_{\mathcal{P}}(1))$. The diagonals total at most yh before the outer factor y, while mixed terms use (18) up to $y/\max(p, q)$. These bounds are independent of χ. Divide by hy, let y grow with $\mathcal{P}$ fixed, then enlarge $\mathcal{P}$ so that h tends to infinity. Prefix differences give the dyadic claim.

The finite prime torus $(\chi(p))_{p\le 2X}$ is the closure of $(p^{it})_{p\le 2X}$ because the logarithms of the primes are rationally independent. This gives equality of the suprema. Prescribe $\chi(p) = D(p)$ for $p \in (X, 2X]$ and choose the other prime phases independently with Haar law. Every composite on the dyad has mean contribution zero and every prime contributes one. This proves the lower bound, asymptotic to $1/\log X$. Finally Plancherel on the full $\log s$ line bounds the convolution operator with discrete coefficients $D(n)/X$ by its Fourier multiplier norm η_X. Restrict the left side to $s \in [a, b]$. □

On $3 \le s < 3 + 1/(2X)$ the floors give the fixed slope exactly. The length of this interval in $\log s$ is of order X^{-1}. Restricting the mean-square bound there costs a factor of order $\sqrt{X}$ in the bound for the absolute bias, already larger than the saving allowed by the prime lower bound. There is also an arithmetic example illustrating the loss.

Corollary 6.2. *There exists a sign sequence D_* which is the product of two consecutive signs, satisfies $\sup_\chi|\mathbb{E}_X D_*(n)\chi(n)| = o(1)$, and has $D_*(n)D_*(3n + 1) = -1$ for every n.*

Proof. Take $D_*(n) = \lambda(2n+1)$. The identity $2(3n+1)+1 = 3(2n+1)$ and $\lambda(3) = -1$ give the prescribed product. The same prime-dilation argument as above, now with $g = \chi$ and $(r,s) = (2p, 2q)$ in (18), proves the full multiplicative supremum. Finally define $a_1 = 1$ and $a_{n+1} = a_n D_*(n)$. Then $a_n \in \{-1, 1\}$ and $D_*(n) = a_n a_{n+1}$. No multiplicativity is asserted for the sequence (a_n). □

The exact identity for D_* shows that orthogonality to all moving multiplicative phases can coexist with perfect correlation at the prescribed slope. It concerns a constructed sign sequence and does not settle the correlation of the true D. For that sequence, the arithmetic content of (17) lies in the interaction of A and $\widehat{b}$ at the resolution of slope three.

7 Which estimates apply to the quartic remainder?

The same four forms support several probability laws and signed divisor sums. For example, opening $\lambda = 1 * g$ at every site preserves the quartic at each integer, but restricting its four divisors is an operation on that expansion. It need not implement a restriction on the m_i, r_i of the Vaughan product. The following table specifies the domain of each conclusion. The proofs concerning the harmonic and four-divisor measures are in the appendices.

Estimate	Domain and weights	Conclusion
Cutoff average (Thms. 1.1, 1.2)	Ordinary dyad at a fixed affine site, with its actual type I coefficient	Exact mixed limit and optimal first-order one-site mean square.
Two Mellin families (Prop. 5.2)	Square over $mn + 1 = aq$. Short n for the joint orthogonal projection in $L^2(\pi_M)$.	Under the binary square estimate, removal to fixed logarithmic precision. The remaining square still sees (12).
Prescribed slope (Prop. 6.1)	Multiplicative phases on the ordinary dyad, then L^2 in $\log s$	Under the fixed-form binary theorem, decay on a fixed interval of slopes.
Common cofactor (Prop. A.1)	Harmonic origin md; separate prime and semiprime catalogues at the same m	Variance bound for the difference of the two conditional means.
Band occupations (§B.1)	Four simultaneous divisors with $\prod_i g(d_i)$ and a smooth test of their sizes	Positive-site occupations cancel in the signed measure after summing at each origin.
Short cofactor (Prop. B.2)	One short cofactor, a second divisor above $\sqrt{X}$, and two small divisors as in (20)	Savings of every fixed logarithmic power for the signed sum.
Two anchored primes (Prop. B.4)	Free exterior divisor at site two, band divisors at sites three and four; weight $\ell_1 \Phi(\ell_2)$	Signed cancellation on the specified tuples. A free table on two divisors can recover the origin (§B.4).
Thin resonances (§B.5)	Balanced divisors and cofactors at a pair of sites, $2\gamma + \rho < 1$	Small absolute mass even for divisor-bounded weights.

At shift two, the cells of *Local Compensation of Vaughan Remainders with Prescribed Divisors* [16, Theorem 1.1 and end of Section 6] consist of the points $am + 2 = bk$ for a fixed pair (a, b). Its signed estimate sums within each cell before taking absolute values over the pairs. Thus a multiplier depending only on (a, b) can be pulled outside the inner sum. An absolute bound on discarded terms also tolerates any pointwise multiplier of modulus at most one. A signed identity for the remaining terms, even when its representative has total variation $O(X)$, controls the cell sum with its prescribed weight. Two points with weights $+1, -1$ already have cell sum zero and give weighted sum two against point values $+1, -1$.

In (11) all three remaining signs vary with the product mr inside the sum. The cell estimate supplies no bound for their joint weighting. Cancellation for the actual coefficients on each fixed power-length grid would suffice for the quartic. Theorems 1.1–1.2 and Proposition 4.4 quantify the preceding one-site approximation.

A A common cofactor for prime and semiprime multipliers

To compare the two signs of a multiplier, keep its cofactor fixed. Here $I = (X, 2X]$, $H_X = \sum_{n\in I} 1/n$, and $c_X = H_X^{-1}\sum_{n\in I} C(n)/n$. Write $Q(n) = \lambda(n)C(n)$. The argument below works for any complex test C with $|C| \le 1$, and so for the actual quartic.

Selberg's symmetry formula already puts primes and products of two primes in a single identity [19]. Its exact harmonic version is particularly simple. Write Λ for the von Mangoldt function and $\Lambda_2 = \mu * (\log)^2$, where the square is pointwise. Then

$$\Lambda_2(d) = \Lambda(d)\log d + (\Lambda * \Lambda)(d) \ge 0, \qquad \sum_{d|n} \Lambda_2(d) = (\log n)^2.$$

For the first identity, convolve both sides with 1, use $1 * \Lambda = \log$, and the derivation rule for multiplication by $\log n$ under convolution. For $X > 1$ the probability on (m, d) given by

$$\mathbf{1}_{md\in I}\,\frac{\Lambda_2(d)}{H_X md\,(\log(md))^2}$$

has exactly the harmonic law on $n = md$. Thus c_X is exactly its expectation of $\lambda(m)\lambda(d)Q(md)$. Prime powers are included in this law. The following separately normalized prime and squarefree semiprime laws give an explicit common marginal on m.

Choose $m \le X/3$ with probability $\nu_X(m)$ proportional to $\log(X/m)/m$. Given m, put $Y = X/m$ and sample $d \in (Y, 2Y]$ from one of two catalogues, using the respective weights

$$w_1(p) = \frac{(\log p)^2}{p}, \qquad w_2(pq) = \frac{2\log p\log q}{pq} \quad (p < q \text{ prime}).$$

Each catalogue has positive mass for $Y \ge 3$. For the second one may use $2p$ with $Y/2 < p \le Y$, $p \ge 3$. Write $B_j(m) = \mathbb{E}_j Q(md)$.

Proposition A.1. *Uniformly for $|C| \le 1$, one has*

$$c_X = \tfrac{1}{2}\mathbb{E}_{\nu_X}\lambda(m)(B_2(m) - B_1(m)) + O(L^{-1}), \qquad |c_X|^2 \le \tfrac{1}{4}\mathbb{E}_{\nu_X}|B_2 - B_1|^2 + O(L^{-1}).$$

Proof. The normalizing mass of m is $L^2/2 + O(L)$ by integral comparison. Write $Z_j(Y) = \sum_d w_j(d)$ on the indicated interval. Both masses satisfy $Z_j(Y) = (\log 2)\log Y + O(1)$ uniformly for $Y \ge 3$. Indeed the prime number theorem with its classical remainder gives $\sum_{t<p\le 2t}\log p/p = \log 2 + O((\log(2t))^{-2})$ for $t \ge 2$, which proves the statement for Z_1 by $\log p = \log Y + O(1)$ on the dyad. For Z_2 first take the smaller prime $p \le \sqrt{Y}$. The inner sum has the preceding estimate with $t = Y/p \ge \sqrt{Y}$, and $2\sum_{p\le\sqrt{Y}}\log p/p = \log Y + O(1)$. Pairs with both primes above $\sqrt{Y}$ contribute $O(1)$ since they lie in $[\sqrt{Y}, 2\sqrt{Y}]$. The bounded initial values are absorbed in the error.

It follows that

$$\log Y\,\mathbb{E}_j Q(md) = (\log 2)^{-1}\sum_d w_j(d)Q(md) + O(1).$$

After weighting by $1/m$ the error totals $O(L)$ before normalization. Extending the raw catalogue sums to $m > X/3$ costs $O(1)$ since then $d < 6$. Now $\lambda(m)Q(md)$ is $-C(md)$ for primes, $+C(md)$ for semiprimes. The two contributions add to a harmonic sum on n weighted by

$$\sum_{p|n}(\log p)^2 + 2\sum_{p<q,\ pq|n}\log p\log q = (\log\operatorname{rad} n)^2.$$

The dyadic harmonic mean of $\log(n/\operatorname{rad} n)$ is $O(1)$ by counting multiples of p^j for $j \ge 2$ and summing $\sum_{p,j\ge 2}\log p/p^j$. The harmonic L^1 difference between $(\log\operatorname{rad} n)^2/L^2$ and one is therefore $O(1/L)$. Since $H_X = \log 2 + O(1/X)$ this proves the identity. Cauchy–Schwarz gives the inequality. □

Matched collections of primes and semiprimes are constructed by Bergelson and Richter [1, Lemma 2.2]; Richter [17, Proposition 2.2] compares primes with products of k primes for sufficiently large k. These comparisons provide a precedent for testing the effect of a multiplier while retaining its cofactor. Here the two catalogues use the displayed logarithmic weights and have exactly the same marginal on m. This is what permits the difference $B_2(m) - B_1(m)$ to be taken before averaging over the cofactor. A related cancellation of a common multiplicative factor occurs in *Arithmetic coding against unknown small-prime signs* [15, Proposition 3.1]. For the quartic application, an estimate for the difference of the two conditional means is still needed.

B Signed divisor regions on the four affine forms

Put $v = n + 1$ and use the four forms in the order $F_1 = v$, $F_2 = v - 1$, $F_3 = 3v - 1$, $F_4 = 3v - 2$. In this appendix $\mathbb{E}_X$ averages over $X < v \le 2X$ and shifting either endpoint by one costs $X^{-1+o(1)}$ in all normalized divisor sums. Write $g = \mu * \lambda$. Then $g(p^a) = 2(-1)^a$ for $a \ge 1$ and $\sum_{d|n} g(d) = \lambda(n)$. The signs and coefficients in this expansion are fixed throughout.

B.1 Band occupation and smooth divisor sizes

Partition a finite set of primes into disjoint bands. At site i and band j let S_{ij} be the Liouville sign of the full prime-power part in that band. For a divisor tuple let K be the number of bands hit by the chosen divisors at all four sites, and K_- the number of bands whose four full signs are negative. For an integer $r \ge 1$,

$$\sum_{d_i | F_i} \prod_i g(d_i)\ \Pr\{\mathrm{Bin}(K, 1/16) \ge r\} = C(v-1)\mathbf{1}_{K_- \ge r}.$$

Indeed restricting one divisor to hit a band replaces its local sum S_{ij} by $S_{ij} - 1$, a relative multiplier $1 - S_{ij}$ in the signed sum. At all four sites this is 16 for a negative band and zero otherwise. Expand the binomial tail in factorial moments, using $\sum_{s=r}^{k} (-1)^{s-r} \binom{s-1}{r-1} \binom{k}{s} x^s$; the resulting factors 16^{-s} cancel exactly. This calculation holds before taking a dyadic mean. Taking absolute values term by term generally loses the identity.

A smooth size restriction inside the divisor expansion. The exact band identity has a quantitative version when the test depends on the four divisor sizes. Put $y = \exp(L^{.49})$, $Z = \exp(L^{.75})$, and take complete disjoint bands of width three in $\log\log p$ between $.50 \log L$ and $.75 \log L$. At site i write $R_i = (F_i)_{>y}$ with full multiplicities, T_{ib} for the prime-power part in band b, and $Q_i = R_i / \prod_b T_{ib}$. For $d_i \mid R_i$ let A be the set of occupations (i, b) for which d_i hits b but $\lambda(T_{ib}) = 1$. Independently mark each band hit by all four divisors with probability $1/16$; call the resulting set M. Let G be the bands with all four $\lambda(T_{ib}) = -1$. Write $\ell_i(d) = \log d_i / L$. For each fixed $\Phi \in C_c^\infty(\mathbb{R}^4)$, the signed law

$$\nu_v = \sum_{d_i | R_i} \Phi(\ell(d)) \prod_i g(d_i)\, \mathcal{L}(M, A \mid d, v)$$

satisfies

$$\mathbb{E}_X \Big\| \nu_v - \lambda\Big(\prod_{ib} T_{ib}\Big) \sum_{e_i | Q_i} \Phi(\ell(e)) \prod_i g(e_i)\, \delta_{(G,\varnothing)} \Big\|_1 \ll_\Phi L^{-1/4}.$$

In particular all tuples with at least one such occupation have signed sum $O_\Phi(L^{-1/4})$ after any bounded test of (M, A, v); the norm is the sum of absolute values *after* the divisor sums at a fixed origin. This also permits a smooth box around $(1/2, 1/2, 1/2, 1/2)$.

For the Fourier test $e^{i\xi \cdot \ell}$ put

$$s_{ib} = \sum_{d | T_{ib}} g(d) d^{i\xi_i / L}.$$

Introduce a formal variable u_{ib} when $\lambda(T_{ib}) = 1$, setting $u_{ib} = 1$ otherwise. With a marker z_b the per-band polynomial is

$$P_b(\xi) = \prod_i (1 + u_{ib}(s_{ib} - 1)) + (z_b - 1)16^{-1} \prod_i u_{ib}(s_{ib} - 1).$$

At frequency zero it is the single monomial $(\prod_i \lambda(T_{ib}))z_b^{\mathbf{1}_{b\in G}}$ of coefficient norm one. Call an origin good if all R_i are squarefree and $\Omega(R_i) \le K \log L$ for each i, with K a large absolute constant. Writing $A_0 = \prod_i \prod_{p^a \| R_i}(1 + 2a)$ bounds the total absolute coefficient mass. For monotone local weights $f(p^a)$, $f(1) = 1$, the elementary affine multiple count gives $\mathbb{E}_X f(R_i) \ll \prod_{y<p\le CX}(1+\sum_{j\ge1}(f(p^j)-f(p^{j-1}))/p^j)$. Apply it to $(1+2a)^8$ and 2^a at p^a, then use Hölder, Markov and the rough-square bound. This gives $\mathbb{E}_X A_0^2 \ll L^{C_0}$ and $\mathbb{E}_X A_0 \mathbf{1}_{\mathrm{bad}} \ll L^{-10}$ on increasing K. On good origins $s_{ib} = \prod_{p|T_{ib}}(1-2p^{i\xi_i/L})$. Put $h = \sum_{ib} \log T_{ib}/L \ll L^{-1/4} \log L$. Telescoping each prime product around its sign at zero, and next the product of the band polynomials in coefficient ℓ^1, gives

$$\Big\| \prod_b P_b(\xi) - \prod_b P_b(0) \Big\|_1 \ll |\xi| h \exp(O(|\xi|h)) \ll |\xi| h \quad (|\xi| \le L^{1/8}).$$

Indeed each band's deviation is bounded by $O(|\xi|h_b)\exp(O(|\xi|h_b))$ with $h_b = \sum_i \log T_{ib}/L$, and $\sum h_b = h$. The expected square of h on good origins is $O((\log Z/L)^2)$ by expanding the prime sums, using $pq \le Z^2 = o(X)$ for the off-diagonal divisibility counts, and summing $(\log p)^2/p$ on the diagonal.

The complementary Fourier factor satisfies

$$G_v(\xi) = \prod_i \sum_{e|Q_i} g(e) e^{i\xi_i \log e/L}, \qquad \mathbb{E}_X \mathbf{1}_{\mathrm{good}} |G_v(\xi)|^2 \ll (2 + |\xi|)^C.$$

On squarefree complements its local modulus is $|1 - 2e^{it}| \ge 1$, so adding the omitted primes only increases the majorant. Use Hölder and the same multiple count with local weight $|1 - 2e^{i\xi_i \log p/L}|^8$ on each prime, constant at its powers. The prime sum of the excess divided by p is $O(1 + \log(2 + |\xi|))$ by $|1 - 2e^{it}|^8 - 1 \ll \min(1, t^2)$ and Mertens. Cauchy–Schwarz with the second moment of h now bounds the low-frequency error by $O(L^{-1/4}(2 + |\xi|)^{C+1})$. Beyond $L^{1/8}$ use the crude A_0 moment; rapid decay of $\widehat{\Phi}$ absorbs it. Fourier inversion proves the assertion. The argument uses the multiplicative Fourier factor of the actual complements; an arbitrary table of complementary divisors need not obey its moment bound.

B.2 A short complementary factor and a modulus above the square root

In §B.1, keep $y = e^{L^{.49}}$ and $Z = e^{L^{.75}}$. Call a divisor tuple admissible if $d_h \mid F_h(v)$ and $P^-(d_h) > y$ at each site. Let W_0 be any function of modulus at most one of the four sets of bands hit by its divisors. Let $B_0(v)$ be any function of modulus at most one of the full parts on primes $\le y$ of a fixed number of fixed primitive affine forms positive and $O(X)$ on the interval.

The required dispersion estimate allows a progression in the short variable. Its uniformity in that progression follows by retaining the corresponding coprimality restriction in the variance of Fouvry–Radziwiłł.

Lemma B.1 (Dispersion with a short progression)**.** *Fix a nonzero integer a_0 and a constant $C_* > 1$. There is an absolute $c_0 > 0$ with the following property. Let $L = \log X$ and suppose*

$$C_*^{-1} X \le MN \le C_* X, \qquad C_*^{-1} e^{L^{.90}} \le N \le C_* X^{.003}, \qquad C_*^{-1} X^{.505} \le Q \le C_* X^{.506}.$$

Let $s_0 \ge 1$ be any integer, let b_0 be any residue class modulo s_0, and let $I \subset (N, 2N]$ be an interval. For $|\alpha_m| \le \tau(m)$ supported on $m \sim M$, set $\beta_n = \mathbf{1}_{n\in I}\mathbf{1}_{n\equiv b_0\ (s_0)}$ and

$$E(q) = \sum_{mn\equiv a_0\ (q)} \alpha_m \beta_n - \frac{1}{\varphi(q)} \sum_{(mn,q)=1} \alpha_m \beta_n.$$

For all sufficiently large X,

$$\sum_{q\sim Q,\ (q,s_0a_0)=1} |E(q)| \ll_{a_0,C_*} MN\,N^{-c_0}. \tag{19}$$

The bound is uniform in s_0, b_0, I and the coefficients α_m; no coprimality between b_0 and s_0 is required.

Proof. We adapt the proof of [5, Theorem 1.1, equations (49)–(51)], with equation numbers referring to arXiv:1811.08672v1. Choose its modulus phases to vanish unless $(q, s_0a_0) = 1$. This is the same permitted restriction of phases used at the start of Section 7 of that source. In its equation (50), a nonzero factor $c_{dh_1}\overline{c_{dh_2}}$ retains $(dh_1h_2, s_0) = 1$. Consequently the Cauchy–Schwarz step leading to (51) bounds the variance term by

$$\mathcal{V} \ll M \sum_{d\ll Q} \frac{1}{d} \sum_{\substack{h\asymp Q/d\\(dh,s_0)=1}} \frac{1}{h} \sideset{}{^*}\sum_{b \bmod d} |E_b^*(\beta;h)|^2,$$

where the star denotes reduced classes and

$$E_b^*(\beta;h) = \sum_{\substack{n\in I\\ n\equiv b_0\ (s_0)\\ n\equiv b\ (d)\\ (n,h)=1}} 1 - \frac{1}{\varphi(d)} \sum_{\substack{n\in I\\ n\equiv b_0\ (s_0)\\ (n,dh)=1}} 1.$$

Open $(n, h) = 1$ by $\sum_{t|(n,h)} \mu(t)$. For reduced b, the contribution vanishes when $(t, d) > 1$. Otherwise $(dh, s_0) = 1$ implies that the conditions $t \mid n$, $n \equiv b \pmod d$ and $n \equiv b_0 \pmod{s_0}$ have one solution class modulo tds_0. Its count is $|I|/(tds_0)+O(1)$. The main term is independent of b and disappears on centering. Thus

$$|E_b^*(\beta;h)| \ll \tau(h), \qquad \sideset{}{^*}\sum_b |E_b^*(\beta;h)|^2 \ll \min\{d\tau(h)^2,\ N + N^2/d\}.$$

For the second bound, centering decreases the squared norm, and the uncentered pair count has at most $O(N + N^2/d)$ pairs in the same class modulo d. Use the first bound when $d \le N$ and the second when $d > N$. The estimates $\sum_{h\asymp H} \tau(h)^2/h \ll (1 + \log(2H))^4$ and $\sum_{h\asymp H} 1/h \ll 1$ give

$$\mathcal{V} \ll MNL^C.$$

This is the only part of the dispersion argument at which distribution of the short sequence is needed.

After division by MN^2, the other terms inside the square root in [5, Theorem 1.1] are bounded by

$$L^C(Q/M + D^{-1/2}) + D^{C_1}X^{\xi}\left[(N^{36}Q^{33}/X^{17})^{1/20} + (N^{15}Q^{15}/X^8)^{1/8}\right].$$

Here $\xi > 0$ is arbitrary, and C_1 may depend on it. The common factor $D^{C_1}X^{\xi}$ multiplies both terms. In the stated ranges the two expressions in square brackets are respectively $O_{C_*}(X^{-.0097})$ and $O_{C_*}(X^{-.045625})$, since

$$36(.003) + 33(.506) - 17 = -.194, \qquad 15(.003) + 15(.506) - 8 = -.365.$$

Choose ξ small relative to these margins, then put $D = N^{\rho}$ with $\rho > 0$ small enough that $10\rho < 1$ and the margins remain positive. The hypotheses $M > Q(MN)^{\xi}$ and $M > N > D^{10}$ then hold for large X. Since $\|\alpha\|_2 \ll \sqrt{M}L^{O(1)}$, the variance contributes $MN \cdot O(N^{-1/2}L^{O(1)})$ and the remaining terms give a fixed power saving in N. The lower bound on N absorbs every fixed logarithmic loss and proves (19). □

Proposition B.2. *Fix distinct sites i, j, denote the other two by r, s, and set $k = F_i(v)/d_i$, $e = d_j$, $w' = d_r d_s$. There is an absolute $\delta_0 > 0$ such that the normalized signed sum*

$$X^{-1} \sum_{X<v\le 2X} B_0(v) \sum_{\mathbf{d}\ \text{admissible}} W_0(\mathbf{d}) \prod_{h=1}^{4} g(d_h)$$

restricted to

$$e^{L^{.90}} \le k \le X^{.003}, \qquad X^{.505} < e \le X^{.506}, \qquad w' \le k^{\delta_0} \tag{20}$$

is $O_A(L^{-A})$ for every fixed $A > 0$, uniformly in both bounded functions with the forms fixed.

Proof. We apply Lemma B.1 on dyadic pieces $MN \asymp X$, $N \asymp k$, $Q \asymp e$, and then evaluate its density. First expand B_0 by coordinatewise Möbius inversion in the small parts. The coefficients at indices a_h supported on primes $\le y$ are bounded by $\prod_h 2^{\omega(a_h)}$. Truncate each $a_h \le e^{L^{.82}}$. If a full small part exceeds this threshold, it has a divisor in $[e^{L^{.82}}/y, e^{L^{.82}}]$. The proportion of such integers is smaller than every L^{-A} by Rankin with exponent $1/\log y$ and the elementary multiple count. Fixed moments of all absolute divisor weights, including the truncated expansion, grow at most polylogarithmically by Hölder on the fixed forms. This proves the same error in the weighted sum. Each retained tuple gives one class of v modulo $q_0 = [a_h]$ or none, since the forms are primitive. Both q_0 and the weighted number of choices are $e^{O(L^{.82})} = N^{o(1)}$.

Fix d_r, d_s. The prime divisors of their product w' exceed y, while those of q_0 do not. Distinct forms have fixed nonzero determinants. Thus for large X the two small divisors on an admissible tuple are coprime and prescribe one class of v modulo w'. With $F_h = A_h v + B_h$, write $u = d_i$, $s_0 = A_i w' q_0$. The remaining conditions are a fixed class for $uk \pmod{s_0}$ and

$$A_j uk \equiv a_0 = A_j B_i - A_i B_j \pmod{e}.$$

On the true support u is a unit modulo s_0 and e is coprime to $s_0 a_0$. Split u into its unit classes modulo s_0, making the constraint on k a single progression. The bound $k \ge (w')^{1/\delta_0}$ further cuts its interval. Prescribing the band occupations of u, e separates their coefficients and costs $e^{O(\log L)}$ choices. Refine the u dyad into multiplicative intervals of ratio $1 + L^{-J}$ to freeze $X < v \le 2X$. Altered tuples lie over an endpoint interval of length $O(XL^{-J} + 1)$; after summing all actual divisors their absolute mass is $O_A(XL^{-A})$ by Cauchy–Schwarz and the same divisor moments, with J large. Apply (19) with long variable $A_j u$ and short variable k. The modulus coefficient is at most $2^{\omega(e)} = e^{O(L^{.51})} = N^{o(1)}$. The weighted choices of d_r, d_s are $N^{\delta_0 + o(1)}$ and the number of long-variable classes is also $N^{\delta_0 + o(1)}$. Taking $3\delta_0 < c_0$ leaves a power saving in N after summing the pieces.

It remains to evaluate the density in (19). For fixed other variables its modulus sum is of the form

$$\sum_{\substack{e\in I_0,\ P^-(e)>y \\ (e, A_j uks_0)=1}} \frac{g(e)}{\varphi(e)} D(\mathcal{P}(e)),$$

where $I_0 \subset [X^{.505}, X^{.506}]$ is an interval, $\mathcal{P}(e)$ is its set of occupied bands and $|D| \le 1$. The cost of replacing $\varphi(e)^{-1}$ by e^{-1} is $O_A(L^{-A})$, since $\omega(e) \ll L/\log y$ and $\sum_{e\ll X} |g(e)|/e \ll L^{O(1)}$. Write $e = lt$ with $P^+(l) \le Z < P^-(t)$. Rankin at $1/\log Z$ for the positive weights $2^{\omega(l)}/l$ removes $l > e^{L^{.85}}$ at the same precision. The band test depends only on l. Primes $p > Z$ dividing the polynomially bounded integer $A_j uks_0$ may be ignored in the t sum at cost $O(L^{O(1)} \sum_{p|A_j uks_0,\ p>Z} 1/p) = O_A(L^{-A})$.

For clarity, cancellation of the remaining harmonic sum of $g(t)$ has a direct proof. At a prime, the generating series for $g(p^a)$ is $(1-x)/(1+x)$. Hence $g\mathbf{1}_{P^->Z} = g * c_Z$ where c_Z has Euler factors $(1+x)/(1-x)$ at $p \le Z$ and 1 otherwise. This positive function has harmonic mass $O((\log Z)^2)$. Between any endpoints in $[X^\eta, X^{1/\eta}]$ with fixed $0 < \eta < 1$, terms of $c_Z(d)/d$

with $d > X^{\eta/2}$ cost $O_A(L^{-A})$ by Rankin, even against the crude harmonic mass of $|g|$. For smaller d the endpoints after division are still polynomial in X. The prime number theorem with its classical remainder for μ and λ gives $\sum_{n\le t} g(n) \ll_B t(\log t)^{-B}$ by the Dirichlet hyperbola method applied to $\mu * \lambda$. For λ the same input follows from $\lambda(n) = \sum_{a^2|n} \mu(n/a^2)$. Partial summation then gives every inverse logarithmic power on these harmonic intervals. Choose η below .505 with room for $l \le e^{L^{.85}}$. This proves the required modulus sum estimate.

After bounding that sum, the remaining weights sum to $O(XL^{O(1)})$. Indeed $uk = F_i(v)$ and the two small divisors and all the small-part indices still divide their actual forms. Sum their absolute weights on the slightly enlarged v interval using divisor moments; recombine the disjoint profiles and classes before doing so. This completes the proof. □

B.3 Two anchored primes and the logarithm of a divisor

Take $\mathcal{Q} = [X^{.5053}, X^{.5057}] \cap \mathbb{P}$. For $v \in (X, 2X]$ put $P(v) = \log p/L$ if a prime $p > \sqrt{2X}$ divides v, and zero otherwise. Write $\sum^*$ for the sum over $q \in \mathcal{Q}$ and m with $v = qm + 1 \in (X, 2X]$. Both p and q, when present, are unique and occur once. Denote by S_i the full part on primes $\le y$, by $T_i = \prod_b T_{ib}$ the part in the selected bands of §B.1, and by $U_i = S_iT_i$, $Q_i = F_i/U_i$. All primes of U_i are $\le Z$.

Lemma B.3. *Let $H(v)$ of modulus at most one depend only on the full parts on primes $\le Z$ of the four forms. For arbitrary $|b|, |h| \le 1$ and each fixed real t, put $f_t(m) = \sum_{e|m} g(e)e^{it\log e/L}$. Then*

$$X^{-1}\left|\sum{}^* b(q)h(m)H(v)(P(v) - 1/2)\right| \ll L^{-1/10}, \tag{21}$$

$$X^{-1}\left|\sum{}^* b(q)f_t(m)H(v)\right| \ll_t L^{-1/10}. \tag{22}$$

Proof. We include details since replacing the weight P by the indicator of a large prime would change the first assertion. Use $\sum_{d|v} \Lambda(d) = \log v = L + O(1)$. After subtracting $LP(v)$, discard proper prime powers at total cost $O(X)$ by their harmonic sum and the uniqueness of q. In the remaining small-prime sum retain only primes

$$e^{L^{.88}} < \ell \le \sqrt{X}\, e^{-L^{.88}}.$$

Summing over all v at the omitted ends costs $O(XL^{.88})$ by Mertens' logarithmic prime sum and the multiple count. The same bound holds when multiplied by the original bounded weights.

Expand H in divisibility indices a_i by coordinatewise inversion and truncate at $Y_0 = e^{L^{.83}}$ per coordinate. As in the proof of Proposition B.2, with Z in place of y, the exception costs $O_A(XL^{-A})$, including fixed divisor moments and any additional logarithmic loss. Every remaining term imposes a single class of v mod s, $s = [a_i] \le Y_0^4$, or is empty. The total weighted number of choices is $e^{O(L^{.83})}$, and their summed absolute coefficients at each fixed s are $\le \tau(s)^8$. Split q into its unit classes modulo s (here $q > s$); then m lies in one class modulo s. The coefficients in m, q are now separated. Divide the q range into intervals of ratio $1 + L^{-J}$ and freeze the endpoints of the m interval. The affected v have distance $O(XL^{-J} + 1)$ from a dyadic endpoint. Reassembling the original weights bounds the total boundary error to any required logarithmic accuracy, since q is unique, both the prime weight and its density are $O(L)$ at each origin, and the expanded local coefficients have fixed divisor moments.

For a rectangular piece $q \asymp Q$, $m \asymp M$, $MQ \asymp X$, and $D < \ell \le 2D$ in the retained range, replace $qm \equiv -1 \pmod{\ell}$ by $\mathbf{1}_{\ell\nmid qm}/(\ell - 1)$. The sum of absolute discrepancies over the prime moduli, before the factor $\log \ell$, is

$$\ll D^{-1}\sqrt{(4D^2 + O(Q))Q}\sqrt{(4D^2 + O(M))M}.$$

Orthogonality of characters followed by the multiplicative large sieve for primitive characters gives the estimate [10, Chapter 7]; all nonprincipal characters here have prime conductor.

Arbitrary bounded separated coefficients and the progression restrictions are absorbed in their squared norms. After dividing by X, the last bound is $O(D^{-1} + M^{-1/2} + Q^{-1/2} + D/\sqrt{X})$. This saves $e^{-cL^{.88}}$, absorbing the $e^{O(L^{.83})}$ choices and logarithmic subdivisions. The density weight summed over retained ℓ is $L/2 + O(L^{.88})$ per pair, uniformly: the losses from primes dividing qm are $O(\log L)$ by splitting at L. Recombine the local indices before applying this bound. Inserting it in the logarithmic divisor identity gives (21).

For the second assertion use $f_t = \lambda * k_t$, where $\lambda^{-1} = \mu^2$ in the convolution algebra. Writing $u = t \log p/L$ gives

$$k_t(p^a) = 2(1 - e^{iu}) \sum_{j=0}^{a-1} (-e^{iu})^j, \qquad \sum_{d \le 7X} \frac{|k_t(d)|\tau(d)^j}{d} \ll_{t,j} 1$$

for every fixed j. The inequality follows from $|k_t(p^a)| \le 2a \min(2, |t| \log p/L)$ and the Euler product. The harmonic mass of nonsquarefree d is $O_t(1/L)$ by summing $\log p/(Lp^2)$ over primes. For squarefree $d > 1$ the harmonic mass in any interval of width Δ for $\log d/L$ in $(0, 2]$ is $O_t(\Delta + 1/L)$. Indeed write d as an ordered product of $j \ge 1$ primes with factor $1/j!$. Once $j-1$ of them are fixed the last has weighted harmonic mass $O_t(\Delta + 1/L)$ by $\sum_{p \le x} \log p/p = \log x + O(1)$. The other factors have total mass $O_t(1)^{j-1}$.

Fix q and write $M = X/q$. Terms of $d \mid m$ with $d > Me^{-L^{.90}}$ in this convolution cost $O_t(ML^{-1/10})$ before expanding H, by the preceding interval estimate and the bound $O(M/d)$ for multiples of $d \le 2M$. For the others set $m = dn$, $N_d = M/d \ge e^{L^{.90}}$. Use the same expansion of H; its exceptional set still costs every inverse logarithmic power after summing over q, d, since $\sum_{d|m} |k_t(d)|$ is divisor-bounded and q unique. On a retained class $v \equiv a \pmod s$, $qdn + 1 \equiv a \pmod s$ gives one class modulo $r = s/(s, d)$ or none. For $Y = Y_0^4$, and each fixed j, B,

$$\sum_{r \le Y} \tau(r)^j \max_{a \bmod r,\ u \le 3N_d} \left| \sum_{n \le u,\ n \equiv a\ (r)} \lambda(n) \right| \ll_{j,B} N_d L^{-B}.$$

For reduced classes this follows from (16) with arbitrary divisor power by the same Cauchy–Schwarz argument. For unrestricted classes factor out $(a, r) = c$, use complete multiplicativity and the reduced-class estimate at $3N_d/c$; the remaining levels Y/c are below $(N_d/c)^{1/2-\varepsilon'}$ with room, since $Y = N_d^{o(1)}$ uniformly. The sum of $\tau(c)^j/c$ is polylogarithmic. For fixed d the bound after grouping $s = cr$, $c \mid d$, is therefore $\ll N_d L^{-B} \sum_{c|d} \tau(c)^8$. The harmonic moment estimate for k_t sums this over d, and $\sum_{q \in \mathcal{Q}} X/q \ll X$. This proves (22). □

Let $\mathcal{S}$ be the origins with both primes present. Let G denote the set of bands whose four full signs are negative, and put $r_0 = \lfloor N_b/64 \rfloor$, where N_b is the number of complete bands. Retain the tuples

$$d_1 = pt_1, \qquad d_2 = qet_2, \qquad d_3 = t_3, \qquad d_4 = t_4, \qquad t_i \mid T_i, \quad e \mid Q_2/q,$$

with no occupation of a band at a site where its full sign is positive. Let K count the bands hit at all four sites, $W = \Pr\{\mathrm{Bin}(K, 1/16) \ge r_0\}$ and

$$a(v, \mathbf{d}) = H_-(v) W \prod_i \lambda(S_i) g(d_i), \qquad \ell_i = \log d_i/L,$$

where $|H_-| \le 1$ is any function of the four full parts on primes $\le Z$.

Proposition B.4. *On these tuples one has for each fixed $t \in \mathbb{R}$*

$$\Sigma_t = X^{-1} \sum_{v \in \mathcal{S},\, \mathbf{d}} a(v, \mathbf{d}) \ell_1 e^{it\ell_2} = O_t(L^{-1/10}).$$

Replacing $e^{it\ell_2}$ by any fixed $\Phi(\ell_2)$ with $\Phi \in C_c^\infty(\mathbb{R})$ gives $o_\Phi(1)$.

Proof. The allowed local divisor sum on a band is $f_t(T_{ib})$ at a negative site with frequency t, and one at a positive site. For $p^a \| m$, the formula $f_t(p^a) = 1 + 2\sum_{j=1}^{a}(-e^{it\log p/L})^j$ gives $|f_t(p^a)| \le \exp(Ca|t|\log p/L)$, by the geometric series when $|t|\log p/L \le 1/2$ and the trivial bound otherwise. In the first case $|f_t(p^a) - (-1)^a| \ll a|t|\log p/L$. Hence $|f_t(m)| \ll_t 1$ on $m \le 7X$ and $|f_t(U_i) - \lambda(U_i)| \ll_t \log U_i/L$. At zero the first derivative from a prime power at site one costs at most $2\lceil a/2\rceil \log p/L$.

The marking polynomial for the good occupations, with local values s_i as above, is $\prod_i s_i + (z_b - 1)\prod_i (s_i - 1)/16$. At zero frequencies it is $(\prod_i \lambda(T_{ib}))z_b^{\mathbf{1}_{b\in G}}$. Insert frequencies at sites one and two, differentiate the first at zero and retain the second equal to t. The product rule and telescoping in the coefficient ℓ^1 norm cost $O_t(\sum_i \log U_i/L)$: each band costs $O_t(\sum_i \log T_{ib}/L)$ and their sum is bounded. Summing the coefficients of at least r_0 marks has norm at most one. The logarithmic derivative of p contributes $P(v)$, and the exterior divisor sum at site two recombines as $f_t(Q_2/q)$. With $m = U_2Q_2/q$ and $g(p)g(q) = 4$, the result is

$$\Sigma_t = 4X^{-1}\sum\nolimits^* P(v)e^{it\log q/L} f_t(m)H_0(v) + O_t(L^{-1/4}), \qquad H_0 = H_-\mathbf{1}_{|G|\ge r_0}\prod_{i\ne 2}\lambda(U_i).$$

Here the dyadic mean of $\log(F_i)_{\le Z}$ is $O(\log Z)$ by the multiple count for prime powers. Apply (21) with f_t divided by its fixed-t bound, then (22). This proves the frequency claim.

For smooth Φ the requisite domination can be seen on the good origins of §B.1. There $h = \sum_{ib}\log T_{ib}/L \ll L^{-1/4}\log L$. Up to $|t| \le L^{1/8}$ the product of the coefficient norms and its first-site derivative have size $O((1+h)e^{C|t|h}) = O(1)$. The remaining squarefree divisor sum in modulus is bounded by $\prod_{p|R_2}|1 - 2e^{it\log p/L}|$. Its mean square grows polynomially in $2 + |t|$ by the positive divisor expansion in §B.1. The bad-origin absolute mass is $O(L^{-10})$; for $|t| > L^{1/8}$ the crude divisor moments are already polynomial in $|t|$. Dominated convergence in Fourier inversion proves the smooth version. □

At $t = 0$, before insertion of ℓ_1, the restricted divisor sum at each origin is exactly $4\lambda(m)H_0(v)$. The factor ℓ_1 supplies the weight $P(v)$ for which Lemma B.3 gives the inter-origin estimate.

B.4 An arbitrary divisor coefficient can recover the origin

Keep the two anchors on $\mathcal{S}$ but allow all other rough divisors at every site. For a tuple with $p \mid d_1$, $q \mid d_2$, the product $D = d_1d_2$ determines v throughout the dyad. Indeed p, q are the only two prime factors of D exceeding $\sqrt{2X}$. If the assignments agree at another origin w, then $pq \mid v - w$. If they are reversed then $pq \mid v + w - 1$. The first forces equality and the second is impossible since $pq > X^{1.0053}$ for large X.

For instance take unrestricted band occupations and the genuine weights $a_0(v, \mathbf{d}) = \prod_i \lambda(S_i)g(d_i)$. At an anchored site the sum restricted to multiples of its prime is $g(p) = -2$ times the Liouville sign of the remaining rough part, giving twice the full sign at that site. Thus $\sum_{\mathbf{d}} a_0(v, \mathbf{d}) = 4C(v-1)$ at each anchored origin. Choose $u(D) = C(v-1)$ on represented products, zero elsewhere. Lemma B.3 with constant tests and Mertens' theorem give

$$X^{-1}\sum_{v\in\mathcal{S},\,\mathbf{d}} P(v)u(d_1d_2)a_0(v,\mathbf{d}) = 4X^{-1}\sum\nolimits^* P(v) = 2\log(.5057/.5053) + o(1) > 0.$$

For the middle equality the imposed prime p is exactly the support of P; the uniqueness of q identifies the pair count with $\sum_{q\in\mathcal{Q}}(X/q + O(1))$. With the good-occupation and binomial filters retained, the same pointwise calculation gives $4C(v-1)\mathbf{1}_{|G|\ge r_0}$.

In fact the latter filter costs $o(1)$ under the normalized bounded weight P. Here are the details. For any bounded function of the full parts at primes $\le Z$, the ordinary dyadic mean differs by $O_A(L^{-A})$ from the model given by independent Haar integers at these primes. Expand and truncate each of the four indices at $e^{L^{.83}}$ as in Lemma B.3. Every joint congruence has discrepancy at most $1/X$, over $e^{O(L^{.83})}$ weighted choices. In the model the same Rankin

truncation follows from the exact densities of divisibility. Cauchy–Schwarz for the truncated coefficients uses fixed moments of the small parts' divisor counts, bounded by powers of L by Euler products and Hölder. In the model distinct bands are independent. At one prime p in a band, the four roots are distinct and the expectation of the product of the signs on any subset of h sites is $1 - 2h/(p+1)$, by the geometric valuation law at each root. A complete band has $\sum 1/p = 3 + o(1)$ uniformly and its least prime tends to infinity. Fourier inversion on four signs gives probability $((1-e^{-6})/2)^4 + o(1) > 1/32$ that all four band signs are negative. As $N_b \to \infty$, Chebyshev's inequality gives $\Pr(|G| < r_0) = o(1)$. Transfer to the ordinary dyad and use $P(v) \le 1 + O(1/L)$ and uniqueness of q. Thus the same positive limit holds with both filters. A free multiplier on D can undo signed cancellation even while all the original g coefficients are retained.

B.5 Rational resonances in a balanced pair

Consider two distinct forms $F_i = A_i v + B_i$, $F_j = A_j v + B_j$ from $v, v-1, 3v-1, 3v-2$, where $v = n+1$. Choose divisors e_i, e_j and their complements, all in $[cX^{1/2-\eta}, CX^{1/2+\eta}]$ for fixed $c, C > 0$, $0 \le \eta < 1/2$. Let $D = e_i e_j$, $H = F_i F_j / D$. If $|bD - aH| \le X^{\rho}$ for some $1 \le a, b \le X^{\gamma}$, $\gamma < 1/2 - \eta$, $\rho \ge 0$, the number of such tuples divided by X is $O_\epsilon(X^{-1+2\gamma+\rho+\epsilon})$. This remains true with any fixed powers of divisor functions as weights. To prove it, write $m = A_i A_j$, $t = 2mv + A_i B_j + A_j B_i$, $\Delta = A_i B_j - A_j B_i \ne 0$, $u = bD - aH$ and $s = bD + aH$. Then

$$(ms)^2 - mab\,t^2 = m^2 u^2 - mab\Delta^2.$$

For a nonzero right side, the uniform binary-quadratic estimate of Vaughan–Wooley [23, Lemma 3.5] gives $O_\epsilon(X^\epsilon)$ choices of (s,t) for each (a,b,u): its coefficients, right side and variable bounds all have polynomial size in X. Here is also the direct counting argument in this setting. If mab is a square, factor the left side and use the divisor bound for the nonzero right side. Otherwise, in the real quadratic field with squarefree part of mab, the algebraic integer $ms + t\sqrt{mab}$ has the fixed nonzero norm $m^2u^2 - mab\Delta^2$. Its principal ideal divides the ideal of that integer. There are $X^{o(1)}$ such ideals, uniformly in the field, since the integer has polynomial size and the degree is two. For each ideal, any two generators differ by a unit. Both embeddings of a represented generator are bounded above in absolute value by a fixed power of X and, since their product is a nonzero integer, below by an inverse power of X. Every real quadratic unit of absolute value greater than one is uniformly bounded away from one. Thus each ideal contributes $O(\log X)$ generators in this range, up to sign. Each resulting (s,t) determines v, $D = (s+u)/(2b)$ and $H = (s-u)/(2a)$. The number of factorizations of D, H, including any fixed divisor weights, is $X^{o(1)}$.

For a zero right side, $q = \sqrt{mab}$ is rational because $q = |mu/\Delta|$, hence is an integer. Since $s, t > 0$, the identity gives $ms = qt$, while $mu = \pm q\Delta$. Therefore

$$mbD = \frac{q(t \pm \Delta)}{2} \in \{qA_iF_j,\ qA_jF_i\}.$$

In the first case $e_i \mid qA_iF_j$ and $\gcd(e_i, F_j) \mid \gcd(F_i, F_j) \mid |\Delta|$, so $e_i \le qA_i|\Delta|$. In the second case the same argument gives $e_j \le qA_j|\Delta|$. Both conclusions contradict $e_i, e_j \ge cX^{1/2-\eta}$ for large X, since $q \le \sqrt{m}\,X^{\gamma}$ and $\gamma < 1/2 - \eta$. Thus the zero-norm case has no solutions. There are $O(X^{2\gamma+\rho})$ choices of (a,b,u); summing the nonzero-norm estimate, with its exponent reduced to absorb the fixed divisor weights, proves the claim. The normalized variation is small when $2\gamma + \rho < 1$; this leaves the complement of the resonance regions.

Acknowledgments and verification. OpenAI language models assisted in developing the arguments. ChatGPT assisted with the cutoff optimization, comparisons with the cited results, exposition and the accompanying finite checks. The checks cover divisor identities, band weights, allocation geometry and the limiting kernel. They also compare endpoint weights against independent integrals of the resulting energy. The accompanying Zenodo source archive gives the program and its output. The analytic limiting arguments are given in the text. They have not been formally verified.

References

[1] V. Bergelson and F. K. Richter, *Dynamical generalizations of the Prime Number Theorem and disjointness of additive and multiplicative semigroup actions*, Duke Math. J. **171** (2022), no. 15, 3133–3200. DOI 10.1215/00127094-2022-0055. Lemma 2.2; see also arXiv:2002.03498v3.

[2] J. Bourgain, P. Sarnak and T. Ziegler, *Disjointness of Moebius from horocycle flows*, in *From Fourier Analysis and Number Theory to Radon Transforms and Geometry*, Developments in Mathematics 28, Springer, 2013, 67–83. DOI 10.1007/978-1-4614-4075-8_5. Theorem 2 of arXiv:1110.0992v1.

[3] E. Carneiro, A. Chirre, H. A. Helfgott and J. Mejía-Cordero, *Optimality for the two-parameter quadratic sieve*, Acta Arith. **203** (2022), 195–226. DOI 10.4064/aa201018-5-1. Theorem 1.2 and Corollary 1.3; arXiv:2005.03162v6.

[4] H. Dette, A. Pepelyshev and A. Zhigljavsky, *Optimal designs in regression with correlated errors*, Ann. Statist. **44** (2016), 113–152. DOI 10.1214/15-AOS1361. Theorem 2.1 and (2.12) in arXiv:1501.01774v2.

[5] É. Fouvry and M. Radziwiłł, *Level of distribution of unbalanced convolutions*, Ann. Sci. Éc. Norm. Supér. (4) **55** (2022), 537–568. DOI 10.24033/asens.2501. Equation numbering for the variance calculation follows arXiv:1811.08672v1.

[6] D. A. Goldston and C. Y. Yıldırım, *Higher correlations of divisor sums related to primes I: Triple correlations*, Integers **3** (2003), A5. Journal archive DOI 10.5281/zenodo.7607508.

[7] D. A. Goldston and C. Y. Yıldırım, *Higher correlations of divisor sums related to primes III: k-correlations*, arXiv:math/0209102v1, 10 September 2002.

[8] S. Graham, *An asymptotic estimate related to Selberg's sieve*, J. Number Theory **10** (1978), 83–94. DOI 10.1016/0022-314X(78)90010-0.

[9] A. Granville, D. Koukoulopoulos and J. Maynard, *Sieve weights and their smoothings*, Ann. Sci. Éc. Norm. Supér. (4) **54** (2021), 1089–1177. DOI 10.24033/asens.2478. Preprint version arXiv:1606.06781v4.

[10] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications 53, 2004. DOI 10.1090/coll/053.

[11] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1986), 223–225. DOI 10.1007/BF01949145.

[12] J. D. Lichtman, *Improved bounds for the binary Chowla conjecture*, arXiv:2611.28392v1, 20 November 2026. DOI 10.48550/arXiv.2611.28392.

[13] OpenAI, *Ordinary two-point correlations of multiplicative functions*, manuscript dated 24 September 2026, Theorem 1.1. OpenAI Math Release, mathematical manuscript, repository snapshot f5702854040e.

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

[15] B. Pouly, *Arithmetic coding against unknown small-prime signs*, Zenodo, 2026. DOI 10.5281/zenodo.23319574.

[16] B. Pouly, *Local Compensation of Vaughan Remainders with Prescribed Divisors*, Zenodo (2026). 10.5281/zenodo.24578879. Theorem 1.1 and Corollary 6.2.

[17] F. K. Richter, *A new elementary proof of the Prime Number Theorem*, Bull. Lond. Math. Soc. **53** (2021), 1365–1375. DOI 10.1112/blms.12503.

[18] A. Sedunova, *A logarithmic improvement in the Bombieri–Vinogradov theorem*, J. Théor. Nombres Bordeaux **31** (2019), 635–651. DOI 10.5802/jtnb.1098. Lemma 1.5.

[19] A. Selberg, *An elementary proof of the prime-number theorem*, Ann. of Math. (2) **50** (1949), 305–313. DOI 10.2307/1969455.

[20] P. Srivastav, *Log-free bounds on exponential sums over primes*, arXiv:2505.07803v2, 27 January 2026. Lemma 2.1. DOI 10.48550/arXiv.2505.07803.

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

[22] R. C. Vaughan, *Sommes trigonométriques sur les nombres premiers*, C. R. Acad. Sci. Paris Sér. A–B **285** (1977), A981–A983.

[23] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240. DOI 10.1007/BF02392467. Lemma 3.5.