hexagon:2611.01205v1

# Polynomial phases in ordinary two-point correlations of multiplicative functions

Brice Pouly

Research and writing with AI assistance

**Abstract**

Can a polynomial phase detect a correlation of two shifted multiplicative functions that no fixed arithmetic progression detects? The coefficients may change with the length of the ordinary average. We evaluate the limiting maximum for a sequence $n^{i\tau}b(n)$ when $b$ is approximable in mean by periodic sequences. It factors into a rational maximum and an oscillatory integral optimized over polynomials on the unit interval. In fixed degree $d$, the latter loses $\tau^2/(2(d+1)^2)+O_d(\tau^4)$ from one at small twist and has order $|\tau|^{-1/(d+1)}$ at large twist. For general shifted products we propose that every persistent peak has a fixed rational center and bounded rescaled coefficient drifts along a subsequence. This would give the progression equivalence. We prove a rational-profile formula under that localization hypothesis and present the signed graph comparisons behind the proposed argument. A second candidate estimate gives a logarithmic saving for the adjacent Liouville pair with a moving quadratic phase. Applications distinguish a common shift with a truncated sign, which uses an established cofactor law, from configurations of eight or ten distinct origins, which require additional polynomial estimates.

## 1 Which polynomial phases can detect a shifted product?

A residue class can be recovered from finitely many rational linear phases. Reversing this observation is much harder. For example, the sequence $\mathrm{e}(-\sqrt{2}\,n^2)$ has mean zero in every fixed residue class by Weyl's theorem, yet one quadratic phase correlates with it perfectly. Now take

$$a(n)=f_1(n)f_2(n+h),\qquad |f_1|,|f_2|\le 1,\qquad h\ge 1,\quad \mathrm{e}(x)=\exp(2\pi i x).$$

Both functions are fixed and multiplicative on coprime arguments. We ask whether this structure rules out the irrational quadratic obstruction, even when the testing polynomial is chosen anew at every length.

Normalize every sum by the full length, including sums in one residue class. For $X\ge 1$ and fixed integer $d\ge 1$ put

$$D_d(a;X)=\sup_{\alpha_1,\dots,\alpha_d\in\mathbb{T}}\left|\frac{1}{X}\sum_{1\le n\le X}a(n)\mathrm{e}\left(\sum_{i=1}^{d}\alpha_i n^i\right)\right|,\qquad \mathbb{T}=\mathbb{R}/\mathbb{Z}.$$

Constant phases disappear on taking the absolute value. Write $\|x\|_{\mathbb{T}}$ for distance to the nearest integer. All logarithms are natural. A sum over positive integers $n\le X$ has denominator $X$, also when $X$ is not an integer.

For $a(n)=n^{i\tau}$, dividing the normalized sum by $X^{i\tau}$ gives the limit $(1+i\tau)^{-1}$ at zero phase. Testing instead with $\mathrm{e}(sn/X)$ gives $\int_0^1 u^{i\tau}\mathrm{e}(su)\,du$. The frequency still tends to zero, yet a suitable $s$ increases the limiting modulus whenever $\tau\neq 0$. Theorems 2.3 and 2.4 evaluate the effect of these drifts. The general question is whether every persistent peak admits a description at such scales.

**Candidate theorem 1.1** (Localization of a persistent correlation)**.** *Suppose $X_j \to \infty$ and $P_j(n) = \sum_{i=1}^{d} \alpha_{i,j} n^i$ gives an absolute normalized sum at least $\delta > 0$ for every $j$. On a subsequence there are fixed rationals $r_1, \dots, r_d$ such that*

$$X_j^i \|\alpha_{i,j} - r_i\|_{\mathbb{T}} = O(1) \qquad (1 \le i \le d). \tag{1}$$

*Consequently, for each fixed $d \ge 1$, we have the equivalence of* (2) *and* (3) *below.*

$$X^{-1} \sum_{\substack{n \le X \\ n \equiv r \ (b)}} a(n) \longrightarrow 0 \qquad \text{for all fixed } b \ge 1 \text{ and } r \bmod b, \tag{2}$$

$$D_d(a; X) \longrightarrow 0. \tag{3}$$

*The subsequence, centers and scales in* (1) *may depend on the pair, $h, d$ and the initial sequence of peaks.*

For instance, $\mathrm{e}((1/3 + t/X^2)n^2)$ consists of a periodic factor and a factor whose total variation is at most $2\pi|t|$ on $[0, X]$. Bounded $t$ can be handled by summation by parts. When $t$ grows, this argument loses control. Assertion (1) says precisely that, after passing to a suitable rational center, a persistent peak has only bounded rescaled drifts. The algebraic converse from (3) to (2) already holds for every bounded sequence.

OpenAI's *Ordinary two-point correlations of multiplicative functions* proves (2) if, for one original factor $f$ and every fixed Dirichlet character $\chi$,

$$\inf_{|t| \le X} \sum_{p \le X} \frac{1 - \Re(f(p)\overline{\chi(p)}p^{-it})}{p} \longrightarrow \infty.$$

Indeed $n = bm + r$ gives two fixed affine forms of determinant $bh$. The same work proves a power-of-logarithm saving for Liouville values at fixed nonproportional forms [26, Theorems 1.1–1.2]. Candidate Theorem 1.1 would turn the resulting qualitative cancellation into a supremum over moving external coefficients. Its localization claim also covers pairs for which the untwisted mean is nonzero.

**Candidate theorem 1.2** (The adjacent Liouville pair)**.** *Let $\lambda(n) = (-1)^{\Omega(n)}$, where $\Omega$ counts prime factors with multiplicity. For some absolute $c_2 > 0$,*

$$\sup_{\alpha, \theta \in \mathbb{R}} \left| \frac{1}{X} \sum_{n \le X} \lambda(n)\lambda(n+1)\mathrm{e}(\alpha n^2 + \theta n) \right| \ll (\log X)^{-c_2} \qquad (X \ge 3). \tag{4}$$

### 1.1 Earlier results and the quantifiers of the question

Products of shifted multiplicative functions against polynomial phases already have a logarithmic theory. Frantzikinakis and Host prove that, for any fixed 1-bounded multiplicative functions $f_1, \dots, f_\ell$, fixed shifts $h_1, \dots, h_\ell$, and fixed real polynomial $P$ with an irrational nonconstant coefficient,

$$\frac{1}{\log X} \sum_{n \le X} \frac{f_1(n + h_1) \cdots f_\ell(n + h_\ell)\mathrm{e}(P(n))}{n} \longrightarrow 0$$

[14, Corollary 1.4 and the following remarks]. Their result also includes appropriate fixed zero-mean weights from totally ergodic nilsystems. The present localization question concerns ordinary averages

at every sufficiently large length and a supremum over coefficients chosen anew at that length. A perturbation of the coefficient of $n^i$ is magnified by $X^i$, so fixed-phase convergence gives no uniform bound on this family.

There is also an existing route to localization when both multiplicative factors pretend to fixed functions $\chi_i(n)n^{it_i}$. The subsequential periodic approximation of Frantzikinakis, Lemańczyk and de la Rue [15, Proposition 5.5], followed by the inverse Weyl estimate, gives that case. We spell out this deduction in Section 2.3; it is an application of their theorem. General pairs need not admit these fixed twists. The exact maximum in Theorem 2.3 concerns the more specific class with a global periodic approximation in mean.

For a *single* multiplicative function, Frantzikinakis and Host decompose it into an approximately periodic part and a part small in a fixed Gowers norm [13]. Polynomial tests also appear in the short-interval uniformity theory of Matomäki, Radziwiłł, Tao, Teräväinen and Ziegler [24], and in the Weyl-sum estimates of Bordignon, Bortolotto and Kerr [3]. These provide related structural and quantitative tools; the shifted product $f_1(n)f_2(n+h)$ usually loses multiplicativity.

The graph comparison is the proposed extension specific to this product. Wang and Yi already treat varying characteristic functions of $\Omega(n)$ and $\Omega(n+1)$ by marked graphs [33, Proposition 3.3 and Lemmas 3.4–3.5]. Here the external polynomial assigns a mark depending on the arithmetic position of an edge as well as on its divisor label. Lichtman's estimate for the untwisted adjacent Möbius product has logarithmic exponent $1-\varepsilon$; the remarks following his theorem also extend the method to Liouville and to fixed nonzero shifts [22, Theorem 1.1 and the following remarks]. That stronger untwisted rate does not supply uniformity in the external coefficients in (4).

### 1.2 Results and route of proof

Theorems 2.3 and 2.4 evaluate the maximum for a periodic approximation times an Archimedean twist and quantify how much a moving phase can recover. Both have direct proofs. The rational-profile formula in Theorem 2.2 is an abstract theorem whose hypothesis is localization. The two general arithmetic assertions above retain candidate status: Sections 4–6 give the proposed marked comparisons and the ensuing arguments. The intermediate growing-modulus, linear, and large-prime marked estimates are stated as candidate propositions with their own quantifiers. The formalized unweighted theorem of [26] is an input to this argument; its formal scope does not include these extensions.

Section 2 explains the shape of a localized peak. Sections 3–5 first reduce the functions and then compare signed graph moments at fixed graph size, taking the limit in $X$ before increasing that size. Section 6 supplies the uniform comparisons needed when the graph grows with $X$. The fourth-moment argument isolates two diagonal families through the convexity of reciprocal powers. Sections 7–8 preserve the different sampling regimes and state the exact inputs needed for each sign law. Section 9 explains why uniformity against external phases still leaves a prescribed affine correlation untreated.

## 2 The shape and size of a localized peak

In this section $a$ is any complex sequence with $|a| \le 1$. Say that $a$ has the localization property in degree $d$ if it satisfies the first assertion of Candidate Theorem 1.1. We give two ways to evaluate $D_d$. One assumes this property. The other verifies it directly in a class of examples.

## 2.1 Rational phases and limits of partial sums

Localization has an equivalent uniform formulation at each positive height. For every $\delta > 0$ there exist finite $B, C$ and a threshold $X_0$ such that, at every $X \geq X_0$, each peak of size $\delta$ lies within $CX^{-i}$ at coefficient $i$ of some rational polynomial of common denominator at most $B$. Indeed failure allows one to take $B = C = j$ and $X_j \geq j$ successively, contradicting (1). The reverse follows by taking a subsequence on the finite set of centers modulo one. This formulation gives no rate for the dependence on $\delta$.

**Lemma 2.1.** *Localization implies the equivalence of* (2) *and* (3).

*Proof.* The indicator of $n \equiv r \bmod b$ is $b^{-1} \sum_{j \bmod b} \mathrm{e}(j(n-r)/b)$, proving one direction. Conversely, if (2) holds and $R$ is a fixed rational polynomial, periodicity gives $A_R(y) = \sum_{n \leq y} a(n)\mathrm{e}(R(n)) = o(y)$. Hence $\sup_{y \leq X} |A_R(y)| = o(X)$. On a localized subsequence, lift each $\alpha_{i,j} - r_i$ to its least absolute representative. The total variation of $\mathrm{e}(P_j - R)$ on $[0, X_j]$ is at most $2\pi \sum_i X_j^i \|\alpha_{i,j} - r_i\|_{\mathbb{T}}$. Partial summation contradicts a persistent peak. □

To keep the information when the progression means are nonzero or fail to exist, draw the partial sums on the scale of the interval. For $R \in \mathbb{Q}[n]$ of degree at most $d$, put $F_{R,X}(k/X) = A_R(k)/X$ at integers $0 \leq k \leq X$. Interpolate linearly and extend constantly over the last fraction of a mesh interval up to 1. These functions vanish at 0 and are 1-Lipschitz. Denote by $\mathcal{F}_R$ their uniform subsequential limits as $X \to \infty$. This set is nonempty by compactness. Every $F \in \mathcal{F}_R$ is absolutely continuous, with $|F'| \leq 1$ almost everywhere.

**Theorem 2.2** (Rational profiles)**.** *Suppose $a$ has localization in degree $d$. Write $p_t(u) = \sum_{i=1}^{d} t_i u^i$. Then*

$$\limsup_{X \to \infty} D_d(a; X) = \sup_{\substack{R \in \mathbb{Q}[n],\ \deg R \leq d \\ F \in \mathcal{F}_R,\ t \in \mathbb{R}^d}} \left| \int_0^1 \mathrm{e}(p_t(u)) F'(u)\, du \right|. \tag{5}$$

*A positive left side is attained by a triple on the right. If all the individual progression means of $a$ exist, set $m_R = \lim_{X \to \infty} A_R(X)/X$. Then the limit exists and*

$$\lim_{X \to \infty} D_d(a; X) = \sup_{R \in \mathbb{Q}[n],\, \deg R \leq d} |m_R|. \tag{6}$$

*If this value is positive, every asymptotically maximizing sequence of polynomials has a subsequence with a fixed maximizing rational center and $X_j^i \|\alpha_{i,j} - r_i\|_{\mathbb{T}} \to 0$ for every $i$.*

*Proof.* Suppose $F_{R,X_j} \to F$ and $t_j \to t$. The exponential $\mathrm{e}(p_{t_j})$ converges uniformly to $\mathrm{e}(p_t)$. Replacing a right endpoint on each mesh interval by its integral mean costs $O_t(X_j^{-1})$ for the fixed test $\mathrm{e}(p_t)$. Integration by parts and uniform convergence of the primitives therefore give

$$\frac{1}{X_j} \sum_{n \leq X_j} a(n)\mathrm{e}(R(n) + p_{t_j}(n/X_j)) \longrightarrow \int_0^1 \mathrm{e}(p_t(u)) F'(u)\, du. \tag{7}$$

This proves the lower bound in (5) using constant $t_j = t$. For the upper bound take a sequence realizing the positive limsup and maximizers at its lengths (the coefficient torus is compact). Localize. The scaled distances with their signs are bounded, so a further subsequence gives a vector $t$ and a limit $F$. Formula (7) attains the limsup. The zero case follows from the lower bound.

If the progression means exist, periodicity gives $A_R(y) = m_R y + o_R(y)$. All the functions in $\mathcal{F}_R$ then equal $u \mapsto m_R u$. This bounds the limsup by the right side of (6), since the oscillatory

integral has modulus at most one. Every fixed $R$ gives the reverse bound for the liminf. For positive equality in an extremizing subsequence, $|m_R|$ must itself be maximal and $\left|\int_0^1 \mathrm{e}(p_t(u))\,du\right| = 1$. The continuous integrand has constant phase, so $p_t$ is constant on $[0,1]$, hence zero. This proves the last claim. □

## 2.2 An unconditional model with nonzero drift

The function $n^{i\tau}$ already distinguishes (5) from (6). Its natural average rotates with $X^{i\tau}$ when $\tau \neq 0$. Changing a polynomial coefficient by $s/X^k$ can improve the limiting modulus even though that coefficient tends to zero.

Call a bounded sequence $b$ *Besicovitch rationally almost periodic* if for each $\varepsilon > 0$ there is a bounded periodic $b_0$ with $\limsup_X X^{-1}\sum_{n\le X}|b(n)-b_0(n)| < \varepsilon$. This is approximation by periodic sequences in mean; for uniformly bounded sequences the $L^1$ and $L^2$ formulations are equivalent. The class and its multiplicative examples belong to the theory of Daboussi and Delange [9, 11]; see also [2, Theorem 2.10 and Corollary 2.11]. We can take $|b_0| \le 1$ if $|b| \le 1$, by projection onto the unit disk. Every rational twist of $b$ has a mean $m_R(b)$ by periodic approximation. Define

$$M_d(b) = \sup_{\substack{R\in\mathbb{Q}[n]\\ \deg R\le d}} |m_R(b)|, \qquad J_d(\tau) = \sup_{t\in\mathbb{R}^d}\left|\int_0^1 u^{i\tau}\mathrm{e}(p_t(u))\,du\right|.$$

**Theorem 2.3** (A periodic factor with an Archimedean twist)**.** *Let $b$ be 1-bounded and approximable by periodic sequences in mean, and let $\tau \in \mathbb{R}$. Then $a(n) = n^{i\tau}b(n)$ has localization in every fixed degree and*

$$\lim_{X\to\infty} D_d(a;X) = M_d(b)J_d(\tau). \tag{8}$$

*For $d \ge 1$ and $\tau \neq 0$ the supremum defining $J_d(\tau)$ is attained and*

$$(1+\tau^2)^{-1/2} < J_d(\tau) < 1, \qquad \lim_{d\to\infty} J_d(\tau) = 1.$$

*For $\tau = 0$ one has $J_d(0) = 1$.*

*Proof.* Fix a peak of size $\delta$. Choose a single periodic approximant with error less than $\delta/8$ in limiting mean. Its finite Fourier expansion leaves a positive peak for one fixed rational translate of the linear coefficient, on a subsequence. It suffices to localize the remaining sum of $n^{i\tau}\mathrm{e}(P_j(n))$. Discard $n \le \varepsilon X_j$ for a sufficiently small fixed $\varepsilon$. The weight $n^{i\tau}$ on the rest has variation at most $|\tau|\log(1/\varepsilon) + O_\tau(1)$. Partial summation forces a prefix exponential sum of size $\gg_{\delta,\tau,b_0} X_j$, at length comparable to $X_j$. The inverse form of Weyl's theorem [16, Proposition 4.3] supplies a bounded positive common multiplier making each nonconstant binomial coefficient of the polynomial $O(X_j^{-i})$ from an integer at degree $i$. Inverting the triangular change of basis at fixed $d$ clears fixed denominators and preserves these scales. Extract the finitely many denominator and residue choices. Undoing the rational translate proves localization.

On any such arc, divide the sum by $X^{i\tau}$. The rationally twisted sequence $b(n)\mathrm{e}(R(n))$ has mean $m_R(b)$. On $[\varepsilon,1]$ the test $u^{i\tau}\mathrm{e}(p_t(u))$ is continuously differentiable, and partial summation as in (7) gives the limiting value $m_R(b)\int_\varepsilon^1 u^{i\tau}\mathrm{e}(p_t(u))\,du$ there. The omitted part costs at most $\varepsilon$. This applies also to $t_j \to t$. Taking localized subsequences for the upper bound and fixed $(R,t)$ at every length for the lower bound proves (8).

The integral defining $J_d(\tau)$ tends to zero as $\max_i |t_i| \to \infty$ at fixed $d$. Indeed, divide $p_t$ by that maximum and pass to a convergent coefficient subsequence with nonconstant limit $q$. Remove $[0,\varepsilon]$ and small neighborhoods of the finitely many zeros of $q'$. On the remaining intervals $|p_t'|$ is bounded

below by a constant times the maximum and $|p_t''|$ is at most a constant times it. Integration by parts proves decay there. The discarded length can be arbitrarily small. Continuity and the nonzero value at $t = 0$ give attainment. Modulus one would require $\tau/u + 2\pi p_t'(u) = 0$ for all $u > 0$ in the interval. Multiplying by $u$ and taking $u \downarrow 0$ contradicts $\tau \neq 0$.

At $t = 0$ the integral is $(1 + i\tau)^{-1}$. With just the coefficient of $u^k$ equal to $s$, the derivative at $s = 0$ of its squared modulus is

$$-\frac{4\pi k\tau}{(k+1+\tau^2)^2 + k^2\tau^2} \qquad (1 \le k \le d).$$

It is nonzero. A drift of the appropriate sign therefore improves the modulus. The convergence in $d$ also follows by approximating $-\tau \log u/(2\pi)$ uniformly on $[\varepsilon, 1]$ by polynomials and then letting $\varepsilon \downarrow 0$. Theorem 2.4 gives an explicit error bound. □

For a shifted multiplicative example let $f_i(n) = n^{it_i} b_i(n)$ with multiplicative 1-bounded $b_i$ approximable by periodic sequences in mean. Examples include bounded multiplicative $b$ with $\sum_p |1 - b(p)|/p < \infty$. Indeed, the convolution $c = b * \mu$ has $\sum_m |c(m)|/m < \infty$, so truncating $b(n) = \sum_{m|n} c(m)$ gives periodic approximants, as in Lemma 3.1. The product $b_1(n)b_2(n+h)$ retains mean approximability by periodic functions. Also $(1 + h/n)^{it_2} \to 1$, so Theorem 2.3 applies with $\tau = t_1 + t_2$. The sharp distinction can be seen already for $f_1(n) = n^{i\tau}$ and $f_2 = 1$. Every fixed rational phase has limiting correlation modulus at most $(1 + \tau^2)^{-1/2}$, whereas a moving linear phase attains a strictly larger limiting modulus. This shows why the existence of the progression means matters in (6).

**How much of the logarithmic phase can a polynomial remove?**

A small twist gives a least-squares problem. A large twist gives a stationary point whose order is limited by the degree. Both questions concern the same factor in (8).

**Theorem 2.4** (Size of the Archimedean factor)**.** *For every integer $d \ge 1$ and every real $\tau$,*

$$0 \le 1 - J_d(\tau) \le \frac{\tau^2}{2(d+1)^2}.$$

*For each fixed $d$ the small- and large-twist estimates are*

$$J_d(\tau) = 1 - \frac{\tau^2}{2(d+1)^2} + O_d(\tau^4) \quad (\tau \to 0), \qquad J_d(\tau) \asymp_d |\tau|^{-1/(d+1)} \quad (|\tau| \ge 1).$$

*In a neighborhood of zero the maximizing vector $t$ is unique and depends real-analytically on $\tau$. In degree one, $2\pi t_1 = -3\tau + O(\tau^3)$.*

*Proof.* Let $U$ be uniform on $(0, 1)$, set $g(u) = \log u$, and write $q_s(u) = \sum_{j=1}^d s_j u^j$, with $s_j = 2\pi t_j$. We first compute the squared error of the orthogonal projection $\Pi_d g$ onto the real polynomials of degree at most $d$, constants included. With $L_k(u) = P_k(2u - 1)$ the shifted Legendre polynomials, one has $\int_0^1 L_k^2 = 1/(2k + 1)$ and $[u(1 - u)L_k']' = -k(k + 1)L_k$ [25, Tables 18.3.1 and 18.8.1]. For $k \ge 1$, integration by parts gives

$$k(k+1)\int_0^1 (\log u)L_k(u)\,du = \int_0^1 (1-u)L_k'(u)\,du = -L_k(0) = (-1)^{k+1}.$$

The boundary term at zero vanishes since $u \log u \to 0$. The same logarithmic moments follow by differentiating the power moments in Sidi [32, (2.6)–(2.7)] at exponent zero. The constant coefficient is $-1$ and $\int_0^1 g^2 = 2$. Hence

$$\|g - \Pi_d g\|_2^2 = 2 - 1 - \sum_{k=1}^{d} \frac{2k+1}{k^2(k+1)^2} = \frac{1}{(d+1)^2}.$$

The summand is $k^{-2} - (k+1)^{-2}$. Choose $q_s$ to be $-\tau \Pi_d g$ up to its constant term. Since $\cos x \geq 1 - x^2/2$, taking the real part after removing the constant gives the bound for all $\tau$.

For sharpness, consider $F(\tau, s) = |\mathbb{E} \exp(i[\tau g(U) + q_s(U)])|^2$. The integral is analytic in the parameters near zero. For $|\Im \tau| \leq 1/2$ the integrand and its fixed-order derivatives are dominated by constants times $u^{-1/2}(1 + |\log u|^k)$ on compact coefficient sets. Thus $F$ is real-analytic. At $\tau = 0$ its only maximizer is $s = 0$, since equality in the triangle inequality forces $q_s$ to be constant. The decay in $s$ proved above and the uniform bound $|\mathbb{E}[(e^{i\tau g(U)} - 1)e^{iq_s(U)}]| \leq |\tau| \, \mathbb{E}|g(U)|$ place every maximizer in any prescribed neighborhood of zero for all sufficiently small $|\tau|$. At the origin the Hessian in $s$ is $-2(\mathrm{Cov}(U^j, U^k))_{1 \leq j,k \leq d}$. It is negative definite because a polynomial of zero variance is constant. The implicit function theorem now gives a unique analytic maximizing branch. Complex conjugation gives $s(-\tau) = -s(\tau)$.

For $Y = \tau g(U) + q_s(U)$ and an independent copy $Y'$, expand $F = \mathbb{E} \cos(Y - Y') = 1 - \mathrm{Var}(Y) + O_d((|\tau| + |s|)^4)$. Differentiating the stationarity equations at zero gives

$$\mathrm{Cov}\big(g(U) + q_{s'(0)}(U),\, U^k\big) = 0 \qquad (1 \leq k \leq d).$$

Thus $q_{s'(0)} = -\Pi_d g$ up to a constant. Substituting $s(\tau) = \tau s'(0) + O_d(\tau^3)$ and taking the square root proves the expansion. In degree one $\mathrm{Cov}(g(U), U) = 1/4$ and $\mathrm{Var}(U) = 1/12$, giving the stated drift.

For the large-twist upper bound put $m = d + 1$. The $m$th derivative of $\tau \log u + q_s(u)$ has absolute value $(m-1)!|\tau| u^{-m}$. The one-dimensional van der Corput lemma on $[\varepsilon, 1]$ [7, Lemma 1.1(a)] gives $O_d(|\tau|^{-1/m})$ uniformly in $s, \varepsilon$. Let $\varepsilon \downarrow 0$.

For the lower bound we may take $\tau > 0$ by conjugation. Choose the polynomial phase in radians to be $\tau \sum_{j=1}^{d} (1-u)^j/j$, up to a constant. It cancels the first $d$ terms of $\tau \log u$ at $u = 1$. With $v = 1 - u$ the resulting phase divided by $\tau$ is

$$\phi(v) = \log(1 - v) + \sum_{j=1}^{d} v^j/j, \qquad \phi'(v) = -v^d/(1 - v).$$

For fixed $R > 0$, the change $v = z\tau^{-1/m}$ gives $\tau\phi(v) \to -z^m/m$ uniformly on $0 \leq z \leq R$. On the rest of the interval $|\phi'|$ is increasing. The first-derivative estimate [7, Lemma 1.1(c)], first stopping short of $v = 1$, gives

$$\tau^{1/m} \left| \int_{R\tau^{-1/m}}^{1} e^{i\tau\phi(v)}\, dv \right| \ll_d R^{-d}.$$

Let $\tau \to \infty$ and then $R \to \infty$. The rescaled integral tends to $m^{1/m}\Gamma(1 + 1/m)e^{-i\pi/(2m)}$. Indeed, put $x = z^m/m$ and use the gamma integral with a damping factor $e^{-\varepsilon x}$ and let $\varepsilon \downarrow 0$ [25, (5.9.1)]. Integration by parts bounds the tail uniformly in $\varepsilon$. This nonzero limit is the constant for the chosen endpoint phase. The two bounds give the decay exponent for the supremum. The positive value at $s = 0$ covers the remaining compact range of $\tau$. □

For $a(n) = n^{i\tau}$ a rational phase of zero drift has limiting modulus at most $(1 + \tau^2)^{-1/2} = 1 - \tau^2/2 + O(\tau^4)$. A linear phase of slope $-3\tau/(2\pi X)$ already attains $1 - \tau^2/8 + O(\tau^4)$ in the limit $X \to \infty$ followed by $\tau \to 0$. The all-degree bound also gives $J_d(\tau_d) \to 1$ whenever $d \to \infty$ and $|\tau_d| = o(d)$.

### 2.3 Localization for two fixed pretentious twists

The localization part has a wider previously available subcase. Suppose, for $i = 1, 2$, that a fixed Dirichlet character $\chi_i$ and a fixed $t_i \in \mathbb{R}$ satisfy

$$\sum_p \frac{1 - \Re(f_i(p)\overline{\chi_i(p)}p^{-it_i})}{p} < \infty.$$

Localization for $f_1(n)f_2(n+h)$ then follows from [15, Proposition 5.5] and the inverse Weyl estimate, as follows. Set $b_i(n) = f_i(n)n^{-it_i}$. From any sequence of lengths tending to infinity, the cited proposition provides a common subsequence $X_j$, periodic 1-bounded sequences $b_{i,m}$, and constants $|z_{i,m}| = 1$ such that

$$\limsup_{j\to\infty} \frac{1}{X_j} \sum_{n\le X_j} |b_i(n) - z_{i,m}b_{i,m}(n)|^2 \le m^{-1}.$$

Apply the proposition successively to the two factors. A fixed shift changes this bound by $O(h/X_j)$, and Cauchy–Schwarz bounds the mean error for the product by $2m^{-1/2}$. Moreover

$$f_1(n)f_2(n+h) = n^{i(t_1+t_2)}b_1(n)b_2(n+h)(1+h/n)^{it_2},$$

where replacing the last factor by 1 has mean error $O_{h,t_2}(\log(2X_j)/X_j)$. For a persistent peak, fix $m$ sufficiently large in terms of its height. A fixed periodic product now retains a positive peak. Its finite Fourier expansion and the argument of Theorem 2.3 give the required rational centers and the scales $X_j^{-i}$.

This deduction uses fixed twists and a periodic approximation along a selected subsequence. It does not identify a single global periodic-mean limit $b$ to which (8) applies. The exact factorization $M_d(b)J_d(\tau)$ is asserted only under the hypotheses of Theorem 2.3.

## 3 Reducing the multiplicative functions

The geometric step below uses identities at both endpoints of a dilated edge. Ordinary multiplicativity at coprime arguments does not give these identities. The following approximation arranges them before the graph is introduced.

**Lemma 3.1** (Reduction to unimodular functions)**.** *It suffices to prove localization for pairs of completely multiplicative functions taking values on the unit circle.*

*Proof.* For a factor $f$, if $\sum_p(1 - |f(p)|)/p$ diverges, use any finite set $S$ of primes. Bound $|f(n)|$ by the product of $|f(p)|$ over $p \in S$ with $p \parallel n$. By the Chinese remainder theorem its limiting mean is bounded by $\prod_{p\in S}(1 - (1 - |f(p)|)(p-1)/p^2)$. Exhausting the primes gives mean zero in absolute value, also after the fixed shift. There can be no peak.

Otherwise choose $u(p) = f(p)/|f(p)|$ when $f(p) \neq 0$ and $u(p) = 1$ otherwise, and extend completely multiplicatively. Set $b = f\bar{u}$ and $c = b * \mu$. Here $c(p) = |f(p)| - 1$ and $|c(p^j)| \le 2$ for $j \ge 2$, so $\sum |c(m)|/m < \infty$ by the Euler product. The error in approximating $b(n) = \sum_{m|n} c(m)$ by its terms $m \le D$ has mean absolute value at most $\sum_{m>D} |c(m)|/m$. Project the approximant onto the closed unit disc to obtain a bounded periodic function with the same error bound. This elementary argument is related to the mean almost-periodicity theory in Daboussi's *Fonctions multiplicatives presque périodiques B* [9]. A fixed translation preserves the limiting mean error.

Approximate both factors in this way before passing to the subsequence of peaks. Their product is approximated by $u_1(n)u_2(n+h)$ times a bounded periodic multiplier with mean absolute error

less than $\delta/3$ in the limit. Expand this one periodic function in its finite Fourier series. Some fixed rational translate of the linear frequency retains a positive peak on a subsequence. Subtracting the translate from the rational center in (1) gives the claim. Cancellation on progressions is required only for the original pair in Lemma 2.1. □

## 4 Where the phase enters the graph

Part I of [26] supplies the following graph construction. This description fixes the normalizations for both proofs. The centered divisibility graph has antecedents in *Expansion, divisibility and parity* and *Improved bounds for the two-point logarithmic Chowla conjecture* [19, 27]. At scale $U$, with an absolute $W$ chosen large and fixed, set $J = \lfloor \log U/(1200W) \rfloor$, $\eta = e^{-J}$ and $K = e^{4J}$. The disjoint prime bands $P_i$ in [26, (2.2)–(2.3)] have harmonic masses $V_i \in [W, 2W]$ and begin above $H_0 = \exp(U^{199/200})$. Choose a single prime from each band, with product $\nu$. The padding label $q$ is squarefree, on disjoint primes, with weight $u(q) = 4^{\omega(q)}$. The cutoffs in [26, Lemma 2.2] give $q\nu \le e^{101U}$ and retained harmonic mass $S_0 \ge S \prod V_i/2$, where $S = \prod_{p \text{ padding}}(1 + 4/p)$. Partition $w = q\nu$ by $\ell\eta \le \log w < (\ell+1)\eta$, and set $T_\ell = Xe^{\ell\eta}$.

More explicitly, the bands use primes $p \equiv 1 \bmod 5$ with successive cuts in $\log p$ at $U^{199/200}e^{6Wi}$. The padding primes are $p \le e^U$, $p \not\equiv 1 \bmod 5$. Both supplies exclude prime divisors of the exterior modulus and of $h$. Only bins with $\ell\eta \le 100U$ and labels with $\omega(q) \le 100 \log U$ are used. The retained harmonic mass is $S_0 = \sum_{\ell,\nu,q} u(q)/(q\nu)$ over these labels.

The edge with lower endpoint $x$ leads to $x + hw$. The original graph includes the divisibility tests selecting multiples of $w$. For a rational polynomial $R$ of denominator $B$, omit the primes dividing $Bh$ and attach the complex factor

$$\overline{u_1(w)u_2(w)}\, \mathrm{e}\big(R(\bar{w}x \bmod B) + (P - R)(x/w)\big) \tag{9}$$

to the upward edge. The opposite orientation has its complex conjugate. The bar on $w$ denotes inversion modulo $B$. The rational phase is 1 if $B = 1$ and $R$ has integral coefficients. In the bin sum, (9) multiplies

$$\frac{1}{T_\ell}\, u(q)\mathbf{1}_{q|x} \prod_{p|\nu} \left(\mathbf{1}_{p|x} - \frac{1}{p}\right) u_1(x)u_2(x + hw), \qquad 1 \le x \le T_\ell.$$

Selecting every divisibility indicator in the expansion forces $x = wn$. Complete multiplicativity recovers $u_1(n)u_2(n+h)\mathrm{e}(P(n))$ exactly. Replacing the upper endpoint $T_\ell/w$ by $X$ costs $O(\eta + X^{-1})$ in normalized means.

There are two places where the argument uses cancellation.

1. The proof of [26, Theorem 5.4, Lemma 5.5 and Sections 6–7] sums bounds for the absolute values of individual signed expectations of numerical walks in a nonbacktracking trace. Each signed expectation must be taken *before* absolute values are inserted into the prime coordinates. The singleton identity $\mathbb{E}[(\mathbf{1}_{Z=a} - 1/p)H(Z)] = p^{-1}\mathbb{E}[H(a) - H(Z)]$ explains why entrywise domination by the unsigned matrix loses the estimate.

2. Centering only part of $\nu$ leaves $\nu = \nu' z$, with $z$ a product of the omitted bands. Put $x = q\nu' v$. Each dyadic piece then tests

$$\mathcal{C} = \frac{1}{\lfloor Y \rfloor} \sum_{v \le Y} F(v) \sum_{D \le z < 2D} c_z G(v + hz)\mathrm{e}(P_z(v)), \quad |c_z| \le 1/z, \tag{10}$$

where $Y = T_\ell/(q\nu') \ge e^{-\eta}XD$, $D \ge H_0/2$, and $P_z(v) = R(\bar{z}v \bmod B) + (P-R)(v/z)$. The labels $z$ have unique factorizations in the prescribed bands. After partial division the remaining factor $\overline{u_1(z)u_2(z)}$ is absorbed in $c_z$. Both endpoint functions are bounded by one.

**Proposition 4.1** (Conditional graph assembly)**.** *Use the graph, cutoffs and normalizations above. Suppose that each partially centered dyadic sum in* (10) *is $O(U^{-21/20})$, uniformly in its labels. Suppose also that, for both Hermitian families defined below, the numerical-word averages admit the signed product-law bounds used in the trace majorant [26, (5.3)], with the same total comparison-error budget. Here each signed expectation is taken before its absolute value, and the average is over block origins. Then*

$$\left| X^{-1} \sum_{n \le X} a(n)\mathrm{e}(P(n)) \right| \ll \eta + 2^J \eta^{-1} U^{-1/20} + 2^J (K^{-1} + U^{-100} + e^{-2WJ}) + K\eta^{-1}(C/\sqrt{W})^J + e^{-U^{0.9}} + X^{-1}. \tag{11}$$

*For fixed $U$, if the two hypotheses hold only up to $o_U(1)$ as $X \to \infty$, the conclusion holds with an additional $o_U(1)$. In particular,* (11) *is conditional on both comparisons; it is not a bound for an arbitrary phase without these hypotheses.*

*Proof.* The summations in [26, (4.8), (8.8)–(8.14)], after division by $S_0$, spend $O(2^J U/\eta)$ for the centered subsets and bins. The exceptional-site estimates in [26, Lemma 4.6] and the weighted row estimates in [26, Section 8.2] use entrywise absolute values. Their integer weights therefore survive multiplication by (9).

For the spectral step, let $C_\nu$ be the strictly upper triangular matrix of retained increasing entries at label $\nu$, including the mark and the two endpoint cutoffs of [26, (5.2)]. Set

$$A_\nu^\zeta = \zeta C_\nu + \bar{\zeta} C_\nu^*, \qquad \zeta = 1, i, \qquad C_\nu = \tfrac{1}{2}(A_\nu^1 - iA_\nu^i).$$

The prime-degree projection $E$ is the same in both matrices. Along each fixed numerical word the extra $\zeta$'s give a scalar of modulus one. Reversing the second half of a squared matrix entry conjugates its weights exactly. Thus each marked-word comparison below applies to both Hermitian families. In the weighted Schur estimate, take absolute entry values on the invariant level sets of the prime-divisor weight as in [26, Section 8.2]. This preserves the single-operator bound $\|A_\nu^\zeta\| \le 2K$ and the projected sum-of-squares bound. The self-adjoint nonbacktracking inequality and the trace bound of that source then control $E\sum_\nu A_\nu^\zeta E$ outside two exceptional sets of total relative size $O(e^{-2\lfloor U \rfloor})$. The displayed identity gives the same norm bound, up to an absolute constant, for $E\sum_\nu C_\nu E$.

With $w_0(x) = 5^{\omega_Q(x)}$ for the padding primes $Q$, use test vectors $\sqrt{w_0(x)}\,\overline{u_1(x)}$ and $\sqrt{w_0(x)}\,u_2(x)$ subject to the degree cutoffs. The first vector is conjugated in the scalar product, recovering the desired product. Both squared norms satisfy the estimate in [26, Proposition 8.4]. The discarded blocks use the same entrywise row bound. A block of $M = \lceil e^{103U} \rceil$ sites beginning just after $t$ contains an edge $(x, x+hw)$ exactly when $x + hw - M \le t \le x - 1$. This gives the multiplicity and endpoint errors of the overlapping-block argument in that proposition. These steps prove (11) under the stated comparisons. At fixed $U$ all expansions are finite, so their $o_U(1)$ errors remain $o_U(1)$ after summation. □

Take $W$ so that $e^5 C/\sqrt{W} < e^{-1}$. The displayed right side tends to zero as $U \to \infty$. For the qualitative argument, take the limit superior as $X \to \infty$ at each fixed $U$ first. For a growing exterior conductor, the $2^J/K$ term acquires the power of $1 + \log U$ described in Section 6.

## 5 Qualitative localization in every degree

Choose bounded representatives of the $\alpha_{i,j}$ and pass to a subsequence on which each converges to $a_i$. In the rational vector space spanned by $1, a_1, \dots, a_d$, take a complement $E$ to $\mathbb{Q}$. Write $a_i = r_i + a_i'$, $r_i \in \mathbb{Q}$, $a_i' \in E$. Set $c_{i,j} = \alpha_{i,j} - r_i$ and $R(n) = \sum r_i n^i$. Extract diagonally once more so every $X_j^k |\sum_i t_i c_{i,j}|$ has a limit in $[0, \infty]$ for rational $t_i$ and $1 \le k \le d$. Choose these subsequences before any graph scale. The denominator $B$ of $R$ is fixed.

**Lemma 5.1** (Averages of the phase of a fixed walk)**.** *Let $Q_j(x) = \sum_{k=0}^{d} v_{k,j} x^k$, where each nonconstant coefficient is a fixed rational linear combination of the $c_{i,j}$. Either every average of $\mathrm{e}(Q_j)$ over each fixed residue class in $[1, N_j]$, $N_j \asymp X_j$, tends to zero, or $v_{k,j} = O(X_j^{-k})$ for all $k \ge 1$. In the latter case each class average differs from the full average by $o(1)$, when normalized by its own length.*

*Proof.* On the class $b + Mn$ the nonconstant binomial coefficients of $Q_j(b + Mn)$ are obtained by an upper triangular rational transformation with diagonal $k!M^k$. Their limits belong to $E$. If a class average has a positive limsup, the circle case of Green and Tao's quantitative polynomial equidistribution theorem [16, Proposition 4.3] gives a common nonzero integer $t$ bounded in terms of the peak, modulus and degree such that the distance of $t$ times the $k$th binomial coefficient to an integer is $O(X_j^{-k})$. Extract $t$ constant. Here $E$ is a $\mathbb{Q}$-vector subspace with $E \cap \mathbb{Q} = \{0\}$. The limiting binomial coefficients are both rational and in $E$, so are zero. Because the common multiplier is fixed, the nearest integers are eventually zero as well. Triangular inversion, starting at degree $d$, gives $v_{k,j} = O(X_j^{-k})$ on this subsequence. The diagonal extraction makes all the scaled coefficients bounded on the original subsequence. In this case the phase has bounded variation, and summation by parts with the $O(M)$ prefix discrepancy for a fixed class proves the assertion. □

**Lemma 5.2** (Signed transfer at fixed graph size)**.** *Use the subsequence and polynomials $Q_j$ of Lemma 5.1. Fix coprime positive integers $B, M$, a bounded signed function $H$ on $\mathbb{Z}/M\mathbb{Z}$, a nonnegative function $I$ on $\mathbb{Z}/B\mathbb{Z}$, and a function $\Phi$ on $\mathbb{Z}/B\mathbb{Z}$ with $|\Phi| \le 1$. For $N_j \asymp X_j$,*

$$\limsup_{j \to \infty} \left| \mathbb{E}_{x \le N_j} H(x \bmod M) I(x \bmod B) \Phi(x \bmod B) \mathrm{e}(Q_j(x)) \right| \le (\mathbb{E}_{r \bmod B} I(r)) \, |\mathbb{E}_{s \bmod M} H(s)|.$$

*All the residue data in this statement are fixed before $j \to \infty$.*

*Proof.* In the first alternative of Lemma 5.1, sum over the finitely many classes modulo $BM$. In the second, each class average is $z_j + o(1)$, where $z_j = \mathbb{E}_{x \le N_j} \mathrm{e}(Q_j(x))$ and $|z_j| \le 1$. The Chinese remainder theorem gives the marked average as

$$z_j \, \mathbb{E}_{r \bmod B} I(r) \Phi(r) \; \mathbb{E}_{s \bmod M} H(s) + o(1).$$

Taking its modulus proves the assertion. The signed expectation of $H$ is preserved. □

For a numerical walk at fixed $U$, the unmarked divisibility factors form such a signed function $H$, with $M$ a product of supply primes. An exterior progression test gives $I$, and the mark is a unit scalar times $\Phi(x \bmod B)\mathrm{e}(Q_j(x))$. Expanding $(x + t)^i / w^i$ at each fixed offset $t$ verifies the coefficient hypothesis. Lemma 5.2 therefore preserves every singleton cancellation. The interval of origins in the trace has length $\lfloor T_\ell \rfloor$ at fixed bin $\ell$ and begins at zero. After fixing the starting block position, every lower endpoint is this origin plus a fixed integer. The two path halves have a common start and end, so reversal gives exactly the downward conjugates. Membership of a path in its block is a deterministic property of these offsets. The nonlinear site cutoffs, including prohibited

paths, remain functions of the prime residues. Their common modulus may be large but is fixed. There are only finitely many words and initial positions at fixed $U$, so the errors can be summed before taking the limit in $j$. This supplies the word comparison in Proposition 4.1.

It remains to center. Translate $v = v_0 + m$, $1 \le m \le D$ and average over $v_0 \le Y$, paying $O(D/Y)$ at the endpoints. The $D$-by-$(2h+1)D$ matrix with entry $c_z \mathrm{e}(P_z(v_0+m))$ in position $(m, m+hz)$ acts between vectors with norms bounded by $\sqrt{D}$ and $\sqrt{(2h+1)D}$. Its operator norm to the fourth power is at most $\mathrm{Tr}((KK^*)^2)$, where here $K$ denotes this rectangular matrix. A rectangle has offsets $x, z, y, w$ in cyclic order with $x + y = z + w$. There are at most $D$ first rows for any triple, each contribution carrying absolute weight at most $D^{-4}$. Explicitly, Cauchy–Schwarz on the two endpoint vectors, followed by Hölder in $v_0$, gives

$$|\mathcal{C}| \;\le\; O(D/Y) + O(\sqrt{2h+1})\,\Big(\mathbb{E}_{v_0 \le Y}\,\mathrm{Tr}((K_{v_0}K_{v_0}^*)^2)\Big)^{1/4}.$$

The normalization $D^{-1}$ of the translation average cancels the product of endpoint norms up to the indicated factor.

If (1) fails for the selected $r_i$, take the largest $s$ for which $X_j^s |c_{s,j}| \to \infty$. All coefficients of higher degree are $O(X_j^{-i})$. In the phase of a rectangle, the coefficient of $v_0^s$ after fixing the rational classes is

$$c_{s,j}\Delta_s + O_U(X_j^{-s-1}), \qquad \Delta_s = x^{-s} + y^{-s} - z^{-s} - w^{-s}.$$

In fact there is a quantitative convexity identity. Write $a = x - z$, $b = y - z$, so $w = z + a + b$, and put $f(u) = u^{-s}$. Twice applying the fundamental theorem of calculus gives

$$\Delta_s = -ab \int_0^1 \int_0^1 f''(z + ta + rb)\, dt\, dr, \qquad |\Delta_s| \asymp_s D^{-s-2}|x - z|\,|y - z|. \tag{12}$$

Every argument of $f''$ is a convex combination of the four corners in $[D, 2D)$. In particular $\Delta_s = 0$ exactly on the two diagonals. On any other rectangle Lemma 5.1 makes the phase average tend to zero. At most $O(D^2)$ triples are diagonal. For clarity, if $K_{v_0}$ denotes the rectangular matrix, each triple contributes $O(D^{-3})$ after summing its at most $D$ row positions. The average over $v_0$ thus gives

$$\limsup_{j \to \infty} \mathbb{E}_{v_0}\,\mathrm{Tr}((K_{v_0}K_{v_0}^*)^2) \ll D^{-1}.$$

The preceding trace inequality now gives $\limsup |\mathcal{C}| \ll_h D^{-1/4}$. The factor $D \ge \exp(U^{199/200})/2$ absorbs any required fixed power of $U$. With the word comparison above, (11) contradicts the peak as $U \to \infty$. This is the proposed proof of the marked graph estimate at fixed graph size and hence of localization. Lemmas 3.1 and 2.1 give Candidate Theorem 1.1.

**Corollary 5.3** (Two fixed affine forms)**.** *Assume the localization assertion of Candidate Theorem 1.1 for completely multiplicative unit-valued functions. It extends to $f_1(An + B)f_2(Cn + D)$ for every pair of fixed 1-bounded multiplicative functions, with $A, C$ positive integers and $AD - BC \neq 0$. Equations* (2)–(3) *and Theorem 2.2 then apply to this affine product under their respective hypotheses.*

*Proof.* Exchange the forms if necessary so that $h = AD - BC > 0$ and discard the finitely many arguments outside $\mathbb{N}$. In Lemma 3.1, an error with full mean at most $\varepsilon$ in absolute value has limiting upper mean at most $A\varepsilon$ on $An + B$, by positivity. The identical assertion holds with $C$ for the second form. This covers both the zero-mean case and approximation by unit-valued completely multiplicative functions times a single periodic function of $n$. Its Fourier expansion costs only a fixed rational shift of the linear coefficient.

For the resulting unit-valued $u_1, u_2$, put $m = ACn + BC$. Then $m = C(An + B)$ and $m + h = A(Cn + D)$. Up to the fixed factor $\overline{u_1(C)u_2(A)}$, the sum is a shifted product in $m$, restricted to one class modulo $AC$. Multiply by $AC$ to pass from the normalization at $ACX + BC$ to that at $X$, with relative error $O(1/X)$. Open the class indicator into its linear characters. Some fixed character retains a peak for the polynomial $Q_j(m) = P_j((m - BC)/(AC)) + \ell m/(AC)$. Localize its nonconstant coefficients and substitute $m = ACn + BC$ back. The error at degree $i$ is a fixed linear combination of errors of degrees $k \geq i$, hence $O(X^{-i})$. Rational centers remain rational. No hypothesis on pretentious distance was used in this reduction. □

# 6 The quadratic rate

The graph size used in Section 5 depends on the desired accuracy. A power of $\log X$ requires a growing graph and comparisons uniform in its walks. We give the proposed proof of Candidate Theorem 1.2 in that regime. The periodic, linear and marked estimates below retain candidate status as parts of this argument; they are not consequences of the fixed-parameter statement of [26] alone. Throughout this section $h = 1$ unless specified. Write $U = (\log X)^{\epsilon}$. The linear argument first fixes its own exponent $\epsilon_F$ and saving $c_F$; the quadratic argument subsequently chooses $\epsilon$ afresh. All constants are fixed before $X \to \infty$.

## 6.1 Rational and linear tests

First we need a periodic estimate with a growing shift and a growing modulus. The fixed-progression theorem in [26] allows its implied constant to depend on both.

**Candidate proposition 6.1** (Uniform periodic tests)**.** *There are absolute constants $c_0, \sigma > 0$, $C$ and $X_0 \geq 3$ such that, for every $X \geq X_0$,*

$$\sup_{b,k \leq (\log X)^{\sigma}} \sup_{|\psi| \leq 1} \left| X^{-1} \sum_{n \leq X} \lambda(n)\lambda(n+k)\psi(n \bmod b) \right| \leq C(\log X)^{-c_0}, \qquad b, k \in \mathbb{N}. \tag{13}$$

*The inner supremum is over all functions $\psi : \mathbb{Z}/b\mathbb{Z} \to \mathbb{C}$ bounded by one; the constants are independent of $\psi, b, k$.*

*Proposed proof.* We track the dependence on $b, k$ in the untwisted argument. Use $U_0 = (\log X)^{1/A}$ with the large absolute $A$ of [26, Part I], and take $\sigma \leq 1/(200A)$ so that $b, k \leq U_0^{1/200}$. Omit prime divisors of $bk$ from all supplies. The removed mass in the high bands is $O(\log(2bk)/H_0)$. The padding anti-concentration estimate in [26, (4.12)–(4.13)] loses at most $\exp(O(\sum_{p \mid bk} p^{-1}))$. This is at most $(1 + \log U_0)^{O(1)}$, even when $bk \leq X$ and the graph scale is a fixed power of $\log X$. Split the prime divisors at $(\log(10bk))^2$ and use Mertens below it. The retained-mass and cutoff moment bounds hold for the reduced padding product, since the positive inclusion probabilities only decrease. The anti-concentration loss multiplies the $2^J/K$ term.

Attach $\psi(\bar{w}x \bmod b)$ to an upward edge and its conjugate downward. At the full divisor it recovers the target. Conditional on an origin class $r \bmod b$, every mark in a fixed word is constant. The product residue model of [26, Lemma 3.1] applies to $(x - r)/b$ on an interval of length $\gg X/b$. Multiplication by $b$ permutes each remaining prime coordinate. Arithmetic constraints on walks use $k$ times sums of steps. Those primes do not divide $k$, the bound on a difference coefficient is $2\lfloor U_0 \rfloor k e^{102U_0+1}$, and the number of its distinct prime factors stays $O(U_0)$. No new choice of configuration occurs for fixed $k$ in a graph. The block length $\lceil e^{103U_0} \rceil$ absorbs the maximum step $ke^{101U_0}$. These observations retain the costs in [26, Lemmas 6.2, 6.4, 6.5 and Section 8.3].

At partial division group the $z$ in (10) by $z \equiv s \bmod b$. The remaining weight is $\psi(\bar{s}v \bmod b)$. Its finite Fourier expansion has $\ell^1$ norm $\le \sqrt{b}$ and norm 1 for a pure rational additive character. In [26, (4.4)–(4.7)] the first endpoint's large-spectrum term therefore loses at most $\sqrt{b}$, while the small-spectrum Parseval bound uses only $|\psi| \le 1$. The rough-shift polynomial $B(\theta) = \sum_z c_z \mathrm{e}(kz\theta)$ satisfies the supremum and fourth-moment bounds of [26, Lemma 4.2] for $|c_z| \le 1/z$. In the fourth moment, character orthogonality imposes one additive equation and absolute values give the same positive count. At the Liouville endpoint the Fourier input is [23, Theorem 1.7 and (1.12), corrected text]. Multiplying the frequencies in the shift polynomial by $k$ preserves Haar measure on the circle. The $L^2$ and $L^\infty$ bounds for the short sum at the second endpoint on the enlarged interval cost $\sqrt{2k+1}$ and $2k+1$, respectively. The bound per residue piece becomes

$$O(D/Y + \sqrt{2k+1}\, U_0^{-1.05} + (2k+1)\sqrt{b}\, U_0^{-1.55}).$$

Summing the at most $b$ classes and the $O(2^J U_0/\eta)$ bins and subsets leaves $O(2^J \eta^{-1} U_0^{-1/30})$ plus negligible endpoints. The assembly argument of Proposition 4.1, with this centering replacement and the stated conductor loss, gives (13) for some absolute $c_0, \sigma > 0$. This argument uses a single rational mark directly and does not require an independent large-modulus progression estimate. □

**Candidate proposition 6.2** (Uniform linear tests)**.** *There are absolute constants $c_F > 0$, $C_F$ and $X_F \ge 3$ such that, for $X \ge X_F$,*

$$\sup_{\theta \in \mathbb{R}} \left| X^{-1} \sum_{n \le X} \lambda(n)\lambda(n+1)\mathrm{e}(\theta n) \right| \le C_F (\log X)^{-c_F}.$$

*Proposed proof.* In this proof take $U = (\log X)^{\epsilon_F}$. For arbitrary $\theta$, Dirichlet approximation gives $\theta = a/b + \beta$ with $(a,b) = 1$, $b \le X^{1/3}$ and $b|\beta| \le 2X^{-1/3}$. Use $a\bar{w}x/b + \beta x/w$ on the upward edge, with the same supply omissions. At a walk the rational phase is constant on a class of $b$. The drift has slope $O(U|\beta|)$. Freeze it on blocks of length $\asymp bX^{1/6}$ at cost $O(UX^{-1/6})$. Each class has length $\asymp X^{1/6}$, sufficient for the finite-law comparison when $\epsilon_F A < 1$. Its error absorbs the $\exp(O(U^5))$ total coefficients of the walk expansion.

Here is the arithmetic of linear centering. Set $B_1 = X|\beta|$. In the fourth moment from Section 4, a rectangle has slope

$$a(\bar{x} + \bar{y} - \bar{z} - \bar{w})/b + \beta\Delta_1, \qquad \Delta_1 = -\frac{(x+y)(x-z)(y-z)}{xyzw}.$$

If the rational term is nonintegral, its distance to an integer is at least $1/b$. Since $|\Delta_1| \ll D^{-1}$ and $b|\beta| \le 2X^{-1/3}$, the perturbation is $|\beta\Delta_1| = o(1/b)$, so the normalized geometric sum is $O(b/Y)$. Otherwise $b \mid (x+y)(x-z)(y-z)$. For $b > 16D^3$ this forces a diagonal. For smaller $b$, distribute the prime-power factors of $b$ among the three factors. One of the resulting moduli is $\ge b^{1/3}$, so the proportion of triples is $O(\tau_3(b)(D^{-1} + b^{-1/3})) \ll D^{-1/2} + b^{-1/4}$. Alternatively, excluding $|x-z|$ or $|y-z| \le DB_1^{-1/4}$ costs $O(B_1^{-1/4} + D^{-1})$ when $B_1 \ge 1$. Else $Y|\beta\Delta_1| \gg B_1^{1/2}$. The fourth root gives, for fixed $h$,

$$|\mathcal{C}| \ \ll_h \ D/Y + (b/Y)^{1/4} + D^{-1/8} + (1 + \max(b, B_1))^{-1/16},$$

where $\mathcal{C}$ is given by (10). If $\max(b, B_1) > U^{200}$, this suffices in (11). If $b, B_1 \le U^{200}$, apply Proposition 6.1 with the character $\mathrm{e}(an/b)$ to every prefix $Y' \ge \sqrt{X}$, then remove the drift by partial summation at cost $O(1 + B_1)$. The shorter prefixes cost $O(X^{-1/2}(1 + B_1))$. Choosing $200\epsilon_F < \min(c_0, \sigma)/3$ gives a uniform linear saving $(\log X)^{-c_F}$, $c_F > 0$. The exponent $c_F$ and its graph scale are now fixed, before applying this estimate to quadratic sums. □

### 6.2 Comparing a growing family of quadratic walks

Take $C_1 \geq 30$ large enough that the limited-independence comparison of [26, Lemma 3.1] inspects at most $\lceil U^{C_1} \rceil$ prime coordinates at a time. Set

$$E_0 = \exp(U^{C_1+5}), \quad W_1 = \exp(U^{C_1+10}), \quad W_2 = \exp(U^{C_1+15}).$$

Decrease $\epsilon$ so $W_2 = X^{o(1)}$ and even length $X^{1/30}$ exceeds the threshold $\exp(U^A/2)$. A word has $\leq 2U$ edges, $w_i \leq e^{101U}$ and offsets $|s_i| \leq e^{104U}$. If orientations are $\varepsilon_i$, its quadratic, quadratic-to-linear, and linear rational coefficients are

$$R_2 = \sum_i \frac{\varepsilon_i}{w_i^2}, \qquad R' = 2\sum_i \frac{\varepsilon_i s_i}{w_i^2}, \qquad R_1 = \sum_i \frac{\varepsilon_i}{w_i}.$$

Their heights are $\leq \exp(O(U^2))$. Each tested product $M$ of primes is at most $\exp(O(U^{C_1+1}))$. The interval of origins has length $N \asymp T_j$, with $X - 1 \leq N \leq Xe^{101U}$.

**Lemma 6.3** (Testing a phase against a residue circuit)**.** *Let $U$ be sufficiently large, and let $\mathcal{P}$ be a set of at most $e^U$ primes, each at most $e^U$ and coprime to $l$. Put*

$$B_{\mathrm{bit}} = \lceil U^{15} \rceil, \qquad n_{\mathrm{bit}} = \max(1, B_{\mathrm{bit}}|\mathcal{P}|) \leq e^{2U}, \qquad t = \lceil U^{C_1} \rceil.$$

*Let $x$ be uniform on the points of one residue class modulo $l$ in an interval of length $N$. Suppose every product $M$ of at most $t$ distinct primes from $\mathcal{P}$ satisfies $M \leq M_*$. Let $Z(x)$ have modulus one and suppose*

$$\left|\mathbb{E}(Z(x) \mid x \equiv r \bmod lM)\right| \leq \delta$$

*on every compatible class for these products, including $M = 1$. Let $B$ be an event represented by an unbounded-fan-in Boolean circuit of depth at most 20 and size at most $\exp(U^6)$ in translated residue equalities. Depth counts AND/OR levels, and size counts gates and input occurrences, as in [26, Lemma 3.1]. If*

$$(t+1)n_{\mathrm{bit}}^t(lM_*/N + \delta) \leq e^{-2U^{11}}$$

*then $|\mathbb{E}\, Z(x)\mathbf{1}_B(x)| \ll e^{-U^{10}}$. The constant and threshold are uniform in the residue class, offsets, $l$ and $Z$ subject to these hypotheses.*

*Proof.* On the exterior class, the marginal law at $M$ is at total variation distance $O(lM/N)$ from product uniformity. For $\omega = 1, i$, reweight the law by $(1 + \Re(\omega Z))/(1 + \mathbb{E}\Re(\omega Z))$. The denominator is $1 + O(\delta)$. In each $M$-class the conditional change in mass is a relative $O(\delta)$ by assumption. The new law is therefore within $O(lM/N + \delta)$ of product uniformity on every such marginal.

For each prime $p$, encode residue $r$ using independent uniform jitter on $[r/p, (r+1)/p)$ and the first $B_{\mathrm{bit}}$ dyadic bits; if $\mathcal{P}$ is empty, add one unused uniform bit. Under the product law all bits are independent. The boundary-cell decoding error is at most $2p\, 2^{-B_{\mathrm{bit}}}$ even conditional on the residue. Summing over $\mathcal{P}$ gives at most $2e^{2U}2^{-B_{\mathrm{bit}}}$. A truth-table expansion of each decoded equality gives depth at most 22 and size at most $\exp(U^{17})$. Every collection of at most $t$ bits belongs to at most $t$ primes, so the positive reweighted law and the original law both have all nonconstant Walsh coefficients of degrees $\leq t$ bounded in modulus by $O(lM_*/N + \delta)$. Let their total absolute mass through degree $t$ be $q_0$. If $f$ denotes the density of one of these laws relative to uniform bits and $L_f$ its nonconstant low-order Walsh polynomial, then

$$(f - L_f + q_0)/(1 + q_0)$$

is a nonnegative density with those coefficients identically zero, at total variation distance $O(q_0)$ from $f$. This is the almost-to-exact independence correction of Alon, Goldreich and Mansour [1], in the Fourier form used in [26, Lemma 3.1]. Choose $C_1$ for Braverman's theorem [5, Corollary 2] at depth 22, size $\exp(U^{17})$ and error $e^{-U^{11}}$. The displayed hypothesis bounds $q_0$ by $O(e^{-2U^{11}})$. Both the original and reweighted probabilities of $B$ are consequently within $O(e^{-U^{10}})$ of the independent probability. Subtract after restoring the normalizing denominator. The extra term is $O(\delta)$ and the two choices of $\omega$ recover the real and imaginary parts. □

**Candidate proposition 6.4** (Signed transfer for growing quadratic walks)**.** *Use the supplies and numerical words above, with $U = (\log X)^{\epsilon}$, $\epsilon A < 1$ and $\epsilon(C_1 + 15) < 1$. Uniformly for $\alpha, \theta \in \mathbb{R}$, one can choose the lifted edge phase and an exterior modulus $l \le X^{1/10+o(1)}$ as below so that the marked word averages admit the unmarked signed product-law majorants. The sum of the comparison errors over all scalar expansions and words in [26, (5.5)] is $O(e^{-U^9})$. This holds for both Hermitian families in Proposition 4.1, after omitting the primes dividing $l$.*

*Proposed proof.* Use Lemma 6.3 with $M_* = \exp(O(U^{C_1+1}))$ and $\delta = E_0^{-1}$. All tested primes belong to the supplies in Section 4, so its size restrictions hold. Its quantitative hypothesis follows from the chosen scales. If a word instead has slow phase on subintervals with at least $X^{1/30}$ points *in its exterior class*, freeze its phase there to error $O(X^{-c})$ for some fixed $c > 0$. The unweighted circuit comparison on each subinterval has the required accuracy. In applying it, the substitution $x = r + ly$ permutes residues at every remaining prime and merely translates the equality gates. Subdivide into comparable lengths to absorb a final short remainder. In this case the word is its signed independent expectation times an average of unit scalars, up to the comparison error. In the oscillatory case Lemma 6.3 bounds the word by expanding it in the circuit indicators of [26, (5.5)]. The total absolute coefficient budget is $\exp(O(U^5))$, which both errors absorb. The factor of a rational phase fixed on an exterior class is a scalar. Averaging over those classes costs no factor $l$. Removing the primes dividing $l$ costs $O(\log X/H_0)$ high-band mass, while the padding loss is the one already accounted for at (13).

The following cases check that one of the two comparisons always applies. By the inverse form of Weyl's estimate [16, Lemma 4.4], a quadratic sum of relative size $> u$ on $K$ consecutive integers has leading coefficient $\xi$ satisfying $K^2\|t\xi\|_{\mathbb{T}} \ll u^{-C}$ for some integer $1 \le t \ll u^{-C}$, with absolute $C$. Changing the initial point preserves the leading coefficient.

**No small denominator at precision $X^{-1}$.** Suppose $X\|b\alpha\|_{\mathbb{T}} > W_2$ for every integer $1 \le b \le W_2$. First suppose also $X\|j\alpha + k\theta\|_{\mathbb{T}} > W_1$ for every nonzero integer pair $|j|, |k| \le W_1$. Use the unaltered phase $\alpha(x/w)^2 + \theta x/w$. If $R_2 = u/v \ne 0$, a walk average $> E_0^{-1}$ on step $M$ forces

$$X^2\|tuM^2\alpha\|_{\mathbb{T}} \ll E_0^C M^2 v < W_1, \qquad 1 \le |tuM^2| < W_1.$$

If $R_2 = 0$ but $(R', R_1) \ne (0, 0)$, the geometric sum similarly yields a forbidden small relation between $\alpha$ and $\theta$. All other walks have constant phase.

If such a small relation exists, its $k$ is nonzero. Change signs if necessary and write $\theta = -(j/k)\alpha + a/k + \beta_0$ with $k > 0$, $kX|\beta_0| \le W_1$. Lift $a/k$ on the edge by the inverse of $w$ modulo $k$. On a fixed class of $k$ the preceding quadratic test now uses $kM$ and heights $W_1^{O(1)} < W_2$. If $R_2 = 0$ the remaining slope is $S\alpha + R_1\beta_0$ with $S = R' - (j/k)R_1$. For $S = u/v \ne 0$ a large geometric sum implies $X\|kMu\alpha\|_{\mathbb{T}} \ll vE_0^{O(1)}kM + vMW_1|R_1| < W_2$, again impossible. For $S = 0$ the drift is constant to $O(X^{-1/3})$ on blocks of length $kX^{1/2}$. This gives the comparison in all cases.

**A small denominator at precision $X^{-1}$.** Write $\alpha = a/b + \beta$ with $(a,b) = 1$, $b \le W_2$ and $bX|\beta| \le W_2$. By Dirichlet take $\theta = d/s + \beta_1$ with $s \le X^{1/10}$ and $s|\beta_1| \le 2X^{-1/10}$. The upward phase on step $w$ is

$$a\bar{w}^2x^2/b + d\bar{w}x/s + \beta(x/w)^2 + \beta_1 x/w.$$

Use $l = \operatorname{lcm}(b,s)$ with inverses at each modulus. On an $l$-class the walk phase has coefficients $\beta R_2$ and $\beta R' + \beta_1 R_1$. If $X^2|\beta R_2| > X^{1/50}$, a large sum on step $lM$ gives $X^2\|t\beta R_2(lM)^2\|_{\mathbb{T}} \ll E_0^C(lM)^2$. Its argument has absolute value $\le X^{-4/5+o(1)}$, hence $X^2|\beta R_2| \ll E_0^C$, impossible. Otherwise freeze the drift on blocks of length $lX^{1/20}$. Its three variations are bounded respectively by $X^{-1+1/50+o(1)}lX^{1/20}$, $X^{-1+o(1)}lX^{1/20}$ and $O(UbX^{-1/20})$. These are all $O(X^{-c})$ for some fixed $c > 0$. Every word therefore falls into the oscillatory or slow alternative. Summing the errors with the coefficient budget above gives $O(e^{-U^9})$, uniformly in the two coefficients and in the exterior classes. □

Partially divided edges, which supply the second hypothesis of Proposition 4.1, are estimated next.

### 6.3 A fourth moment with quadratic reciprocals

In the first case a rectangle has leading coefficient $\alpha\Delta_2$. Writing $S = x+y = z+w$, $u = xy$, $v = zw$, the following identity gives its numerator.

$$\Delta_2 = \frac{(v-u)[S^2(u+v) - 2uv]}{u^2v^2}.$$

Equation (12) gives $|\Delta_2| \asymp |x-z||y-z|/D^4$. A nondiagonal average exceeding $U^{-12}$ at length $Y \gg XD$ forces an approximation to $\alpha$ of denominator $\le U^{O(1)}(2D)^{O(1)} < W_2$ and error in the multiple $\ll U^{O(1)}(2D)^{O(1)}/Y^2 < W_2/X$. This contradicts the hypothesis since $D \le e^{2U}$. Diagonals cost $O(D^{-1})$ before the fourth root.

In the second case put $B_2 = X^2|\beta|$. The leading coefficient is $a\mathcal{D}/b + \beta\Delta_2$, where $\mathcal{D} = \bar{x}^2 + \bar{y}^2 - \bar{z}^2 - \bar{w}^2 \pmod b$. A rectangle average exceeding $U^{-12}$ forces, for some $1 \le t \le U^{C_0}$,

$$b \mid t\mathcal{D}, \qquad B_2|x-z||y-z|/D^2 \ll U^{C_0}. \tag{14}$$

Indeed $tb|\beta\Delta_2| = o(1)$ and $bU^{C_0}/Y^2 = o(1)$, so the nearest integer after clearing the denominator must account exactly for the rational part. When $B_2 > U^{K_0}$ for $K_0 > C_0 + 32$, remove $|x-z|$ or $|y-z| \le DU^{-15}$ at cost $O(U^{-15} + D^{-1})$. The second condition in (14) then fails.

For a large denominator we use the following specialized consequence of the polynomial congruence theory of Konyagin and Steger [21]. We include a direct Vandermonde proof: it records the reciprocal reduction and the content bound needed for composite moduli and repeated prime factors.

**Lemma 6.5** (Even reciprocal powers modulo a composite integer)**.** *Fix an even integer $s \ge 2$, $q \ge 1$ and integers $x, y$ coprime to $q$. Write $S = x + y$, $w = S - z$. In an interval of length $D \ge 1$, the number of $z$ for which $(zw, q) = 1$ and*

$$\bar{x}^{\,s} + \bar{y}^{\,s} \equiv \bar{z}^{\,s} + \bar{w}^{\,s} \pmod q$$

*is $O_s(1 + Dq^{-1/[s(2s+1)]})$. Bars denote modular inverses.*

*Proof.* Clear denominators. Every admissible $z$ is a root of the integer polynomial

$$(x^s + y^s)z^s(S-z)^s - (xy)^s\{z^s + (S-z)^s\}.$$

Its degree is at most $2s$. Its leading coefficient is $x^s + y^s$ and its constant coefficient is $-(xy)^s S^s$. An odd prime dividing $q$ and both coefficients would divide $S$, forcing $x^s + y^s \equiv 2x^s \neq 0$ at this prime. If $2 \mid q$, $x, y$ are odd and $x^s + y^s \equiv 2 \bmod 4$. The gcd of $q$ with the content is therefore at most two. For $2s+1$ distinct roots the adjugate identity for their Vandermonde matrix forces $q$ to divide the determinant times each coefficient. Thus the determinant is divisible by an integer at least $q/2$. Its absolute value is at most the diameter of the roots raised to $s(2s+1)$. Cover the interval by intervals of diameter strictly less than $(q/2)^{1/[s(2s+1)]}$, using the trivial bound for bounded $q$. Each contains at most $2s$ roots. □

For $b > U^{K_0}$ apply Lemma 6.5 with $s = 2$ and $q = b/(b,t)$ to the first congruence in (14). At fixed $x, y, t$ the proportion is $O(D^{-1} + (2t/b)^{1/10})$. Summing over $t \le U^{C_0}$ gives $O(U^{-12})$ when $K_0 > 20C_0 + 200$, since $D \ge H_0/2$. The fourth root of the full estimate is $O(U^{-3} + D^{-1/4})$, sufficient in (11). The linear coefficient of a rectangle played no role in this test.

For the remaining $b, B_2 \le U^{K_0}$ open $\mathrm{e}(an^2/b)$ in linear characters with total coefficient mass at most $b$. Partial summation removes $\mathrm{e}(\beta n^2)$ at cost $O(1 + B_2)$. Apply the linear estimate with the shifted frequency on prefixes $\ge \sqrt{X}$ and use the trivial bound below. Taking $2K_0\epsilon < c_F/2$ proves the candidate rate. This last choice is made after $c_F$ has been obtained at its own, independent graph scale.

**Candidate proposition 6.6** (Large-prime marks fixed across prefixes)**.** *There are absolute constants $c_m, \sigma_m > 0$, $C_m$ and $V_0$ with the following property. Fix $V \ge V_0$ and completely multiplicative functions $s_i$ with $|s_i| \le 1$ and*

$$s_i(p) = 1 \quad \text{for every prime } p \notin [V^{0.39}, 3V].$$

*Set $g_i = \lambda^{e_i} s_i$, where $e_i \in \{0,1\}$ and $e_1 + e_2 > 0$. For every $\sqrt{V} \le Y \le 2V$,*

$$\sup_{\alpha,\theta\in\mathbb{R}} \left| Y^{-1} \sum_{n\le Y} g_1(n) g_2(n+1) \mathrm{e}(\alpha n^2 + \theta n) \right| \le C_m (\log V)^{-c_m}, \tag{15}$$

$$\sup_{\substack{b,k\in\mathbb{N}\\ b,k\le(\log Y)^{\sigma_m}}} \sup_{|\psi|\le 1} \left| Y^{-1} \sum_{n\le Y} g_1(n) g_2(n+k) \psi(n \bmod b) \right| \le C_m (\log V)^{-c_m}. \tag{16}$$

*Here $\psi : \mathbb{Z}/b\mathbb{Z} \to \mathbb{C}$. The same pair of marks, chosen using $V$, is used for every prefix $Y$; all constants are independent of these marks and of the displayed parameters. The linear estimate is the case $\alpha = 0$ of (15). Subtracting the prefixes at $V$ and $2V$ gives the corresponding dyadic bounds.*

*Proposed proof.* For graph scales attached to $\sqrt{V} \le Y \le 2V$, every dilation $w \le e^{101U}$ is less than $V^{0.39}$ for sufficiently large $V$. Thus $g_i(w) = \lambda(w)^{e_i}$ exactly. Their product is compensated by the edge sign, and partial division leaves a sign depending only on $z$. The trace comparison in Proposition 6.4 and the rectangle argument are unchanged.

For the periodic input, one endpoint in the short-interval Fourier estimate of [26, (4.4)–(4.7)] is $\lambda s_i$. Replacing $\lambda$ by this function changes the squared pretentious distance to any test by at most

$$2 \sum_{V^{0.39}\le p\le 3V} p^{-1} = O(1),$$

uniformly on every prefix. Hence the saving from [23, Theorem 1.7 and (1.12), corrected text] is preserved. If the Liouville endpoint is second, exchange the two vectors. Under $v' = v + kz$ the test

becomes $\psi(\bar{z}v' - k \bmod b)$ and the other endpoint moves by $-kz$, with the same enlarged interval, Haar and Parseval bounds. All residual signs in a fixed word have modulus one. The proofs of Propositions 6.1 and 6.2 therefore apply with the frozen marks; the quadratic argument then gives (15). Shrinking the positive exponents to common $c_m, \sigma_m$ gives the stated simultaneous bounds. The relation $\log Y \asymp \log V$ makes their constants uniform over the prefix range. □

**Order of constants.** The choices used above can be made in the following order. The linear saving is fixed before the quadratic graph scale is chosen.

| Choice | Required condition and use |
|---|---|
| $C_1, A$, then $W$ | $C_1 \geq 30$ is sufficient for Braverman's theorem at depth 22; $A > C_1 + 20$; then $e^5 C/\sqrt{W} < e^{-1}$. |
| $c_0, \sigma$ | Obtain the periodic saving from Proposition 6.1 at scale $U_0 = (\log X)^{1/A}$. |
| $\epsilon_F$, then $c_F$ | $\epsilon_F A < 1$ and $200\epsilon_F < \min(c_0, \sigma)/3$; apply Proposition 6.2. |
| $C_0$, then $K_0$ | $C_0$ is the absolute Weyl exponent in (14); choose $K_0 > 20C_0 + 200$. |
| Quadratic $\epsilon$ | $\epsilon A < 1$, $\epsilon(C_1 + 15) < 1$, and $2K_0\epsilon < c_F/2$. These ensure the finite-law threshold, $W_2 = X^{o(1)}$, and absorption of the final rational drift. |

All restrictions are upper bounds on a positive exponent chosen last, so they are compatible. The marked version uses the same order of choices.

# 7 Sign laws under translations

Let $V$ tend to infinity through the positive integers, $L = \log V$ and $T = \log\log V$. An origin is an integer $V < v \leq 2V$. We first choose a population of origins and two signs attached to each origin, called its anchors. We then ask when signs at translated sites become independent of those anchors. In an assertion for almost all choices from a specified finite family, the exceptional proportion tends to zero. Equivalently, for every fixed accuracy the proportion failing that accuracy tends to zero. The order of sampling matters even for two signs.

The four origin colors in (17) have equal limiting frequencies. On the full dyad this follows from fixed two-point cancellation. On the two thinner populations the cofactor theorem cited below gives a quantitative error. The inputs for the shifted signs depend on how the shift is sampled.

| Sampling and shifted signs | Input beyond the anchor law | Status |
|---|---|---|
| Independent short shifts; two full signs | Candidate Proposition 6.1 and Bessel's inequality | Conditional on the periodic estimate |
| Short common shift; one full and one truncated sign | [23, Proposition 5.3 and (5.5)], with the origin majorant | Proved on the full dyad and on $E_1, E_2$ |
| Long common shift; two full signs | Candidate Proposition 6.2 and Parseval's identity | Conditional on the linear estimate |
| Prime-power shifts averaged at almost every origin; two full signs | Fixed progression means [26] and prime exponential sums [8] | Unconditional on the full dyad |
| Origins averaged at almost every prime-power shift; two full signs | Candidate Proposition 6.2 and the same prime multiplier | Conditional on the linear estimate |

### 7.1 The origin populations

On the full dyadic interval take the anchor color to be $(\lambda(v), \lambda(v-1))$. The means of its three nontrivial sign characters tend to zero, by the prime number theorem for each single sign and the ordinary two-point theorem for the product [26]. We also use two thinner populations. Write $v = wt$ and $v - 1 = qm$, with $q$ prime and $\log q/L \in [0.5053, 0.5057]$. For $E_1$, $w = p$ is prime with exponent $\log p/L \in [1 - 2/T, 1 - 1/T]$. For $E_2$, $w = p_1p_2$ with one prime exponent in each half of $[1/2 - 2/T, 1/2 - 1/T]$, assigning the midpoint only to the upper half. The anchor color is $(\lambda(t), \lambda(m))$. The prime windows make $w, q$ unique for large $V$. The anchor signs are the two full signs with fixed sign switches.

Write $\mathcal{A}_S(v)$ for the indicated anchor pair on a population $S$. The origin input to the sign laws is

$$\frac{\#\{v \in S : \mathcal{A}_S(v) = (\epsilon_1, \epsilon_2)\}}{|S|} = \frac{1}{4} + o(1) \qquad ((\epsilon_1, \epsilon_2) \in \{\pm 1\}^2). \tag{17}$$

For $S = E_1, E_2$ the signed input is available on the same windows and under uniform sampling of the integers in $S$. We first record their mass.

**Lemma 7.1** (Mass of the two windows)**.** *For $a = 1, 2$ one has $|E_a|/V \sim c_Q T^{-a}$, where $c_Q = \log(0.5057/0.5053)$.*

*Proof.* Use Fouvry and Radziwiłł's *Level of distribution of unbalanced convolutions* [12, Corollary 1.1(i)]. On dyadic boxes write $w \asymp M$, $t \asymp N$, $q \asymp Q$. The coefficient at $w$ is the indicator of its prescribed prime or prime-product window. The short coefficient at $t$ is one, so satisfies the Siegel–Walfisz condition by counting residue classes (with coprimality restrictions, use inclusion–exclusion). The shift is 1, $V < wt \le 2V$, $\log N \asymp L/T$, and $Q \ll V^{0.5057}$. For small fixed $\varepsilon > 0$ these imply

$$\exp(L^\varepsilon) \le N \le Q^{-11/12} V^{17/36 - \varepsilon}, \qquad 0.5057 < 17/33.$$

The corollary sums the absolute discrepancy over a dyadic modulus interval. We may retain just the primes in the required range. Its product cutoff is already $V < wt \le 2V$. Sum the errors over $O(L^3)$ dyadic boxes at an arbitrarily large fixed logarithmic saving. All $t$ are smaller than $q$, and the prime windows at $w$ are disjoint from the $q$ window. The remaining density term is therefore

$$\sum_q \frac{1}{q-1} \sum_w \#\{t : V < wt \le 2V\}.$$

Counting $t$ costs $O(1)$ per $w$, with $\#\{w\} \le V^{1-1/T}$. Partial summation with the prime number theorem gives $\sum_q (q-1)^{-1} \sim c_Q$ and reciprocal prime mass $\sim 1/T$ in the first window for $E_1$ and in each of the two windows for $E_2$. Since each origin occurs only once, the formula follows. □

For the signs we use Theorem 1.1 of *Joint cofactor signs in prime-factor windows on consecutive integers* [29]. Its dyad, prime windows (including the half-open split), short cofactor $t$ and opposite cofactor $m$ are precisely those above. For $a = 1, 2$, a sign pair $c$, and complex marks $A(w), B(q)$ of modulus at most one, that theorem gives

$$\left| \sum_{v \in E_a} A(w) B(q) (\mathbf{1}_{\mathcal{A}_{E_a}(v) = c} - \tfrac{1}{4}) \right| \ll V L^{-\kappa}$$

with an absolute $\kappa > 0$. Take $A = B = 1$ and divide by the mass in Lemma 7.1 to obtain (17) with error $O(T^a L^{-\kappa})$. The marks depend separately on the two selected factors. The Fourier coefficients in Section 8 can depend on the whole origin.

### 7.2 Independent shifts in a short interval

For populations of density bounded below and an interval of shifts whose length tends to infinity, the qualitative independent-shift conclusion is already covered by Matomäki, Radziwiłł and Tao [23, Theorem 1.6], which allows the other factors to be arbitrary bounded functions. The estimate below records the density loss explicitly and permits any set of distinct shifts of the stated diameter.

Assume Candidate Proposition 6.1. Let $S$ be any population of density $\rho = |S|/V > 0$, and let $A_v$ have modulus at most one. For distinct shifts $i$ in a set $I$ of size $H$, put $b_i(v) = \lambda(3v + s + i)$, assuming $3v + s + i \in [V, 10V]$ throughout the dyad. Suppose $\operatorname{diam} I \le (\log V)^{\sigma'}$, where $0 < \sigma' < \sigma$ in (13). With normalized counting measure on the full dyad, the Gram matrix of the $b_i$ has diagonal one and off-diagonal entries $O(L^{-c_0})$. For a distinct pair set $n = 3v + s + \min(i, j)$ and subtract two prefixes in one class modulo three in (13). The positive shift is $|i - j|$ and all cutoff lengths are comparable to $V$. The Gram matrix has operator norm at most $1 + O(HL^{-c_0})$ by its maximum row sum. Apply its dual inequality to $\mathbf{1}_S\overline{A_v}$, whose squared norm is at most $\rho$, to obtain

$$\frac{1}{H}\sum_{i\in I}\left|\frac{1}{|S|}\sum_{v\in S}A_v b_i(v)\right|^2 \ll \frac{1}{\rho}(H^{-1} + L^{-c_0}). \tag{18}$$

If $\sigma' < \min(c_0, \sigma)/2$, the first term has coefficient $1 + o(1)$ and absorbs the second. To test two independently chosen shifts, fix one and include its sign in $A_v$ before averaging over the other. The estimate is uniform in that choice. Thus almost all ordered pairs of shifts give vanishing mixed averages when $H\rho \to \infty$ in this range. Under the anchor law (17), character inversion gives the sixteen frequencies $1/16 + o(1)$. The only candidate input used here is Proposition 6.1. For two signs at the same shifted pair of sites the weight would vary with the averaging index.

### 7.3 A short common shift with a truncated fourth sign

Write $n_y = \prod_{p\le y} p^{\nu_p(n)}$. Let $S$ be the full dyad ($a = 0$) or $E_a$ with $a = 1, 2$, with the origin law (17) just established. Put $N_h = 3v - 2 + h$. For $0 \le h_0 \le V$, take $h_0 < h \le h_0 + H$, $H \le \exp(T^2)$, and put $P = \min(T, \log H)$. If $y \to \infty$ and

$$a\log T + (\log y)(\log\log(3y))^2 = o(P),$$

then for almost all these shifts the two anchors, $\lambda(N_h)$ and $\lambda((N_h + 1)_y)$ have their sixteen patterns with frequencies $1/16 + o(1)$. On $E_1, E_2$ one may take $H = \lceil\exp((\log T)^2)\rceil$ and $y = \exp((\log T)^{3/2})$. On the full dyad $H$ may tend to infinity arbitrarily slowly, with $y$ chosen accordingly. The last sign counts only primes up to $y$.

Let $B_a$ be $E_a$ without the condition on $v - 1$, and $B_0 = S$ on the full dyad. The prescribed product $w$ at $v$ is unique and $w \le V^{1-1/T}$. Writing $v = wt$, counting by $w$ and using the prime number theorem in the indicated intervals gives $|B_a| \asymp V/T^a$. The harmonic mass of $w$ is $\asymp T^{-a}$ and the total endpoint error is at most $V^{1-1/T}$. For $b \le \exp(3L^{0.7})$ all selected primes are coprime to $b$. Counting the classes of $t$ gives

$$\Pr_{B_a}(v \equiv r \bmod b) = 1/b + O(V^{1-1/T}/|B_a|). \tag{19}$$

By Lemma 7.1 (and trivially for $a = 0$), $|B_a|/|S| = O(1)$. Nonnegative errors pass from $B_a$ to $S$ at bounded cost.

Put $Z = y^{(\log\log(3y))^2}$, $A_0 = 1 + a\log T + \log Z + 100\log\log(3H)$ and $W_0 = \exp(\sqrt{PA_0})$. Thus $A_0 = o(P)$ and $A_0 = o(\log W_0)$. Apply [23, Definition 2.1, Proposition 5.3 and (5.5)] to the

multiplicative function $\lambda$ at $X = 10V$. Write $g = \lambda \mathbf{1}_{\mathcal{S}} \mathbf{1}_{[1,10V]}$ for its sifted restriction. The parameters satisfy $(\log H)^{20} \le W_0 \le \min(H^{1/250}, (\log(10V))^{1/125})$ and the distance condition by [23, (1.12)]. The mean-square estimate for its correlations yields, for all $|b_h| \le 1$,

$$\sum_n \left| \sum_{h=h_0+1}^{h_0+H} b_h g(n+h) \right|^2 \ll HV + H^2 V W_0^{-1/10}. \tag{20}$$

Translation by $h_0$ disappears in the full square sum. Truncating both copies of the sifted sequence at $10V$ changes an individual lag correlation by $O(H)$, negligible in this range. The diagonal costs $O(HV)$ and the sum of squared positive-lag correlations up to $H$ is $O(HV^2 W_0^{-1/5})$. Removing integers missing a required sieve band costs $O(\log W_0 / \log H) + o(1)$ in $B_a$, uniformly in $h$. Here the sieve set $\mathcal{S}$ consists of integers having a prime in every retained band $[P_i, Q_i]$. For $i \ge 2$ the definition uses $\log P_i = i^{4i} (\log Q_1)^{i-1} \log P_1$ and $\log Q_i = i^{4i+2} (\log Q_1)^i$, up to the cutoff of [23, Definition 2.1]. Use the upper Bonferroni sum of even order $k = 2\lfloor L^{0.1}/2 \rfloor$ for each band in that definition. Each prime is at most $\exp(\sqrt{\log(10V)})$ and exceeds three. These $\exp(O(L^{0.6}))$ products give single residue classes by (19). The next independent term is at most $(CT)^k/k!$ by Mertens. The independent missing-band probabilities are $\ll \log P_i / \log Q_i$, whose sum is $O(\log W_0 / \log H)$ since their ratios equal $i^{-2} \log P_1 / \log Q_1$, $P_1 = W_0^{200}$ and $Q_1 = H/W_0^3$.

For each $h$, let $G_h$ be a function of $((N_h)_y, (N_h+1)_y)$ of modulus at most one. Double Möbius inversion expresses it through $b_1 \mid N_h$, $b_2 \mid N_h + 1$ with $y$-smooth $b_i$ and coefficients bounded by $\tau(b_1)\tau(b_2)$, where $\tau(b)$ is the divisor count. Truncation to $b_i \le Z$ costs $o(1)$ in mean, uniformly in $h$. Indeed, if $M = 3v + c \in [1, 10V]$, $s = 1/\log y$ and $k$ is fixed, then

$$\mathbb{E}_{B_a} \tau(M_y)^k M_y^s \ll_k (\log y)^{C_k}.$$

Expand by positive divisor coefficients $u$ supported on $y$-smooth integers, $u(p^i) \le (i+1)^k p^{is}$. For $b \le \exp(L^{0.6})$, (19) bounds the divisibility mass by $3/b$ plus its negligible remainder. For larger $b$, inject $v \mapsto M$ to bound it by $O(T^a/b)$, then use Rankin with $b^s$ to bound their cost by $O(T^a e^{-sL^{0.6}} \sum_b u(b) b^{-1+s})$. Both Euler products cost at most powers of $\log y$. On the exceptional set where either smooth part exceeds $Z$, the truncated sum is bounded by the product of squared divisor counts. Cauchy–Schwarz and the moment bound give $O((\log y)^C e^{-s \log Z/2}) = o(1)$.

For fixed $b_1, b_2$ split $3v - 2$ modulo $q = \operatorname{lcm}(b_1, b_2)$. On each class the divisibility tests depend on $h$ only. Align the phases of the sums over $v$, apply (20), then use Cauchy–Schwarz within and across the $q$ classes. With any $|A_v| \le 1$ fixed before $h$, each divisor pair costs at most

$$\ll \tau(b_1)\tau(b_2) \sqrt{qV/|S|}\, (H^{-1/2} + W_0^{-1/20}).$$

The total is $O(Z^{3+o(1)} T^{a/2} (H^{-1/2} + W_0^{-1/20})) = o(1)$. Hence $|\mathbb{E}_S A_v \lambda(N_h) G_h|$ tends to zero in mean over shifts. This controls the sign characters containing the entire shifted sign.

For the shifted smooth-part sign alone, align phases in the shift average and apply Cauchy–Schwarz in $v$, extending the square to $B_a$. For $j = |h - h'| > 0$, truncation and (19) compare the two smooth-part signs with their expectation under independent uniform prime-adic coordinates to $o(1)$ uniformly. For $p \ge 5$, $p \nmid j$, the local factor is $1 - 4/(p+1)$ since the two sites cannot both have positive valuation. For unweighted integers, Cassaigne, Ferenczi, Mauduit, Rivat and Sárközy [6, Theorem 5] treat the two truncated signs on $n \le x$ at lags $j \le y \le (\log x / \log\log x)^2$. Daboussi and Sárközy [10, (1.10), Corollaries 1–2] give this local factor at lag one. Here the affine slope is three and the lag can exceed $y$. At $p = 3$ or $p \mid j$ the local expectation has modulus at most one, as does the factor at $p = 2$. The product is bounded by $O((\log y)^{-4} \prod_{p \mid j,\ p \ge 5} (1 + C/p)) + o(1)$. Its

mean in $j$ is $o(1)$ by $\prod_p(1+C/p^2)<\infty$. Diagonal shifts cost $O(1/H)$. Inversion over the four signs proves the law without any polynomial estimate for the pair.

**Corollary 7.2** (A common shift with separate prime observations)**.** *In the short common-shift regime above let $a=1,2$, let $c,c'\in\{\pm1\}^2$, and put $\mathcal{B}_h(v)=(\lambda(N_h),\lambda((N_h+1)_y))$. Choose the same two prime marks for all shifts in the interval. Then*

$$\sup_{|A|,|B|\le1}\frac{1}{H}\sum_{h_0<h\le h_0+H}\Big|\mathbb{E}_{v\in E_a}\Big[A(w)B(q)\Big(\mathbf{1}_{\mathcal{A}_{E_a}=c,\,\mathcal{B}_h=c'}-\tfrac{1}{16}\Big)\Big]\Big|=o(1).$$

*Proof.* Write $W(v)=A(w)B(q)$. The difference inside the expectation is the sum of $W\mathbf{1}_{\mathcal{A}=c}(\mathbf{1}_{\mathcal{B}_h=c'}-1/4)$ and $W(\mathbf{1}_{\mathcal{A}=c}-1/4)/4$, with $\mathcal{A}=\mathcal{A}_{E_a}$. For the first term expand the shifted indicator in its three nonconstant sign characters. The two preceding mean-over-shift estimates apply with the fixed origin weight $W\mathbf{1}_{\mathcal{A}=c}$ and are uniform over its unit ball. For the second term the separate-marks bound above costs $O(T^aL^{-\kappa})$. Both bounds are absolute. Taking $A=B=1$ gives all sixteen frequencies for almost all shifts. For $W\ge0$ with $\liminf_V\mathbb{E}_{E_a}W>0$, division by $\mathbb{E}_{E_a}W$ gives the corresponding weighted frequencies. □

## 7.4 A long common shift with two full signs

Let $S\subset(V,2V]\cap\mathbb{Z}$ have $|S|=\rho V>0$. Choose complex numbers $|A_v|\le1$ before the shift. For a bounded $g$ supported on $[V,8V]$, put $\delta=V^{-1}\sup_\theta|\sum_n g(n)\mathrm{e}(\theta n)|$. For any $H$ distinct integer shifts $h\in[-V,V]$,

$$\left(\frac{1}{H}\sum_h\left|\frac{1}{|S|}\sum_{v\in S}A_v g(3v-2+h)\right|^2\right)^{1/2}\ll\delta\sqrt{\frac{V}{H\rho}}.\tag{21}$$

Embed the sums in $\mathbb{Z}/p\mathbb{Z}$ for a prime $16V<p<32V$ and set $f(3v-2)=pA_v/|S|$ on $S$, zero elsewhere. For the normalized Fourier transform the full shift mean square is $\sum_\xi|\widehat{f}(-\xi)|^2|\widehat{g}(\xi)|^2\ll\delta^2p/|S|$ by Parseval. Restriction to $H$ shifts costs $p/H$. There is no wraparound on the stated supports.

Take $g$ to be each of $\lambda(n)$, $\lambda(n+1)$ and $\lambda(n)\lambda(n+1)$ restricted to this interval. For the first two, Green and Tao's *The Möbius function is strongly orthogonal to nilsequences* [17] and $\lambda(n)=\sum_{j^2m=n}\mu(m)$ give any logarithmic saving. For the third, assume Candidate Proposition 6.2. The indicator of any prescribed pair of signs is one quarter times the sum of its four sign characters. Under (17), applying (21) to each of the four origin characters gives all sixteen joint frequencies $1/16+o(1)$ for almost all shifts whenever $\delta^2V/(H\rho)=o(1)$. With these inputs, this works in particular with $\rho\asymp(\log\log V)^{-a}$ for fixed $a$ and $H\ge V/(\log V)^\kappa$ if $0<\kappa<c_F$. It uses the entire signs at both shifted sites and admits arbitrary origin weights independent of $h$.

## 7.5 Prime-power shifts and the order of conditioning

A sparse family of translations on the full dyad admits two entire shifted signs using only the fixed-progression two-point theorem [26]. Fix $d\ge1$, $|b|\le V$, $Y=V^{1/d}/2$, and sample a prime $Y<p\le2Y$. For all but $o_{d,\varepsilon}(V)$ origins $v\in(V,2V]$, each pair

$$(\lambda(3v+1+b+p^d),\lambda(3v+2+b+p^d))$$

has frequency within $\varepsilon$ of $1/4$ over $p$. The exceptional set can depend on $b$, with a uniform size bound. Adjoining the origin signs $(\lambda(v),\lambda(v+1))$ gives sixteen frequencies $1/16+o(1)$ in the joint sampling of $(v,p)$.

Write $m(\alpha)=\sum_{Y<p\le 2Y}(\log p)\mathrm{e}(\alpha p^d)/\sum_{Y<p\le 2Y}\log p$ and fix $0<\eta<1/2$. The minor-arc and major-arc estimates of Chapman and Chow, *Arithmetic Ramsey theory over the primes* [8, Lemmas 2.2, 4.2 and 4.3], imply that $|m(\alpha)|>\eta$ requires

$$\alpha=a/q+\beta,\quad (q,a)=1,\quad q\ll_d \eta^{-2d},\qquad V|\beta|\ll_d \eta^{-1}.$$

Use their polynomial $t^d$ on the unrestricted progression, subtracting two prime prefixes for the dyad. Outside the polylogarithmic major arcs the normalized sum tends to zero. Inside them, Siegel–Walfisz and partial summation give the product of the normalized complete sum of $\mathrm{e}(au^d/q)$ over units and the normalized integral of $\mathrm{e}(\beta t^d)$ over $[Y,2Y]$, up to $o(1)$. For $d=1$ the complete sum divided by $\varphi(q)$ is the Ramanujan sum $\mu(q)/\varphi(q)$. For $d\ge 2$, [8, Lemma 2.2] gives $O_d(q^{-1/(2d)})$ after absorbing $C_d^{\omega(q)}q/\varphi(q)$ into a power of $q$. In degree two this also follows by inclusion–exclusion for the unit restriction and the ordinary quadratic Gauss sum bound. In all degrees integration by parts gives $O_d((1+V|\beta|)^{-1})$ for the second factor. Both factors also have absolute bound $1+o(1)$.

Embed in a cyclic group of prime order $32V<M<64V$. There are $O_d(\eta^{-(4d+1)})$ exceptional frequencies on the grid $\xi/M$. For each of the three sign characters $g=\lambda,\lambda(\cdot+1),\lambda(\cdot)\lambda(\cdot+1)$ restricted to $[V,9V]$, the progression theorem and partial summation give $\widehat{g}(\xi)=o_{d,\eta}(1)$ on this set. Parseval gives $\|\mathbb{E}_p^{\log}g(\cdot+p^d)\|_2^2\ll\eta^2+o_{d,\eta}(1)$. Let $V\to\infty$ then $\eta\to 0$. Sampling $3v+1+b$ costs the bounded density factor $M/V$. The logarithmic and uniform prime measures differ in total variation by $O(1/\log Y)$ on this dyad. Sign inversion and Chebyshev prove the claim.

The opposite conditioning, fixing almost every prime and averaging in $v$, assumes Candidate Proposition 6.2. For $|A_v|\le 1$ put $R(h)=V^{-1}\sum_{V<v\le 2V}A_v g(3v+1+h)$ on this group. As in (21), Cauchy–Schwarz and Parseval give $\|\widehat{R}\|_1\ll 1$ and $\|R\|_2^2\ll\delta^2$ for the Fourier supremum $\delta$. Thus $D=|R|^2$ satisfies $\|\widehat{D}\|_1\ll 1$, $\|\widehat{D}\|_\infty\ll\delta^2$. Splitting against $m$ at height $\eta$ bounds $\mathbb{E}_p^{\log}|R(b+p^d)|^2$ by $O_d(\eta+\eta^{-(4d+1)}\delta^2)$. Taking the two limits in the same order suffices. Choosing $A_v$ to be the origin characters proves the reverse almost-everywhere claim if $\delta\to 0$.

# 8 Equal sums and equal sums of squares

Two independently shifted origins need not satisfy any additive relation. We now sample the origins subject to one or two exact moment equations. Polynomial Fourier coefficients enter directly through orthogonality. We first keep the color function arbitrary, so that the counting bounds depend only on its Fourier coefficients, and then apply these bounds with the precise Liouville inputs. Tuples are ordered and their origins are pairwise distinct. For colors built from $(\lambda(v-1),\lambda(v))$, distinct origins can still give overlapping adjacent sites; the distinctness condition is imposed on the origins.

## 8.1 An elementary count at four and ten origins

Let $E\subset(V,2V]$ have density $\rho$ and color $C(v)\in\{\pm1\}^2$. For $j=1,2$ write $\eta_j$ for the largest normalized polynomial Fourier coefficient of the three sequences $\mathbf{1}_E\chi(C)$, over nonconstant sign characters $\chi$ and polynomials of degree at most $j$. The indicator $I_c$ of each color satisfies $\sup_{\deg P\le j}|\sum(I_c-\mathbf{1}_E/4)\mathrm{e}(P)|\le V\eta_j$ by character inversion.

For ordered parallelograms $v_1+v_2=v_3+v_4$ of distinct points in $E$, the frequency of any specified color tuple is

$$4^{-4}+O(\eta_1/\rho^2+1/(V\rho^4)),\qquad V\rho^4\longrightarrow\infty. \tag{22}$$

Indeed on replacing four indicators successively, bound the centered Fourier sum in supremum norm by $V\eta_1$, one other by $\rho V$, and the two remaining sums in $L^2$ by $(\rho V)^{1/2}$. At least $\gg\rho^4V^3$

parallelograms exist by Cauchy–Schwarz on the sums $v_1 + v_2$. Repetitions cost $O(V^2)$. The probability of four distinct colors is $24/256 + o(1) = 3/32 + o(1)$ when the errors vanish.

The quadratic counterpart counts ordered, pairwise distinct $v_1, \dots, v_{10} \in E$ with

$$\sum_{i=1}^{5} v_i = \sum_{i=6}^{10} v_i, \qquad \sum_{i=1}^{5} v_i^2 = \sum_{i=6}^{10} v_i^2. \tag{23}$$

Every specified ten-color tuple has relative frequency

$$4^{-10} + O\Big(\rho^{-10}(\eta_2 + V^{-1})\Big), \qquad V\rho^{10} \longrightarrow \infty. \tag{24}$$

For the proof put $S_f(\theta, \alpha) = \sum_{V<v\le 2V} f(v)\mathrm{e}(\theta v + \alpha v^2)$. For $|f| \le 1$ we have $\|S_f\|_{L^8(\mathbb{T}^2)}^8 \ll V^5$. Here are the details of the eighth moment bound. By orthogonality and taking absolute values, it suffices to count four-tuples $x, y$ with equal sums and equal sums of squares. The case $x = y$ costs $O(V^4)$. Otherwise write $x - y = gd$ with $g > 0$, $d$ a primitive integral vector, $\sum d_i = 0$ and $m = \max |d_i|$. At fixed $m$ there are $O(m^2)$ vectors and $O(V/m)$ choices of $g$. The squares give $\sum d_i y_i = -g \sum d_i^2/2$. Solve for an index with $|d_i| = m$. Primitivity makes the map from the other three residue coordinates onto $\mathbb{Z}/m\mathbb{Z}$ surjective, so there are exactly $m^2$ admissible triples modulo $m$ for an integral right side (and no solutions for a nonintegral one). Extending their ranges to complete boxes costs a bounded factor since $m \ll V$. Hence $O(V^3/m)$ solutions occur per $d, g$. The sum for $m \ll V$ is $O(V^5)$.

In the tenfold integral replace each $I_c$ by $\mathbf{1}_E/4$ in turn, bounding the changed sum by $V\eta_2$, another by $V$, and the eight remaining factors by Hölder's inequality and the eighth moment. Each error is $O(V^7\eta_2)$. Cauchy–Schwarz on the $O(V^3)$ possible pairs of sums and sums of squares gives $\gg \rho^{10}V^7$ solutions with repetitions. Fixing any two coordinates equal leaves a factor of supremum norm $O(V)$ and eight eighth-moment factors, hence costs $O(V^6)$. Division proves (24).

## 8.2 Eight origins by restriction to the parabola

The tuples $(1, 4, 6, 7)$ and $(2, 3, 5, 8)$ each have sum 18 and sum of squares 102. Adding $V$ to every entry preserves both equalities and places all eight integers in the dyadic interval for $V \ge 8$. A restriction estimate gives the frequency of color patterns on all such disjoint pairs of quadruples. Use the same population $E$, its density $\rho$, its four colors, and the uniform coefficient bound $\eta_2$ just defined.

**Proposition 8.1** (Eight distinct origins)**.** *Consider ordered $v_1, \dots, v_8 \in E$, pairwise distinct, with matching sums and matching sums of squares across the first and last four positions. If $\rho^4 V^{4/7} \to \infty$, the frequency of every assigned color tuple is*

$$4^{-8} + O\Big(\rho^{-9/2}\eta_2 + \rho^{-4}V^{-4/7}\Big). \tag{25}$$

*Proof.* The one-dimensional periodic restriction estimate for the parabola goes back to Bourgain [4]. We use the Strichartz formulation of Killip and Vişan, *Scale-invariant Strichartz estimates on tori and applications* [20, Theorem 1.1], at $p = 7 > 6$:

$$\|S_f\|_{L^7(\mathbb{T}^2)} \ll V^{1/2-3/7}\|f\|_{\ell^2} \le \rho^{1/2}V^{4/7} \qquad (|f| \le \mathbf{1}_E).$$

Indeed take the standard one-dimensional torus, a dyadic frequency cutoff equal to one on $[V, 2V]$, and Fourier coefficients $f(v)$. Reversing the time coordinate changes the sign of the quadratic term. The space-time norm is over one period in each variable.

Cauchy–Schwarz applied to fourfold sums and sums of squares gives at least $\gg |E|^8/V^3 = \rho^8 V^5$ solutions with repetitions. In the integral for any given color assignment, telescope the eight indicators. Put the centered factor in supremum norm and the other seven in $L^7$. The error is $O(\eta_2 \rho^{7/2} V^5)$. If two coordinates are equal, their shared variable contributes a combined factor of supremum norm at most $\rho V$, whether the two signs in the moment equations agree or cancel. For an upper bound on repeated tuples we may drop the colors. The remaining six factors contribute at most $\|S_{\mathbf{1}_E}\|_7^6$ by Hölder on a probability space. Summing over the at most 28 pairs costs $O(\rho^4 V^{31/7})$. Dividing by the count of distinct solutions proves (25). □

**Liouville and cofactor colors.** Color the full dyadic interval by $(\lambda(v-1), \lambda(v))$. The two single-site characters have uniformly small polynomial coefficients by [17]. For the pair character, Candidate Theorem 1.1, together with fixed-progression cancellation from [26], would give $D_2 = o(1)$. This qualitative input already suffices for the limiting color frequencies in (22), (24) and (25), since $\rho = 1$. Candidate Theorem 1.2 gives the stronger logarithmic saving. In either application the pair estimate on the dyad follows by shifting the polynomial and subtracting two prefixes.

For the cofactor colors on $E_a$, the density is supplied by Lemma 7.1, and the signed input is Candidate Proposition 6.6. Presence of a prime in an interval $I$ is $1 - s_I(n)$, where $s_I$ is completely multiplicative, zero at primes in $I$ and one at other primes. All the selected prime windows lie in $[V^{0.39}, 3V]$ for sufficiently large $V$. Expand the finitely many masks. Up to fixed sign switches each nonconstant color character on $E_a$ becomes a bounded sum of the marked tests, possibly after replacing $n$ by $n-1$. The candidate proposition therefore gives $\eta_2 \ll L^{-c}$. The zero-phase instance is covered by the separate cofactor theorem above. The estimate for coefficients chosen with $V$ still uses Candidate Proposition 6.6. Equations (24) and (25) give respectively $\gg V^7/T^{10a}$ and $\gg V^5/T^{8a}$ ordered distinct tuples for *each* assigned pattern. These rare-population conclusions use the marked estimate in addition to the unsigned density.

## 9 An affine obstruction and the nilsequence question

The product $F(n) = \lambda(n)\lambda(n+1)$ is usually tested here against an external phase. The prescribed quartic product is $F(n)F(3n+1)$, whose two factors have a common arithmetic origin. For a sharp distinction, take $G(n) = \lambda(2n+1)$. For each fixed $d, A$,

$$\sup_{\deg P \le d} \left| \sum_{n \le X} G(n) \mathrm{e}(P(n)) \right| \ll_{d,A} X(\log X)^{-A}, \qquad G(n)G(3n+1) = -1.$$

Indeed $2(3n+1)+1 = 3(2n+1)$. For cancellation set $m = 2n+1$, open the odd-class indicator into two rational linear phases and use $\lambda(m) = \sum_{j^2 s = m} \mu(s)$. The terms $j > (\log X)^{A+2}$ cost $O(X(\log X)^{-A-2})$, and [17, Theorem 1.1] on the circle with polynomial filtration treats the remaining $\mu$ sums uniformly in the polynomial. There is also an unconditional Gowers-norm version of this example. The bounded multiplicative function $b(m) = \mathbf{1}_{m \text{ odd}} \lambda(m)$ is aperiodic, meaning its mean on every fixed arithmetic progression is zero. Indeed opening a fixed progression into Dirichlet characters after separating the common divisor reduces this to the prime number theorem for $\lambda$ in fixed unit classes (or to the square-divisor identity and the corresponding theorem for $\mu$). By Frantzikinakis and Host [13, Theorem 2.5], $\|b\|_{U^s[N]} \to 0$ for each fixed $s \ge 2$. In the unnormalized interval cube sums, every nonzero cube of $b$ has odd base and even increments. The affine change $m = 2n+1$, $k_i = 2h_i$ identifies those cubes on $1 \le m \le 2N+1$ with the cubes of $G$ on $0 \le n \le N$. Removing the cubes incident to $n = 0$ costs $O_s(N^s)$, and each interval normalization is comparable to its length to the power $s+1$. This gives $\|G\|_{U^s[N]} \to 0$ as well.

Moreover $G(n) = a_n a_{n+1}$ for signs defined by $a_1 = 1$, $a_{n+1} = a_n G(n)$. This recursive construction gives an arbitrary sign sequence, with no multiplicativity asserted for $a_n$. Polynomial tests and a neighboring-product representation are therefore compatible with perfect affine anticorrelation. Consequently, uniformity against external polynomial phases alone does not establish cancellation of the prescribed product $\lambda(n)\lambda(n+1)\lambda(3n+1)\lambda(3n+2)$. The harmonic bound in *An explicit bound below one half for a quartic Liouville correlation* [28] concerns that prescribed quartic itself.

The order of the frequency choice also matters in stationary models. In *An isolated quartic correlation under signed prime dilations* [30, Proposition 7.1], a stationary sign process has fair independent marginals on every seven distinct sites. For its adjacent product the second moment of a normalized length-$H$ Fourier sum is $1/H$ at every preassigned frequency, yet the expected squared supremum over frequencies has a positive limit. In *Bernoulli components and single-pair tests in logarithmic Liouville systems* [31, Theorem 6.1], the spectral criterion allows the frequency to be chosen on each block before integration over origins. Candidate Theorem 1.2 chooses one polynomial for the entire ordinary interval. A deduction of that block criterion would require control after the frequency is chosen locally.

A stronger question for the shifted multiplicative product $a(n) = f_1(n)f_2(n+h)$ is whether (2) forces $\|a\|_{U^3[X]} \to 0$ for the normalized interval Gowers norm. The inverse theorem tests moving two-step nilsequences [18]. Ordinary quadratic phases supply only part of that family. In particular the central character on a Heisenberg nilmanifold introduces bracket expressions and rational covers. The comparison of a marked walk and the vanishing of the nondiagonal rectangle integral would both need to hold for those moving parameters.

There is already a logarithmic antecedent for this direction. Frantzikinakis and Host [14, Corollary 1.4 and the following remarks] treat products of shifted multiplicative functions against fixed irrational polynomial phases and certain weights arising from nilsystems. The question here concerns ordinary averages and the moving nilsequence tests required by the inverse theorem. These changes of averaging and quantifiers remain to be addressed.

**Scope of the results and acknowledgments.** The rational-profile formula, Theorem 2.2, is proved under its stated localization hypothesis. The evaluation in Theorem 2.3 for $n^{it}$ times a rationally almost periodic sequence and the bounds of Theorem 2.4 are unconditional. Corollary 7.2 uses the separate-marks cofactor theorem and the common-shift estimate with one truncated sign. The affine reduction, the reciprocal-power bounds and the configuration counts under their stated Fourier hypotheses have direct proofs. The marked comparisons at fixed and growing graph size provide the proposed arguments for Candidate Theorems 1.1 and 1.2. The periodic, linear and large-prime marked estimates retain the same candidate status; their uses are identified in Sections 7 and 8. The formalized ordinary two-point result in [26] covers the unweighted input and does not include these extensions.

AI systems assisted with the mathematical arguments, the exposition, the bibliographic comparisons and the finite checks supplied with the sources. The checks cover finite arithmetic identities, the enumerations of moment configurations, and the projection and oscillatory-integral calculations in Theorem 2.4. Formal verification of the unweighted input is described in [26]. The proposed marked graph extensions are analytical arguments.

## References

[1] N. Alon, O. Goldreich and Y. Mansour, *Almost k-wise independence versus k-wise independence*, Inform. Process. Lett. **88** (2003), no. 3, 107–110. 10.1016/S0020-0190(03)00359-4.

[2] V. Bergelson, J. Kułaga-Przymus, M. Lemańczyk and F. K. Richter, *A structure theorem for level sets of multiplicative functions and applications*, Int. Math. Res. Not. IMRN **2020** (2020), no. 5, 1300–1345. 10.1093/imrn/rny040. arXiv:1708.02613v2, Theorem 2.10 and Corollary 2.11.

[3] M. Bordignon, C. Bortolotto and B. Kerr, *Weyl sums with multiplicative coefficients and joint equidistribution*, Algebra Number Theory **19** (2025), no. 8, 1549–1580. 10.2140/ant.2025.19.1549.

[4] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations*, Geom. Funct. Anal. **3** (1993), 107–156. 10.1007/BF01896020.

[5] M. Braverman, *Polylogarithmic independence fools $AC^0$ circuits*, J. ACM **57** (2010), no. 5, article 28. 10.1145/1754399.1754401.

[6] J. Cassaigne, S. Ferenczi, C. Mauduit, J. Rivat and A. Sárkózy, *On finite pseudorandom binary sequences III: The Liouville function, I*, Acta Arith. **87** (1999), 367–390. 10.4064/aa-87-4-367-390.

[7] A. Carbery, M. Christ and J. Wright, *Multidimensional van der Corput and sublevel set estimates*, J. Amer. Math. Soc. **12** (1999), 981–1015, Lemma 1.1(a),(c). 10.1090/S0894-0347-99-00309-4.

[8] J. Chapman and S. Chow, *Arithmetic Ramsey theory over the primes*, Proc. Roy. Soc. Edinburgh Sect. A **156** (2026), 966–1012. doi:10.1017/prm.2024.96.

[9] H. Daboussi, *Fonctions multiplicatives presque périodiques B*, Astérisque 24–25 (1975), 321–324. https://www.numdam.org/item/AST_1975__24-25__321_0/.

[10] H. Daboussi and A. Sárkózy, *On the correlation of the truncated Liouville function*, Acta Arith. **108** (2003), 61–76. 10.4064/aa108-1-6.

[11] H. Daboussi and H. Delange, *On multiplicative arithmetical functions whose modulus does not exceed one*, J. London Math. Soc. (2) **26** (1982), no. 2, 245–264. 10.1112/jlms/s2-26.2.245.

[12] 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. arXiv edition, Corollary 1.1(i).

[13] N. Frantzikinakis and B. Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc. 30 (2017), 67–157. DOI 10.1090/jams/857. arXiv:1403.0945v2.

[14] N. Frantzikinakis and B. Host, *Furstenberg systems of bounded multiplicative functions and applications*, Int. Math. Res. Not. IMRN **2021** (2021), no. 8, 6077–6107. 10.1093/imrn/rnz037. arXiv:1804.08556v3, Corollary 1.4 and the following remarks.

[15] N. Frantzikinakis, M. Lemańczyk and T. de la Rue, *Furstenberg systems of pretentious and MRT multiplicative functions*, Ergodic Theory Dynam. Systems **45** (2025), 2765–2844. 10.1017/etds.2024.140. arXiv:2304.03121v4, Proposition 5.5.

[16] B. Green and T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) 175 (2012), 465–540. doi:10.4007/annals.2012.175.2.2. Corrected text: arXiv:0709.3562v6, August 2015, Proposition 4.3 and Lemma 4.4.

[17] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2) 175 (2012), 541–566. doi:10.4007/annals.2012.175.2.3.

[18] B. Green, T. Tao and T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) 176 (2012), 1231–1372. doi:10.4007/annals.2012.176.2.11.

[19] H. A. Helfgott and M. Radziwiłł, *Expansion, divisibility and parity*, preprint (2021). DOI 10.48550/arXiv.2103.06853.

[20] R. Killip and M. Vişan, *Scale-invariant Strichartz estimates on tori and applications*, Math. Res. Lett. 23 (2016), 445–472. DOI 10.4310/MRL.2016.v23.n2.a8.

[21] S. V. Konyagin and T. Steger, *On polynomial congruences*, Math. Notes **55** (1994), no. 6, 596–600. 10.1007/BF02110354.

[22] J. D. Lichtman, *Improved bounds for the binary Chowla conjecture*, preprint (2026), arXiv:2611.28392v1. DOI 10.48550/arXiv.2611.28392.

[23] K. Matomäki, M. Radziwiłł and T. Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory 9 (2015), 2167–2196. doi:10.2140/ant.2015.9.2167. Corrected text: arXiv:1503.05121v3.

[24] K. Matomäki, M. Radziwiłł, T. Tao, J. Teräväinen and T. Ziegler, *Higher uniformity of bounded multiplicative functions in short intervals on average*, Ann. of Math. (2) **197** (2023), 739–857. 10.4007/annals.2023.197.2.3.

[25] National Institute of Standards and Technology, *NIST Digital Library of Mathematical Functions*, Table 18.3.1 (orthogonality), Table 18.8.1 (differential equations), and equation 5.9.1 (gamma integral).

[26] OpenAI, *Ordinary two-point correlations of multiplicative functions*, preprint, 24 September 2026. Repository edition, Part I, Sections 2–8; commit f5702854040e97fa7675c21ba37a9608f5f92dd3. Proof sources at this commit. Scope of the formalized theorem.

[27] C. Pilatte, *Improved bounds for the two-point logarithmic Chowla conjecture*, J. Amer. Math. Soc., published online 2026. DOI 10.1090/jams/1084. arXiv:2310.19357v3, 25 August 2026.

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

[29] B. Pouly, *Joint cofactor signs in prime-factor windows on consecutive integers*, Zenodo (2026), Theorem 1.1. 10.5281/zenodo.24578855.

[30] B. Pouly, *An isolated quartic correlation under signed prime dilations*, Zenodo (2026), Proposition 7.1. 10.5281/zenodo.24578889.

[31] B. Pouly, *Bernoulli components and single-pair tests in logarithmic Liouville systems*, Zenodo (2026), Theorem 6.1. 10.5281/zenodo.24578887.

[32] A. Sidi, *Asymptotic expansions of Legendre series coefficients for functions with endpoint singularities*, Asymptot. Anal. **65** (2009), 175–190. 10.3233/ASY-2009-0950.

[33] B. Wang and S. Yi, *Dynamical generalizations of the two-point Chowla conjecture*, preprint (2026), arXiv:2610.16394v1. DOI 10.48550/arXiv.2610.16394.