Bernoulli components and single-pair tests in logarithmic Liouville systems

Brice Pouly*

hexagon:2611.01198v1

Abstract

Fix a logarithmic limiting law of the Liouville signs. An observer sees the products of two signs at one fixed lag. We determine the optimal probability of predicting the present product from its own past, however distant. Beyond the length of the pair the answer depends only on the two conditional biases. At shorter gaps it depends on the absolute bias of a single sign. Unless the law is fair Bernoulli, the advantage over one half is positive uniformly over all nonzero lags. The proof combines the nilpotent–Bernoulli structure of Frantzikinakis and Host with multiplicative recurrence. An entropy argument gives the optimal predictor for every finite product of signs and every gap. In the same fixed law, the fair Bernoulli law is the only possible atom of the ergodic decomposition, and its mass is determined by the least frequencies of finite words. For a particular quartic Liouville correlation, we relate its bias to the discrete spectra of two pairs and transfer a rectangle of prime translates to a conditioned eight-sign mean on actual prime factors. Both the existence of an arithmetic predictive advantage and the vanishing of that conditioned mean remain open.

1 What can one pair of signs detect?

If the signs at the vertices of a line are independent and fair, so are the products across its edges. Indeed, each finite assignment of edge signs has exactly two lifts on each connected component. What happens if the signs come from the Liouville function and we can observe the edges only? Tao's logarithmically averaged two-point theorem [26] gives mean zero for the product across every nonzero lag. Its mean alone therefore carries no information about possible higher correlations. Its own past can carry such information.

Write $\lambda(n) = (-1)^{\Omega(n)}$, where Ω counts prime factors with multiplicity. In a window $I_\nu = (U_\nu, V_\nu]$, $1 \le U_\nu < V_\nu$, give n mass proportional to $1/n$ and assume $\log(V_\nu/U_\nu) \to \infty$. Fix any subsequence along which all finite distributions of $(\lambda(n+i))_{i\in\mathbb{Z}}$ converge. Its law μ on $\{-1,1\}^{\mathbb{Z}}$ is stationary under the shift T because translating a harmonic window by a fixed integer costs bounded unnormalised mass. Values at nonpositive arguments can be filled arbitrarily. Write $X_i(x) = x_i$, $f = X_0$, $T^j g = g \circ T^j$ and B for the fair product law.

For $h \neq 0$ let $Y_j^{(h)} = X_j X_{j+h}$. Let $s_h(N)$ be the supremum of $\mathbb{P}(\widehat{Y} = Y_0^{(h)})$ over sign-valued predictions measurable in $\sigma(Y_{-j}^{(h)} : j \ge N)$. Entropies are measured in bits. For $s \ge 1$ let $\mathcal{Z}_s$ denote the componentwise Host–Kra factor of order s and $\mathcal{Z}_\infty = \bigvee_{s\ge1} \mathcal{Z}_s$. It contains the ergodic-component label. Set

$$H = \mathbb{E}_\mu(f \mid \mathcal{Z}_\infty), \qquad H_i = T^i H.$$

Thus H_i is the conditional bias at site i.

At gap N the observer knows every pair product at times $j \le -N$. Write $b_h = \max(0, h)$. The last input site used in these observations is $b_h - N$. If $N > |h|$, both sites of the target lie beyond it. If $1 \le N \le |h|$, only the right endpoint does. This distinction persists after passing from the underlying signs to the products that the observer actually sees.

*Prepared with AI assistance.

Theorem 1.1 (Prediction at one lag)**.** *For every $h \neq 0$ and $N \geq 1$,*

$$s_h(N) = \frac{1}{2} + \frac{1}{2} \begin{cases} \displaystyle\int |H_0 H_h| \, \mathrm{d}\mu, & N > |h|, \\ \displaystyle\int |H| \, \mathrm{d}\mu, & 1 \leq N \leq |h|. \end{cases} \tag{1.1}$$

If $\mu \neq \mathsf{B}$, the advantage over $1/2$ is positive uniformly in h and N. A lower bound depends only on $1 - h_\mu(T)$. For fixed h, N, predictions using finite blocks ending at $-N$ approach the optimum.

Thus observing a single nonzero lag, however far into its past one has to look, gives a criterion for the fair Bernoulli law. For each fixed $h \neq 0$ and $N \geq 1$, $\mu = \mathsf{B}$ if and only if $s_h(N) = 1/2$. Equivalently, $\mathbb{E}(Y_0^{(h)} \mid \mathcal{Z}_s(Y^{(h)})) = 0$ for every $s \geq 1$. The latter test uses factors of the observed pair process. The long-gap value in (1.1) depends on the lag through the joint sizes of the two biases. The short-gap value is independent of the lag.

Theorem 1.2 (Mass of the Bernoulli component)**.** *Write $\mu = \int \omega \, \mathrm{d}\pi(\omega)$ for its decomposition into ergodic sign laws, and $[w]$ for the cylinder of a contiguous word w. The decomposition measure π has no atom other than possibly B, whose mass is*

$$b = \pi(\{\mathsf{B}\}) = \mu\{H_i = 0 \text{ for all } i\} = \lim_{L \to \infty} 2^L \min_{w \in \{-1,1\}^L} \mu([w]). \tag{1.2}$$

This is also the total mass of components with no irrational eigenvalue. If the point spectra of almost all components lie in a common countable set, $\mu = \mathsf{B}$.

Here b is a mass in an ergodic decomposition. It is distinct from the entropy rate and from the amount of Bernoulli noise in an extension. Frantzikinakis proved that ergodicity of a logarithmic Liouville system along a sequence of initial intervals forces logarithmic Chowla along that sequence [6, Theorem 1.1]. Górska, Lemańczyk and de la Rue obtained a related conclusion under pure atomicity of *all* logarithmic Möbius systems [9, Remark 6.22]. Theorem 1.2 concerns one fixed law, which may also have a non-atomic part. An ergodic component of that law need not itself be an arithmetically realised logarithmic limiting law, so the ergodic theorem cannot simply be applied to each component.

Our structural input is the nilpotent–Bernoulli description of the system of prime progressions due to Frantzikinakis and Host [7, Proposition 3.9, (29), Lemma 4.11 and §4.5]. Together with naturality and the past-tail description of the Pinsker factor, it yields the intrinsic conditional law used here. Kanigowski, Kułaga-Przymus, Lemańczyk and de la Rue give the corresponding Pinsker and nilfactor criteria for logarithmic Chowla, as well as the relative-independence principle needed to adjoin a zero-entropy clock [15, Corollary 1.12 and Proposition 2.2]. We also give a direct division proof of the conditional law. It retains the arithmetic information used in the finite-order statement and the uniform recurrence argument.

The prediction result uses two further ingredients. Once the preceding input signs are fixed, each new product determines its rightmost sign. Equality of the input and output entropy rates then identifies their conditional laws for future products. A quantitative adaptation of Hildebrand's stable-set mechanism [10, 11] shows that division forces the supports of H_0 and H_h to meet at every lag. The observed products determine the conditional mean of the present pair even across an arbitrarily long gap. This relates the formula to the study of Sarnak processes [15, 18], while specifying both the observed pair and its exact predictive success. For the quartic correlation we identify the discrete projections that retain a particular four-site moment, then transfer a spectral rectangle to signed tests on actual prime factors. The numerical quartic bound comes from *An explicit bound below one half for a quartic Liouville correlation* [21].

Sections 2 and 3 prove the prediction formula and give its extension to finite products. Section 4 recovers the atom from word frequencies. For the quartic $\lambda(n)\lambda(n+1)\lambda(3n+1)\lambda(3n+2)$, Sections 5–6 relate a possible bias to the two individual pairs, then to a rectangle sampled by four primes. The last section records two symmetries under arithmetic conditioning.

2 The conditional distribution of a sign word

Write $\mathcal{A} = \sigma(H_i : i \in \mathbb{Z})$ and let $\Pi(X)$ be the Pinsker factor, the largest invariant factor of entropy zero. All sigma fields below are completed for the law in use. We shall occasionally adjoin the profinite origin $J \in \widehat{\mathbb{Z}}$ by joint extraction with all finite residues of n. The marginal of J is Haar, and $J \circ T = J + 1$. Further extraction leaves μ intact.

The following formulation makes the structural input observable inside the sign system. Sections 2.1–2.3 prove it directly and explain its relation to the progression-system formulation.

Theorem 2.1 (Intrinsic conditional sign law)**.** *For every finite set $S \subset \mathbb{Z}$ and signs ϵ_i, the joined law satisfies*

$$\mathbb{P}(X_i = \epsilon_i,\ i \in S \mid J, (H_j)_{j\in\mathbb{Z}}) = \prod_{i\in S} \frac{1+\epsilon_i H_i}{2} \quad \text{almost surely.} \tag{2.1}$$

The same formula holds conditional on $\mathcal{A}$ alone. In the sign system,

$$\Pi(X) = \mathcal{Z}_\infty(X) = \mathcal{A}, \qquad h_\mu(T) = \int h_2((1+H)/2)\,\mathrm{d}\mu, \tag{2.2}$$

where $h_2(v) = -v\log_2 v - (1-v)\log_2(1-v)$, extended continuously to $[0,1]$.

2.1 What division preserves

For a factor map, the projection of a lifted observable onto $\mathcal{Z}_s$ is the lift of its projection downstairs. To see this, the downstairs nilfactor lifts into the upstairs nilfactor, and the Host–Kra seminorm of any pullback equals its downstairs seminorm. The latter follows by induction from the recursive cube formula, starting with the mean ergodic theorem. Apply the seminorm criterion for orthogonality to the residual after projection [13, 17]. These statements apply componentwise and integrate over the decomposition. Increasing s and taking L^2 limits gives the same naturality for $\mathcal{Z}_\infty$. In particular the clock leaves H unchanged as an observable of the signs.

The profinite origin and partial division maps are part of Sawin's dynamical model for Liouville [25, Definition 1.1]. In the joined sign law the map on $\{d \mid J\}$ has the concrete form

$$U_d(x, J) = ((\lambda(d)x_{di})_{i\in\mathbb{Z}}, J/d), \qquad d \geq 1.$$

This sends the conditional law to the full joined law. For $n = dm$, complete multiplicativity gives the sign identity at each site, and $1/n = 1/(dm)$. Replacing I_ν/d by I_ν costs $O_d(1)$ in unnormalised harmonic mass. These identities on cylinders give the assertion. Also $U_d T^d = T U_d$ on its domain. For $s \geq 1$ the factors $\mathcal{Z}_s(T^d)$ and $\mathcal{Z}_s(T)$ agree [17, Theorem 2 in the separate text]. Restricting to a T^d-invariant domain of positive mass restricts the componentwise projection. Apply naturality to $f \circ U_d = \lambda(d) f$. For s finite or infinite, writing $H_s = \mathbb{E}(f \mid \mathcal{Z}_s)$, we obtain

$$H_s \circ U_d = \lambda(d) H_s \quad \text{on } \{d \mid J\}. \tag{2.3}$$

2.2 Removing one residual sign

Let $R = f - H$. We prove that the product of any nonempty set of distinct translates of R has zero conditional expectation given $(J, (H_i)_i)$. Test it against a clock class $M = \mathbf{1}_{J=a \bmod q}$ and a monomial in finitely many H_i. These tests span a dense algebra by polynomial approximation on $[-1, 1]$ and the cylinder construction. Fuse all factors at each site. Each resulting g_i is H^m or RH^m. At least one has the second form and has zero projection on $\mathcal{Z}_\infty$ because H^m is measurable there. By (2.3), $g_i \circ U_p = (-1)^{e_i} g_i$ on $p \mid J$, with $e_i = m$ or $m + 1$. If b_i are the distinct sites and $p \equiv 1 \bmod q$, exact transport gives

$$\mathbb{E}\left[p\mathbf{1}_{p|J} M \prod_i T^{pb_i} g_i\right] = (-1)^{\sum e_i} \mathbb{E}\left[M \prod_i T^{b_i} g_i\right]. \tag{2.4}$$

Average over such primes with weights $1/p$, normalised by their sum up to P. Tao and Teräväinen's entropy-decrement estimate [28, Proposition 2.2] removes $p\mathbf{1}_{p|J}$ at a cost $o(1)$ in mean absolute error as $P \to \infty$. Its inputs are bounded observables of a stationary process with profinite origin. Here one includes the clock observable at site zero and rescales the observables to the unit disk. For every fixed tolerance the exceptional dyadic indices m have $\sum 1/m < \infty$. The reciprocal-prime mass of $[2^m, 2^{m+1})$ is $O(1/m)$ and the selected class has positive relative prime density. These facts justify the restricted harmonic average.

The remaining normalised prime average tends to zero. In the average over all primes, the condition $p \equiv 1 \bmod q$ is represented by the sum of products $\sum_{v \bmod q} \mathbf{1}_{J=v} T^p \mathbf{1}_{J=v+1}$, using the same modulus throughout. At a coincident position the additional factor is measurable on the rational Kronecker factor and preserves residual orthogonality. By stationarity translate all positions simultaneously into the positive integers and insert constant observables at missing positions. On each ergodic component Theorem 4.3 of Frantzikinakis and Host [7] gives zero for the Cesàro prime average because one observable is orthogonal to $\mathcal{Z}_\infty$. Integration over components and partial summation give zero also for the normalised reciprocal-prime average. The right side of (2.4) is independent of p, so every test integral vanishes. Expanding $X_i = H_i + T^i R$ gives the product formula for conditional moments. Finally $\mathbf{1}_{X_i=\epsilon_i} = (1 + \epsilon_i X_i)/2$ gives (2.1). The tower property removes J.

2.3 Recovering the factor from the past

Each finite-step nilfactor has entropy zero on components. Entropy is affine under ergodic decomposition and continuous along increasing generating factors, so $\mathcal{A} \subseteq \mathcal{Z}_\infty \subseteq \Pi(X)$. For the reverse inclusion the Rokhlin–Sinai theorem identifies $\Pi(X)$ with $\bigcap_{N \geq 1} \sigma(X_j : j \leq -N)$ for this finite generating partition [12, §11.5]. Disintegrate over $\mathcal{A}$. For a fixed event of this completed tail choose, for each N, a representative depending only on coordinates at most $-N$. Their limsup is a single uncompleted tail event agreeing with the original modulo a null set. By (2.1) almost every conditional law is a product law. Kolmogorov's zero–one law makes the conditional probability of this event either zero or one. The event consequently belongs to $\mathcal{A}$ modulo null sets. This proves the first equality in (2.2) and both inclusions are equalities.

In particular H is measurable in the strict past of X. Given $\mathcal{A}$, the past and X_0 are independent, so the conditional mean of X_0 given the strict past is H. The conditional-entropy formula for a stationary finite generator gives (2.2). It follows also that $H = 0$ almost surely if and only if $\mu = \mathsf{B}$.

Relation with the progression-system structure. The same kernel can be read from the progression-system structure of [7, Proposition 3.9, (29), Lemma 4.11 and §4.5], using the transport identity above to obtain the sign factor for the windows considered here. In that

extension, retain the sigma field generated by the coordinate maps to the full base nilfactor, together with the invariant base-component label. On each ergodic fibre the progression system is the product of an infinite-step nilsystem and a Bernoulli system. This sigma field is its infinite-step nilfactor. The sign coordinates produced by the factor, which includes the global reversal of signs, are conditionally independent over it. Naturality under that factor identifies their conditional means with the lifts of H_i. The tower property then gives the product kernel conditional on $\mathcal{A}$. The tail argument just given identifies $\mathcal{A} = \mathcal{Z}_\infty = \Pi$.

The clock has zero entropy. The relative-independence statement of [15, Proposition 2.2], applied to the class of zero-entropy systems, says that a joining with the clock is relatively independent over the Pinsker factor of the sign system. Thus adjoining J leaves this conditional product kernel unchanged. This identifies Theorem 2.1 as an intrinsic consequence of the existing structure theory. The direct proof above also records the division identities and the treatment of the harmonic windows needed below.

2.4 A finite-order version

The preceding kernel uses all nilpotent orders. If only a bounded number of signs is observed, a finite-order factor suffices. The number of tested amplitude coordinates is unrestricted; it is the number of residual sign coordinates that determines the required order.

Proposition 2.2. *For $s \geq 1$, conditional on $\sigma(J, (T^i H_s)_{i\in\mathbb{Z}})$, any at most $s+2$ distinct signs have product law with means $T^i H_s$.*

We isolate the cancellation on progression spaces that is needed for this assertion and later for the quartic moment. If (X, η, T) is an ergodic system, $q \geq 1$ and $(a, q) = 1$, denote by $\widetilde{\eta}_{a,q}$ its law of progressions with prime step in the class a mod q. Thus, for bounded observables at finitely many integer positions,

$$\int \prod_i g_i(x_{b_i})\, \mathrm{d}\widetilde{\eta}_{a,q} = \lim_{P\to\infty} \frac{1}{\pi(P; q, a)} \sum_{\substack{p\leq P \\ p\equiv a \ (q)}} \int \prod_i T^{pb_i} g_i \, \mathrm{d}\eta, \tag{2.5}$$

where $\pi(P; q, a)$ counts the indicated primes. The law is invariant under the coordinate shift S, and every coordinate has marginal η. Existence follows from the prime convergence theorem by adjoining a finite cyclic factor and selecting the step class. The same construction can be integrated over an ergodic decomposition.

Lemma 2.3 (Vertical cancellation with a prescribed step class)**.** *Let (X, η, T) be an ergodic finite-dimensional nilsystem and let $\widetilde{\eta}_{a,q}$ be as in (2.5). Fix $s \geq 1$. Suppose that, at distinct positions, a finite product has ordinary observables measurable on $\mathcal{Z}_s(X)$ and j distinguished observables orthogonal to $\mathcal{Z}_s(X)$, where $1 \leq j \leq s+2$. Its integral against $\widetilde{\eta}_{a,q}$ is zero. Finite cyclic masks may be included among the ordinary observables.*

At every descent stage of step $k > s$, if V_k is the last vertical torus, the progression law is invariant under

$$(x_i)_{i\in\mathbb{Z}} \longmapsto \big(z^{(i)}_{(\ell)} x_i\big)_{i\in\mathbb{Z}}, \qquad z \in V_k, \quad 0 \leq \ell \leq k. \tag{2.6}$$

These actions preserve every finite cyclic mask, including an initial-origin class.

Proof. We first justify the step restriction in (2.6). Use a minimal presentation $X = G/\Gamma$, with G generated by its identity component and the translating element τ. Let v be a multiple of q and of the number of connected components of X. Split the selected primes into the finitely many reduced classes t mod v with $t \equiv a$ mod q, and then split the initial point among its connected components. For the restriction to one class, adjoin a uniform independent cyclic coordinate C mod v advanced by one at each iterate. At a step n the sum

$$\sum_{\ell \bmod v} \mathbf{1}_{C=\ell}\, T^n \mathbf{1}_{C=\ell+t} = \mathbf{1}_{n\equiv t \bmod v}$$

selects exactly this class. The return map to each connected component is totally ergodic, including after adjoining the clock and taking the vth power. Apply [7, Theorem 4.4] with these two masks, retaining the chosen initial component in the integral. The average over reduced classes then has a single term of weight $1/\varphi(v)$. Dividing by the same prime density gives the integer-step average along $t+v\ell'$, with its conditional normalization. On a connected return component, this is the progression law for T^v, followed at coordinate i by the fixed translation T^{it}. Proposition 4.7 and Appendix B of [7] identify that law with Haar measure on its Hall–Petresco progression space.

The top vertical group does not shrink on taking this return map. Indeed, put $G(v) = \langle G^0, \tau^v \rangle$ and take the closures of the lower central groups when needed. Every vth power in G lies in $G(v)$. At the last nontrivial level $k \geq 2$, commutators are multilinear, so the v^kth power of every weight-k commutator lies in $G(v)_k$. The connected abelian quotient $G_k/G(v)_k$ therefore has bounded exponent. Multiplication by any positive integer on a connected abelian Lie group is surjective, so this quotient is trivial. Thus $G(v)_k = G_k$.

The Hall–Petresco group contains the central polynomial sequence $\left(z^{\binom{i}{\ell}}\right)_i$ for $z \in V_k$ and $0 \leq \ell \leq k$. Haar invariance gives (2.6). Centrality makes the action commute with the fixed translations T^{it}; connectedness makes it trivial on finite cyclic factors. Averaging over initial components and admissible step classes preserves the invariance.

We now use it to cancel the distinguished observables. If the system has order at most s, the distinguished observables already vanish. Otherwise expand each observable in vertical characters, initially by bounded Fejér approximation. Every ordinary observable is constant on V_k, since $k > s$. Let $b_1, \dots, b_j$ be the distinguished positions and $\chi_1, \dots, \chi_j$ their vertical characters, written additively. An integral can survive (2.6) only if

$$\sum_{r=1}^{j} \binom{b_r}{\ell} \chi_r = 0 \qquad (0 \leq \ell \leq k).$$

The first j equations are available because $j \leq s+2 \leq k+1$. Their determinant is

$$\det\left[\binom{b_r}{\ell}\right]_{\substack{0\leq \ell<j \\ 1\leq r\leq j}} = \frac{\prod_{r<t}(b_t - b_r)}{\prod_{\ell=0}^{j-1} \ell!} \neq 0.$$

The character group of a torus is torsion free. Consequently all χ_r vanish. Integrating vertically therefore replaces every distinguished observable by its projection on the quotient of step at most $k-1$. This projection remains orthogonal to the order-s factor, by the tower property. Iterate until step s, where each distinguished observable is zero. Bounded approximation, together with the coordinate marginals η, removes the initial continuity and finite Fourier assumptions. □

Proof of Proposition 2.2. Put $R_s = f - H_s$ and repeat the transport identity (2.4), testing against a finite clock class and a monomial in any finite collection of the amplitudes $T^i H_s$. After fusing observables at equal positions, every nonempty residual term has $j \leq s+2$ distinguished observables $R_s H_s^m$, orthogonal to $\mathcal{Z}_s$, and ordinary observables H_s^m. Entropy decrement removes the divisibility weight exactly as in Section 2.2.

In the resulting prime average, first project all observables onto $\mathcal{Z}_\infty$ by [7, Theorem 4.3]. On each ergodic component approximate them by conditional expectations on a common finite nilfactor that includes the finite clock. The ordinary projections are measurable on its order-s factor, and the distinguished projections are orthogonal to that factor. To check this compatibility, test against a function on the finite factor and use naturality of the order-s projection; duality gives both assertions. Lemma 2.3 now makes every approximating integral zero.

Each coordinate of the progression law has the original base marginal. A telescoping product estimate therefore bounds the error by a constant times the sum of the L^1 approximation errors. This justifies passage through the inverse limit and then through the ergodic decomposition. Every nonempty residual term vanishes. Expansion of the signs followed by finite Walsh inversion gives the claimed conditional product law. □

3 A predictor using only its observed products

For a finite nonempty $D \subset \mathbb{Z}$, write $a = \min D$, $b_D = \max D$, $r = b_D - a$ and

$$Y_j^D = \prod_{i \in D} X_{j+i}, \qquad G_D = \prod_{i \in D} H_i, \qquad \mathcal{P}_N^D = \sigma(Y_j^D : j \le -N).$$

The last input site used by this past is $c_N = b_D - N$. Write $\mathcal{F}_c^X = \sigma(X_i : i \le c)$. Although $\mathcal{P}_N^D$ contains only products, it gives the same conditional law for the next block of products as $\mathcal{F}_{c_N}^X$.

Proposition 3.1 (Prediction of a finite product)**.** *For $N \ge 1$ the conditional mean, pulled back to the sign system, is*

$$M_{D,N} := \mathbb{E}(Y_0^D \mid \mathcal{P}_N^D) = \prod_{\substack{i \in D \\ i \le c_N}} X_i \prod_{\substack{i \in D \\ i > c_N}} H_i. \tag{3.1}$$

Its optimal prediction succeeds with probability

$$s_D(N) = \frac{1}{2} + \frac{1}{2} \int \prod_{\substack{i \in D \\ i > c_N}} |H_i| \, \mathrm{d}\mu.$$

For $N > r$, the mean equals G_D and is also the pullback of $\mathbb{E}(Y_0^D \mid \mathcal{Z}_\infty(Y^D))$. Empty products are one.

Proof. We first work with any stationary binary input. Write $V = (Y_{1-N}^D, \dots, Y_0^D)$ for the block between the past and the prediction time. Given $\mathcal{F}_{c_N}^X$, its values are in bijection with the input block $(X_{c_N+1}, \dots, X_{b_D})$. Indeed, at time j all sites in $j + D$ except $j + b_D$ are strictly earlier than $j + b_D$. Multiplying the specified output by those earlier signs recovers the new endpoint, one site at a time. In the language of cellular automata the local rule is permutative in its rightmost coordinate. The adjacent group-product rule is a classical example [20].

Let H denote Shannon entropy in bits and $h(X), h(Y^D)$ the entropy rates. Counting possible initial boundaries gives $h(Y^D) = h(X)$ by Proposition 4.1, whose proof is independent of this section. The chain rule and the entropy-rate formula for a stationary finite alphabet [1, Theorem 4.2.1] now give

$$\begin{aligned} \mathsf{H}(V \mid \mathcal{F}_{c_N}^X) &= \mathsf{H}(X_{c_N+1}, \dots, X_{b_D} \mid \mathcal{F}_{c_N}^X) = N h(X), \\ \mathsf{H}(V \mid \mathcal{P}_N^D) &= N h(Y^D). \end{aligned}$$

The smaller sigma field is contained in the larger one. Its conditional probability vector for V is the conditional expectation of the larger field's probability vector. Strict concavity of Shannon entropy on the finite simplex forces the two vectors to agree almost surely. In particular the two fields give the same conditional mean of the last entry Y_0^D.

Return to the logarithmic sign law. By (2.2) and the past-tail description in Section 2.3, $\mathcal{A} \subseteq \mathcal{F}_{c_N}^X$. Given $\mathcal{A}$, the input law is a product by (2.1). Hence, conditional on $\mathcal{F}_{c_N}^X$, each target site at or before the cut contributes its observed sign, and the joint conditional mean

of the remaining target signs is the product of their amplitudes. Formally this follows first against cylinders in the past and then by the monotone class theorem. This proves (3.1), including the measurability of its full right-hand side in the observed product past.

For any sign Y and information $\mathcal{G}$, a most likely sign succeeds with probability $(1+\mathbb{E}|\mathbb{E}(Y \mid \mathcal{G})|)/2$. Taking absolute values in (3.1) gives the asserted success probability. When $N > r$ every site is beyond the cut. On the other hand (2.1) gives $\mathbb{E}(Y_0^D \mid \mathcal{Z}_\infty(X)) = G_D$, and naturality of the nilfactor projection identifies this with the projection in the observed factor. This proves the last assertion. □

At $N = 1$ the bijection has a useful consequence for every stationary binary input. Only the new right endpoint is unknown given the raw past. Thus every nonempty product filter has the same one-step optimal success as predicting a single input sign from its own strict past. If $m = \mathbb{E}(X_0 \mid \mathcal{F}_{-1}^X)$, that success is $(1 + \mathbb{E}|m|)/2$, at least $\frac{1}{2} + \frac{1}{2}(1 - h(X))$ since $1 - h_2((1+t)/2) \le |t|$.

For a pair in the logarithmic sign law, the short-gap conditional mean is $X_a H_{b_h}$, for $b_h = \max(0, h)$ and $a = \min(0, h)$. Passing from gap $|h|$ to $|h|+1$ costs exactly $\frac{1}{2}\int |H_{b_h}|(1-|H_a|)\,\mathrm{d}\mu$ in success probability. For a general D the value can change only when the cut passes a site of D. Thus a filter with $|D|$ sites has at most $|D|$ distinct success probabilities.

Using a finite past. Let $\mathcal{P}_{N,L}^D = \sigma(Y_{-N-L+1}^D, \dots, Y_{-N}^D)$, let $m_L = \mathbb{E}(Y_0^D \mid \mathcal{P}_{N,L}^D)$, and write $s_{D,L}(N)$ for the corresponding success probability. The martingale convergence theorem gives $m_L \to M_{D,N}$ in L^1. One can also express the loss through conditional entropy. Put $\psi(t) = h_2((1+t)/2)$ and

$$e_L = \mathsf{H}(Y_0^D \mid \mathcal{P}_{N,L}^D) - \int \psi(M_{D,N})\,\mathrm{d}\mu.$$

Then $e_L \downarrow 0$ and

$$0 \le s_D(N) - s_{D,L}(N) \le \sqrt{(\ln 2) e_L / 2}.$$

Indeed, $m_L = \mathbb{E}(M_{D,N} \mid \mathcal{P}_{N,L}^D)$ and $\psi(t) + t^2/(2\ln 2)$ is concave on $[-1, 1]$, as its second derivative on the interior shows. Conditional Jensen gives $\mathbb{E}(M_{D,N} - m_L)^2 \le 2(\ln 2) e_L$. The Bayes formula and Cauchy–Schwarz give the bound. The excess depends on the law and the chosen gap. This estimate entails no uniform length for a prescribed accuracy.

Example 3.2 (A pair with zero mean)**.** Choose a phase uniformly modulo four. Given the phase, take independent signs with repeating means $(1, 1, -1, 1/2)$. The four conditional means of the adjacent products are $(1, -1, -1/2, 1/2)$, so the unconditional mean is zero. The phase is measurable in every raw past, and also in every product past: the average of products spaced four apart converges to one of the four distinct means, by conditional independence. The preceding proof therefore applies directly to this stationary model.

For an adjacent pair, the average absolute value of one amplitude is $7/8$ and the average absolute product of two is $3/4$. The short-gap advantage exceeds the long-gap advantage by $1/16$. The probabilities for the last sixteen products can be computed without estimating the phase. For each observed word w, sum $\mathbb{P}(w, Y_0 = \epsilon)$ by enumerating the input signs and averaging their product probabilities over the four phases. Summing the larger of the two joint masses over w gives the middle column below.

last observed time	sixteen products	entire product past
-1	1919/2048	15/16
-2	3581/4096	7/8

The same environment gives three levels for $D = \{0, 1, 3\}$. Formula (3.1) gives success $15/16$ at gaps $1, 2$, then $7/8$ at gap 3 and $13/16$ at every gap $N \ge 4$. The conditional product law and the position of the cut suffice for these calculations.

The gain at short gaps alone need not imply a gain at long gaps for a general stationary sign law. For example, alternate a deterministic $+1$ and a fair sign, choosing the parity of the deterministic sites uniformly and the fair signs independently. The past identifies the parity almost surely. The same product-law calculation gives $s_1(1) = 3/4$, while $s_1(N) = 1/2$ for every $N \geq 2$, since every adjacent pair contains a zero amplitude. For Liouville, multiplicativity supplies the extra recurrence that forces $\int |H_0 H_h| \, \mathrm{d}\mu > 0$ at every nonzero lag whenever H is nonzero. The argument is a quantitative adaptation of Hildebrand's stable-set mechanism [10, 11]. Its form here includes the profinite origin and applies to level sets of the nilprojection. We keep the constants to exhibit the uniformity in the lag. Stability alone does not force analogous intersections at arbitrary collections of three or more shifts; see [19].

Lemma 3.3. *Let $A = \{|H| > \eta\}$ and $\alpha = \mu(A) > 0$. It is independent of every finite residue of J. Choose $k \geq 2$ with $k\alpha > 1$, and integers $0 = a_1 < \cdots < a_k$ such that each difference $a_j - a_i$ divides both endpoints. Let M be the least common multiple of the positive differences. Then for every $h \neq 0$,*

$$\mu(A \cap T^{-h} A) \;\geq\; \frac{k\alpha - 1}{M\binom{k}{2}}. \tag{3.2}$$

Proof. Fix q and write $m(a) = \mathbb{P}(A, J = a \bmod q)$. From (2.3), for $p \nmid q$ prime, $p\mathbb{P}(A, J = a \bmod q, p \mid J) = m(ap^{-1})$. The variables $(p\mathbf{1}_{p|J} - 1)/\sqrt{p-1}$ are orthonormal by Haar independence. Bessel's inequality applied to $\mathbf{1}_A \mathbf{1}_{J = a \bmod q}$ gives

$$\sum_{p \nmid q} \frac{|m(ap^{-1}) - m(a)|^2}{p - 1} < \infty.$$

The reciprocal primes in each reduced residue class have divergent sum, so m is invariant under multiplication by units. Also $\mathbb{P}(A, d \mid J) = \alpha/d$ for $d \mid q$ by (2.3). Inclusion–exclusion determines the total of m on each gcd stratum, and units act transitively on that stratum. Hence $m(a) = \alpha/q$.

Fix the integer construction deterministically: start with $\{0, 1\}$; to increase its size, translate the set by the least common multiple of its positive differences, adjoin zero, and list the result in increasing order. Hence the final modulus $M = M(k)$ depends only on k. Old differences still divide their endpoints, and each new difference equals its nonzero endpoint. Condition on $J = 0 \bmod M$ and count the visits $Z = \sum_i \mathbf{1}_A \circ T^{ha_i}$. Independence of every finite residue gives $\mathbb{E}(Z \mid J = 0 \bmod M) = k\alpha$. Since $\binom{Z}{2} \geq Z - 1$, some pair $i < j$ has simultaneous conditional mass at least $(k\alpha - 1)/\binom{k}{2}$. Shift the selected event by ha_i. Its first site is now zero, its gap is hd with $d = a_j - a_i$, and the shifted clock is divisible by d since $d \mid a_i, M$. Its unconditioned mass is at least $(k\alpha - 1)/(M\binom{k}{2})$. On $d \mid J$, division by d preserves membership in A and sends gap hd to h. The conditional normalisation multiplies the last lower bound by $d \geq 1$. This proves (3.2). □

Proof of Theorem 1.1. Proposition 3.1 gives (1.1). If $\Delta = 1 - h_\mu(T) > 0$, then $\mathbb{E}H^2 \geq \Delta$. Indeed, for $|t| \leq 1$,

$$1 - h_2((1+t)/2) = \frac{1}{\log 2} \sum_{\ell \geq 1} \frac{t^{2\ell}}{2\ell(2\ell - 1)} \leq t^2.$$

The coefficients are positive and sum to one. Taking $\eta = \sqrt{\Delta/2}$ gives $\alpha \geq \Delta/(2 - \Delta)$ by splitting $\mathbb{E}H^2$ at that threshold. With $k = \lceil 4/\Delta \rceil$ and M from Lemma 3.3, $k\alpha - 1 \geq 1$. Consequently, using $|H_a| \leq 1$ for the short-gap branch, for every $h \neq 0$ and $N \geq 1$,

$$s_h(N) - \tfrac{1}{2} \;\geq\; \frac{\Delta}{4M\binom{k}{2}}. \tag{3.3}$$

The bound is intended only to establish strict positivity. Equation (2.2) shows that $\Delta = 0$ exactly when $H = 0$, equivalently $\mu = \mathsf{B}$. Finally the projections onto $\mathcal{Z}_s(Y^{(h)})$ converge to the projection onto $\mathcal{Z}_\infty(Y^{(h)})$, identified in the last assertion of Proposition 3.1. This proves the asserted equivalences. The finite block length that approximates the supremum may depend on the lag, the gap to the past, the law and the accuracy. □

4 How much Bernoulli mass is present?

A positive Bernoulli mass b forces every word w of length L to have frequency at least $b\,2^{-L}$. The converse follows from all lengths together. Some features of this calculation hold for every stationary binary law.

Proposition 4.1. *For a stationary binary law ν, the largest a with $\nu \geq a\mathsf{B}$ equals its Bernoulli ergodic mass $b(\nu)$. The numbers $d_L(\nu) = 2^L \min_{|w|=L} \nu([w])$ decrease to $b(\nu)$. For every nonempty finite D, the map $x \mapsto Y^D$ preserves both entropy and $b(\nu)$.*

Proof. Ergodic uniqueness identifies the maximal dominated multiple, since the difference of the two invariant measures is invariant. Refining a word by its final sign gives $d_{L+1} \leq d_L$. At the limit the cylinder inequalities give domination first on the finite-coordinate algebra, by stationarity and subdivision, then on its generated sigma field.

Translate D so its endpoints are 0 and r. Every specified output word of length L has exactly 2^r input words of length $L + r$. Choose the first r signs and solve successively for each right endpoint. Thus Haar measure maps to Haar, and the Shannon block entropies satisfy

$$H_L(Y^D) \leq H_{L+r}(X) \leq H_L(Y^D) + r.$$

Division by L gives equality of entropy rates. A stationary binary law of entropy one has every length-L block uniform, since $h \leq H_L/L \leq 1$. Thus a stationary lift of B is B. Apply this to the images of the ergodic components to obtain the mass assertion. □

There is no rate in the word formula valid for all stationary laws. For any k the periodic process obtained by choosing a random phase of a binary de Bruijn cycle of order k [3] gives every word of length k frequency 2^{-k}, but has Bernoulli mass zero. Such a cycle is the Euler circuit in the binary word graph with vertices the $(k-1)$-words and edges the k-words. Mixing its law with mass b of B gives $d_k = 1$ and $d_{k+1} = b$, since the cycle omits at least one $(k+1)$-word. This example concerns the class of all stationary laws; it does not establish the absence of a rate within the narrower class of logarithmic Liouville laws.

Proof of Theorem 1.2. The event $E = \{H_i = 0 \text{ for all } i\}$ is invariant and belongs to $\mathcal{A}$. By (2.1), its conditional sign law, when it has positive mass, is B. Conversely on a Bernoulli component the Pinsker factor is trivial and the sign has mean zero. The componentwise nilprojection therefore vanishes there. This gives $b = \mu(E)$, and Proposition 4.1 gives the last expression in (1.2).

To identify which components can be atoms we use two spectral facts. First, the full sign system has no irrational eigenvalue. For completeness let $P(\omega)$ be the law of the signs sampled along a prime progression in an ergodic base ω, as in [7, §4]. With $q = 1$ and sign coordinates, (2.4) and the ensuing prime average identify μ with the global sign reversal of $\int P(\omega)\,\mathrm{d}\pi(\omega)$. A sign reversal commutes with the shift. The whole law is thus a factor of the system of prime progressions, which has no irrational eigenvalue by [7, Theorem 3.11]. This argument works for the windows used here. Second, Tao's logarithmic two-point theorem [26] gives Lebesgue spectral measure for f. On almost every component the spectral measure of f therefore charges no rational frequency, by disintegration and countability.

On an ergodic component with only rational eigenvalues, $\mathcal{Z}_\infty$ is an inverse limit of finite rotations. To see this, take an ergodic finite-dimensional nilfactor in its minimal presentation G/Γ, generated by G^0 and the translating element. Its horizontal rotation is $G/(G_2\Gamma)$, with G_2 connected and rational [7, §4.2.1, Appendix A]. A positive-dimensional horizontal rotation has an irrational eigenvalue. If the horizontal rotation is finite, then $G^0 = G_2$, by discreteness of the projected lattice. The closed group generated by $[G, G^0]$ is G_2, as the quotient is generated by a central subgroup and a single element. Hence $G_3 = G_2$, and nilpotence forces $G_2 = G^0 = \{1\}$. This nilfactor is finite. The second spectral fact now gives $H = 0$ on almost every such component. Conversely B itself has no irrational eigenvalue. The total mass in question is therefore exactly b.

For clarity, the set of components involved is measurable. A positive weighted sum of spectral measures of countably many bounded dense cylinder observables gives a measurable kernel K_ω of maximal spectral type. Its atom function $(\omega, t) \mapsto K_\omega(\{t\})$ is measurable by decreasing balls. Integrating the indicator of positive atom mass over irrational t against this kernel detects the components in question. For any *fixed* irrational t, the global absence of this eigenvalue gives $K_\omega(\{t\}) = 0$ almost everywhere. Hence a component with an irrational eigenvalue cannot be an atom of π. Taking a countable union proves the final assertion of the theorem. The common-countable-spectrum condition also appears in the study of logarithmic Sarnak orthogonality under the name almost countable spectrum [14]; here it is applied to the components of the fixed arithmetic law. □

5 A quartic bias makes each of its pairs predictable

Let

$$A(n) = \lambda(n)\lambda(n+1), \qquad D(n) = \lambda(3n+1)\lambda(3n+2), \qquad C(n) = A(n)D(n).$$

Write $c_{a,q}$ for the limiting harmonic mean of $C(n)$ conditional on $n \equiv a \bmod q$, and $c = c_{0,1}$. Take a common extraction for the countably many finite residues and shifts used here. An auxiliary extraction is needed if their joint laws have not already been specified. The unmarked sign law on $3I_\nu$ is the same μ, since dilating the endpoints by three changes unnormalised harmonic mass by $O(1)$. In its clock extension put $F_h = X_0X_h$. At $m = 3n$, F_3 reads $A(n)$ because $\lambda(3)^2 = 1$, and TF_1 reads $D(n)$. Write $\mathbb{E}_{a,q}$ for expectation on $J = 3a \bmod 3q$ in this extension. Time for the sampled pair process at that origin advances by T^{3q}.

The next result combines a conditional-law identity, an imported numerical bound, and a specialised reduction to the discrete spectrum. The reduction is proved in Lemma 5.3; the bound comes from [21].

Theorem 5.1. *For every fixed $q \geq 1$ and $a \bmod q$,*

$$c_{a,q} = \mathbb{E}_{a,q}(H_0H_1H_2H_3), \qquad |c_{a,q}| \leq \frac{196567331}{402653184}(1-b). \tag{5.1}$$

In the joint law of the two pair processes sampled at $n + qj$, conditional on $n \equiv a \bmod q$, let P_{d} be projection onto the discrete spectrum on each ergodic component. With A_0, D_0 the two present signs,

$$c_{a,q} = \int (P_{\mathrm{d}}A_0)(P_{\mathrm{d}}D_0). \tag{5.2}$$

Each pair separately admits, for every $N \geq 1$ and $\epsilon > 0$, a predictor from a finite block of its own past ending at $-N$ with success at least $(1 + |c_{a,q}|)/2 - \epsilon$.

The two integrands in (5.1) and (5.2) serve different purposes. The first uses four one-site biases on an extension. The second involves linear spectral information inside the observed pair processes. A nonzero c would therefore be visible to each observer separately.

5.1 From the four sites to two observers

The first identity of (5.1) follows by conditioning (2.1) on the finite origin. The same formula bounds $|c_{a,q}|$ by each of $\mathbb{E}_{a,q}|H_0H_3|$ and $\mathbb{E}_{a,q}|H_1H_2|$. Each product is accessible to the corresponding observer. Indeed the conditional mean of F_3 on the unmarked $\mathcal{Z}_\infty$ is H_0H_3 by (2.2), and the same holds for TF_1 and H_1H_2. Naturality lifts these identities to the clock system. Taking the power T^{3q} and restricting to the indicated invariant class preserves these projections by [17, Theorem 2 in the separate text]. Projecting to either pair process identifies the bias there as a $\mathcal{Z}_\infty$-measurable function, hence accessible from every remote past. Once the past is separated by at least two time steps, the corresponding input sites are disjoint. Conditional on all the amplitudes and the clock, (2.1) then gives the product's conditional mean even after that past is added. Projecting back proves the prediction assertion, including $N = 1$ by allowing the predictor to ignore the most recent sign.

The same four-site identity shows what the bias requires of a single sign. Lemma 3.3, applied at every threshold, makes the distribution of $|H_j|$ on the selected clock class equal to the unconditioned distribution of $|H|$. Hölder's inequality therefore gives

$$|c_{a,q}| \le \int |H|^4 \,\mathrm{d}\mu. \tag{5.3}$$

There is also a direct link with all the unmarked pair observers. Theorem 9.1 of *An explicit bound below one half for a quartic Liouville correlation* [21] gives $\Delta = 1 - h_\mu(T) \ge (439/250)|c_{a,q}|$ for these same fixed progressions. If $c_{a,q} \ne 0$, take $\delta = (439/250)|c_{a,q}|$ in the construction leading to (3.3). Since $\mathbb{E}H^2 \ge \Delta \ge \delta$, the threshold $\sqrt{\delta/2}$ and $k = \lceil 4/\delta \rceil$ give the lower bound $\delta/(4M\binom{k}{2})$ at every unmarked lag. The larger bound $(1 + |c_{a,q}|)/2$ on success in Theorem 5.1 applies to the two particular observers in the selected class.

For the numerical bound write $m_0 = 3q$. On the joined sign and finite-clock system take the prime progression law restricted to steps $p \equiv 1 \bmod m_0$. Formula (2.4), with all observables signs, identifies the observed signs and initial clock with the sign reversal of the progression observations and their initial clock. This holds on every initial class by the same prime average. Within the progression law the clock increases by one per coordinate, so this is an identity of stationary systems. In particular every pair and the four-site moment are unchanged by the reversal.

On each ergodic base component, [7, Theorem 4.3] permits projection of every observable to $\mathcal{Z}_\infty$ and approximation by bounded real observables on finite nilfactors that include the clock. Tao's theorem [26] makes the spectral projection of f onto the rational frequencies zero on almost every component, including in the clock extension because its spectral measure there is unchanged. Every approximating observable thus has integral zero on each connected component of its nilsystem. These components form a finite cycle. For a cycle of length v including the clock, split the prime step into units $t \bmod v$ compatible with $1 \bmod m_0$. The prime progression integral for each class is the integer progression integral for steps $t + v\ell$ [7, Theorem 4.4], with the class selected by clock functions if necessary. On each initial connected component this uses the totally ergodic map T^v with four translated, real, centred observables. The four-term progression has characteristic factor of step at most two. The horizontal and vertical estimates of *An explicit bound below one half for a quartic Liouville correlation* [21, Propositions 4.2, 5.1 and 9.6] bound its absolute value by $B = 196567331/402653184$. Related geometric inequalities are developed in *Sharp inequalities and stability for real profiles on arithmetic progressions* [24]. Each selected origin class has mass $1/m_0$ on every ergodic base component. On components projecting to the Bernoulli sign law the joining with a finite cycle is a product, by total ergodicity of the Bernoulli shift. Every nonzero step there has zero four-site moment. Average the bound over the remaining mass $1 - b$, then remove the finite approximation. This proves the inequality in (5.1).

5.2 Reading spectral atoms on each component

We use the following consequence of uniform Wiener–Wintner theory and finite eigenfunction approximation. It records the largest spectral atom on each component, rather than the total mass of the discrete part.

For a bounded observable F, write $e(t) = \exp(2\pi i t)$ and define

$$M_H F(x) = \sup_{\theta\in\mathbb{T}} \left| \frac{1}{H} \sum_{j=1}^{H} F(T^j x) e(j\theta) \right|, \qquad L_H(F) = \|M_H F\|_4. \tag{5.4}$$

For an ergodic component ω let $\sigma_{F,\omega}$ be the (unnormalised) spectral measure with Fourier coefficients $\langle T^j F, F\rangle_\omega$.

Lemma 5.2. *For bounded complex F, $M_H F$ converges almost surely and in L^p, $1 \le p < \infty$, to the invariant function*

$$a_F(\omega) = \sqrt{\sup_{t\in\mathbb{T}} \sigma_{F,\omega}(\{t\})}.$$

In particular $\lim_H L_H(F)^4 = \int a_F^4 \,\mathrm{d}\pi$ and it is zero exactly when $P_{\mathrm{d}} F = 0$.

Proof. Work first on one ergodic component. The vanishing assertion for continuous spectrum is the classical uniform Wiener–Wintner theorem [4, Theorem 21.8]. For the part g of F with continuous spectral measure, g is bounded because the discrete projection is conditional expectation on the Kronecker factor. Apply van der Corput with differences up to J_0 to each modulated block, before taking the supremum in frequency. For a fixed difference d the phase contributes only $e(d\theta)$. The pointwise ergodic theorem applied simultaneously to the countably many bounded products $T^d g\,\bar g$ gives, almost everywhere,

$$\limsup_H |M_H g(x)|^2 \ll \frac{\|g\|_\infty^2}{J_0} + \frac{1}{J_0} \sum_{d=1}^{J_0} |\langle T^d g, g\rangle_\omega|.$$

Wiener's lemma and Cauchy–Schwarz make the last mean tend to zero as $J_0 \to \infty$.

For a finite sum of normalised eigenfunctions at distinct frequencies the limit is the largest absolute coefficient. Indeed their moduli are one on this component. Outside small disjoint arcs around the finitely many resonant frequencies all normalised geometric sums tend uniformly to zero. Inside one arc all the other terms tend uniformly to zero, and at its centre the corresponding term is exact. Approximate the discrete projection by a sequence of such sums in L^2. For each approximation error u, the ergodic theorem gives $\limsup_H M_H u(x) \le \int |u| \,\mathrm{d}\omega$ almost everywhere. The error in the largest absolute coefficient is at most $\|u\|_2$. Squared coefficients are the spectral atom masses, giving almost sure convergence on this component. The limit a_F is measurable, for example as the limit of the component integrals of $M_H F$ by bounded convergence. Disintegration now gives almost sure convergence globally, and $|M_H F|, a_F \le \|F\|_\infty$ gives all the stated norm convergences. □

5.3 Which vertical frequencies can contribute?

The identity (5.2) needs more than the conditional sign law: it says that the relevant pair information already lies in the componentwise Kronecker factor. The following progression-space calculation identifies exactly which vertical modes remain there. We retain the notation $\widetilde{\eta}_{a,q}$ and S from (2.5).

Lemma 5.3 (Discrete projection of a pair on progression space)**.** *Let (X, η, T) be ergodic, let f be a bounded real observable, and put $g = \mathbb{E}_\eta(f \mid \mathcal{Z}_\infty(X))$. In the progression system $(X^{\mathbb{Z}}, \widetilde{\eta}_{a,q}, S)$, let P_{d} denote componentwise discrete projection. For every $h \neq 0$,*

$$P_{\mathrm{d}}(f(x_0) f(x_h)) = P_{\mathrm{d}}(g(x_0) g(x_h)).$$

On a finite nilfactor, all vertical levels above two can be removed from this projection. At level two it is the simultaneous vertical average of the two observables: only pairs of opposite vertical characters survive. In particular,

$$\int \prod_{i=0}^{3} f(x_i)\,\mathrm{d}\widetilde{\eta}_{a,q} = \int P_{\mathrm{d}}(f(x_0)f(x_3))P_{\mathrm{d}}(f(x_1)f(x_2))\,\mathrm{d}\widetilde{\eta}_{a,q}. \tag{5.5}$$

The identity remains valid after multiplying both integrands by a fixed finite cyclic mask of the initial point, and after integrating over base ergodic components.

Proof. A local quadratic estimate. For $|u_j| \le 1$, Bernstein's inequality and fourth-moment expansion give

$$\int_{\mathbb{T}} \sup_{\beta\in\mathbb{T}} \left| \frac{1}{H} \sum_{j=1}^{H} u_j e(\theta j^2 + \beta j) \right|^4 \mathrm{d}\theta \ll H^{-1}. \tag{5.6}$$

Indeed, a trigonometric polynomial of degree at most H stays above half its maximum on an arc of length $\gg H^{-1}$. Its fourth-power supremum is therefore at most $O(H)$ times its fourth-power integral. In the subsequent double integral, the surviving indices satisfy $i+j=k+l$ and $i^2+j^2=k^2+l^2$. These equations identify the two unordered pairs, leaving $O(H^2)$ terms. The factor $H \cdot H^{-4}$ gives (5.6).

Removing the residual observable. Normalize $\|f\|_\infty \le 1$, and set $V_h = f(x_0)f(x_h) - g(x_0)g(x_h)$. Apply the same supremum bound in the linear frequency to $M_H V_h$, and expand the fourth power. The linear-frequency integral imposes one relation among four time indices. If the four input pairs are disjoint, expansion of each V_h leaves a residual $f-g$ at a site that occurs only once. The prime characteristic-factor theorem annihilates that term. This remains true with a prescribed step class, using finite cyclic masks as in Section 2.2. A collision between two input pairs imposes an additional relation of the form $i-k \in \{0, h, -h\}$, or its counterpart. There are $O_h(H^2)$ such terms. Hence

$$\int (M_H V_h)^4\,\mathrm{d}\widetilde{\eta}_{a,q} \ll_h H^{-1}.$$

Lemma 5.2 gives $P_{\mathrm{d}} V_h = 0$. The same residual argument removes $f-g$ from the four-site moment, whose sites are distinct.

Eliminating levels above two. Approximate g in L^2 on finite nilfactors and use bounded vertical Fejér sums there. Consider a pair of vertical characters χ, χ' at positions $0, h$ on a step-k factor, $k \ge 2$. The symmetry (2.6) also permits $x_j \mapsto z^{j^2} x_j$, since j^2 is an integer combination of the binomial polynomials of degrees at most two. Along the observed pair process this changes the modulation by a quadratic phase whose leading character is $\chi+\chi'$. If that character is nonzero, its Haar image is uniform on $\mathbb{T}$. The frequency supremum absorbs the accompanying linear phase, so integrating (5.6) over the vertical symmetry gives zero discrete projection by Lemma 5.2.

If $\chi' = -\chi \neq 0$ and $k \ge 3$, use instead $x_j \mapsto z^{j^3} x_j$. The cubic coefficients cancel, and the quadratic coefficient is $3h\chi'$. It is nonzero because $h \neq 0$ and the character lattice is torsion free. The same argument applies. Thus, above level two, the only pair terms with a discrete projection have both characters zero. They descend to the next quotient, and this removes all higher levels.

Identifying the level-two projection. At level two the terms with opposite characters have componentwise pure point spectrum. For continuous observables $u_\chi, v_{-\chi}$, write a trajectory in the step-two Hall–Petresco space as

$$x_j = g_0 g_1^j g_2^{\binom{j}{2}} \Gamma, \qquad g_2 \in G_2.$$

With $a_h = g_0 g_1^h g_0^{-1}$, centrality gives

$$x_{j+h} = g_2^{hj+\binom{h}{2}} a_h x_j.$$

Consequently $u_\chi(x_j)v_{-\chi}(x_{j+h})$ is a linear phase in j times the value at x_j of $u_\chi(x)v_{-\chi}(a_h x)$. The latter function is constant on each vertical fibre and hence descends to the horizontal rotation. Its orbit sequence is uniformly almost periodic. On a trajectory generic for the countably many pair autocorrelations, the resulting pure point autocorrelation measure is the component spectral measure. The pair term therefore has pure point spectrum on almost every component.

The terms with nonopposite characters have already been eliminated. Thus, on a two-step factor, P_{d} of a pair is precisely its average under simultaneous vertical translation of its two coordinates. At step one the progression components are rotations, so the whole pair observable has componentwise pure point spectrum. Continuous approximation within the character spaces and bounded Fejér summation extend these statements to bounded measurable observables.

The four-site identity. For the four sites $0, 1, 2, 3$, the symmetries of degrees $0, 1, 2, 3$ eliminate every nonzero vertical character at any level $k \geq 3$, by the Vandermonde argument of Lemma 2.3. At level two the three equations of degrees $0, 1, 2$ leave exactly the character vectors

$$(-\chi, 3\chi, -3\chi, \chi).$$

Both the outer pair and the inner pair therefore have opposite characters. Averaging each pair separately under simultaneous vertical translation preserves the four-site integral. By the preceding paragraph these averages are exactly their discrete projections, proving (5.5) on finite nilfactors.

All vertical actions fix finite cyclic factors, so the same calculation permits the asserted mask. The coordinate marginals control approximation errors in L^2 for pair observables and in L^1 for their products. Naturality identifies the discrete projections of lifted finite-factor observables with the lifts of their discrete projections. Contractivity of conditional expectation therefore permits passage to the inverse limit. Finally, the label of a base ergodic component is the same at every progression coordinate and is invariant under S. Disintegrating with this label proves the last assertion. □

We now apply the lemma with the finite clock modulo $m_0 = 3q$ and prime steps $p \equiv 1 \bmod m_0$. The transport identity (2.4) identifies the sign-and-clock factor of this progression system, after global sign reversal, with the joined arithmetic system used in Section 5.1. The two pair observables are $f(x_0)f(x_3)$ and $f(x_1)f(x_2)$; both are unchanged by the reversal. Apply (5.5) with the mask of the initial class $J = 3a \bmod 3q$.

To pass to the observers' time, first replace S by S^{3q}. The componentwise Kronecker factors of S and S^{3q} agree by the power identity for $\mathcal{Z}_1$. Next restrict to the indicated S^{3q}-invariant clock class and normalize its measure; componentwise projection restricts to that class. Finally take the factor generated by the two sampled pair sequences. Naturality makes each projection the lift of the discrete projection in that joint pair factor. After these three operations, (5.5) is exactly (5.2).

6 The rectangle of four primes

The frequency that detects an atom can depend on the ergodic component. A rectangle average gives a way to test all these frequencies using only translates of the observed signs. The proof combines the algebra of a box average with spectral theory and the W-tricked prime estimate of [27, Lemma 5.3]. For related spectral tests along primes, see [2]; that

work is a comparator, not an additional input to the proof. Let F be real and $|F| \le 1$ on a stationary system. For $R \ge 2$ an integer, take p_1, p_2, q_1, q_2 independently and uniformly in $P_R = (R, 2R] \cap \mathbb{P}$. Define

$$B_R(F) = \int \mathbb{E}_{p_1,p_2,q_1,q_2} \prod_{i,j=1}^{2} T^{p_i+q_j} F. \tag{6.1}$$

Theorem 6.1. *For every R, $B_R(F) \ge \int a_F^4 \, \mathrm{d}\pi$. Moreover,*

$$P_{\mathrm{d}} F = 0 \quad \Longleftrightarrow \quad L_H(F) \longrightarrow 0 \quad \Longleftrightarrow \quad \liminf_{R\to\infty} B_R(F) = 0.$$

Whenever these conditions hold, $B_R(F) \to 0$. For the adjacent pair $F = A$ on its ordinary sign-shift process, each condition implies $c = 0$.

Proof. At a fixed origin put $M_{pq} = T^{p+q}F$, and let M act by averaging on $\ell^2(P_R)$ with its probability norm. The pointwise box average b_R equals $\mathrm{tr}((M^*M)^2) \ge 0$. Set $\delta_R = \sup |\mathbb{E}_{p,q} u_p v_q M_{pq}|$ over real $|u_p|, |v_q| \le 1$. Then

$$\delta_R^4 \le b_R \le \delta_R.$$

The first bound follows from the operator norm, bounded by $b_R^{1/4}$, and also applies to complex unit tests. For the second hold the opposite row and column fixed, use them as bounded tests, and average. Translating a modulated length-H block by $p + q$ changes its normalised sum, after correcting the phase, by at most $8R/H$. Average this identity in p, q. Their correcting phases separate, so each resulting matrix coefficient is bounded by $b_R^{1/4}$ at the shifted origin. Minkowski and stationarity give

$$L_H(F) \le B_R(F)^{1/4} + 8R/H. \tag{6.2}$$

Let $H \to \infty$ with R fixed and use Lemma 5.2.

For the other direction put $W = \prod_{p \le w} p$. Denote the supremum distance between the Fourier transforms of the prime law on P_R and the uniform law on integers in $(R, 2R]$ coprime to W by $e(w, R)$. The U^2 case of the W-tricked von Mangoldt estimate [27, Lemma 5.3] gives $\lim_{w\to\infty} \limsup_{R\to\infty} e(w, R) = 0$. In each reduced class subtract the estimates at the two ends of the interval, use U^2 to dominate linear Fourier coefficients, remove prime powers and use $\log p = \log R + O(1)$ on this dyad. The finitely many reduced classes have their usual masses by the prime number theorem for fixed W.

Replacing one index in the integrated product costs at most $e(w, R)$ by the spectral theorem, since with the other three fixed the variable enters through $\int G\, T^p K$ for bounded products G, K. For the resulting coprime-integer matrix use $b_R \le \delta_R$ and extend the tests to the full interval by zero at the omitted integers. For an interval of R consecutive indices, Plancherel bounds the resulting bilinear form by $R^{-1} \sup_\theta |\sum_{s=2}^{2R} T^{2R+s} F\, e(s\theta)|$ because the Fourier transforms of the two test sequences in the convolution have L^2 norms at most $\sqrt{R}$. Restoring the coprime normalisation, integrating and using stationarity gives

$$B_R(F) \le 4e(w, R) + \big(W/\varphi(W) + o_W(1)\big)^2 \big(2L_{2R}(F) + O(1/R)\big). \tag{6.3}$$

First send $R \to \infty$ at fixed w, then $w \to \infty$. This proves the equivalence. For $F = A$, naturality of $\mathcal{Z}_1$ lifts the vanishing of $P_{\mathrm{d}} A$ from the single pair to the joint pair system in (5.2) with $q = 1$, giving $c = 0$. □

The order of the frequency supremum. On the Haar two-torus take $T(x,t) = (x+t, t)$, $F(x,t) = \cos(2\pi x)$. All nonzero global autocorrelations vanish by integration in t. The global spectral measure is Lebesgue measure times $1/2$. On almost every fibre t is irrational and F has two atoms of mass $1/4$. Thus $L_H(F) \to 1/2$ and $B_R(F) \geq 1/16$. A frequency fixed or optimised outside the integral over origins can miss exactly the componentwise information used in (5.4). The proposed quadratic estimate in *Polynomial phases in ordinary two-point correlations of multiplicative functions* [23, Candidate Theorem 1.2 and §6] concerns a global ordinary sum with a single polynomial over the full interval. Passing to (5.4) would require control of frequencies selected separately on length-H blocks, after the logarithmic law is taken. The local Fourier-uniformity problem for small sets of frequencies is studied in [16]. The estimates there concern, in particular, the single Liouville sequence; they do not supply the required componentwise cancellation for $A(n) = \lambda(n)\lambda(n+1)$ over all frequencies.

6.1 Sampling four pairs of actual prime factors

The preceding criterion uses translates of the adjacent-pair process. We now show that the same rectangle can be read after selecting actual prime factors from each integer. The unsigned density calculation belongs to the Kubilius-model tradition; see [8] for a recent treatment for sieve-theoretic sequences. Its use with the Liouville signs requires the weak-convergence and block arguments below.

For an integer $u > 6$ and $m > 1$, let $p_u(m)$ be the smallest prime factor of m exceeding $m^{1/u}$. Put

$$T_u(m) = \frac{u \log p_u(m)}{\log m}, \qquad E_u(n) = \{1 < T_u(n) \leq 2, \quad 3 < T_u(n+1) \leq 6\},$$

with $T_u(m) = \infty$ when no such prime exists. On $E_u(n)$ write $n = pr$ and $n+1 = qs$, using the selected factors $p = p_u(n)$ and $q = p_u(n+1)$. Complete multiplicativity gives

$$A(n) = \lambda(n)\lambda(n+1) = \lambda(r)\lambda(s).$$

The two removed prime signs cancel.

Take a common extraction of the harmonic windows for the signs and all translates of the masks $\mathbf{1}_{E_u}$, for every integer $u > 6$. This countable extension is stationary. All expectations below are in this fixed limiting law, with the additional independent prime draws used in (6.1). In particular, each u, R and each translation are fixed when the window limit is taken. Set

$$P_{u,R} = \int \mathbb{E}_{p_1,p_2,q_1,q_2} \prod_{i,j=1}^{2} T^{p_i+q_j} \mathbf{1}_{E_u},$$

and let $\Gamma_{u,R}$ be the conditional mean of $\prod_{i,j} T^{p_i+q_j} A$ given that all four masks equal one. It is defined whenever $P_{u,R} > 0$.

Proposition 6.2. *In every common limiting law just specified,*

$$\lim_{M\to\infty} \sup_{\substack{u,R\geq M \\ u,R\in\mathbb{Z}}} |P_{u,R} - 12^{-4}| = 0, \qquad \lim_{M\to\infty} \sup_{\substack{u,R\geq M \\ u,R\in\mathbb{Z}}} |B_R(A) - \Gamma_{u,R}| = 0.$$

The conditional means are therefore defined for all sufficiently large u, R. In particular, $\Gamma_{u_j,R_j} \to 0$ along any sequence with $u_j, R_j \to \infty$ implies $c = 0$.

The window limit precedes these two-parameter limits. The proposition provides no effective rate against sign-dependent tests and no assertion in which the window index, u and R increase simultaneously.

Proof. Step 1: count masks at one and two scales. Fix finitely many sites. On a dyadic interval $(X, 2X]$, replace each moving threshold $(n+b)^{a/u}$ by $X^{a/u}$, for $a = 1, 2, 3, 6$. At fixed u and fixed shifts b, Mertens' theorem bounds the reciprocal-prime sum over each difference band by $o(1)$. Counting divisibility before summing over these primes introduces an additional $O_u(X^{6/u-1})$ in normalised density. Thus the replacement has error $o(1)$, also in harmonic mean, because $u > 6$.

Each first-hit event is a difference of no-hit events. A product of masks at t disjoint adjacent pairs is consequently a linear combination, with coefficients depending only on t, of events forbidding specified prime residues in bands $(X^{1/u}, X^{a/u}]$. The same reduction works at two fixed scales $u' > 6u$: their full prime bands are disjoint, even when both scales are prescribed at one site.

Use the Bonferroni truncations at consecutive orders $k, k+1$, where $k = \lfloor u/24 \rfloor$. For all sufficiently large u, every squarefree product occurring in these truncations is at most

$$X^{6(k+1)/u} < X^{1/2}.$$

There are at most $(4t)^{k+1}$ residue assignments per product in the two-scale calculation. The Chinese remainder theorem therefore gives a total counting error $O_{u,u',t}(X^{1/2})$, uniformly over prefixes of the dyad; omit u' in the one-scale case. Partial summation gives the corresponding harmonic estimate. The total sum of forbidden residue densities is bounded by a constant C_t. Hence the Bonferroni remainder in the comparison model is at most $C_t^{k+1}/(k+1)!$.

For fixed shifts, no distinct sites coincide modulo a marked prime once X is large enough. The error in separating their Euler factors is

$$O_t\left(\sum_{p > X^{1/u'}} p^{-2}\right) = o(1),$$

where $u' = u$ in the one-scale case. Mertens' theorem gives $1/a$ for a single no-hit factor through $X^{a/u}$. The model mass of one adjacent-pair mask is therefore

$$v = (1 - 1/2)(1/3 - 1/6) = 1/12.$$

It follows that there is a function $\rho_t(u) \to 0$ such that the mass of t disjoint masks in the limiting law is $v^t + O_t(\rho_t(u))$. At two scales $u' > 6u$, the joint mass is $v^{2t} + O_t(\rho_t(u))$, after enlarging ρ_t if needed. These bounds are uniform in the fixed shifts *after* the window limit; the threshold for X may depend on the shifts and the scales.

To pass from dyads to the harmonic windows, first discard all integers below a fixed threshold X_0. Their harmonic mass is bounded independently of the window. At most two partial endpoint dyads have bounded harmonic mass as well. On every remaining full dyad the normalised error is bounded by $O_t(\rho_t(u))$ plus a quantity uniformly small once X_0 is large. Sum these bounds with the dyads' harmonic masses as weights, divide by the window mass, and then let X_0 increase. This gives the asserted estimates for arbitrary windows with $\log(V_\nu/U_\nu) \to \infty$.

For $R > 2$ all four sums $p_i + q_j$ are even. Distinct sums therefore select disjoint adjacent pairs. Repeated rows or columns, or an equality between two otherwise distinct sums, have probability $O(|P_R|^{-1})$: in the last case three prime draws determine the fourth. Applying the preceding estimate to each fixed rectangle yields

$$P_{u,R} = v^4 + O(\rho_4(u) + |P_R|^{-1}). \tag{6.4}$$

Step 2: retain the density in the presence of fixed sign tests. For fixed disjoint adjacent pairs with origins $b_1, \dots, b_t$, put

$$Z_u = \prod_{r=1}^{t} T^{b_r} \mathbf{1}_{E_u} - v^t.$$

The one- and two-scale estimates imply

$$|\langle Z_u, Z_{u'}\rangle| \le C_t \rho_t(u) \qquad (u' > 6u),$$

after another enlargement of ρ_t. Thus $Z_u \rightharpoonup 0$ in L^2. Indeed, if a fixed test $Q \in L^2$ contradicted this conclusion, choose a subsequence u_j with $u_{j+1} > 6u_j$ and, after a fixed complex rotation, $\mathrm{Re}\langle Z_{u_j}, Q\rangle \ge \varepsilon > 0$. Expansion of the squared norm, boundedness of Z_u, and the displayed covariance bound give

$$\left\| \frac{1}{J} \sum_{j=1}^{J} Z_{u_j} \right\|_2^2 \le \frac{C_t}{J} + \frac{2C_t}{J^2} \sum_{i<j} \rho_t(u_i) \longrightarrow 0,$$

a contradiction. The test Q may depend on the arithmetic signs. No independence between those signs and one fixed mask has been assumed.

Step 3: obtain local Fourier cancellation uniformly in the block length. Set

$$G_u = A(v^{-1}\mathbf{1}_{E_u} - 1), \qquad |G_u| \le 11.$$

Fix a block length K. The fourth-power supremum bound used in (5.6), followed by integration in the linear frequency, reduces $L_K(G_u)^4$ to a constant times K^{-3} times the sum of moments with $i + j = k + l$. If the four adjacent pairs are disjoint, expansion of the centred masks and Step 2 show that their moment tends to zero as $u \to \infty$. In fact each subset of the masks may be tested against the fixed product of the four signs $A_i A_j A_k A_l$, and the resulting constants cancel. An overlap imposes an additional relation between the indices, leaving $O(K^2)$ terms with uniformly bounded moments. Consequently,

$$\limsup_{u\to\infty} L_K(G_u)^4 \ll K^{-1}.$$

Partition a length-H block into length-K blocks and a remainder. Stationarity and Minkowski's inequality imply

$$L_H(G_u) \le L_K(G_u) + O(K/H).$$

Given a tolerance, choose K first, then u large enough for the fixed-K bound, and finally require H sufficiently large compared with K. It follows that $L_H(G_u) \to 0$ jointly as $\min(u, H) \to \infty$.

For clarity, applying (6.3) to the uniformly bounded family $G_u/12$ uses the following order. Given $\varepsilon > 0$, choose w so that $\limsup_R e(w, R)$ is small enough, keeping $W = \prod_{p\le w} p$ fixed. Choose K so that the contribution of $K^{-1/4}$, multiplied by the now fixed factor $(W/\varphi(W))^2$, is smaller than ε. Next choose u large enough for this K, and choose R large enough both for the Fourier discrepancy bound and for K/R to be small. Homogeneity of B_R then gives

$$B_R(G_u) \longrightarrow 0 \qquad (\min(u, R) \to \infty).$$

All these choices take place in the already fixed limiting law.

Step 4: replace the four entries and cancel the normalisation. For real G_u, the pointwise rectangle is nonnegative and the bilinear supremum from Theorem 6.1 satisfies

$$\delta_R(G_u)^4 \le b_R(G_u), \qquad \int \delta_R(G_u) \le B_R(G_u)^{1/4} = o(1).$$

Telescope the four replacements $A \mapsto v^{-1}A\mathbf{1}_{E_u} = A + G_u$. In each difference, fixing the opposite row and column leaves two separate tests against the matrix $T^{p+q}G_u$. The other

three entries have modulus at most 12, so the integrated difference is at most $12^3 \int \delta_R(G_u)$. Summing the four differences gives

$$\left| B_R(A) - v^{-4} P_{u,R} \Gamma_{u,R} \right| \le 4 \cdot 12^3 B_R(G_u)^{1/4} = o(1).$$

Since $v^{-4} = 12^4$, equation (6.4) gives $v^{-4} P_{u,R} \to 1$. Together with $|\Gamma_{u,R}| \le 1$, this proves both limits in the proposition. If a sequence of conditional means tends to zero, then $\liminf_R B_R(A) = 0$ because $B_R(A) \ge 0$. Theorem 6.1 gives $c = 0$. □

6.2 Seven independent sites can leave a positive rectangle

We compare the unconditioned rectangle $B_R(A)$ of (6.1) for two stationary laws whose marginals agree on every seven sites. Under the fair product law the adjacent products are themselves independent and fair, by the word count in Proposition 4.1. If the two rows or the two columns coincide, the four factors cancel in pairs. Otherwise the smallest origin $p_i + q_j$ occurs just once, making the expectation zero. Thus for $m = |P_R|$ this law has $B_R(A) = 2/m - 1/m^2 \to 0$.

Finite-order independence with nontrivial global structure is classical [5]. The second law is a periodic variant of the matching construction in *An isolated quartic correlation under signed prime dilations* [22, Proposition 7.1]. Choose b uniformly in $\{0, 1\}$ and match the vertices of $\mathbb{Z}$ at $(2k + b, 2k + b + 1)$. Let ξ_k be independent fair signs, let ϕ, α be independent Haar points of $\mathbb{T}$, and take these choices independently. With $s(t) = \operatorname{sgn} \cos(2\pi t)$, set

$$X_{2k+b} = \xi_k, \qquad X_{2k+b+1} = \xi_k s(\phi + k\alpha).$$

The null-set convention at zeros of the cosine is immaterial. The law is stationary since shifting the matching changes b and may reindex k, which translates ϕ by α. Any product at distinct sites vanishes on integrating out the ξ_k unless the sites form complete matched edges. Up to seven sites, at most three such edges occur. For one or three, a half-turn in ϕ reverses the sign. For two distinct edges, the two phases are independent Haar, since the corresponding integer matrix has full rank. Thus every seven-site marginal is fair and independent. In particular every fixed-frequency length-H sum of the adjacent pairs has normalised second moment $1/H$.

Nevertheless for $A_j = X_j X_{j+1}$,

$$\lim_{R \to \infty} B_R(A) = \tfrac{1}{6}. \tag{6.5}$$

The four prime sums are even for $R > 2$. When $b = 1$, the even origins select unmatched edges. Conditional on (ϕ, α) their products are deterministic signs times successive adjacent products of the fair ξ_k. They are therefore independent and fair, and their conditional rectangle has the value $2/|P_R| - 1/|P_R|^2$ just computed. When $b = 0$, every entry instead equals $s(\phi + (p_i + q_j)\alpha/2)$. Approximate s by a bounded Fejér polynomial. A Fourier term with indices a, b', c, d survives integration in α only if

$$(a + b')p_1 + (c + d)p_2 + (a + c)q_1 + (b' + d)q_2 = 0.$$

If at least one coefficient is nonzero, conditioning on the other three draws costs at most $1/|P_R|$. The identically vanishing choices are $(k, -k, -k, k)$, which also survive the integration in ϕ. For any fixed degree take $R \to \infty$. The error upon removing the approximation is bounded by four times its one-site L^1 error. The resulting limit on $b = 0$ is $\sum_k |\widehat{s}(k)|^4 = 1/3$, since the nonzero coefficients have modulus $2/(\pi|k|)$ for odd k. Averaging the two phases of the matching proves (6.5).

In the aperiodic model of [22, Proposition 7.1], the density of matched edges is $0 < d < 1/2$. Four independent integer draws with common distribution ν give limiting rectangle $d^4/3$ when $\|\nu\|_\infty \to 0$. The periodic law here uses the parity shared by the four prime sums to give (6.5). The two stationary examples establish non-identifiability of the *unconditioned* rectangle from seven-site marginals. The arithmetic masks E_u and their conditional rectangle $\Gamma_{u,R}$ require actual prime factors. Their relation to $B_R(A)$ for Liouville is the separate transfer in Proposition 6.2.

7 Reflection and conditioning on a fifth integer

Write $\langle\cdot\rangle_I$ for harmonic expectation, $t = 2n+1$, and $L(n) = (n, n+1, 3n+1, 3n+2)$. Two consequences of the affine isotypy theorem of Tao and Teräväinen [29, Theorem A.1] apply without passing to an extraction. For fixed $q, K \geq 1$, a mod q and r mod K,

$$\langle C(n)(\mathbf{1}_{n\equiv a(q)} - \mathbf{1}_{n\equiv -1-a(q)})\rangle_I = o_q(1), \tag{7.1}$$

$$\langle \mathbf{1}_{n\equiv a(q)} C(n)(\mathbf{1}_{\Omega(t)\equiv r(K)} - 1/K)\rangle_I = o_{q,K}(1). \tag{7.2}$$

For odd K, (7.2) also holds with C replaced by the indicator of any prescribed four-sign pattern at $L(n)$. Thus fixing the fifth residue preserves each of the sixteen frequencies, whatever their values are.

For (7.1) put $m = 6n+3$. The four signs become those at $m + h_i$, $h = (-3, 3, -1, 1)$, with constant multipliers of total product $\lambda(6)^2\lambda(2)^2 = 1$. An origin weight antisymmetric in $n \mapsto -1-n$ becomes an odd periodic weight $W(m)$ supported on $m \equiv 3 \bmod 6$, of period $6q$. The harmonic change of variable has constant multiplier six and error $O_q(1/\log(V/U))$. Separate $d = (m, 6q)$ and write $m = du$. On the units modulo $6q/d$ the character expansion of $W(du)$ contains only odd characters. For a term containing $\chi(u)$ let $F(s)$ be the generalised limiting correlation of this observable and $\lambda(du + sh_i)$ at the four shifts. Affine isotypy in [29, Theorem A.1] gives $F(-1) = \chi(-1)F(1)$ because the product of multiplicative functions agrees with χ at all but finitely many primes. Symmetry of the shifts also gives $F(-1) = F(1)$. The multiplier in [29, Theorem A.1(iii)] is an integer and may equal -1; no evaluation of λ at a negative argument is required, since only the fixed shifts are reflected. This removes each odd term, including those arising on nonunit strata of the original modulus. The argument holds for every generalised limit, giving (7.1).

For (7.2) expand the centred residue indicator over nontrivial $z^K = 1$. Since $m = 3t$, $z^{\Omega(t)} = z^{-1}z^{\Omega(m)}$. On the same gcd stratum the central observable is a constant times $z^{\Omega(u)}\chi(u)$. For a product of j of the four signs the product of the functions at almost every prime is $(-1)^j z\chi(p)$. If $(-1)^j z \neq 1$ it cannot weakly pretend to any character. Primes congruent to one modulo the common character modulus already contribute a positive multiple of $\log\log x$ to the squared distance. Theorem A.1(ii) of [29] gives zero for that term. Every K works for $j = 4$, and all j work for odd K. The case $j = 0$ gives frequency $1/K$ within the origin class. Expansion of a four-sign cylinder completes the proof.

Reflection also preserves the joint law of the finite family $C(2^j n+b)$, $0 \leq j \leq d$, $0 \leq b < 2^j$, while reversing every branch $b \mapsto 2^j - 1 - b$ and the origin class $a \mapsto -1-a$. To see this, expand a finite product of these quartics into an even number of affine signs $\lambda(a_i n + b_i)$. Choose D_0 a multiple of all a_i and put $m = 2D_0 n + D_0$. The corresponding shifts are $2D_0 b_i/a_i - D_0$. Negation replaces b_i by $a_i - b_i$ with the same multiplicative constants. In each quartet this interchanges the two outer and the two inner sites and reverses its branch. After the gcd decomposition of the periodic origin weight, isotypy at $s = -1$ gives precisely the character's parity, hence the reflected weight. Equal shifts may be combined by multiplying their observables. Finite Walsh inversion gives the joint-law statement. All moduli and the depth are fixed before the window limit.

The mass b and the bias c remain undetermined for the actual arithmetic law. A proof that the conditional eight-sign mean in Proposition 6.2 tends to zero would settle the indicated quartic correlation. Determining the full logarithmic Chowla law would require the one-lag prediction in Theorem 1.1 to be exactly at chance for every limiting law.

AI assistance. OpenAI language models assisted with the mathematical exploration and exposition. OpenAI Codex contributed to the division and conditional-law arguments, the prediction and rectangle calculations, verification of bibliographic statements, finite checks and preparation of the manuscript. OpenAI ChatGPT and Codex also assisted with the entropy proof for overlapping observations, the exact finite prediction examples and the comparison of the matching models. The numerical four-function bound is imported from the cited work; its computational certificate was not rerun in this revision.

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley, 2006. Chapter 4, Entropy rates of a stochastic process, Theorem 4.2.1. 10.1002/047174882X.ch4.

[2] C. Cuny, T. Eisner and B. Farkas, Wiener's lemma along primes and other subsequences, *Adv. Math.* **347** (2019), 340–383. 10.1016/j.aim.2019.02.005.

[3] N. G. de Bruijn, A combinatorial problem, *Proc. Sect. Sci. K. Ned. Akad. Wet. Amsterdam* **49** (1946), no. 7, 758–764. Published version.

[4] T. Eisner, B. Farkas, M. Haase and R. Nagel, *Operator Theoretic Aspects of Ergodic Theory*, Graduate Texts in Mathematics, vol. 272, Springer, 2015. 10.1007/978-3-319-16898-2. Theorem 21.8; author version.

[5] L. Flaminio, Mixing k-fold independent processes of zero entropy, *Proc. Amer. Math. Soc.* **118** (1993), no. 4, 1263–1269. 10.1090/S0002-9939-1993-1154245-2.

[6] N. Frantzikinakis, Ergodicity of the Liouville system implies the Chowla conjecture, *Discrete Anal.* (2017), Paper no. 19. 10.19086/da.2733.

[7] N. Frantzikinakis and B. Host, The logarithmic Sarnak conjecture for ergodic weights, *Ann. of Math.* (2) **187** (2018), 869–931. 10.4007/annals.2018.187.3.6.

[8] O. Gorodetsky, A Kubilius model for sieve-theoretic sequences, *Anal. Math.* **52** (2026), no. 3, 999–1022. 10.1007/s10476-026-00178-w.

[9] M. Górska, M. Lemańczyk and T. de la Rue, On orthogonality to uniquely ergodic systems, *J. Anal. Math.* **159** (2026), 581–674. 10.1007/s11854-026-0449-9. Remark 6.22, also in arXiv:2404.07907v3.

[10] A. Hildebrand, On a conjecture of Balog, *Proc. Amer. Math. Soc.* **95** (1985), no. 4, 517–523. 10.1090/S0002-9939-1985-0810155-4.

[11] A. Hildebrand, On integer sets containing strings of consecutive integers, *Mathematika* **36** (1989), 60–70. 10.1112/S0025579300013577.

[12] M. Hochman, *Notes on ergodic theory*, Hebrew University of Jerusalem, 5 July 2017, §11.5, Theorem 11.5.2. Author's notes.

[13] B. Host and B. Kra, Nonconventional ergodic averages and nilmanifolds, *Ann. of Math.* (2) **161** (2005), 397–488. 10.4007/annals.2005.161.397.

[14] W. Huang, M. Tan and L. Xu, *Almost countable spectrum and logarithmic Sarnak conjecture*, preprint, arXiv:2511.04419v1, 6 November 2025.

[15] A. Kanigowski, J. Kułaga-Przymus, M. Lemańczyk and T. de la Rue, On arithmetic functions orthogonal to deterministic sequences, *Adv. Math.* **428** (2023), 109138. 10.1016/j.aim.2023.109138. Proposition 2.2 and Corollary 1.12 in arXiv:2105.11737v3.

[16] A. Kanigowski, M. Lemańczyk, F. K. Richter and J. Teräväinen, On the local Fourier uniformity problem for small sets, *Int. Math. Res. Not. IMRN* **2024** (2024), no. 15, 11488–11512. 10.1093/imrn/rnae134.

[17] A. Leibman, Host–Kra and Ziegler factors, and convergence of multiple averages, Appendix A to V. Bergelson, *Combinatorial and Diophantine applications of ergodic theory*, Handbook of Dynamical Systems, vol. 1B, Elsevier, 2006, 745–869 (chapter). 10.1016/S1874-575X(06)80037-8. Author's separate text, Theorem 2 (Theorem A.2 in the chapter).

[18] M. Lemańczyk, M. D. Lemańczyk and T. de la Rue, A note on Sarnak processes, *Ann. Inst. Henri Poincaré Probab. Stat.* **62** (2026), 1912–1925. 10.1214/25-AIHP1573.

[19] R. McNamara, *A counterexample to Hildebrand's conjecture on stable sets*, preprint, arXiv:2312.08544v2, 25 September 2025.

[20] M. Pivato, Invariant measures for bipermutative cellular automata, *Discrete Contin. Dyn. Syst.* **12** (2005), no. 4, 723–736. 10.3934/dcds.2005.12.723.

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

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

[23] B. Pouly, *Polynomial phases in ordinary two-point correlations of multiplicative functions*, Zenodo (2026). 10.5281/zenodo.24578857. Candidate Theorem 1.2 and Section 6.

[24] B. Pouly, *Sharp inequalities and stability for real profiles on arithmetic progressions*, Zenodo preprint, 2026. 10.5281/zenodo.24443242.

[25] W. Sawin, Dynamical models for Liouville and obstructions to further progress on sign patterns, *J. Number Theory* **213** (2020), 1–15. 10.1016/j.jnt.2020.01.012.

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

[27] T. Tao and J. Teräväinen, Odd order cases of the logarithmically averaged Chowla conjecture, *J. Théor. Nombres Bordeaux* **30** (2018), 997–1015. 10.5802/jtnb.1062.

[28] T. Tao and J. Teräväinen, The structure of correlations of multiplicative functions at almost all scales, with applications to the Chowla and Elliott conjectures, *Algebra Number Theory* **13** (2019), 2103–2150. 10.2140/ant.2019.13.2103.

[29] T. Tao and J. Teräväinen, The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures, *Duke Math. J.* **168** (2019), 1977–2027. 10.1215/00127094-2019-0002.