

Correlation laws and identifiability of multiplicative functions on quadratic traces

Brice Pouly

AI-assisted research and preparation; see the closing declaration.

Abstract

We study which completely multiplicative sign functions can be identified from the statistics of $h(n^2 + 1)$, $h(n^2 + 2)$ and $h(n^2 - 2)$. In a class of finite products of quadratic characters of odd prime conductor, completed without zeros and modified at primes with summable inverse square roots, diagonal correlations of orders two and four determine the function. The second-order data recover the conductor and prime corrections, leaving exactly the ramified local signs undetermined. Fourth-order detectors recover those signs, including at the small primes; the inverse also extends to quadratic components at two. No bounded horizon suffices uniformly: for every prescribed window, two finite-conductor sources can share its complete law and disagree with density arbitrarily close to one half on each trace. For an arbitrary prescribed source h , we construct factors $fg = h$ such that, for each factor separately, the three traces jointly have the uniform word law at every horizon. The choice is continuous and computable relative to h , with a finite bound uniform in h for approximation of a canonical synchronous coupling. The same choice treats every horizon sequence $o(\log \log N)$. The factors need not be independent. Their synchronous empirical law projects exactly to the product's empirical word law; when a limiting canonical coupling exists, it retains the product's word law, which need not identify h as a function.

Contents

1	What is lost when a trace is replaced by its law?	2
2	Local prime laws and the existence of correlations	6
3	What second-order data determine	8
4	A fourth-order detector for every ramified sign	11
5	A whole window can hide the ramified signs	15
6	An arbitrary source as a product of normal factors	19
7	Selecting the factors continuously and computably	24
8	Product word laws in the synchronous coupling	28
9	Scope of the identification results	32

1 What is lost when a trace is replaced by its law?

A sequence of signs contains two kinds of information. It tells us the sign at each position, and it tells us how often finite patterns occur. Passing to pattern frequencies forgets the origin of each occurrence. For a completely multiplicative function, it is not evident how much arithmetic information survives that loss: the signs at different integers are constrained by their prime factorizations, but these constraints need not be visible in separate pattern laws.

Our observations are the three quadratic traces

$$h(n^2 + 1), \quad h(n^2 + 2), \quad h(n^2 - 2).$$

A concrete example already separates pointwise and statistical information. Write χ_{17} for the Legendre character modulo 17, and set

$$h_\epsilon(m) = (-1)^{v_2(m)} \epsilon^{v_{17}(m)} \chi_{17}\left(m/17^{v_{17}(m)}\right), \quad \epsilon \in \{\pm 1\}. \quad (1.1)$$

These are nonvanishing completely multiplicative functions. They differ at the prime 17 and nowhere else on the primes. Their traces disagree with density $1/9$ on each of the three curves. Nevertheless, every joint law of at most three scalar readings, at arbitrary fixed shifts and with any choice of traces, is the same for the two functions. Four readings can distinguish them:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=2}^{N+1} \prod_{a \in \{0,2,17,34\}} h_\epsilon((n+a)^2 + 1) = \frac{51 - 4\epsilon}{289}. \quad (1.2)$$

The pointwise observation at $n = 4$ already reads $h_\epsilon(17) = \epsilon$. The failure is therefore a failure of low-order statistical identification, not an equality of the traces themselves.

This example has two continuations. Within a structured class, one can isolate a local prime contribution and recover the missing sign by a fourth-order correlation. Without that structural restriction, even the complete marginal word laws can lose an arbitrarily prescribed multiplicative function. Both conclusions concern the same three observations, but different classes of sources.

1.1 Observations and conventions

Let $\mathcal{M}_\pm$ denote the group of completely multiplicative functions from the positive integers to $\{\pm 1\}$. We call its elements *sources*. Put

$$\mathcal{D} = \{-1, -2, 2\}, \quad Q_d(X) = X^2 - d.$$

For an integer $H \geq 1$, the word read at origin $n \geq 2$ is

$$W_h^H(n) = (h(Q_d(n+a)))_{0 \leq a < H, d \in \mathcal{D}} \in \{\pm 1\}^{3H}. \quad (1.3)$$

Coordinates are ordered first by a , then by $d = -1, -2, 2$. Labels remain distinct even when two labels happen to give the same integer. All arguments in (1.3) are positive. For finite-valued observations $A(n), B(n)$, our empirical-law notation is

$$\mathcal{L}_N(A, B) = \frac{1}{N} \sum_{n=2}^{N+1} \delta_{(A(n), B(n))}, \quad \mathcal{L}_N(A) = \frac{1}{N} \sum_{n=2}^{N+1} \delta_{A(n)}.$$

Here δ_z denotes unit point mass at z . These are laws for a uniform integer origin, not for random prime signs. We write $\mu_{N,h}^H = \mathcal{L}_N(W_h^H)$ and ν_h^H for its limit when that limit exists.

The uniform law on $\{\pm 1\}^m$ is U_m . A source is *normal on the three traces* if $\nu_h^H = U_{3H}$ for every H .

The diagonal correlations used for reconstruction are

$$C_{d,h}(a) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=2}^{N+1} h(Q_d(n))h(Q_d(n+a)), \quad (1.4)$$

$$D_{d,h}(a, b, c) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=2}^{N+1} h(Q_d(n))h(Q_d(n+a))h(Q_d(n+b))h(Q_d(n+c)). \quad (1.5)$$

All shifts are nonnegative integers. Repetition is allowed. These limits, and the other correlations considered in the structured class, will be proved to exist. *Order* counts scalar factors; *horizon* bounds their shifts. Order four with all shifts is an infinite family of numbers, not four numerical measurements.

On a finite space, $d_{TV}(\alpha, \beta) = \frac{1}{2} \sum_x |\alpha(x) - \beta(x)|$. Empirical laws use an ordinary average unless stated otherwise. A logarithmic average uses weights $1/n$, divided by $\sum_{n=2}^{N+1} 1/n$. All fixed-word ordinary limits below give the same logarithmic limits by summation by parts. Unsubscripted logarithms are natural; $\log_2$ is used for base two.

1.2 Identification in a structured class

For an odd prime q , define the completed quadratic character

$$G_{q,\epsilon}(m) = \epsilon^{v_q(m)} \chi_q(m/q^{v_q(m)}), \quad \chi_q = (\cdot/q), \quad \epsilon \in \{\pm 1\}. \quad (1.6)$$

The local sign at the conductor is free, rather than zero. These are known character completions: the form $\xi^{v_p(m)} \chi(m/p^{v_p(m)})$ also occurs in Allouche–Goldmakher’s description of nonvanishing mock characters [1, Theorem 2.11 in the cited preprint]. We study the statistical identification of these local factors, not their introduction as new objects. Let $\mathcal{C}$ consist of the sources

$$h(m) = \sigma^{v_2(m)} \prod_{q \in \Lambda} G_{q,\epsilon_q}(m) \prod_{q \in A} (-1)^{v_q(m)}, \quad \sum_{q \in A} q^{-1/2} < \infty, \quad (1.7)$$

where $\sigma \in \{\pm 1\}$, Λ is a finite set of odd primes, and A is a set of odd primes disjoint from Λ . The last product is finite at each integer. We refer to Λ as the conductor-prime set; its product is the odd conductor of the character on units. The local signs ϵ_q must not be confused with the actual values $h(q)$: the other characters also contribute at q . The reference to mock characters concerns the local building blocks; it does not assert that all sources in $\mathcal{C}$ are automatic in one common base.

Theorem 1.1 (Second-order ambiguity and fourth-order reconstruction). *Every finite correlation of a source in $\mathcal{C}$ exists. The three families $C_{d,h}(a)$ determine σ , Λ and A . Their fibre inside $\mathcal{C}$ consists exactly of the $2^{|\Lambda|}$ choices of the signs ϵ_q . Adding the three families $D_{d,h}(a, b, c)$ determines every ϵ_q , and hence h .*

Order three is insufficient on the whole class, even if mixed-trace correlations and all shifts are supplied. The reconstruction remains valid after replacing σ^{v_2} by any completed quadratic component at two, as specified in Corollary 4.3.

The second-order step averages correlations along multiples of carefully chosen integers. The resulting quantity is a squared conditional mean. It has three different local values: an ordinary prime, a prime correction, and a conductor prime cannot have the same signature. At a conductor prime, however, orthogonality of valuation layers removes the sign ϵ_q from every pair correlation. A fourth-order configuration makes three readings enter the same first valuation layer; their product contains $\epsilon_q^3 = \epsilon_q$. This is the mechanism behind (1.2).

1.3 A bounded horizon cannot suffice uniformly

Theorem 1.2 (A complete finite window can be ambiguous). *Fix integers $H \geq 1$, $B \geq 2$, a source h_* with $h_*(2) = -1$, and $0 < \eta < 1/2$. There exist sources $f, g \in \mathcal{C}$ with finite conductor-prime and correction sets such that*

$$f(m) = g(m) = h_*(m) \quad (1 \leq m \leq B), \quad \nu_f^H = \nu_g^H,$$

and, for every $d \in \mathcal{D}$,

$$\text{dens}\{n : f(Q_d(n)) \neq g(Q_d(n))\} > \frac{1}{2} - \eta.$$

Each trace of each source has mean zero. All word frequencies exist, with the same ordinary and logarithmic limits. Given $\zeta > 0$, the common H -word law can additionally be chosen within ζ of U_{3H} . Each of the two sources still has zero word-entropy rate and zero topological entropy for its three-trace sequence.

Here the sources depend on H . This does not contradict Theorem 1.1, which uses unbounded shifts. A conductor prime $q \equiv 1 \pmod{8}$, $q > 17$, hides its local sign throughout every three-trace window $H \leq q^{1/3}$. In a different comparison, the word law of a one-conductor source changes from close to far from U_{3H} around $H = (\log_2 q)/3$. Section 5 proves both statements and keeps their different limits separate.

1.4 Complete marginal laws can hide an arbitrary product

For a probability law ν on $\{\pm 1\}^m$, define

$$\mathcal{J}_m(\nu)(x, y) = 2^{-m} \nu(x \odot y), \tag{1.8}$$

where $\odot$ denotes coordinatewise multiplication. Equivalently, choose $X \sim U_m$ independently of $V \sim \nu$, and set $Y = X \odot V$. Both marginals are uniform, while $X \odot Y$ has law ν .

Theorem 1.3 (Continuous selection of normal factors). *Fix a finite set S of primes and signs ξ_q on S . There are maps $F, G : \mathcal{M}_\pm \rightarrow \mathcal{M}_\pm$, with $G(h) = hF(h)$, such that $F(h)(q) = \xi_q$ on S and*

$$\sup_{h \in \mathcal{M}_\pm} d_{\text{TV}}\left(\mathcal{L}_N(W_{F(h)}^H, W_{G(h)}^H), \mathcal{J}_{3H}(\mu_{N,h}^H)\right) = O_{S,H,\varepsilon}(N^{-1/2+\varepsilon}) \tag{1.9}$$

for every fixed $H \geq 1$ and $\varepsilon > 0$. The maps are continuous in the product topology of prime signs and computable relative to the input h : every output prime sign is decided permanently by a finite computation using finitely many input signs. In particular, a computable h has computable factors.

The same maps work for every sequence $H_N = o(\log \log N)$, without a monotonicity requirement or a prior choice of that sequence. The bound then has a constant allowed to depend on (H_N) . An explicit finite bound valid for all H, N is given in (7.8).

Thus every $h \in \mathcal{M}_\pm$ is the pointwise product of two sources normal on the three traces at all horizons. No limiting law of h is assumed. The law on the right of (1.9) follows its empirical law at the same N , including any oscillation. The factors depend on h ; a single factor valid for all inputs would fail on the input equal to that factor. The uniform law concerns all three traces of each factor separately. The two factors need not be independent in their synchronous observation. The target $\mathcal{J}_{3H}(\mu_{N,h}^H)$ encodes the product's word law, not an identification of the arbitrary function h .

1.5 Context and route through the proof

The first contribution is an arithmetic inverse with an exact description of its second-order fibre. Its closest direct-correlation context is Klurman’s work on polynomial correlations of multiplicative functions [11, Theorem 1.3 in the cited preprint]. There, local prime factors describe global correlations with an error controlling the relevant arithmetic tails. A direct formula does not by itself invert the data: here one must isolate local factors, distinguish their types without dividing by possibly zero energies, and recover the ramified signs left invisible by pair correlations. The inverse question itself also has an arithmetic precedent. Klurman–Mangerel [12, Question 1.11 and Proposition 1.12 in the cited preprint] study multiplicative functions whose binary correlations resemble those of a Dirichlet character. Their observations are linear shifts of the function, rather than our three quadratic traces. Our source class consists of nonzero completions and summable prime corrections, and it has the specific second-order ambiguity of Theorem 1.1.

Higher-order reconstruction also has a precise history outside this arithmetic setting. Grünbaum–Moore [9] proved reconstruction up to translation for rational-valued functions on cyclic groups from correlations through order six, and through order four when the group has odd order; see the accounts in Jaming–Kolountzakis [10, Section 3.4] and Casper–Orozco [3, Introduction]. The latter extend the problem to finite abelian groups and give an explicit reconstruction. Grimm–Baake [8] discuss the ambiguities of autocorrelation in homometry. The threshold four is therefore not a new abstract reconstruction principle. In our inverse, the conductor is unknown, valuation layers are retained, and the correction may be infinite. Periodic approximation of the data does not automatically pass a finite-group inverse to this setting; Sections 2–4 establish that passage directly.

The finite-window result combines exact invisibility, a prescribed prefix, and macroscopic disagreement, rather than merely small correlations. Approximation by congruence data belongs to the context of rationally almost periodic sequences [2]. The comparison with the uniform word law uses the classical Weil bound in the form recalled by Ding [5, Lemma 1]. The logarithmic scale for that comparison and the much longer sufficient range $H \leq q^{1/3}$ for exact invisibility of a ramified sign answer different questions; neither scale is inferred from the other.

For the second main construction, almost-sure word normality and a near-square-root error already have a direct linear precedent in Najnudel [13, Proposition 6.3 in the cited preprint]. The arithmetic input for our simultaneous quadratic words is instead the square-class estimate of de la Bretèche–Wang–Xu [4, Lemmas 3.2–3.3], with the fixed-field dependence supplied by Evertse–Silverman [6, Theorem 1(b)]. We do not import a central limit theorem from those works. Once the set of normal sources has full Haar measure, existence of two normal factors follows by intersecting that set with a translate. The additional assertions here are the arithmetic realization of the canonical coupling, its finite bound uniform in the input, and a continuous, relatively computable selection that also works at the stated growing horizons.

Conditional expectations and constructive typical-point selection are established tools; see Galatolo–Hoyrup–Rojas [7]. More specifically, the companion paper on selective normality [15, Section 5] uses the same positive-potential selection scheme: exact input-independent normalizers, a summable conditional tail, and irrevocable rational decisions. Its free primes are constrained to preserve observations. Here the costs use square-class multiplicities with all unprescribed prime signs free; these estimates give the near-square-root rate and the growing-window conclusion. We keep the specialized proof autonomous. The coupling identities, entropy formulas and compact-family statement are consequences of this construction, not separate new principles of probability or compactness.

Sections 2–4 prove the inverse and its exact ambiguity. Section 5 treats the bounded-horizon obstruction. Sections 6–7 give the probabilistic factorization and deterministic selector; the former also yields statements on fixed arithmetic progressions and summable cosets. Section 8 describes the product word laws retained in the coupling. None of these conclusions asserts that a source with almost everywhere positive traces must be constant, or gives a new lower bound for its number of negative readings.

2 Local prime laws and the existence of correlations

To reconstruct a source from its correlations, we first need a legitimate decomposition of those correlations. Finite prime products are handled by congruences. The summability in (1.7) is used to pass from finite products to the actual source; it is not merely a condition on a formal Euler product. Local-factor formulas for polynomial correlations have the arithmetic precedent of Klurman [11, Theorem 1.3 in the cited preprint]. The argument below proves the particular approximation and uniform tail bounds needed by our inverse.

2.1 The correction tail on a quadratic trace

For A as in (1.7), let $u_A(m) = \prod_{q \in A} (-1)^{v_q(m)}$, and let $u_{A, \leq Y}$ retain only $q \leq Y$.

Lemma 2.1 (A tail bound in the observation variable). *For every $d \in \mathcal{D}$, every $Y \geq 2$ and every $N \geq 1$,*

$$\frac{1}{N} \#\{2 \leq n \leq N+1 : u_A(Q_d(n)) \neq u_{A, \leq Y}(Q_d(n))\} \leq 8 \sum_{\substack{q \in A \\ q > Y}} q^{-1/2}. \quad (2.1)$$

For a fixed collection of r shifted readings, the corresponding upper limit of the disagreement proportion is at most $8r$ times the same tail.

Proof. A disagreement requires an odd valuation at some omitted prime. For odd q , the congruence $n^2 \equiv d \pmod{q}$ has at most two roots, so its contribution is at most $2/q + 2/N$. A contributing prime satisfies

$$q \leq (N+1)^2 + 2 \leq 6N^2.$$

Consequently $2/q + 2/N \leq (2 + 2\sqrt{6})q^{-1/2} < 8q^{-1/2}$. Sum over omitted primes. A fixed translation of the interval changes at most twice the translation length in its count; this boundary vanishes after division by N . A union bound gives the final assertion. $\square$

Multiplication by another sign source does not change the positions of disagreement. The lemma therefore applies to the correction part of any $h \in \mathcal{C}$ and to a normal factor multiplied by a summable correction later on.

2.2 Valuation layers and their cancellation

We use normalized Haar measure on $\mathbb{Z}_q$ to abbreviate limits of uniform residue averages. Its use entails no randomness assumption about a deterministic source. Let $P \in \mathbb{Z}[X]$ have only simple roots modulo the odd prime q . Extend χ_q by $\chi_q(0) = 0$, and define

$$U_{P,j}(x) = \mathbf{1}_{q^j | P(x)} \chi_q(P(x)/q^j), \quad j \geq 0. \quad (2.2)$$

The support is precisely $v_q(P(x)) = j$, and the function is periodic modulo q^{j+1} . At the finitely many q -adic roots its value may be defined arbitrarily without changing any integral.

Lemma 2.2 (Conditional orthogonality of layers). *For $j \geq 1$, $U_{P,j}$ has mean zero on every parent class modulo q^j . If P, Q both have simple roots modulo q , $a \in \mathbb{Z}$, and $j \neq k$, then*

$$\mathbb{E}[U_{P,j}(x)U_{Q,k}(x+a) \mid x \bmod q] = 0. \quad (2.3)$$

Moreover,

$$G_{q,\epsilon}(P(x)) = \sum_{j \geq 0} \epsilon^j U_{P,j}(x) \quad (2.4)$$

converges in $L^2(\mathbb{Z}_q)$, also conditionally on each residue modulo q .

Proof. If the parent class is a root modulo q^j , write $x = r + q^j t$. Then

$$P(r + q^j t)/q^j \equiv P(r)/q^j + P'(r)t \pmod{q}.$$

The slope is invertible. The sum of its Legendre character over $t \bmod q$ is zero. A nonroot parent has an identically zero layer. If $j > k$, the factor $U_{Q,k}(x+a)$ is constant modulo q^j , so summing the higher layer over its children proves (2.3). Interchange the factors for $k > j$.

At almost every point, exactly one summand in (2.4) is nonzero. The error after keeping $j < K$ is supported on $q^K \mid P(x)$, of measure $\rho_P(q)/q^K$, where $\rho_P(q)$ is the number of roots modulo q . The same support estimate within a residue class proves conditional L^2 convergence. $\square$

In a product of two completed readings, different layers vanish and equal layers have coefficient $\epsilon^{2j} = 1$. Therefore the joint law of two readings, conditional on $x \bmod q$, is independent of the common sign ϵ . Its first moments and its product moment determine that two-sign law. The two readings need not be independent.

2.3 Passing from congruences to the source

Each Q_d has simple roots at every odd prime at which it has a root. For finitely many prime components, truncate their valuations below a depth K . The observations are then periodic. Uniform integers become uniform residues in a complete period, and the Chinese remainder theorem factors the limiting averages into independent local residue averages. At a conductor or correction prime, the omitted set is a union of at most two root classes for each reading and has measure $O_r(q^{-K})$. Letting K tend to infinity gives the actual finite-prime law.

For the infinite correction set, Lemma 2.1 makes the empirical correlations uniformly Cauchy after truncation. The discrepancy between products of signs is at most twice the proportion of altered readings. Hence every finite correlation exists. The same argument followed by Fourier inversion proves every finite word law exists. Summation by parts gives the logarithmic limits.

There is a second, different tail estimate which will allow shifts to vary after these limits are formed. At a visible correction prime, the probability of an odd valuation is

$$\mathbb{P}(v_q(Q_d(x)) \text{ is odd}) = 2(1 - q^{-1}) \sum_{j \geq 0} q^{-(2j+1)} = \frac{2}{q+1}. \quad (2.5)$$

If $c_{q,A}$ denotes its local r -reading moment, then, uniformly in all shifts,

$$|1 - c_{q,A}| \leq \frac{4r}{q+1}. \quad (2.6)$$

The series of these deficits converges since $\sum_{q \in A} q^{-1} < \infty$. Thus local products can be approximated uniformly in the shifts by finitely many components. The finite- N bound (2.1) justifies passage to the source; the local bound (2.6) justifies subsequent averages and limits in the shifts. We use them for these separate tasks.

2.4 The component at two

The quadratic valuations at two are particularly simple:

$$(-1)^{v_2(Q_d(n))} = c_d(-1)^n, \quad c_{-1} = 1, \quad c_{-2} = c_2 = -1. \quad (2.7)$$

For odd n , the three valuations are $(1, 0, 0)$; for even n , they are $(0, 1, 1)$. If $\sigma = -1$, any product of an odd number of scalar readings has a factor $(-1)^n$ times a constant depending on the shifts. After truncation, all odd-prime components have odd period. Averaging over the parity therefore gives zero. The tail bounds pass this cancellation to the source.

This parity cancellation explains why the pair (1.1) loses third-order information as well as second-order information. It does not by itself explain the missing second-order sign; that comes from Lemma 2.2.

3 What second-order data determine

A correlation at one shift mixes all primes. Averaging over shifts constrained to be multiples of M turns this mixture into a product of nonnegative local energies. By forcing high divisibility at every prime except one, we can then expose that one prime without dividing by an Euler product which might vanish.

3.1 An energy extracted from the autocorrelations

Define, using only the data in (1.4),

$$E_{d,h}(M) = \lim_{R \rightarrow \infty} \frac{1}{R} \sum_{a=1}^R C_{d,h}(Ma), \quad M \geq 1. \quad (3.1)$$

Proposition 3.1 (Local energy factors). *The limit in (3.1) exists and*

$$E_{d,h}(M) = \prod_q B_{q,d,h}(v_q(M)). \quad (3.2)$$

For an odd prime put $\rho = \rho_d(q) = 1 + \chi_q(d) \in \{0, 2\}$. Its factor is as follows:

Prime type	$B_{q,d,h}(0)$	$B_{q,d,h}(e), e \geq 1$
$q \notin A \cup \Lambda$	1	1
$q \in A$	$(1 - 2\rho/(q+1))^2$	$1 - \frac{4\rho q}{q^e(q+1)^2}$
$q \in \Lambda$	q^{-2}	$1 - \rho/q^e$

At two, $B_{2,d,h}(0)$ is 1 if $\sigma = 1$ and 0 if $\sigma = -1$; $B_{2,d,h}(e) = 1$ for $e \geq 1$.

Proof. For a finite-period local observation $z(x)$, averaging $z(x)z(x+Ma)$ over a and then x gives

$$\mathbb{E}(\mathbb{E}[z(x) \mid x \bmod q^e]^2), \quad e = v_q(M). \quad (3.3)$$

Indeed, conditional on its class modulo q^e , the second point becomes uniform in that class. The same assertion for a q -adic observation follows by valuation truncation. The Chinese remainder theorem gives the product for finitely many primes; (2.6) passes both the average and the product to the infinite correction set.

For a correction prime, the mean at depth zero is $1 - 2\rho/(q+1)$ by (2.5). At positive depth, the sign is constant on every nonroot class. On a root class modulo q^e , the mean is

$$(-1)^e \frac{q-1}{q+1}.$$

There are ρ such classes, each of measure q^{-e} . Substitution in (3.3) gives

$$1 - \frac{\rho}{q^e} \left(1 - \frac{(q-1)^2}{(q+1)^2} \right) = 1 - \frac{4\rho q}{q^e(q+1)^2}.$$

For a conductor prime, a root class has conditional mean zero by Lemma 2.2, and a nonroot class has a determined sign. This gives $1 - \rho/q^e$ at positive depth. At depth zero, the mean is $-1/q$. To see this, count solutions of $y^2 = x^2 - d$ over $\mathbb{F}_q$: the substitution $(x-y)(x+y) = d$ gives $q-1$ solutions, so $\sum_x \chi_q(x^2 - d) = -1$. Completion contributes mean zero on roots, leaving that same mean. Its square is q^{-2} . Finally, (2.7) gives the factor at two. $\square$

The conductor factor at depth zero is q^{-2} even on a trace with no roots modulo q . A correction at such a prime, in contrast, is invisible on that trace.

3.2 Isolating a prime without a denominator

For any prime q , let

$$M_{j,q} = j!/q^{v_q(j!)}, \quad T_{d,h}(q) = \lim_{j \rightarrow \infty} E_{d,h}(M_{j,q}). \quad (3.4)$$

The q -valuation is zero, whereas the valuation at any other fixed prime tends to infinity. Proposition 3.1 and the summable correction deficits yield

$$T_{d,h}(q) = B_{q,d,h}(0). \quad (3.5)$$

For completeness, truncate the set of odd prime components first. On this finite set all factors except the q -factor tend to one. The omitted factors contribute a uniformly small sum of deficits, independently of j , and can then be restored. This proves the limit in its stated order.

Every odd prime is visible on at least one of the three traces. A definite choice is

$$d(q) = \begin{cases} -1, & q \equiv 1 \pmod{4}, \\ -2, & q \equiv 3 \pmod{8}, \\ 2, & q \equiv 7 \pmod{8}. \end{cases} \quad (3.6)$$

The quadratic character identities behind this choice are the supplementary laws in [14, §27.9]. At this trace the three possible type signatures are

$$1, \quad \left(\frac{q-3}{q+1} \right)^2, \quad \frac{1}{q^2}, \quad (3.7)$$

respectively. They are distinct: at $q=3$ they are $1, 0, 1/9$, while for $q \geq 5$ the middle value is at least $1/9$ and the last is at most $1/25$. The isolated factor at two recovers σ .

This identifies σ , A and Λ from the three pair families. Conversely, Lemma 2.2 shows that changing any ϵ_q leaves every local pair law, hence every global pair correlation, unchanged. The conclusion holds even for mixed-trace pairs. The different choices really give different sources, since at a conductor prime

$$h(q) = \epsilon_q \prod_{\ell \in \Lambda \setminus \{q\}} \chi_\ell(q). \quad (3.8)$$

Thus the second-order fibre has exactly $2^{|\Lambda|}$ elements. In particular, the parameterization in (1.7) is unique once all its parameters have been recovered.

3.3 An explicit inverse without conductor primes

The subclass

$$\mathcal{S}_{1/2} = \left\{ h \in \mathcal{M}_{\pm} : \sum_{q:h(q)=-1} q^{-1/2} < \infty \right\}$$

has no conductor components, so its pair correlations alone determine h . Here one can replace the denominator-free construction by particularly simple ratios. The following formulas also give a directly reproducible example of how the energy is used.

Proposition 3.2 (Energy ratios for summable sign changes). *For $h \in \mathcal{S}_{1/2}$, $E_{d,h}(6) > 0$ for each d . Moreover,*

$$\frac{E_{d,h}(3)}{E_{d,h}(6)} = \begin{cases} 0, & h(2) = -1, \\ 1, & h(2) = 1, \end{cases} \quad (3.9)$$

$$\frac{E_{-2,h}(18)}{E_{-2,h}(6)} = \begin{cases} 5/3, & h(3) = -1, \\ 1, & h(3) = 1, \end{cases} \quad (3.10)$$

$$\frac{E_{d(q),h}(6q)}{E_{d(q),h}(6)} = \begin{cases} (q^2 + 2q - 7)/(q - 3)^2, & h(q) = -1, \\ 1, & h(q) = 1, \end{cases} \quad q \geq 5. \quad (3.11)$$

If $f, g \in \mathcal{S}_{1/2}$ differ at a prime p , then

$$\max_{d \in \mathcal{D}} \sup_{a \geq 0} |C_{d,f}(a) - C_{d,g}(a)| \geq \begin{cases} 1, & p = 2, \\ 8(p-1)/(p+1)^2, & p \text{ odd}. \end{cases} \quad (3.12)$$

Proof. The only possible zero factors at depth zero are at two and, on a visible trace, at three. The modulus 6 gives positive depth at both, so every factor is positive. The sum of the remaining deficits is finite, and therefore the product is positive. In the ratios, all unchanged prime factors cancel. At three the relevant positive-depth factors are $1/2$ and $5/6$; at $q \geq 5$ they are $1 - 8/(q+1)^2$ and $(q-3)^2/(q+1)^2$. This gives the formulas.

For (3.12), use (3.4) at p for both sources. The difference of the isolated energies is 1 at two, and

$$1 - \left(\frac{p-3}{p+1} \right)^2 = \frac{8(p-1)}{(p+1)^2}$$

at a visible odd prime. Each energy difference is an average of differences of the original correlations, so it cannot exceed their supremum. $\square$

These are inverses from exact limiting data. Neither the ratios nor the factorial isolation gives a common finite largest shift, a common sample size, or a uniform precision requirement on the entire class.

3.4 Why three readings still leave a genuine ambiguity

For the pair (1.1), every odd moment of scalar readings is zero by (2.7). Every pair moment is independent of ϵ by the layer lemma. Fourier inversion on a sign cube,

$$\mathbb{P}(Z = z) = 2^{-r} \sum_{J \subseteq \{1, \dots, r\}} \left(\mathbb{E} \prod_{j \in J} Z_j \right) \prod_{j \in J} z_j, \quad (3.13)$$

therefore proves equality of all joint laws with $r \leq 3$. On each trace its pointwise quotient is $(-1)^{v_{17}(Q_d(n))}$, so (2.5) gives disagreement density $1/9$.

The same construction with a finite set Λ of primes $1 \pmod{8}$ gives $2^{|\Lambda|}$ sources with identical laws through order three. If two sign assignments differ exactly on $J \subseteq \Lambda$, the disagreement density on every trace is

$$\frac{1}{2} \left(1 - \prod_{q \in J} \left(1 - \frac{4}{q+1} \right) \right). \quad (3.14)$$

This follows by the Chinese remainder theorem and then geometric valuation tails. The density concerns parity of the total changed valuation; a sum of individual incidences would not give this exact answer.

There is also a uniform finite-origin version of the three-reading example.

Proposition 3.3 (Finite comparison for the pair at seventeen). *For any at most three scalar readings, with arbitrary nonnegative integer shifts, and $N \geq 17^2$, their empirical laws under h_+ and h_- satisfy*

$$d_{\text{TV}}(\mathcal{L}_{N,h_+}, \mathcal{L}_{N,h_-}) \leq \min\{1, 770N^{-1/2}\}.$$

The shifts may depend on N ; the sampled polynomial arguments are not required to lie in a fixed integer prefix.

Proof. Take $17^K \leq \sqrt{N} < 17^{K+1}$, and replace a reading by zero whenever its valuation at 17 is at least K . The truncated moments are periodic modulo $2 \cdot 17^K$, and their means agree for the two signs by the finite layer argument and parity cancellation. For a product of $r \leq 3$ readings, the two truncation errors and the two period-boundary errors give a discrepancy at most

$$\frac{4r}{17^K} + \frac{4r}{N} + \frac{4 \cdot 17^K}{N} \leq \frac{208}{\sqrt{N}} + \frac{12}{N}.$$

Apply (3.13). There are at most seven nonempty subsets and total variation is one half the ℓ^1 distance, so the result is at most $728N^{-1/2} + 42N^{-1} \leq 770N^{-1/2}$. $\square$

4 A fourth-order detector for every ramified sign

The second-order fibre has now been identified exactly. We do not need a general reconstruction theorem on a large finite group to resolve it. We need one scalar test at each conductor prime, and a way of removing all other primes from that test.

4.1 Three readings at one root and one away from it

Fix an odd prime $q \geq 7$ and a trace d visible at q . Let $r, -r$ be the roots of $x^2 = d$ modulo q . Choose

$$a \not\equiv 0, 2r, -2r \pmod{q}, \quad u \not\equiv 0, 1 \pmod{q},$$

with representatives $1 \leq a < q$ and $2 \leq u < q$. We use the shifts $(0, a, q, uq)$. The three readings at $0, q, uq$ can enter the first valuation layer together near a root, producing $\epsilon^3 = \epsilon$. The reading at a then stays away from a root and weights the two root contributions; we will choose a to prevent their cancellation. To calculate this signal, define the finite integer sums

$$B_q(a) = \sum_{x \bmod q} \chi_q((x^2 - d)((x + a)^2 - d)), \quad (4.1)$$

$$S_q(u) = \sum_{t \bmod q} \chi_q(t(t+1)(t+u)), \quad (4.2)$$

$$W_q(a) = \sum_{r^2 = d \bmod q} \chi_q(2r) \chi_q((r+a)^2 - d). \quad (4.3)$$

Lemma 4.1 (The local fourth-order formula). *For the local completed character,*

$$K_{q,\epsilon} := \int_{\mathbb{Z}_q} \prod_{b \in \{0, a, q, uq\}} G_{q,\epsilon}(Q_d(x+b)) dx = \frac{B_q(a)}{q} + \epsilon \frac{S_q(u)W_q(a)}{q^2}. \quad (4.4)$$

For this configuration the integral depends only on the shifts modulo q^2 .

Proof. Separate the residues of x modulo q . Outside the roots of $Q_d(x)$ and $Q_d(x+a)$, all four valuations vanish. The three factors at shifts $0, q, uq$ have the same character, so their product reduces to the character at shift zero. These residues contribute $B_q(a)/q$, since the omitted residues give zero in that character sum.

At a root of $Q_d(x+a)$, none of the other three polynomials vanishes modulo q . They are fixed locally, and the mean of the remaining completed reading is zero. Its entire class contributes zero.

At a root r of $Q_d(x)$, let $\alpha \in \mathbb{Z}_q$ be its unique root lift and write $x = \alpha + qt$. Then

$$Q_d(x+vq)/q \equiv 2r(t+v) \pmod{q}, \quad v = 0, 1, u.$$

For $t \not\equiv 0, -1, -u \pmod{q}$, all three valuations equal one, producing the coefficient $\epsilon^3 = \epsilon$. Their unit characters multiply to $\chi_q(2r)\chi_q(t(t+1)(t+u))$. The fourth reading has character $\chi_q((r+a)^2 - d)$.

At each excluded value of t , exactly one of the three readings is deeper. All the others are determined on that class. The deeper reading has conditional mean zero, so no term is lost by assigning zero to it. Multiplying the class measures q^{-1} and q^{-1} and summing over both roots yields the second term of (4.4). This argument also shows that perturbing the shifts by multiples of q^2 does not change the integral: the first two residue levels and the unique-deep-reading cancellation are unchanged. $\square$

It remains to choose a, u so that $S_q(u)W_q(a) \neq 0$.

4.2 Nonzero detectors without an exceptional odd prime

Put $s = \chi_q(-1)$. A direct rearrangement gives

$$W_q(a) = \chi_q(2r)\chi_q(a)(\chi_q(a+2r) + s\chi_q(a-2r)). \quad (4.5)$$

For admissible a , this is nonzero exactly when $\chi_q(a^2 - 4d) = s$, and then $|W_q(a)| = 2$. Since the quadratic sum is -1 , removing $a = 0$ removes s , and the two roots contribute zero. Hence

$$\#\{a \not\equiv 0, \pm 2r : W_q(a) \neq 0\} = \frac{q-4-s}{2} > 0. \quad (4.6)$$

For S_q , put $b_t = \chi_q(t(t+1))$. The identity

$$\sum_{u \bmod q} \chi_q(t+u)\chi_q(v+u) = \begin{cases} q-1, & t = v, \\ -1, & t \neq v \end{cases}$$

and $\sum b_t = -1$, $\sum b_t^2 = q-2$ give

$$\sum_{u \bmod q} S_q(u)^2 = q(q-2) - 1.$$

Since $S_q(0) = -1$ and $S_q(1) = -s$, we obtain

$$\sum_{u \neq 0,1} S_q(u)^2 = (q-3)(q+1) > 0. \quad (4.7)$$

We may choose the smallest admissible a with $W_q(a) \neq 0$, and then the smallest $u \in \{2, \dots, q-1\}$ with $S_q(u) \neq 0$. For these values $S_q(u)$ is a nonzero even integer: its sum has $q-3$ nonzero terms, an even number. Consequently

$$|K_{q,+} - K_{q,-}| = \frac{2|S_q(u)W_q(a)|}{q^2} \geq \frac{8}{q^2}. \quad (4.8)$$

The two smaller odd primes require different configurations, not an exception in the theorem. At $q = 3$ take $d = -2$ and shifts $(0, 1, 2, 3)$; at $q = 5$ take $d = -1$ and shifts $(0, 1, 2, 5)$. Their conditional integrals are

q	Residues modulo q	Conditional fourth-order integrals
3	0, 1, 2	$-5/9, 4\epsilon/9, 4\epsilon/9$
5	0, 1, 2, 3, 4	$0, -1/5, 2\epsilon/5, 1/5, -1$

Thus

$$K_{3,\epsilon} = \frac{-5 + 8\epsilon}{27}, \quad K_{5,\epsilon} = \frac{-5 + 2\epsilon}{25}. \quad (4.9)$$

These small calculations can be performed without truncation error. Average over $x \bmod 27$ for $q = 3$, or $x \bmod 25$ for $q = 5$. In every class at most one reading remains unresolved; integrate it as zero by Lemma 2.2. For example, in the class $0 \bmod 3$, two subclasses give -1 , while the third contains a deeper pair of mean $1/3$, so the conditional average is $(-2 + 1/3)/3 = -5/9$. This gives an exact finite recipe for every entry of the table. The resulting gaps are $16/27$ and $4/25$.

For the introductory example, $q = 17$, $d = -1$, $a = 2$ and $u = 2$ give

$$B_{17}(2) = -3, \quad S_{17}(2) = -2, \quad W_{17}(2) = -2.$$

The local integral is $(-51 + 4\epsilon)/289$. The four dyadic factors multiply to -1 , because the sum of the shifts is 53. This proves (1.2) with its stated sign.

4.3 Removing every prime except the one being tested

For a conductor prime q , let $(0, a_1, a_2, a_3)$ be its chosen detector above. Take $M_{j,q}$ from (3.4) and choose $u_{j,q} \in \{1, \dots, q^3 - 1\}$ such that

$$L_{j,q} = M_{j,q}u_{j,q} \equiv 1 \pmod{q^3}. \quad (4.10)$$

At every prime $r \neq q$, $v_r(L_{j,q}) \rightarrow \infty$. For large j , $L_{j,q}$ is even. The q -adic detector is unchanged by this dilation: Lemma 4.1 uses only shifts modulo q^2 , and the two small-prime calculations use at most q^3 .

Proposition 4.2 (Isolation of the fourth-order signal). *For $h \in \mathcal{C}$ and $q \in \Lambda$,*

$$\lim_{j \rightarrow \infty} D_{d(q),h}(L_{j,q}a_1, L_{j,q}a_2, L_{j,q}a_3) = K_{q,\epsilon_q}. \quad (4.11)$$

At $q = 3, 5$, use the traces specified in (4.9). More quantitatively, if L is even, $L \equiv 1 \pmod{q^3}$, and every other conductor prime divides L , then the error relative to K_{q, ϵ_q} is at most

$$4 \sum_{\substack{r \in (A \cup \Lambda) \setminus \{q\} \\ r|L}} r^{-v_r(L)} + 16 \sum_{\substack{r \in A \\ r \nmid L}} \frac{1}{r+1}. \quad (4.12)$$

Proof. At another prime r with $r^e \mid L$, all four arguments agree modulo r^e . Outside the at most two root classes at that depth, their local signs are equal. The product of four equal signs is one. The local deficit is consequently at most $4r^{-e}$. At an omitted correction prime use (2.6) with four readings, giving $16/(r+1)$. At two, even dilation gives a product one. Multiplying the local factors and using $|\prod z_i - 1| \leq \sum |z_i - 1|$ for factors of modulus at most one proves (4.12).

For (4.11), first truncate the correction set. Every non-target factor in this finite product tends to one. The tail of (4.12) is uniformly summable, so the truncation can then be removed. $\square$

The orders of limits matter. The empirical limit defining each D is taken first. Only then do we pass to the sequence of dilated shifts. This is an inverse of exact limiting data, not a claim that one may use enormous moving shifts in the same empirical estimate without an additional error analysis.

4.4 Completing the reconstruction

After Section 3 recovers σ, A, Λ , Proposition 4.2 gives a number K at each conductor prime. The missing local sign is

$$\epsilon_q = \begin{cases} (27K + 5)/8, & q = 3, \\ (25K + 5)/2, & q = 5, \\ (q^2K - qB_q(a))/(S_q(u)W_q(a)), & q \geq 7. \end{cases} \quad (4.13)$$

The denominators are nonzero by (4.6)–(4.7). Recompose the actual prime signs as

$$\begin{aligned} h(2) &= \sigma \prod_{\ell \in \Lambda} \chi_\ell(2), \\ h(q) &= \epsilon_q \prod_{\ell \in \Lambda \setminus \{q\}} \chi_\ell(q) & (q \in \Lambda), \\ h(q) &= (-1)^{1_{q \in A}} \prod_{\ell \in \Lambda} \chi_\ell(q) & (q \text{ odd, } q \notin \Lambda). \end{aligned} \quad (4.14)$$

Multiplicativity now determines every value of h . This proves the reconstruction assertion of Theorem 1.1; Section 3 already proved the exact second-order fibre and the failure through order three.

The construction also gives a quantitative separation of ramified local signs. If $f, g \in \mathcal{C}$ both have q in their conductor-prime sets and opposite ϵ_q , then

$$\sup_{a, b, c \geq 0} |D_{d(q), f}(a, b, c) - D_{d(q), g}(a, b, c)| \geq \begin{cases} 16/27, & q = 3, \\ 4/25, & q = 5, \\ 8/q^2, & q \geq 7. \end{cases}$$

The other parameters may differ: the common dilation removes them from both sources. This statement concerns opposite *local* signs; opposite values of $f(q)$ and $g(q)$ are not equivalent without accounting for the other conductor characters.

4.5 The quadratic components at two

The factor σ^{v_2} is not the most general completed quadratic component at two. If $o = m/2^{v_2(m)}$ is odd, put

$$\theta_{\sigma,u,v}(m) = \sigma^{v_2(m)}(-1)^{u(o-1)/2+v(o^2-1)/8}, \quad u, v \in \{0, 1\}. \quad (4.15)$$

On our three traces this is periodic modulo eight. To check that period, note that the valuations are only zero or one, and that changing n by eight changes n^2 by a multiple of sixteen. Thus the odd part needed modulo eight is preserved.

Corollary 4.3 (No additional quadratic ambiguity at two). *Replace σ^{v_2} by (4.15) in the definition of $\mathcal{C}$. The pair correlations recover (σ, u, v) , A and Λ , with the same free ramified odd signs as before. The fourth-order correlations determine the source in this enlarged class.*

Proof. The three autocorrelations at shift one of the dyadic component alone have the following signatures, obtained by averaging its eight residue classes:

σ	u	v	Signature in trace order $(-1, -2, 2)$
+1	0	0	(1, 1, 1)
+1	0	1	(0, 0, 1)
+1	1	0	(1, 0, 0)
+1	1	1	(0, 1, 0)
-1	0	0	(-1, -1, -1)
-1	0	1	(0, 0, -1)
-1	1	0	(-1, 0, 0)
-1	1	1	(0, -1, 0)

They are all distinct. Let $J_j = j!/2^{v_2(j!)}$, and choose a multiple L_j of J_j congruent to one modulo eight. At each fixed odd prime its valuation tends to infinity, so the odd-prime pair factor tends to one outside root classes of vanishing measure. The correction tail is uniformly bounded by $8 \sum_{q \in A, q > Y} (q+1)^{-1}$. Hence $\lim_j C_{d,h}(L_j)$ gives precisely the displayed signature.

In the odd-prime type isolation, the power of two in $M_{j,q}$ eventually exceeds 2^3 , so its conditional energy equals one. The odd signatures (3.7) remain unchanged. For fourth-order isolation require the dilations to be divisible by eight; then the four dyadic readings coincide and multiply to one. The preceding reconstruction recovers every odd component. Finally multiply the recovered dyadic character into (3.8)–(4.14), as appropriate. This gives all actual prime signs. $\square$

5 A whole window can hide the ramified signs

The order-four inverse uses all required shifts. It does not assert a conductor-independent observation range. To show why such a range is impossible, we now protect every reading in a prescribed window at once, rather than select only three readings.

5.1 At most two roots in each residue class

The layer calculation has a useful conditional form. Suppose a finite family of polynomials has simple roots modulo an odd prime q , and at every residue at most two members vanish. Conditional on that residue, all other completed readings are fixed. The remaining one or two readings have a joint law independent of the common completion sign by Lemma 2.2. It follows that the *entire* finite-dimensional local law is independent of that sign. We have not

declared the root readings independent; their first and second moments determine their joint law.

On a single trace, $L \leq q$ consecutive shifts give at most two roots in any residue, since each of the two roots determines at most one shift. For the three traces together, roots from different quadratics must also be controlled.

Lemma 5.1 (No triple root collision in a short three-trace window). *If $q \equiv 1 \pmod{8}$, $q > 17$, and $H \leq \lfloor q^{1/3} \rfloor$, then at every $x \bmod q$ at most two of the polynomials $Q_d(x+a)$, $0 \leq a < H$, $d \in \mathcal{D}$, vanish.*

Proof. At a single shift the three polynomials cannot share a root modulo such a q , since their pairwise differences belong to $\{\pm 1, \pm 3, \pm 4\}$. A triple collision therefore has distinct shifts $j, j+a, j+a+b$, with $a, b \geq 1$ and $s = a+b < H$.

Write their three right-hand sides as $d_1, d_2, d_3 \in \mathcal{D}$. Subtract the squared equations and eliminate the common variable. One obtains

$$q \mid F := ab(a+b) + (a+b)d_2 - bd_1 - ad_3. \quad (5.1)$$

Its absolute value is at most $s^3/4 + 4s < q$, so $F = 0$. The linear part of F has magnitude at most $4s$, hence $ab \leq 4$. Eliminating the variable from the first two equations also gives

$$q \mid R := (d_2 - d_1 - a^2)^2 - 4a^2d_1. \quad (5.2)$$

There are only eight ordered positive pairs with $ab \leq 4$. Substitution of $d_1, d_2, d_3 \in \{-1, -2, 2\}$ into $F = 0$ gives the following full list:

a	b	d_1	d_2	d_3	R
1	1	-1	-2	-1	8
1	1	-2	-1	2	8
1	1	2	-1	-2	8
1	2	-1	-2	2	8
1	2	2	-1	-1	8
1	3	2	-1	2	8
1	3	2	-2	-2	17
1	4	2	-2	2	17
2	1	-1	-1	2	32
2	1	2	-2	-1	32
2	2	2	-2	2	32
3	1	-2	-2	2	153
3	1	2	-1	2	72
4	1	2	-2	2	272

Every listed R is nonzero and has prime factors only in $\{2, 3, 17\}$, contradicting (5.2). All enumerations in this proof are finite substitutions into the displayed identities; no asymptotic distribution of roots is assumed. $\square$

The exponent $1/3$ is a sufficient range furnished by this collision argument. We do not claim that it is the optimal horizon at which a completion sign becomes visible.

5.2 Preserving a prefix while making the disagreement large

Choose a finite set Λ of primes

$$q \equiv 1 \pmod{8}, \quad q > \max\{B, H^3, 17\}.$$

For each assignment of signs ϵ_q , define

$$h_\epsilon(m) = (-1)^{v_2(m)} \prod_{q \in \Lambda} G_{q, \epsilon_q}(m) \prod_{r \in A} (-1)^{v_r(m)}, \quad (5.3)$$

where A is a suitable subset of the odd primes at most B . All primes in Λ exceed B , so their completion signs do not affect that prefix. Choose A to make the actual prime signs up to B agree with h_* . At two, $\chi_q(2) = 1$ for $q \equiv 1 \pmod{8}$, so the prescribed sign $h_*(2) = -1$ already agrees. Complete multiplicativity then preserves every integer up to B .

Lemma 5.1, the conditional-law observation preceding it, and the Chinese remainder theorem show that the full H -word law is the same for all ϵ . Comparing all positive with all negative completion signs gives disagreement density (3.14), with $J = \Lambda$, on every trace. The divergence

$$\sum_{\substack{q \text{ prime} \\ q \equiv 1 \pmod{8}}} \frac{1}{q} = \infty$$

is a special case of the prime-progression formula in [14, §27.11, equation 27.11.9]. Hence a finite Λ can make that density larger than $1/2 - \eta$. All first moments vanish by the dyadic factor. This proves the main assertions of Theorem 1.2. Its near-uniform and entropy assertions follow from the next two subsections.

5.3 The scale at which a finite conductor looks Bernoulli

Consider a single-conductor source

$$h_{q, \sigma, \epsilon}(m) = \sigma^{v_2(m)} G_{q, \epsilon}(m), \quad q \equiv 1 \pmod{8}.$$

For a word of length L put $m = 3L$, and denote its limiting law by ν_q^L ; the following bounds are uniform in σ, ϵ . Two distinct quadratic factors at shift difference a have resultant

$$a^4 - 2(d + e)a^2 + (d - e)^2. \quad (5.4)$$

It is nonzero because the two distinct monic quadratics are irreducible over $\mathbb{Q}$. For shifts in a window L , its absolute value is at most $(L^2 + 4)^2$. Therefore, if

$$q > (L^2 + 4)^2, \quad (5.5)$$

the $2m$ roots of all these factors are distinct modulo q .

In this regime the word law has an elementary finite description. Choose parity and a residue $x \bmod q$ independently and uniformly. If none of the factors vanishes, the word is determined by its character values and the parity. If exactly one factor vanishes, replace that one character reading by a fair sign, multiplied by its prescribed dyadic sign. Lemma 2.2 shows that this is the exact integrated local law, not a simulation of a new random multiplicative function. All masses are multiples of $1/(4q)$, and the support has at most $2q + 4m$ words.

Proposition 5.2 (A conductor-dependent Bernoulli window). *Under (5.5),*

$$\max \left\{ 0, 1 - \frac{2q + 4m}{2^m} \right\} \leq d_{\text{TV}}(\nu_q^L, U_m) \leq \min \left\{ 1, \frac{1}{2} \sqrt{\frac{(m^2 - m + 1)2^m - 1}{q}} \right\}. \quad (5.6)$$

Consequently, for each fixed $0 < \delta < 1$, as $q \rightarrow \infty$ through primes $1 \pmod{8}$,

$$\begin{aligned} L \leq \frac{1-\delta}{3} \log_2 q &\implies d_{\text{TV}}(\nu_q^L, U_{3L}) = O_\delta((\log q)q^{-\delta/2}) \rightarrow 0, \\ L \geq \frac{1+\delta}{3} \log_2 q &\implies d_{\text{TV}}(\nu_q^L, U_{3L}) \geq 1 - O_\delta(q^{-\delta}) \rightarrow 1. \end{aligned}$$

In these statements the empirical limit defining ν_q^L is taken before the conductor tends to infinity.

Proof. The support bound gives the lower inequality by testing the complement of the support against U_m . For the upper bound, the Walsh coefficient at a nonempty set J of k coordinates has the form

$$\hat{\nu}_q^L(J) = \frac{c_J}{q} \sum_{x \bmod q} \chi_q \left(\prod_{i \in J} P_i(x) \right), \quad |c_J| \leq 1.$$

Here c_J is the dyadic parity average. Under (5.5), the product has $2k$ distinct linear roots over $\mathbb{F}_q$. The classical Weil character-sum bound, in the form recalled in [5, Lemma 1], bounds its sum by $(2k - 1)\sqrt{q}$ in absolute value. Parseval and Cauchy-Schwarz on the finite cube therefore give

$$d_{\text{TV}}(\nu_q^L, U_m)^2 \leq \frac{1}{4} \sum_{J \neq \emptyset} |\hat{\nu}_q^L(J)|^2 \leq \frac{1}{4q} \sum_{k=1}^m \binom{m}{k} (2k - 1)^2.$$

The binomial sum is $(m^2 - m + 1)2^m - 1$, proving (5.6).

For the first limiting assertion, $2^{3L} \leq q^{1-\delta}$ and $m = O(\log q)$; condition (5.5) holds eventually. For the second, apply the lower bound at $L_0 = \lceil (1 + \delta) \log_2 q / 3 \rceil$. It also satisfies (5.5) eventually. For every larger L , projection to its first L_0 positions contracts total variation and proves the claimed lower bound. No assertion about the critical window itself is needed. $\square$

There are thus two different scales. A local completion sign can remain invisible through a three-trace horizon of order $q^{1/3}$, although the word law is already far from Bernoulli at horizons just above $(\log_2 q)/3$. Detecting arithmetic structure and identifying its ramified sign are different questions.

A finite-origin bound also follows from the residue description:

$$d_{\text{TV}}(\mu_{N, h_q, \sigma, \epsilon}^L, \nu_q^L) \leq \min \left\{ 1, \frac{2m}{q} + \frac{2q + 2m}{N} \right\} \quad (5.7)$$

under (5.5). The union of root classes has empirical mass at most $2m/q + 2m/N$; away from it, the word is determined modulo $2q$, with period-boundary cost at most $2q/N$. The bound is useful when the conductor and sample size vary. It does not by itself prove convergence at fixed q , since $2m/q$ remains. That convergence was established separately by deeper valuation truncation.

5.4 Finite-conductor entropy and the completion of Theorem 1.2

Let $p_h(L)$ count the distinct three-trace words of length L occurring at integer origins. For the one-conductor source, take the least $K \geq 1$ with $q^K \geq L$. Within a word, each of the

six root progressions modulo q^K contains at most one unresolved coordinate. The residue modulo $2q^K$ determines all other readings. Thus

$$p_h(L) \leq 2q^K 2^6 \leq 128qL. \quad (5.8)$$

This estimate does not require (5.5).

More generally, if a source of the form (1.7) uses a finite set R of s odd conductor or correction primes, choose a separate depth K_q with $q^{K_q} \geq L$. There are at most six unresolved readings per prime. Therefore

$$p_h(L) \leq 2^{6s+1} \left(\prod_{q \in R} q \right) L^s. \quad (5.9)$$

Its topological entropy $\lim_{L \rightarrow \infty} L^{-1} \log p_h(L)$ is zero. Since a limiting word law is supported on these occurring words, its Shannon entropy is at most $\log_2 p_h(L)$, and its entropy per position is also zero.

To make the common fixed-window law in Theorem 1.2 nearly uniform, add a common completed character at a further prime $q_0 \equiv 1 \pmod{8}$, larger than all previous primes and satisfying both the root-separation and the short-collision conditions. Choose q_0 so large that its local word law is within ζ of U_{3H} by (5.6). Its local law without the dyadic parity has the same upper bound, since the proof used only $|c_J| \leq 1$. The full word law is its convolution, on the sign cube, with the product of the other local laws. Convolution contracts total variation and fixes the uniform law. Rechoose the finite correction set at primes at most B to preserve the prefix; this only changes the second convolution factor. Use the same completion sign at q_0 in both sources, so their quotient and disagreement density are unchanged. The enlarged conductor is still finite, and (5.9) still gives zero entropy. This completes Theorem 1.2.

6 An arbitrary source as a product of normal factors

The preceding constructions hide finitely many conductor signs, but a sufficiently rich correlation family still identifies their sources. We now leave that class. The input is an arbitrary $h \in \mathcal{M}_\pm$, and the aim is to retain it exactly as a product fg while making the separate word laws of f and g uniform. The coupling in (1.8) describes the intended joint law. The arithmetic task is to realize that law using two completely multiplicative functions, rather than unrelated signs at the observation sites. Section 8.3 will return to the quotient $(-1)^{v_{17}}$, now between two sources normal at every horizon rather than the structured pair (1.1).

Najnudel [13, Proposition 6.3 in the cited preprint] proves almost-sure uniform word frequencies, with error $O(N^{-1/2+\varepsilon})$, for consecutive linear observations of a random completely multiplicative function whose prime values are uniform roots of unity. In our sign-valued quadratic setting, the products P_J below and their square-class multiplicities replace that linear arithmetic input. The rate alone is not the additional assertion: the weights record the words of an arbitrary prescribed h , and the target is its canonical coupling.

We first give the probabilistic construction, including its progression-wise and prescribed-horizon variants. Section 7 makes the choice continuous and relatively computable, with a bound uniform in the input. The probabilistic argument also supplies the simultaneous summable-coset result in Section 8.

6.1 The arithmetic quantity that controls the second moment

For a positive integer a , let $\text{core}(a)$ be the product of the primes having odd valuation in a . For a finite set S of primes, define

$$\text{core}_{S^c}(a) = \prod_{q \notin S: v_q(a) \text{ odd}} q.$$

This is a square class, not the radical: even valuations disappear. For a nonempty subset J of the labels in a word of length H , put

$$P_J(X) = \prod_{(a,d) \in J} ((X+a)^2 - d), \quad \emptyset \neq J \subseteq \{0, \dots, H-1\} \times \mathcal{D}. \quad (6.1)$$

Each factor is irreducible over $\mathbb{Q}$. Two factors are equal only when their labels are equal. Thus P_J is separable, and its degree is $2|J| \geq 2$.

The largest multiplicity of an exterior square class is

$$M_{H,S}(T) = \max_{J \neq \emptyset} \max_{b \geq 1} \#\{2 \leq n \leq T+1 : \text{core}_{S^c}(P_J(n)) = b\}. \quad (6.2)$$

For integers $H, T \geq 1$, this is an integer between 1 and T , computed by finitely many factorizations and comparisons. It does not depend on h or on the prescribed signs on S . For a single polynomial, write $M_{P,S}(T)$ for the corresponding maximum over b .

This multiplicity measures the collisions relevant to the second moment. After the signs on S are prescribed, origins in the same exterior square class use the same character of the free prime signs, up to a prescribed sign. Characters from distinct classes are orthogonal, though not necessarily jointly independent. For deterministic weights of modulus at most one, grouping by class costs at most $TM_{P,S}(T)$ in the second moment (6.8); dyadic blocks then control all prefixes through (6.9).

Lemma 6.1 (Square-class multiplicities). *For fixed H and S and every $\eta > 0$,*

$$M_{H,S}(T) \ll_{H,S,\eta} T^\eta. \quad (6.3)$$

Moreover, if $H = H(T) \geq 1$ and $H = o(\log \log T)$, then

$$M_{H,S}(T) = T^{o(1)}. \quad (6.4)$$

The latter assertion is uniform over the label subsets in that window, not over arbitrary polynomials of growing degree.

Proof. The square-class estimates used here are those of de la Bretèche–Wang–Xu [4, Lemmas 3.2–3.3]. Their proof of Lemma 3.2 traces the quadratic estimate to Vaughan–Wooley, *Further improvements in Waring’s problem*, Lemma 3.5, building on Hua’s *Introduction to number theory*, Chapter 11; we use the formulation in [4]. For a fixed separable quadratic P , their degree-two estimate gives

$$\#\{1 \leq n \leq T : P(n) = au^2 \text{ for some integer } u\} \ll_{P,\eta} T^\eta,$$

uniformly in the positive square class a . In degree at least three, the corresponding bound is $2^{O_P(1+\omega(a))}$, where $\omega(a)$ is the number of distinct prime divisors of a .

An exterior class b in (6.2) splits into at most $2^{|S|}$ complete square classes: $\text{core}(P_J(n)) = ba$, where a ranges over squarefree products of primes of S . Every complete class actually

encountered is at most $\max_{n \leq T+1} P_J(n) \ll_{P_J} T^{\deg P_J}$. The elementary estimate $B^{\omega(t)} \ll_{B,\delta} t^\delta$, with δ chosen after P_J and η , proves (6.3).

For (6.4), the dependence on P_J must be retained. All roots of every P_J lie in the same field

$$K_0 = \mathbb{Q}(i, \sqrt{2}), \quad [K_0 : \mathbb{Q}] = 4.$$

We use Evertse–Silverman [6, Theorem 1(b)] in the form also applied in the proof of [4, Lemma 3.3]. Here is the specialization, so that the uniformity in the number of factors is explicit. For a positive squarefree a , the equation $P_J(x) = au^2$ maps into $z^2 = aP_J(x)$ by $z = au$. Work over $\mathbb{Q}$ with the set of places consisting of infinity and the prime divisors of $a \operatorname{disc} P_J$. The polynomial aP_J has unit discriminant outside these places, and K_0 contains at least three of its distinct roots when $|J| \geq 2$. If s is the number of these places, the theorem bounds the number of possible x by

$$7^{4(4+9s)} 4^{\kappa_2(K_0)},$$

where $\kappa_2(K_0)$ is the fixed two-rank of its ideal class group. Since $s \leq 1 + \omega(a \operatorname{disc} P_J)$, this gives

$$\#\{x \in \mathbb{Z} : P_J(x) = au^2 \text{ for some } u \geq 1\} \leq A_0(7^{36})^{\omega(a \operatorname{disc} P_J)}. \quad (6.5)$$

The constant A_0 depends only on the fixed field. In particular, neither the degree nor the shifts introduce a new field or an unspecified degree-dependent base in this bound.

Write $r = |J|$. The discriminants of the quadratic factors have absolute value at most 8, and the pairwise resultants are (5.4). The product-discriminant identity therefore gives

$$0 < |\operatorname{disc} P_J| \leq 8^r (H^2 + 4)^{2r(r-1)}, \quad \log |\operatorname{disc} P_J| \ll r^2 \log(2H). \quad (6.6)$$

For $H = o(\log \log T)$ and large T , each realized class a is at most $(5T^2)^r$. The elementary bound

$$\omega(t) \ll \frac{\log(2t)}{\log \log(3t)}$$

follows, for example, by comparing t with $\omega(t)!$. Together with (6.5)–(6.6) and the $2^{|S|}$ possible protected parts, it implies, for $2 \leq r \leq 3H$,

$$\log M_{P_J,S}(T) \ll_S 1 + H^2 \log(2H) + \frac{H \log T}{\log \log T} = o(\log T). \quad (6.7)$$

For $r = 1$, translate n to $n + a$ and apply the degree-two estimate to one of the three fixed polynomials Q_d up to $T + H + 1$. Its constant no longer depends on the shift. Since $T + H + 1 \ll T$, for every fixed $\eta > 0$ the same bound $O_{S,\eta}(T^\eta)$ holds uniformly for these single-label subsets. This suffices for (6.4). The logarithmic Pell estimate recorded after [4, Lemma 3.2] also gives a logarithmic bound in this single-factor case. $\square$

The two roles of this lemma will remain separate. It supplies an asymptotic estimate for a second moment, but the selection algorithm below uses the exact integer (6.2), not an unevaluated constant from the lemma.

6.2 Orthogonality with deterministic weights

Prescribe signs ξ_q of f on S , and choose all other prime signs independently and fairly. Extend them completely multiplicatively; in particular, f is not set to zero on integers with a square factor. If

$$b_n = \operatorname{core}_{S^c}(P(n)), \quad s_n = \prod_{q \in S} \xi_q^{v_q(P(n))},$$

then $f(P(n)) = s_n X_{b_n}$, where X_b is the product of the free signs at the prime divisors of b . These characters satisfy $\mathbb{E} X_b X_c = \mathbf{1}_{b=c}$. They need not be jointly independent.

For deterministic complex weights $|w_n| \leq 1$ and any $I \subseteq \{2, \dots, T+1\}$, orthogonality and Cauchy–Schwarz give

$$\begin{aligned} \mathbb{E} \left| \sum_{n \in I} w_n f(P(n)) \right|^2 &= \sum_b \left| \sum_{\substack{n \in I \\ b_n = b}} w_n s_n \right|^2 \\ &\leq M_{P,S}(T) \sum_{n \in I} |w_n|^2 \leq M_{P,S}(T) |I|. \end{aligned} \quad (6.8)$$

The class $b = 1$ is included. No centering by an expectation is being suppressed.

When $T = 2^k$, every prefix is a union of at most $k+1$ blocks from its dyadic partitions. Taking Cauchy–Schwarz over those blocks, then enlarging to all blocks at all levels, yields

$$\mathbb{E} \max_{1 \leq t \leq T} \left| \sum_{n=2}^{t+1} w_n f(P(n)) \right|^2 \leq (k+1)^2 T M_{P,S}(T). \quad (6.9)$$

At each dyadic level the blocks partition the T indices, so (6.8) pays only $T M_{P,S}(T)$ at that level.

For each fixed P, S, w and $\varepsilon > 0$, take $\eta < 2\varepsilon$ in the square-class estimate. Markov’s inequality applied to (6.9) at threshold $T^{1/2+\varepsilon}$ has summable probabilities over $T = 2^k$. Borel–Cantelli, with the maximum already present, gives

$$\sum_{n=2}^{N+1} w_n f(P(n)) = O_{f,P,S,w,\varepsilon}(N^{1/2+\varepsilon}) \quad \text{almost surely.} \quad (6.10)$$

A single event of probability one serves a countable family of fixed polynomials and weights, and all positive rational ε . The weights must be fixed before the prime signs of f are chosen. Choosing $w_n = f(P(n))$ afterwards would instead make the sum equal to N .

6.3 Turning weighted cancellation into a coupling

Fix the input h , an integer H , and $m = 3H$. Put $X(n) = W_f^H(n)$ and $V(n) = W_h^H(n)$. For $J \neq \emptyset$ and $v \in \{\pm 1\}^m$, define

$$A_{J,v}(N) = \sum_{n=2}^{N+1} \mathbf{1}_{V(n)=v} f(P_J(n)). \quad (6.11)$$

Its weight depends only on h , so (6.10) applies. Fourier inversion on the sign cube gives the exact identity

$$\mathcal{L}_N(X, V)(x, v) - 2^{-m} \mu_{N,h}^H(v) = \frac{2^{-m}}{N} \sum_{J \neq \emptyset} x_J A_{J,v}(N), \quad x_J = \prod_{j \in J} x_j. \quad (6.12)$$

Consequently,

$$d_{\text{TV}}(\mathcal{L}_N(X, V), U_m \otimes \mu_{N,h}^H) \leq \frac{1}{2N} \sum_{J \neq \emptyset, v} |A_{J,v}(N)|. \quad (6.13)$$

There are finitely many tests at each H , and only countably many over all H . Set $g = fh$. The bijection $(x, v) \mapsto (x, x \odot v)$ preserves total variation and sends (X, V) to (W_f^H, W_g^H) . It also sends the product target to $\mathcal{J}_m(\mu_{N,h}^H)$.

Theorem 6.2 (Almost-sure factorization and fixed progressions). *For each prescribed $h \in \mathcal{M}_\pm$ and finite prescription on S , almost every choice of the other prime signs of f satisfies, with $g = fh$,*

$$d_{\text{TV}}(\mathcal{L}_N(W_f^H, W_g^H), \mathcal{J}_{3H}(\mu_{N,h}^H)) = O_{f,h,S,H,\varepsilon}(N^{-1/2+\varepsilon}) \quad (6.14)$$

for all fixed H and $\varepsilon > 0$. The same choice can satisfy the corresponding statement after restricting the origin to each fixed arithmetic progression, normalized by the number of its origins in $\{2, \dots, N+1\}$. Both factors are normal on the three traces, also along all these fixed progressions. Their fixed-word logarithmic frequencies have the same limits.

Proof. Equations (6.10) and (6.13) prove (6.14). For a fixed progression, insert its indicator into each weight of (6.11), and replace N in Fourier inversion by its cardinality. This is asymptotic to N divided by the fixed modulus, so the same exponent follows. Intersect the probability-one events over the countably many horizons, progressions and positive rational ε .

Every marginal of $\mathcal{J}_m(\nu)$ is U_m , for every ν . Thus neither the normality of f and g nor its progression-wise variant requires any limiting law for h . Summation by parts gives the logarithmic statements. $\square$

The qualitative factorization alone. Let $\mathcal{N} \subseteq \mathcal{M}_\pm$ be the set of sources normal on the three traces, and let $\mathbf{m}$ be Haar probability on the prime signs. The theorem with no prescription gives $\mathbf{m}(\mathcal{N}) = 1$. For every fixed h , translation invariance gives

$$\mathbf{m}(\mathcal{N} \cap h\mathcal{N}) = 1.$$

Choosing f in this intersection gives $g = hf \in \mathcal{N}$ and $fg = h$, since every sign source squares to one. A finite prime prescription may also be imposed by intersecting with its positive-measure cylinder. This observation supplies neither a canonical coupling law nor a continuous computable selector or a rate uniform in h . Those are the stronger conclusions of the weighted argument and Section 7.

All coefficients of the empirical law of h remain in the target. The argument has canceled the nontrivial Fourier coefficients of X within each observed word class of h ; it has not canceled the frequencies of those classes themselves.

Proposition 6.3 (A prescribed increasing horizon in the probabilistic construction). *Fix, before choosing f , an integer nondecreasing sequence $H_N = o(\log \log N)$. Almost every f in Theorem 6.2 also satisfies (6.14) with H replaced by H_N , with a constant allowed to depend on (H_N) .*

Proof. For $T = 2^k$, work at $H = H_T$, $m = 3H$, and use all tests (J, v) for this window. Their number is $R = (2^m - 1)2^m < 4^m$. Combining (6.13), Cauchy–Schwarz over these tests, and (6.9) gives

$$\mathbb{E} \sup_{T/2 \leq t \leq T} d_{\text{TV}}(\mathcal{L}_t(W_f^H, W_h^H), U_m \otimes \mu_{t,h}^H)^2 \leq \frac{16^m (k+1)^2}{T} M_{H,S}(T) = T^{-1+o(1)}.$$

The last equality uses Lemma 6.1; the number of tests is explicitly paid, not hidden in a constant depending on H . Markov at threshold $T^{-1/2+\varepsilon}$ again gives a summable dyadic series. For $T/2 \leq N \leq T$, monotonicity gives $H_N \leq H_T$; projection of the words contracts total variation and projects the target to the target at H_N . The same bijection as above returns the coupling of f, g . Intersect over positive rational ε . $\square$

This route has realization-dependent constants and requires the moving horizon to be prescribed. The next construction removes both restrictions without changing the square-class estimate.

7 Selecting the factors continuously and computably

The almost-sure theorem does not by itself tell us how to choose the prime signs of a valid factor. Searching for a sign assignment that passes more and more finite tests can also fail: later decisions might destroy earlier guarantees, or a test might require an unknown rate in a Diophantine estimate. We instead normalize each finite test by its exactly computable square-class multiplicity, combine all tests into a positive potential, and bound the uncomputed tail before each decision.

The selection scheme is shared with [15, Section 5, especially Proposition 5.1]. That argument also uses exact normalizers, a positive weighted potential, the conditional tail factor $2^s 2^{1-R}$, and a total drift of $1/4$. The arithmetic costs here are different: the second-moment fibres of Section 6 replace constrained-prime moment counts. We give the specialization in full to keep its common normalization and growing-horizon guarantees explicit.

7.1 One normalized energy for an entire scale

Let $T = 2^k$, $m = 3H$, and $R_H = 2^m - 1$. Write $\mathcal{I}_k$ for all blocks at the $k + 1$ levels of the dyadic partitions of $\{2, \dots, T + 1\}$. For the fixed input h , put $V(n) = W_h^H(n)$ and define

$$F_{H,k}(f; h) = \sum_{J \neq \emptyset} \sum_{v \in \{\pm 1\}^m} \sum_{I \in \mathcal{I}_k} \left(\sum_{n \in I} \mathbf{1}_{V(n)=v} f(P_J(n)) \right)^2. \quad (7.1)$$

It is a nonnegative integer depending on finitely many prime signs of f and finitely many values of h . The normalizing denominator is

$$D_{H,k} = R_H(k + 1) T M_{H,S}(T) > 0. \quad (7.2)$$

Let $\mathbb{E}_S$ denote expectation with the signs on S fixed and all other prime signs fair and independent.

Lemma 7.1 (Uniform normalization). *For every input h and every prescription on S ,*

$$\mathbb{E}_S F_{H,k}(f; h) \leq D_{H,k}.$$

The expectation conditional on any finite assignment of additional prime signs is a computable rational number.

Proof. Use (6.8) with $P = P_J$ and $w_n = \mathbf{1}_{V(n)=v}$. Summing over v gives exactly

$$\sum_v \mathbf{1}_{V(n)=v}^2 = 1.$$

Thus the expectation costs $M_{H,S}(T)|I|$ for each J, I , not 2^m times that quantity. At each dyadic level the block sizes sum to T . There are $k + 1$ levels and R_H nonempty subsets J , giving (7.2).

Expanding (7.1) gives a finite polynomial of the form

$$F_{H,k}(f; h) = \sum_B c_B \prod_{q \in B} f(q), \quad c_B \in \mathbb{Z},$$

where B ranges over finite subsets of the free primes. Under a partial assignment, a monomial with an unassigned prime has conditional mean zero; every other monomial is determined. All coefficients, factorizations and input queries are finite. This gives the exact conditional expectations. $\square$

The saving in the sum over v is useful: it avoids charging every word separately before using that the word classes partition the origins.

7.2 Paying the infinite tail in advance

Define the positive potential

$$\mathcal{U}_h(f) = \sum_{H \geq 1} \sum_{k \geq 0} \frac{2^{-H}}{(k+1)(k+2)} \frac{F_{H,k}(f; h)}{D_{H,k}}. \quad (7.3)$$

The weights sum to one, so Lemma 7.1 gives $\mathbb{E}_S \mathcal{U}_h \leq 1$, uniformly in h . If we retain only the rectangle

$$1 \leq H \leq R, \quad 0 \leq k \leq 2^R - 2,$$

its total weight is $(1 - 2^{-R})^2$. Hence the expected omitted tail is at most

$$1 - (1 - 2^{-R})^2 = 2^{1-R} - 2^{-2R} \leq 2^{1-R}. \quad (7.4)$$

After s free prime signs have been fixed, their cylinder has probability 2^{-s} . Positivity therefore gives, for every such cylinder,

$$\mathbb{E}_S[\text{omitted tail} \mid \text{these } s \text{ signs}] \leq 2^s 2^{1-R}. \quad (7.5)$$

This is a conditional-expectation bound before choosing the next sign. It is not a claim that the uncomputed tail is pointwise small for every source.

A finite pilot may be included without discarding any horizon. Fix integers $H_0 \geq 1$, $k_0 \geq 0$, and a dyadic rational $\delta = 2^{-D_0}$, $D_0 \geq 0$. Put

$$P_h(f) = \frac{F_{H_0, k_0}(f; h)}{D_{H_0, k_0}}, \quad \mathcal{V}_h(f) = (1 - \delta)P_h(f) + \delta \mathcal{U}_h(f). \quad (7.6)$$

Then $\mathbb{E}_S \mathcal{V}_h \leq 1$. Taking $\delta = 1$ is the construction without a pilot. A small δ allows more early decisions to be made using only the pilot, at the eventual cost of a factor $\delta^{-1/2}$ in the error bound.

7.3 Irrevocable rational decisions

Choose a computable enumeration of all primes outside S , independent of h . The natural order suffices. For a pilot, one may instead list its finitely many relevant primes first, then the remaining primes in natural order.

At step $s \geq 1$, the first $s - 1$ free signs have been fixed. Set $\varepsilon_s = 2^{-s-2}$. For each candidate value of the next sign, compute the conditional expectation of the pilot and a finite part of the universal potential. If $D_0 \geq 2s + 2$, omit the universal part entirely: its conditional expectation is at most $\delta 2^s \leq \varepsilon_s$. Otherwise take

$$R = \max\{1, 2s + 3 - D_0\}$$

and use the rectangle of (7.4); its weighted omitted tail is at most $\delta 2^s 2^{1-R} \leq \varepsilon_s$.

The two computed expectations are rational numbers A_+ and A_- . Assign the sign with the smaller value, using $+1$ in case of equality. Each sign is decided once and is never revised.

Lemma 7.2 (Summable drift). *This procedure defines a source f_h satisfying the prescription and*

$$\mathcal{V}_h(f_h) \leq \frac{5}{4}, \quad \mathcal{U}_h(f_h) \leq \frac{5}{4\delta} < \frac{2}{\delta}. \quad (7.7)$$

Every output prime sign is computable relative to finitely many input prime signs. The map $h \mapsto f_h$ is continuous in the product topology.

Proof. Let C_{s-1} be the conditional expectation of the full potential after $s-1$ choices, and C_+, C_- its two child expectations. These are finite because each cylinder has positive probability and $\mathbb{E}_S \mathcal{V}_h \leq 1$. Their average is C_{s-1} . The tail bound gives

$$A_{\pm} \leq C_{\pm} \leq A_{\pm} + \varepsilon_s.$$

For the selected child,

$$C_s \leq \min\{A_+, A_-\} + \varepsilon_s \leq C_{s-1} + \varepsilon_s.$$

Since $C_0 \leq 1$ and $\sum_{s \geq 1} \varepsilon_s = 1/4$, every C_s is at most $5/4$.

A finite subsum of (7.6) is permanently determined once its finitely many prime signs have been assigned. Its value is then at most C_s by positivity of the remaining terms. Increasing the finite subsums proves the first inequality of (7.7); the second follows from $\delta \mathcal{U}_h \leq \mathcal{V}_h$.

At each step the rectangle, the polynomial values, their factorizations, and the finite rational comparisons are computable. Only finitely many values of h are used, each reducible to finitely many prime-sign queries. The bounds on the queried arguments depend on the decision rank and on the fixed parameters, not on the unknown rest of h . Thus every output coordinate depends on finitely many input coordinates. This proves relative computability and continuity. No comparison of two unknown real numbers or undecided asymptotic threshold is needed. $\square$

The source exists at all primes because the enumeration eventually reaches each one. The argument bounds every finite part of the potential along that single source, rather than choosing a different source at each observation scale.

What computability costs in this prescription. Without a pilot, $\delta = 1$ and $D_0 = 0$. Already at the first decision, $s = 1$, the prescribed rectangle has $R = 5$ and reaches $k = 2^5 - 2 = 30$, hence includes tests with $T = 2^{30}$ origins. This is a calculation of the stated parameters, not a runtime measurement or a lower bound for every possible implementation. It explains why small pilot decisions, whose omitted universal tail is paid for separately, are not a run of the full no-pilot rectangle.

7.4 The finite bound and its asymptotic uses

Theorem 7.3 (Uniform finite coupling bound). *For the selector just constructed, set $g_h = hf_h$. For all $H, N \geq 1$, let $k = \lceil \log_2 N \rceil$ and $T = 2^k$. Then*

$$\begin{aligned} d_{\text{TV}}(\mathcal{L}_N(W_{f_h}^H, W_{g_h}^H), \mathcal{J}_{3H}(\mu_{N,h}^H)) &\leq \min\{1, B_{S,\delta}(H, N)\}, \\ B_{S,\delta}(H, N) &= 2^{5H+1} \delta^{-1/2} (k+2)^2 \sqrt{\frac{M_{H,S}(T)}{T}}. \end{aligned} \quad (7.8)$$

The bound is common to all inputs h , including noncomputable inputs. Its square is a computable rational number, independent of h .

Proof. Keep the sums $A_{J,v}(N)$ from (6.11), now evaluated at f_h . There are $R_H 2^m$ tests, where $m = 3H$. Cauchy–Schwarz in (6.13) and the dyadic decomposition of a prefix give

$$d_{\text{TV}}(\mathcal{L}_N(W_{f_h}^H, W_h^H), U_m \otimes \mu_{N,h}^H)^2 \leq \frac{R_H 2^m (k+1)}{4N^2} F_{H,k}(f_h; h).$$

The (H, k) term of (7.3), together with (7.7), gives

$$F_{H,k}(f_h; h) \leq \frac{2}{\delta} 2^H (k+1)(k+2) D_{H,k}.$$

Substituting (7.2) yields the decisive estimate

$$d_{\text{TV}}^2 \leq \frac{2^{m+H-1}}{\delta} R_H^2 (k+1)^3 (k+2) \frac{TM_{H,S}(T)}{N^2}. \quad (7.9)$$

Now $N \geq T/2$, $R_H < 2^m$, and $m = 3H$. These inequalities imply the square of the upper bound in (7.8), with harmless slack in its power of two. Finally, the bijection $(x, v) \mapsto (x, x \odot v)$ preserves total variation and gives the asserted coupling of f_h, g_h . $\square$

For fixed H , Lemma 6.1 turns (7.8) into $O_{S,\delta,H,\varepsilon}(N^{-1/2+\varepsilon})$. Both target marginals are uniform. Together with Lemma 7.2, this proves the fixed-horizon part of Theorem 1.3, taking $\delta = 1$. Logarithmic word frequencies again follow by summation by parts. The uniform bound concerns the two selected factors for each input; the factors themselves vary with the input.

Corollary 7.4 (The same selection at all slowly growing horizons). *Fix the selector, including S , the prescription and any pilot parameters. For every integer sequence $H_N \geq 1$ with $H_N = o(\log \log N)$ and every $\varepsilon > 0$,*

$$\sup_{h \in \mathcal{M}_{\pm}} d_{\text{TV}}(\mathcal{L}_N(W_{f_h}^{H_N}, W_{g_h}^{H_N}), \mathcal{J}_{3H_N}(\mu_{N,h}^{H_N})) = O_{S,\delta,(H_N),\varepsilon}(N^{-1/2+\varepsilon}).$$

No monotonicity or prior prescription of (H_N) is needed.

Proof. The finite inequality (7.8) already holds for every pair (H, N) along the same selected source. For $T = 2^{\lceil \log_2 N \rceil}$, $\log \log T \sim \log \log N$. The moving-window part of Lemma 6.1, and the factor $2^{5H_N} (\log N)^2 = N^{o(1)}$, make its right-hand side $N^{-1/2+o(1)}$. There is no new selection or uncountable intersection of probability-one events. $\square$

The constant and the threshold may depend on the chosen sequence. This is not a single rate uniform over every possible sequence satisfying a little- o condition.

There is also a horizon certified entirely by finite arithmetic. For $k \geq 1$, let $\widehat{H}(k)$ be the largest integer $1 \leq H \leq k$ satisfying

$$\left[2^{10H+2} \delta^{-1} (k+2)^4 M_{H,S}(2^k) \right]^2 \leq 2^k, \quad (7.10)$$

and set $\widehat{H}(k) = 0$ when none qualifies. This is a rational test on finitely many exact multiplicities. For each fixed H , (6.3) eventually makes the test true. Therefore $\widehat{H}(k) \rightarrow \infty$. For $2^{k-1} < N \leq 2^k$ and $\widehat{H}(k) \geq 1$, Theorem 7.3 gives

$$\sup_h d_{\text{TV}}(\mathcal{L}_N(W_{f_h}^{\widehat{H}(k)}, W_{g_h}^{\widehat{H}(k)}), \mathcal{J}_{3\widehat{H}(k)}(\mu_{N,h}^{\widehat{H}(k)})) \leq 2^{-k/4} \leq N^{-1/4}. \quad (7.11)$$

No monotonicity of $\widehat{H}$ is asserted. The value zero corresponds to the empty word. Computability here does not imply practical efficiency: computing all the multiplicities, or the rectangles required by the selector, can be enormous. The asymptotic theorem ensures eventual acceptance of every fixed horizon without supplying a practical small-size threshold.

8 Product word laws in the synchronous coupling

The marginal word laws of the selected factors are all the same. Their joint law is not generally the product of those marginals. For any pointwise factorization $fg = h$, multiplying the readings at each known origin gives the corresponding reading of h ; pushing forward the synchronous empirical measure therefore gives its empirical word law exactly. Our construction adds quantitative approximation by the canonical coupling. We now describe what this means for product *word laws*, without identifying that information with the arbitrary source itself.

8.1 Convergence, subsequential limits and entropy

For any two laws ν, ν' on the sign cube,

$$d_{\text{TV}}(\mathcal{J}_m(\nu), \mathcal{J}_m(\nu')) = d_{\text{TV}}(\nu, \nu'). \quad (8.1)$$

Indeed, each value of $v = x \odot y$ has 2^m preimages, each with weight 2^{-m} . Summing the absolute differences proves the identity.

Proposition 8.1 (Preservation of product word laws and their subsequential limits). *For either factorization in Theorem 6.2 or Theorem 7.3, and each fixed H , the empirical law of (W_f^H, W_g^H) converges if and only if $\mu_{N,h}^H$ converges. If its product law converges to ν_h^H , its synchronous law converges to $\mathcal{J}_{3H}(\nu_h^H)$. The set of subsequential limits of the synchronous sequence is exactly the image under $\mathcal{J}_{3H}$ of the set of subsequential limits of $\mu_{N,h}^H$.*

Proof. At every origin and sample size,

$$W_f^H(n) \odot W_g^H(n) = W_h^H(n).$$

Thus projection by coordinatewise product sends the synchronous empirical law exactly to $\mu_{N,h}^H$. The error to $\mathcal{J}_{3H}(\mu_{N,h}^H)$ tends to zero. The convergence assertions follow from this projection and (8.1); the assertion about subsequential limits follows by taking subsequences in the compact finite-dimensional probability simplex. $\square$

This statement includes inputs with no limiting word law. Their nonconvergence remains visible in the sequence of synchronous empirical laws, even though the two marginal laws converge to Bernoulli.

The product word law need not identify its source. Take distinct normal sources h_1, h_2 , for instance the pair constructed in Section 8.3. Then $\nu_{h_1}^H = \nu_{h_2}^H = U_{3H}$ for every H . Applying the selected factorization to each input gives, in both cases, the same synchronous limiting law

$$\mathcal{J}_{3H}(U_{3H}) = U_{3H} \otimes U_{3H},$$

even though the pointwise products are the different functions h_1 and h_2 . Thus the canonical coupling is injective as a function of a word law, by (8.1), but not as a function of an arbitrary source. Origin-labelled synchronous readings and their unlabelled limiting distribution are different data.

When all ν_h^H exist, they form a consistent stationary word law: dropping a first or a last position changes the relevant finite averages by at most a boundary term. With Shannon

entropy H in bits, the bijection $(X, V) \mapsto (X, X \odot V)$ and the independence in the definition of $\mathcal{J}$ give

$$\begin{aligned} H(\mathcal{J}_{3H}(\nu_h^H)) &= 3H + H(\nu_h^H), \\ I(W_f^H; W_g^H) &= 3H - H(\nu_h^H). \end{aligned} \quad (8.2)$$

Here the mutual information is evaluated in the limiting synchronous law. Dividing by H and using the entropy rate of the stationary input law gives joint entropy rate $3 + h(\nu_h)$ and mutual-information rate $3 - h(\nu_h)$. These are identities for the described coupling, not assumptions that the two marginal processes are independent.

8.2 A parity product: pairwise independence with a four-reading constraint

Let $\beta_q(m) = (-1)^{v_q(m)}$ and take $h = \beta_2$. Its three-trace vector is

$$(\beta_2(Q_{-1}(n)), \beta_2(Q_{-2}(n)), \beta_2(Q_2(n))) = (-1)^n(1, -1, -1).$$

Its H -word takes two opposite values $v_H, -v_H$, each with limiting probability $1/2$. The corresponding synchronous law is therefore

$$\nu_{f,g}^H(x, y) = \begin{cases} 2^{-(3H+1)}, & y = x \odot v_H \text{ or } y = -x \odot v_H, \\ 0, & \text{otherwise.} \end{cases} \quad (8.3)$$

It has $2 \cdot 8^H$ states rather than 64^H . Each marginal nevertheless has all 8^H words with equal probabilities. Every pair of distinct scalar coordinates in the two words is independent and fair: different coordinates inherit independent signs from X , while the same coordinate across the two factors has an independent fair parity mask. Yet, for every $d \in \mathcal{D}$ and every origin,

$$f(Q_d(n))g(Q_d(n))f(Q_d(n+1))g(Q_d(n+1)) = -1. \quad (8.4)$$

This is a pointwise constraint on four readings across two sources. It is different from the diagonal single-source correlations used in the structured inverse.

For $H = 1, 2, 3$, the numbers of synchronous states are respectively 16, 128, 1024, and their entropies are 4, 7, 10 bits. In general the input word carries one bit per block and zero bits per position in the entropy-rate limit. Both marginal entropy rates are three, while their joint entropy rate is three rather than six. The common dyadic prime cannot be prescribed identically in this example: $f(2)$ and $g(2)$ are opposite.

8.3 The same prime discrepancy at three different information levels

For $h = \beta_{17}$, prescribe $f(2) = -1$ in the deterministic selector. It gives computable sources

$$g = f\beta_{17}, \quad f(2) = g(2) = -1,$$

whose word laws equal U_{3H} at every fixed horizon. Their prime signs differ only at 17. On each of the three traces their disagreement density is

$$2(1 - 17^{-1}) \sum_{j \geq 0} 17^{-(2j+1)} = \frac{1}{9}.$$

This is the same quotient as that of the explicit pair (1.1), but not the same pair of sources. The explicit finite-conductor pair is indistinguishable at order at most three and is separated

at order four. The new pair has the same complete Bernoulli marginal laws at every order and horizon, and is not subject to the structured-class hypothesis of Theorem 1.1.

At horizon one, the quotient law has mass $2/3$ on $(+, +, +)$ and mass $1/9$ on each of the three words with one negative coordinate. The six roots of the three quadratics are distinct modulo 17, so at most one trace can have a positive 17-adic valuation; the odd-valuation formula gives these masses. Its synchronous coupling has 32 states, with two uniform marginals on eight states.

There is a stronger common-prefix variant. Given $B \geq 2$, any $h_* \in \mathcal{M}_\pm$ and $0 < \eta < 1/2$, choose a finite set Γ of primes $q > B$, $q \equiv 1 \pmod{8}$, such that

$$\delta_\Gamma = \frac{1}{2} \left(1 - \prod_{q \in \Gamma} \frac{q-3}{q+1} \right) > \frac{1}{2} - \eta. \quad (8.5)$$

The prime-progression divergence used in Section 5 permits this choice. Apply the selector to $h = \prod_{q \in \Gamma} \beta_q$, prescribing $f = h_*$ at primes at most B . Since $h = 1$ throughout that prefix, both factors agree with h_* on $[1, B]$. One and the same pair satisfies

$$\nu_f^H = \nu_g^H = U_{3H} \quad \text{for every } H,$$

and has disagreement density δ_Γ on each trace. Here $h_*(2)$ need not be negative. Only its finitely many prescribed prime values are used, so the resulting sources are computable once those finite data are given. This differs from Theorem 1.2: the common law is now exactly Bernoulli at every horizon, but the sources are no longer finite-conductor, zero-entropy sources.

Taking the computable input $h(n) = \lambda(n) = (-1)^{\Omega(n)}$ likewise gives computable factors $f_\lambda g_\lambda = \lambda$, each normal on the three traces. This identity provides no cancellation estimate for Liouville itself. Its empirical product law is the very law retained in the target coupling, not an error term that has been removed.

8.4 A whole summable coset in a single normal fibre

The uniform selector adapts the factor to the input. A different simultaneous statement follows from the probabilistic construction: a single factor can accommodate every summable prime perturbation of the identity.

Proposition 8.2 (A common normal coset). *There exists a source f , with any prescribed finite set of prime signs, such that fu is normal on the three traces for every*

$$u \in \mathcal{S}_{1/2} = \left\{ u \in \mathcal{M}_\pm : \sum_{q: u(q)=-1} q^{-1/2} < \infty \right\}.$$

The same statement can hold along every fixed arithmetic progression. For a fixed prime set Γ with $\sum_{q \in \Gamma} q^{-1/2} < \infty$, write $u_A = \prod_{q \in A} \beta_q$. Then this f satisfies, for each fixed H ,

$$\sup_{A \subseteq \Gamma} d_{\text{TV}}(\mu_{N, fu_A}^H, U_{3H}) \longrightarrow 0. \quad (8.6)$$

Proof. There are only countably many sources u supported at finitely many negative primes. The weighted proof of Theorem 6.2 allows a single f for which all such fu are normal, at all horizons and on all fixed progressions.

For arbitrary $u \in \mathcal{S}_{1/2}$, truncate its negative-prime support at Y to get $u_{\leq Y}$. Multiplication by f does not change the positions where the two perturbations disagree. Lemma 2.1, followed

by a union bound over the $3H$ readings and fixed-shift boundary terms, gives

$$\limsup_{N \rightarrow \infty} d_{\text{TV}}(\mu_{N,fu}^H, \mu_{N,fu \leq Y}^H) \leq 24H \sum_{\substack{q > Y \\ u(q) = -1}} q^{-1/2}. \quad (8.7)$$

At fixed Y the truncated product is normal. Let $Y \rightarrow \infty$ after $N \rightarrow \infty$. On a fixed progression, restricting the count and renormalizing multiplies the upper limit by at most its fixed modulus, which is harmless.

For all $A \subseteq \Gamma$, the tail in (8.7) is bounded by the same tail of Γ . Below Y there are only finitely many possible assignments, and their normality is already known. This proves (8.6). No uncountable intersection of probability-one events has been taken. $\square$

If Γ is infinite, this gives an uncountable family of distinct sources with the same complete word laws. Choosing Γ beyond the prescribed prefix preserves that prefix throughout the family. If its primes are all $1 \pmod{8}$, the disagreement density of fu_A and fu_B on each trace is

$$\frac{1}{2} \left(1 - \prod_{q \in A \Delta B} \frac{q-3}{q+1} \right).$$

The product converges by summability, and the same truncation argument proves the density formula. It is strictly positive when $A \neq B$. There is no rate common to all of $\mathcal{S}_{1/2}$, whose tails need not be uniformly small.

This also illustrates why the positive inverse in that class is not contradicted. The inverse assumes that the source itself is in $\mathcal{S}_{1/2}$ (or in $\mathcal{C}$ after a finite character component). The present functions are translates fu by a normal source. In the progression-wise construction, every scalar trace of f disagrees with every periodic sign sequence with density $1/2$, since it is balanced on each residue class of the period. It cannot be approximated in density by periodic sign sequences, unlike the structured traces obtained from the truncations in Section 2.

8.5 A compact uniformly normal family whose square is the whole group

Give $\mathcal{M}_{\pm} \cong \{\pm 1\}^{\{\text{primes}\}}$ the product topology. With no prescription and no pilot, the maps $F(h) = f_h$ and $G(h) = hf_h$ from Section 7 are continuous. Define

$$K = F(\mathcal{M}_{\pm}) \cup G(\mathcal{M}_{\pm}).$$

Corollary 8.3 (Uniform normality does not restrict products). *The set K is compact, and for every fixed H ,*

$$\sup_{u \in K} d_{\text{TV}}(\mu_{N,u}^H, U_{3H}) \leq \min\{1, B_{\emptyset,1}(H, N)\} \rightarrow 0.$$

Nevertheless, $KK = \mathcal{M}_{\pm}$.

Proof. Continuous images of the compact group $\mathcal{M}_{\pm}$ are compact. The displayed estimate is the marginal projection of Theorem 7.3. Every h is $F(h)G(h)$, and a product of two elements of $\mathcal{M}_{\pm}$ remains in $\mathcal{M}_{\pm}$. $\square$

With a prescription ξ on S , restrict inputs to the subgroup $\mathcal{M}_{\pm,S} = \{h : h(q) = 1 \ (q \in S)\}$. The union $K_{S,\xi}$ of the two restricted images is compact, lies in the prescribed cylinder, is uniformly normal, and satisfies $K_{S,\xi}K_{S,\xi} = \mathcal{M}_{\pm,S}$.

The estimate remains valid even when an element of K is selected differently at each N . It says nothing about independence between two elements of K . The constant function, a parity perturbation, or Liouville may still be their product. A factor independent of the input cannot replace the selector: if f worked for every h , the input $h = f$ would give $fh = 1$.

Comparison with qualitative compactness. Existence of some uniformly normal compact set whose square is $\mathcal{M}_\pm$ also follows from full Haar measure. Apply Egorov’s theorem separately to the countably many fixed horizons, assigning summable measure losses of total less than $1/4$. Their intersection has measure greater than $3/4$ and uniform convergence for each horizon. By inner regularity it contains a compact set K' with $\mathfrak{m}(K') > 1/2$. For every h ,

$$\mathfrak{m}(K' \cap hK') \geq 2\mathfrak{m}(K') - 1 > 0,$$

so $K'K' = \mathcal{M}_\pm$. Corollary 8.3 gives the specific compact defined by the selector images and the common finite bound (7.8); it does not require a separate claim of a new qualitative compactness principle.

9 Scope of the identification results

The three kinds of statistical data considered here have different strengths. In $\mathcal{C}$, second-order correlations identify the nonramified structure, while a local fourth-order detector recovers each remaining conductor sign. A bounded horizon is still insufficient uniformly over this class, even when its full word distribution is given exactly. Without the structural hypothesis, complete separate word laws need not identify even a single prime discrepancy. The synchronous empirical law, unlike either marginal alone, projects to the empirical word law of the prescribed product. Its canonical limiting coupling retains exactly that limiting word-law information, which can itself fail to identify the product as a function.

The inverse uses limits in a specified order: first the average over integer origins, then averages or isolated sequences of shifts. It is not a reconstruction algorithm with a uniform finite observation budget. Conversely, the selection theorem has a finite computable bound but does not claim practical efficiency. Neither result deduces a structured model from small pointwise trace errors. In particular, the equal-law examples are not sources whose three traces are almost always positive, and the factorization of an input does not change or estimate that input’s errors.

The square-class estimate in Lemma 6.1 is the Diophantine input to the factorization, and the character-sum bound in Proposition 5.2 controls the finite-conductor comparison. The accompanying finite checks test selected identities, residue laws and initial selector decisions, not the external theorems or the limiting conclusions.

Reproducibility

The source-and-check package for this edition is `N4_correlation_laws_publication.zip`. It contains the LaTeX sources, frozen bibliography, exact verification code, recorded outputs, commands and a SHA-256 manifest. From its root, `sh build.sh` compiles the manuscript and `sh verify.sh` reruns the checks in ordinary and optimized Python modes and compares their outputs.

The finite-check bundle `N4_checks_v1_2.zip` is supplied inside the published source package on Zenodo, <https://doi.org/10.5281/zenodo.23318918>. From the extracted bundle’s root, run `sh run_checks.sh`; this also compares the new outputs with the recorded ones. Its `check_exact.py` is the unchanged checker from the initial edition, identified by its hash in the bundle manifest. It checks the stated finite residue identities, collision cases, dyadic signatures, coupling identities and finite energies. Its fifteen pilot models contain 39 free-prime decisions in total, with $\delta = 2^{-64}$ paying the omitted universal tail; they do not evaluate the full no-pilot selector. The separate `check_revision.py` checks the finite coupling illustration and the first-decision parameter calculation. Both use only the Python standard library. The

bundle README separates exact identities, certified pilot decisions, and claims not tested. No finite experiment is used to infer an infinite normality theorem, and no independent review code is represented as part of this verification suite.

AI assistance and responsibility

GPT-6 Astra assisted the exploration, writing, finite computations, reference checking and code preparation of the source notes and initial manuscript. GPT-6 Astra Pro assisted subsequent revisions, bibliographical comparisons and reproducibility checks. The author retains responsibility for the manuscript and any submission. These AI-assisted readings and finite checks are neither independent human peer review nor formal verification.

References

- [1] Jean-Paul Allouche and Leo Goldmakher. Mock characters and the Kronecker symbol. *Journal of Number Theory*, 192:356–372, 2018. doi:10.1016/j.jnt.2018.04.022. Theorem 2.11 is cited from arXiv:1608.03957v2, 4 January 2017, <https://arxiv.org/html/1608.03957v2>.
- [2] Vitaly Bergelson, Joanna Kułaga-Przymus, Mariusz Lemańczyk, and Florian K. Richter. Rationally almost periodic sequences, polynomial multiple recurrence and symbolic dynamics. *Ergodic Theory and Dynamical Systems*, 39(9):2332–2383, 2019. arXiv:1611.08392v2, <https://arxiv.org/abs/1611.08392v2>; doi:10.1017/etds.2017.130.
- [3] W. Riley Casper and Bobby Orozco. Reconstructing rational functions on finite abelian groups with higher autocorrelations, 2026. arXiv:2503.21022v2, 4 July 2026, <https://arxiv.org/html/2503.21022v2>. Journal version: doi:10.1016/j.acha.2026.101908. Comparisons use the cited preprint.
- [4] Régis de la Bretèche, Victor Y. Wang, and Max Wenqiang Xu. Random multiplicative functions and making squares from polynomial values, 2026. arXiv:2607.06398v1, 7 July 2026. Lemmas 3.2–3.3 and the explicit proof of Lemma 3.3. <https://arxiv.org/html/2607.06398v1>.
- [5] Cunsheng Ding. Pattern distributions of Legendre sequences. *IEEE Transactions on Information Theory*, 44(4):1693–1698, 1998. Author copy, <https://www.cse.ust.hk/faculty/cding/JOURNALS/it982.pdf>.
- [6] Jan-Hendrik Evertse and Joseph H. Silverman. Uniform bounds for the number of solutions to $Y^n = f(X)$. *Mathematical Proceedings of the Cambridge Philosophical Society*, 100(2):237–248, 1986. doi:10.1017/S0305004100066068. Original article: <https://ir.cwi.nl/pub/1788/1788D.pdf>.
- [7] Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. A constructive Borel–Cantelli lemma. Constructing orbits with required statistical properties. *Theoretical Computer Science*, 410(21–23):2207–2222, 2009. doi:10.1016/j.tcs.2009.02.010; arXiv:0711.1478v2, <https://arxiv.org/abs/0711.1478v2>.
- [8] Uwe Grimm and Michael Baake. Homometric point sets and inverse problems. *Zeitschrift für Kristallographie*, 223:777–781, 2008. arXiv:0808.0094, <https://arxiv.org/abs/0808.0094>.
- [9] F. Alberto Grünbaum and Calvin C. Moore. The use of higher-order invariants in the determination of generalized Patterson cyclotomic sets. *Acta Crystallographica Section A*, 51(3):310–323, 1995. doi:10.1107/S0108767394009827. The reconstruction result is used here for historical context via the accounts in Jaming–Kolountzakis, Section 3.4, and Casper–Orozco, Introduction.
- [10] Philippe Jaming and Mihail N. Kolountzakis. Reconstruction of functions from their triple correlations. *New York Journal of Mathematics*, 9:149–164, 2003. <https://nyjm.albany.edu/j/2003/9-12p.pdf>.
- [11] Oleksiy Klurman. Correlations of multiplicative functions and applications. *Compositio Mathematica*, 153(8):1622–1657, 2017. doi:10.1112/S0010437X17007163. Theorem 1.3 is cited from arXiv:1603.08453v1, <https://arxiv.org/html/1603.08453v1>.

- [12] Oleksiy Klurman and Alexander P. Mangerel. Rigidity theorems for multiplicative functions. *Mathematische Annalen*, 372:651–697, 2018. doi:10.1007/s00208-018-1724-6. Question 1.11 and Proposition 1.12 are cited from arXiv:1707.07817v2, <https://arxiv.org/html/1707.07817v2>.
- [13] Joseph Najnudel. On consecutive values of random completely multiplicative functions. *Electronic Journal of Probability*, 25:1–28, 2020. Paper no. 59; doi:10.1214/20-EJP456. Proposition 6.3 is cited from arXiv:1702.01470v2, <https://arxiv.org/html/1702.01470v2>.
- [14] National Institute of Standards and Technology. NIST Digital Library of Mathematical Functions. §27.9, Quadratic Characters, and §27.11, Asymptotic Formulas: Partial Sums, especially equation 27.11.9. Consulted 27 September 2026. <https://dlmf.nist.gov/27.9>; <https://dlmf.nist.gov/27.11>.
- [15] Brice Pouly. Selective normality and algebraic rigidity of multiplicative traces, 2026. Author’s manuscript, 38 pp. doi:10.5281/zenodo.23318326. Section 5, pp. 16–19, especially Proposition 5.1 and the marked-law bound in Section 5.3. Edition consulted 30 September 2026; comparison of the shared selection method, not a proof dependency.