

Rare events in multiplicative sequences

Poisson laws, clusters, and observation

Brice Pouly

Version 1.2

This article received substantial assistance from GPT-6 Astra;
see the declaration at the end of the text.

Abstract

A multiplicative identity can persist at every scale while disappearing from a rare-event count. Another identity can survive as a cluster whose points lie far apart. Which distinction matters depends on what is observed. We explain this through one completely calculated model: a uniform germ can trigger two occurrences in disjoint spatial windows. Its two marginal counts become Poisson, but their sum need not do so. A coupling identifies the correct same-grid target and separates the mean number of points from the mean number of active germs.

We then isolate the additional work required by multiplicative sources: private prime coordinates, weighted relations between windows, corrections to prime-site intensities, and relations visible at the target scale. A final passage from marked configurations to records explains the roles of ties, recognition and observation clocks. Selected results are stated with their hypotheses; a short concordance leads to their proofs. The elementary calculations are proved here, while the arithmetic and functional limit theorems are taken from the four cited research articles.

1 What does a rare occurrence see?

Choose independent fair signs at the primes and extend them completely multiplicatively:

$$f(1) = 1, \quad f(n) = \prod_{p|n} f(p)^{v_p(n)}, \quad \mathbb{P}(f(p) = 1) = \mathbb{P}(f(p) = -1) = \frac{1}{2}. \quad (1.1)$$

Here $v_p(n)$ is the exponent of p in n . The prime values are the independent random inputs; the integer values are not. In particular, $f(n^2) = 1$ and $f(2n) = f(2)f(n)$. Even distant observations can therefore share an exact constraint. This is the completely multiplicative sign model, not the model that vanishes off squarefree integers.

A long constant stretch is rare in a sequence of independent signs. Count it only at its beginning: at $x \geq 2$, require a change of sign from $x - 1$ to x , followed by L equal signs. In the independent model this event has probability 2^{-L} . Thus N candidate starts suggest the intensity $N2^{-L}$ and the critical length $L = \log_2 N + O(1)$. Counting every constant window would count one long run several times; the choice of occurrence already changes the appropriate target [4].

The arithmetic model (1.1) retains the independent Poisson form at this critical scale, with locations, signs and excess lengths recorded [8, Theorem 1.1]. But the reason cannot be that the entire sequence resembles independent signs. Squares rule that out immediately. The useful question is narrower: *how much of the dependence can this particular rare event detect?*

Rarity need not erase dependence. In the example below, with $a = 1$ and $c = 1/2$, two disjoint windows each have a limiting Poisson count of mean one. Yet the probability that both are empty tends to $e^{-3/2}$, rather than e^{-2} . Some arrivals create one point in each window; the calculation will identify and count those shared arrivals. Persistent overlap makes the sum non-Poisson.

1.1 Source, configuration, observation

We distinguish three objects. The *source* consists of the random inputs and the identities they satisfy. A *configuration* records the occurrences retained from that source, including their positions and any marks. An *observation* is a function of that configuration: a total count, counts in windows, a maximum, or an ordered history. A mark is simply extra recorded information, such as a sign or an excess length.

The law $\text{Pois}(\lambda)$ assigns probability $e^{-\lambda}\lambda^r/r!$ to $r \in \mathbb{N}_0 = \{0, 1, \dots\}$. For probability laws on one measurable space, put

$$d_{\text{TV}}(P, Q) = \sup_A |P(A) - Q(A)|.$$

It bounds the discrepancy in the probability of *every* event on that space. If P and Q are compared and the same measurable observation Φ is applied to each, then

$$d_{\text{TV}}(\mathcal{L}(\Phi(X)), \mathcal{L}(\Phi(Y))) \leq d_{\text{TV}}(\mathcal{L}(X), \mathcal{L}(Y)). \quad (1.2)$$

Indeed, an output event has the input event $\Phi^{-1}(A)$. The same conclusion holds for a common randomized observation, by averaging over its extra randomness. This elementary contraction is the connection between arithmetic comparison and probabilistic observation, not a substitute for either.

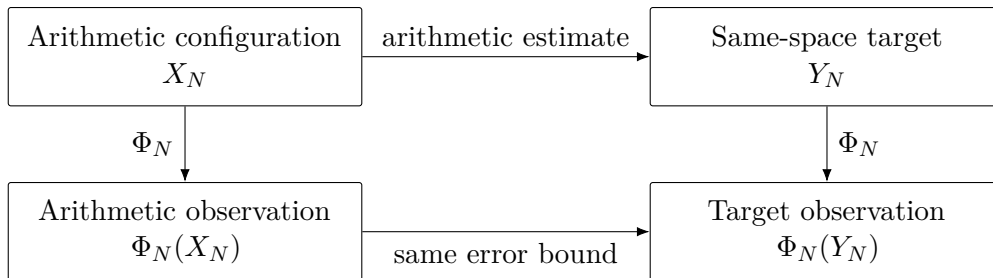


Figure 1: Two tasks, one comparison. The target may be Poisson or a cluster law. Its dependence on N remains after the lower arrow; identifying a limiting observation requires a further argument.

The target first uses the same spatial grid: a nonempty configuration supported on the grid distinguishes it from a diffuse law. Passing to continuous positions is a separate limiting argument. Likewise, (1.2) controls probabilities, not unbounded moments.

The research sources are cited by short titles: *Long runs* [8], *Arithmetic clusters* [10], *Moving prime environments* [9] and *Records and scale flows* [11]. All unlabelled logarithms are natural.

A first route. Sections 1–2 give the central calculation; Section 3 explains what the arithmetic adds, and Section 5 is a reference guide. For a quick first pass, continue from the model to the record example in Section 4.2.

2 One germ, two possible occurrences

2.1 A common input that does not create dependence

Fix $a > 0$ and $c \in \mathbb{R}$. Work on the additive circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, with uniform probability measure m . Let H have any distribution on $\mathbb{T}$, and independently choose $U_1, \dots, U_k$ uniform and independent. For an integer $k > 2(a + |c|)$ set

$$\begin{aligned} \Theta_i &= U_i + H \pmod{1}, & J_k &= \left(-\frac{a}{2k}, \frac{a}{2k}\right) \pmod{1}, \\ A_i &= \mathbf{1}_{\{\Theta_i \in J_k\}}, & B_i &= \mathbf{1}_{\{\Theta_i + c/k \in J_k\}}. \end{aligned} \tag{2.1}$$

The index i labels a *germ*: one input capable of producing one or several recorded points. A germ is active when it produces at least one point.

For every fixed $H = h$, translation preserves the uniform measure and the variables Θ_i remain independent. Their joint conditional law is always $m^{\otimes k}$, regardless of h . Hence the whole array of pairs (A_i, B_i) is independent of H . A shared symbol in a representation is not by itself a proof of statistical dependence. What decides the issue here is the complete conditional law.

Within one pair, however, the same Θ_i is tested twice. Write

$$b = (a - |c|)_+ \quad \text{where } t_+ = \max(t, 0). \tag{2.2}$$

The two activation arcs are J_k and $J_k - c/k$. Our choice of k ensures that their union can be read without a circular wrap, so their overlap has mass b/k . Consequently

Category (A_i, B_i)	(1, 0)	(0, 1)	(1, 1)	(0, 0)
Probability	$(a - b)/k$	$(a - b)/k$	b/k	$1 - (2a - b)/k$

The three nonempty categories are mutually exclusive for one germ. Different germs are independent. This is the independence we will use.

To retain geometry, put

$$x_i = \frac{1}{4} + \frac{i - 1/2}{12k}, \quad \mathcal{X}_k = \sum_{i=1}^k (A_i \delta_{x_i} + B_i \delta_{2x_i}). \tag{2.3}$$

Here δ_x denotes a unit point at x , with multiplicities counted when points are added. The two bands are $(1/4, 1/3)$ and $(1/2, 2/3)$. These positions are a convention of this model; no completely multiplicative function has been constructed.

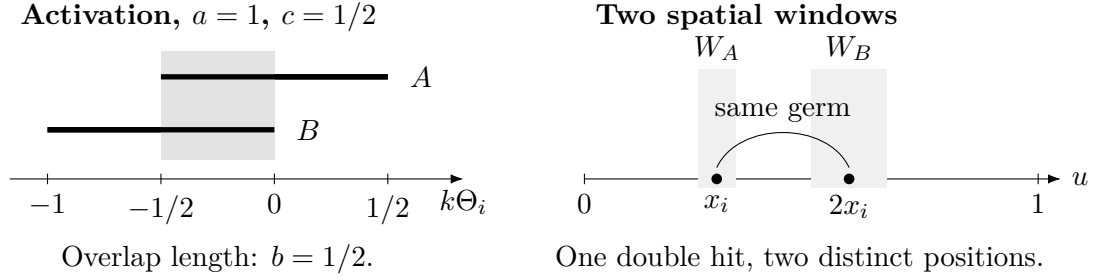


Figure 2: Angular overlap and spatial separation are different data. The left panel uses the enlarged angular coordinate; the right uses $W_A = (1/4, 1/3)$ and $W_B = (1/2, 2/3)$, with one pair from (2.3).

2.2 The right Poisson object lives on categories

At each germ i , replace the three category indicators by independent Poisson variables of means $(a - b)/k, (a - b)/k, b/k$, also independent between germs and of H . Each occurrence of type 10 produces x_i , type 01 produces $2x_i$, and type 11 produces both. Call the resulting configuration $\mathcal{X}_k^P$. The germ counts are Poisson; the offspring configuration can be a *Poisson cluster field* rather than a field of independent points.

Proposition 2.1 (Same-grid comparison). *For the preceding fixed a, c and every integer $k > 2(a + |c|)$,*

$$d_{\text{TV}}(\mathcal{L}(\mathcal{X}_k), \mathcal{L}(\mathcal{X}_k^P)) \leq \frac{(2a - b)^2}{k}. \quad (2.4)$$

The same bound holds jointly with the retained variable H .

Proof. The active-germ probability is $q = (2a - b)/k < 1$. Before recording the category, the source has one Bernoulli(q) count and the target has one Poisson(q) count. For $0 \leq q \leq 1$,

$$d_{\text{TV}}(\text{Bern}(q), \text{Pois}(q)) = q - qe^{-q} = q(1 - e^{-q}) \leq q^2. \quad (2.5)$$

To check the equality, the Bernoulli mass exceeds the Poisson mass only at the value 1: at 0, $1 - q \leq e^{-q}$, and at integers at least two the Bernoulli mass is zero. The total positive excess is therefore the expression displayed. Keeping the common masses at zero and one, and coupling the remaining masses arbitrarily, realizes this discrepancy as a coupling failure probability.

When the counts agree at one, use the same type, with probabilities proportional to $a - b, a - b, b$. In the target these are precisely the types obtained by independent Poisson splitting. The two category vectors can thus be coupled with failure probability at most q^2 . Couple the k germs independently. The probability of any failure is at most kq^2 . Applying the common offspring rule (2.3) cannot create a disagreement when the category vectors agree, proving (2.4). The entire construction can be carried out independently of H , whose conditional occurrence law is constant. $\square$

This is the elementary rare-Bernoulli mechanism behind Poisson approximation [7]; Poisson marking and mapping provide its general language [6, Chapter 5]. The proof also explains an apparent mismatch: the target permits several category arrivals at a germ, whereas the source permits at most one. Those collisions are part of the error just bounded. Nothing requires pretending that the mutually exclusive source categories were independent.

2.3 Two Poisson marginals whose sum is not Poisson

Write $Z_{A,k} = \sum_i A_i$, $Z_{B,k} = \sum_i B_i$. Summing the independent target categories over germs gives

$$(Z_A, Z_B) = (P_{10} + P_{11}, P_{01} + P_{11}), \quad \begin{cases} P_{10}, P_{01} \sim \text{Pois}(a - b), \\ P_{11} \sim \text{Pois}(b), \end{cases} \quad (2.6)$$

with the three P 's independent. This target law no longer depends on k , so (2.4) also proves total-variation convergence of the count vector to (2.6). Its probability generating function is

$$\mathbb{E}[s^{Z_A} t^{Z_B}] = \exp((a - b)(s - 1) + (a - b)(t - 1) + b(st - 1)).$$

Each marginal is $\text{Pois}(a)$, but $\text{Cov}(Z_A, Z_B) = b$. For the total,

$$Z_A + Z_B = S + 2D, \quad S \sim \text{Pois}(2a - 2b), \quad D \sim \text{Pois}(b), \quad S \perp D. \quad (2.7)$$

This is the classical Hermite family [5]. What matters here is why it appears: a single nonempty category can create two distinct observed points.

The mean number of points is $2a$, whereas the mean number of active target germs is $2a - b$. Their difference b counts the expected extra descendants beyond the first. In particular,

$$\mathbb{P}(Z_A + Z_B = 0) = e^{-(2a-b)}, \quad \text{Var}(Z_A + Z_B) = 2a + 2b. \quad (2.8)$$

A false target $\text{Pois}(2a)$ would match the mean but not the void probability. Testing the void event in the finite source yields

$$\liminf_{k \rightarrow \infty} d_{\text{TV}}(\mathcal{L}(Z_{A,k} + Z_{B,k}), \text{Pois}(2a)) \geq e^{-2a}(e^b - 1). \quad (2.9)$$

This is positive for $b > 0$. The conclusion uses a bounded event, not an inference of moment convergence from total variation.

For completeness, the finite-source calculations are just as direct. Independence between germs and the four-category table give

$$\begin{aligned} \mathbb{P}(Z_{A,k} + Z_{B,k} = 0) &= \left(1 - \frac{2a - b}{k}\right)^k, \\ \text{Cov}(Z_{A,k}, Z_{B,k}) &= b - \frac{a^2}{k}, \quad \text{Var}(Z_{A,k} + Z_{B,k}) = 2a + 2b - \frac{4a^2}{k}. \end{aligned} \quad (2.10)$$

At $a = 1, c = 1/2$, the limiting marginals both have mean one, the covariance is $1/2$, and the total is $\text{Pois}(1) + 2\text{Pois}(1/2)$ with independent summands. At $c = 0$, every hit is double. When $|c| \geq a$, the overlap is zero and the target becomes a simple Poisson field. Even in this last case the finite source pair is exclusive, not independent; its covariance $-a^2/k$ disappears only in the limit.

The deficit also gives an upper bound. The simple Poisson target with the same point intensity would have independent $\text{Pois}(a/k)$ counts at each of x_i and $2x_i$. Starting from $\mathcal{X}_k^P$, keep only the first descendant of every type-11 germ. The anchors now form independent Poisson counts of mean a/k at x_i and $(a - b)/k$ at $2x_i$. Fill the missing second-position intensity by independent $\text{Pois}(b/k)$ counts.

The resulting field has the desired simple Poisson law. It agrees with the cluster field whenever there are no double germs and no filling points. These two independent absence events each have probability e^{-b} . Writing $\mu_k = \frac{a}{k} \sum_i (\delta_{x_i} + \delta_{2x_i})$, we obtain

$$d_{\text{TV}}(\mathcal{L}(\mathcal{X}_k), \text{PPP}(\mu_k)) \leq \frac{(2a - b)^2}{k} + 1 - e^{-2b}. \quad (2.11)$$

Here $\text{PPP}(\mu_k)$ denotes the independent Poisson counts just specified. Together with (2.9), this shows why the point-minus-germ deficit b , rather than the point intensity alone, controls the return to a simple target. The same argument underlies the first-active construction in *Arithmetic clusters* [10, Section 6.4] and the deficit bound in *Moving prime environments* [9, Theorem 3.3].

2.4 A count forgets where the descendants went

Restricting (2.3) to the first band gives a Poisson limiting count of mean a . Observing both bands recovers their covariance b . A prefix-count path therefore retains dependence between disjoint spatial increments when $b > 0$.

Now move every second descendant from $2x_i$ to $3x_i$, in $(3/4, 1)$. The two-family count vector and the total have exactly the same laws as before. The spatial configuration does not: in one construction the second band can contain points; in the other it is always empty. Counts cannot identify this change of geometry.

Finally, the raw displacement c/k tends to zero for every fixed c , including values giving no overlap at all. Its size relative to the shrinking target, not its unscaled limit, determines the cluster law. We will meet precisely this distinction in a moving prime environment. The toy calculation has done its job: it chooses the target and analyzes its observations. It has not supplied the independence of arithmetic germs.

3 What the multiplicative source must supply

3.1 Private coordinates and exact conditional forcing

The uniform translation in (2.1) has a binary counterpart used in *Long runs*. Let V be any finite-valued random variable and let U be independent and uniform on $\mathbb{F}_2^r$. For any map a into $\mathbb{F}_2^r$, set $Y = U + a(V)$. Given V , Y is uniform. To impose $Y = w$, keep V fixed and replace

$$U \quad \text{by} \quad U' = w - a(V).$$

For every v , the conditioning event has probability 2^{-r} . Thus the conditional law of V is unchanged, and (V, U') has exactly the law of (V, U) conditional on $Y = w$. The assertion concerns the whole pair and hence every function of it, not merely the marginal of Y .

For (1.1), write $f(p) = (-1)^{X_p}$ with fair bits X_p . The bit encoding $f(n)$ is a linear expression in the prime bits, modulo two. If each integer in a selected support has a prime of odd valuation dividing no other integer in that support, these bits are private coordinates. Keeping all other bits fixed, one can prescribe the values by fixing the private bits, just as above. The exact conditional law is constructed in the companion to *Long runs* [8, technical companion, Proposition F.7].

The qualification “in that support” matters. A private prime can divide an integer in another window, which may therefore change when the prime is forced. The construction chooses the largest prime of odd valuation, not necessarily the ordinary largest prime factor, and counts the other windows those choices can affect. This directed set need not be a dependency graph. A shared prime which is not selected need not be changed.

There are consequently two arithmetic costs: supports without the required private coordinates, and occurrences affected by the forcing when the coordinates exist. The reciprocal estimates for the selected primes control the second cost; the weighted relation estimate below controls joint occurrences inside it. The classical reciprocal asymptotic is attributed to Erdős–Ivić–Pomerance [3]; that companion supplies both a direct proof adapted to the application and the square-part reduction to that antecedent. The new arithmetic comparison is not supplied by the finite forcing calculation alone.

3.2 Why relations have to be counted with their weights

A left-maximal run imposes L affine binary constraints: one inequality across its left endpoint and $L - 1$ equalities inside it. For two disjoint windows, stack these constraints into $AX = b$. If A has $2L$ rows and relation dimension $\rho = \dim \ker A^\top$, elementary linear algebra gives

$$\mathbb{P}(AX = b) = 2^{-2L} \eta 2^\rho, \quad \left| \mathbb{P}(AX = b) - 2^{-2L} \right| \leq 2^{-2L} (2^\rho - 1), \quad (3.1)$$

where η is one or zero according as the system is compatible or incompatible. Indeed, a compatible system has rank $A = 2L - \rho$ independent conditions. Incompatibility requires a nonzero relation, so the second inequality also holds when $\eta = 0$.

The relations have arithmetic content. Write each run through the L products of consecutive values. Multiplying selected constraints cancels the values that occur twice and leaves an even number of values in each window. A homogeneous relation occurs when the product of their combined integer arguments is a square; compatibility with the prescribed signs is the separate factor η in (3.1) [8, Lemmas 2.1–2.2].

For the windows $x = 100, y = 67, L = 3$ in *Long runs*, the pairs $(99, 66)$ and $(102, 68)$ have ratio $3 : 2$, and

$$99 \cdot 102 \cdot 66 \cdot 68 = (99 \cdot 68)(102 \cdot 66) = 6732^2.$$

The individual affine systems have rank three, but the stack has rank five and is compatible: the joint probability is $1/32$, not $1/64$ [8, Section 3.1].

Knowing only which pairs of windows carry a relation is insufficient: a pair of relation dimension ρ carries $2^\rho - 1$ nonzero relations. At the critical scale, *Long runs* proves that, for each fixed $C_0 > 0$ and $\varepsilon > 0$, uniformly when $|L - \log_2 N| \leq C_0$,

$$\sum_{\substack{x, y \in [N, 2N) \cap \mathbb{Z} \\ |x - y| > L}} (2^{\rho_L(x, y)} - 1) \ll_{C_0, \varepsilon} N^{5/3 + \varepsilon}. \quad (3.2)$$

The probabilities in (3.1) carry the factor $2^{-2L} \asymp N^{-2}$, so these relations contribute $O_{C_0, \varepsilon}(N^{-1/3 + \varepsilon})$. Marginal defects, local overlaps and shared-prime costs are handled separately in that article. Its theorem also gives the macroscopic, varying-length profile needed beyond this critical example [8, Theorem 3.1]; the growing-prefix comparison does not follow from (3.2) alone.

Where the gain comes from. There are two resources: a small number of pairs of windows, or a small relation space on each pair. Exact rational identities are separated first; treating them as accidental prime collisions would lose their organization. Residual relations are then encoded by components linked by large primes. Small dimensions allow many pairs. Large dimensions must force a more rigid configuration, which can be counted using bounded components, defects or a Runge exclusion.

The last case shows why a plausible bound can still be too weak. Restrict the larger start to $[X, 2X)$ and write $B = L + 1 \asymp \log X$. A single small odd kernel can occur in a set of size $X^{3/4+o(1)}$. The gain comes not from shrinking this container but from counting windows in which relations force *several* candidates. The kernel $K_B(n)$ is the squarefree part formed by the primes above B whose valuations in n are odd. Set

$$K_B(n) = \prod_{\substack{p > B \\ v_p(n) \text{ odd}}} p, \quad A_T(x) = \#\{-1 \leq i < L : K_B(x+i) \leq T\}.$$

For $T \ll (XB)^{1/2}$, the shifted energy estimate is

$$\sum_{X \leq x < 2X} \binom{A_T(x)}{2} \ll_{\varepsilon} X^{2/3+\varepsilon} B^2. \quad (3.3)$$

A window containing two candidates contributes at least one, so this bounds the number of such windows. A separate Pell argument bounds the possible partners of a fixed start by $X^{o(1)}$. Where only one small kernel is forced, the relation weight is lower and the larger container is affordable. This is the division in *Long runs*, Lemmas 3.20–3.22 and Proposition 3.23 [8, Section 3.5]. Pell counting at polynomial height is a classical input [2]; the gain in this application comes from how these counts are combined with the relation weights. No optimality of the exponent $5/3$ is asserted.

3.3 A representative arithmetic conclusion, and its boundary

Return to the run indicator

$$J_{x,L} = \mathbf{1}_{\{f(x-1)=-f(x), f(x)=\dots=f(x+L-1)\}}.$$

On a start let ℓ_x be its full rightward length and set $K_{x,e}^{\sigma} = \mathbf{1}_{\{J_{x,L}=1, \ell_x=L+e, f(x)=\sigma\}}$ for $e \geq 0$ and $\sigma = \pm 1$. Only the start is restricted to the block below; the full run may continue beyond it.

Result from the corpus (*Long runs*, Theorem 1.1). *For (1.1), fix $C_0 > 0$. Uniformly over integer lengths $|L - \log_2 N| \leq C_0$, the array on $x \in [N, 2N) \cap \mathbb{Z}$, $e \in \mathbb{N}_0$, $\sigma = \pm 1$ satisfies*

$$d_{\text{TV}} \left(\mathcal{L}((K_{x,e}^{\sigma})_{x,e,\sigma}), \bigotimes_{x,e,\sigma} \text{Pois}(2^{-L-e-2}) \right) \longrightarrow 0. \quad (3.4)$$

In particular $\sum_x J_{x,L}$ is close in total variation to $\text{Pois}(N2^{-L})$ with its exact, possibly nonconvergent parameter.

The result cited is stronger than the displayed convergence: its error is at most $\exp[-(1/\sqrt{2} + o_{C_0}(1))\sqrt{\log N \log \log N}]$. The factor 2^{-e-2} separates into a geometric excess probability 2^{-e-1} and a fair sign. These are genuine arithmetic signs in the

comparison, not colours assigned afterwards. Continuous spatial limits follow weakly after normalizing x by N and choosing a subsequence on which the intensity converges [8, Theorem 1.1].

The left boundary is a different source. Since $f(1) = 1$,

$$\begin{aligned} \{f(1) = \cdots = f(L)\} &= \{f(p) = 1 \text{ for every prime } p \leq L\}, \\ \mathbb{P}(f(1) = \cdots = f(L)) &= 2^{-\pi(L)}. \end{aligned} \quad (3.5)$$

Here $\pi(L)$ counts primes at most L . The equality follows both ways: the initial block contains those primes, and all its integers factor over them. The much harder assertion is that this single event and the bulk account for all long runs at the required precision. If R_M is the longest run fully visible in $[1, M]$, $L \rightarrow \infty$, $L = O(\log M)$, and $M2^{-L} \rightarrow 0$, *Long runs* proves

$$\mathbb{P}(R_M \geq L) = 2^{-\pi(L)} + M2^{-L} + o(2^{-\pi(L)} + M2^{-L}). \quad (3.6)$$

A mere $o(1)$ error would say little about these vanishing probabilities. The relative precision in (3.6), supplied by its boundary localization and relative bulk theorem, is what identifies their competition [8, Theorem 1.2; Section 7.5 and technical companion, Appendix E].

3.4 A sparse prime source can retain a macroscopic intensity

Change now to circle-valued multiplicative functions F , identifying $\mathbb{T}$ with the unit complex numbers by $u \mapsto e^{2\pi i u}$. Its uniform probability is Haar measure. Suppose the prime values are independent with densities g_p relative to uniform probability m , and

$$g_p \geq 0, \quad \int g_p dm = 1, \quad \sup_p \int g_p^2 dm \leq K_2 < \infty.$$

Products can smooth these densities, but a prime site has only one input. Its true rare-event probability is $q_p = \int_{J_{N,p}} g_p dm$, not the uniform mass $a_p = m(J_{N,p})$.

Result from the corpus (*Arithmetic clusters*, Theorem 3.1). *Fix $C_0 \geq 1$, $A > 0$ and K_2 . Let $S_N \subset \{2, \dots, \lfloor C_0 N \rfloor\}$ be any deterministic mask, with deterministic Borel targets $J_{N,n}$ of masses $a_n \leq A/N$. For*

$$\Xi_N = \sum_{n \in S_N} \mathbf{1}_{\{F(n) \in J_{N,n}\}} \delta_{(n, F(n))},$$

the Poisson target has intensity

$$\mu_N = \sum_{\substack{n \in S_N \\ n \text{ composite}}} \delta_n \otimes \mathbf{1}_{J_{N,n}} m + \sum_{\substack{p \in S_N \\ p \text{ prime}}} \delta_p \otimes \mathbf{1}_{J_{N,p}} g_p m. \quad (3.7)$$

For some $\kappa = \kappa(K_2) > 0$, with $Y = (\log N)^4$ and $\mathcal{F}_Y = \sigma(F(p) : p \leq Y)$,

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\Xi_N \mid \mathcal{F}_Y), \text{PPP}(\mu_N)) \ll_{C_0, A, K_2} (\log N)^{-\kappa}.$$

There is no boundedness assumption on the total prime-site intensity.

Here $\text{PPP}(\mu)$ means a Poisson point measure with intensity μ : disjoint sets have independent Poisson counts with means prescribed by μ . Thus the count parameter in this theorem is

$$\lambda_N = \sum_{n \in S_N} a_n - \sum_{p \in S_N} a_p + \sum_{p \in S_N} q_p, \quad (3.8)$$

where the last two sums run only over primes. This is a joint field result, not just an expectation calculation [10, Theorem 3.1].

A useful example takes the same density $g(u) = \log(1/(2|u|))$ at every prime, in angular turns $0 < |u| < 1/2$. Substitution $t = 2|u|$ gives $\int g \, dm = 1$ and $\int g^2 \, dm = 2$. For a centred arc of mass a ,

$$q(a) = 2 \int_0^{a/2} \log \frac{1}{2u} \, du = a(1 + \log(1/a)).$$

At $a = \tau/N$, the $\sim N/\log N$ primes in $[N, 2N)$ contribute asymptotic intensity τ , even though their proportion tends to zero. In the comparison target, the composites supply another τ . The theorem therefore gives $\text{Pois}(2\tau)$, not $\text{Pois}(\tau)$ [10, Sections 1.1 and 3.6]. Deleting a population on the basis of its density in the integers would have erased half the limiting intensity.

The proof needs joint smoothing, not only one-site smoothing. Two independent prime-coordinate groups of rank two supply two L^2 densities on $\mathbb{T}^2$; their convolution is bounded by the product of their L^2 norms. Under the retained-site hypotheses, the failure of two such disjoint bases is classified by ratios $n/m = p^r$. Those exceptional pairs are fewer and are paid with a weaker bound. This is the purpose of *Arithmetic clusters*, Section 2.7 and Lemmas 3.2–3.3. It is an arithmetic verification of marked Chen–Stein comparison [1], not a substitution of a new density into the binary theorem.

3.5 When the dependence belongs in the target

Fix instead a finite set of exceptional primes $\mathcal{E}$ and their phases ξ_p . Let the remaining prime values be independent and uniform. Every integer has a unique decomposition

$$n = sd, \quad s \in \mathcal{D} := \{s \geq 1 : p \mid s \Rightarrow p \in \mathcal{E}\}, \quad (d, \prod_{p \in \mathcal{E}} p) = 1, \quad F(sd) = \chi_\xi(s)F(d), \quad (3.9)$$

where $\chi_\xi(s) = \prod_{p \in \mathcal{E}} \xi_p^{v_p(s)}$. The regular core d is not a squarefree kernel. For a shrinking target J_N , a germ phase u activates all descendants satisfying $\chi_\xi(s)u \in J_N$ and $sd \leq N$. The comparison kernel assigns independent Poisson germs to the regular cores, with all their activated descendants. This is a target construction, not a claim that the original $F(d)$ are independent. The geometry from Section 2 reappears, now with many possible descendants and arithmetically specified rotations.

For $d = 1$, there is no regular randomness left. These environment-only points are removed exactly when defining the corrected count. For instance, if $\mathcal{E} = \{2\}$ and $\xi_2 = 1$, every power of two up to N hits an arc about one. This is not centering the whole field by its expectation: it is removing its exact environment-only subconfiguration.

For fixed $\mathcal{E}$, fixed ξ , a common arc of mass τ/N about one and fixed $\tau > 0$, *Arithmetic clusters* proves that the corrected count has a compound-Poisson limit. It is simple Poisson of mean τ exactly when

$$L_\xi := \{v \in \mathbb{Z}^\mathcal{E} : \xi^v = 1\} = \{0\}.$$

Ordering the multipliers in each equal-phase class gives the jump intensities [10, Theorems 4.1 and 5.1]. This classification fixes the phases, although the uniform kernel comparison already allows them to vary with N .

Moving prime environments classifies limits for arbitrary ξ_N with fixed $\mathcal{E}$. Write W_N for the count on $2 \leq n \leq N$, $n \notin \mathcal{D}$, with independent uniform regular prime values, and retain the same centred arc of mass τ/N . Write $\xi_N^v = \prod_{p \in \mathcal{E}} \xi_{N,p}^{v_p}$. For each nonzero $v \in \mathbb{Z}^{\mathcal{E}}$ define

$$\begin{aligned} A(v) &= \prod_{p \in \mathcal{E}} p^{(v_p)_+}, & B(v) &= \prod_{p \in \mathcal{E}} p^{(-v_p)_+}, & H(v) &= \max(A(v), B(v)), \\ \beta_N(v) &= \left(\tau - N \|\xi_N^v\|_{\mathbb{T}} \right)_+. \end{aligned} \quad (3.10)$$

The circular distance is in turns. The analogue of the toy overlap b is now an entire array $\beta_N(v)$. The reciprocal height weights $1/H(v)$ come from counting the admissible cores for each channel, and their sum is finite. The precise criterion is

$$\begin{aligned} d_{\text{TV}}(\mathcal{L}(W_N), \text{Pois}(\tau)) \rightarrow 0 &\iff \sum_{v \neq 0} \frac{\beta_N(v)}{H(v)} \rightarrow 0 \\ &\iff \liminf_{N \rightarrow \infty} N \|\xi_N^v\|_{\mathbb{T}} \geq \tau \quad \text{for every fixed } v \neq 0. \end{aligned} \quad (3.11)$$

This criterion is proved in [9, Theorem 3.1]. The direct Haar comparison first reaches the moving germ kernel with a uniform error; the overlaps then decide whether that kernel is simple Poisson. A persistent overlap is part of the target, not a failure of the comparison. Summability controls channels of growing height, so the last equivalence does not interchange pointwise and uniform convergence without an estimate.

On subsequences, the finite limits of the enlarged displacements form an integer subgroup L , with an additive displacement map $h : L \rightarrow \mathbb{R}$. These effective profiles describe the marked cluster limits [9, Proposition 5.2 and Theorem 5.4]. Their geometry is visible through the ratios $A(v)/B(v)$, just as the change from $2x_i$ to $3x_i$ was visible in the toy field but not in its total count. The full arithmetic theorem requires both the germ comparison and the control of infinitely many multipliers. Neither was built into our independent-germ experiment.

4 Reading more than a count

4.1 Spatial restriction and marks

In the toy field, a window meeting only the first band sees a simple Poisson limit; the union of both bands sees shared descendants. *Moving prime environments* makes this interaction exact. In a deterministic effective profile (L, h) with interval target of length τ , an interior spatial window (r, s) , $0 < r < s \leq 1$, sees a nontrivial pair channel only if

$$\left| \log \frac{A(v)}{B(v)} \right| < \log \frac{s}{r}, \quad |h(v)| < \tau, \quad v \in L \setminus \{0\}. \quad (4.1)$$

The first condition allows both descendant positions to fit; the second allows their angular activation sets to overlap. The restricted limit is Poisson exactly when no such channel exists [9, Theorems 10.1–10.2]. A cluster present on the whole prefix can therefore disappear from a narrower observation without any change to the source.

More information can also identify more of the law. For deterministic interval profiles, that article recovers the whole unmarked spatial law from its complete spatial covariance measure, by subtracting the known single-point diagonal, separating primitive multiplicative rays and reconstructing finite interval activations. It is an identity from an entire measure in a specified family, not from a few measured covariances. For a known bounded Borel target, the corresponding theorem uses *both* positions and enlarged angular marks, which reveal signed shifts. Arbitrary mixtures of profiles are not identified by averaging those covariance data [9, Theorems 8.4 and 12.2; Section 11]. These restrictions explain why the toy's two scalar marginal laws cannot provide a spatial reconstruction.

4.2 A record needs an order and a recognition rule

A configuration of long runs carries a start, a full length and a sign. Consider the successive run lengths and signs used in the introduction to *Records and scale flows*:

$$4^+, \quad 2^-, \quad 7^+, \quad 7^-, \quad 11^+.$$

Strict records have lengths 4, 7, 11; the second 7 does not replace the first holder. An independent fractional addition to each integer length can change which tied candidate wins. It is a useful interpolation, but a different record observable [11, Section 1.1].

Recognition adds another distinction. Let a run start at integer a_i , have full length q_i , and face a historical maximum H_{i-1} . At physical prefix m its visible contribution is

$$\left(\min(q_i, m - a_i + 1)\right)_+.$$

If $q_i > H_{i-1}$, the episode is recognized as a strict record at $m = a_i + H_{i-1}$. In this list, a length-11 run following a maximum 7 is recognized at its eighth symbol. It gives one record episode, although its visible length raises the maximum through 8, 9, 10, 11.

This is why counting record episodes is not counting every increase of a censored maximum. It also explains a topological distinction in that article: several unit increases can converge to a single larger jump. The scalar M_1 topology allows this approximation through completed jump graphs; a J_1 time change cannot simply remove the intermediate values. The first-holder sign uses the episode order and admits a separate scalar J_1 conclusion [11, Section 4, especially Theorem 4.5]. The common field comparison transfers the input configuration, while the recognition analysis proves which path it represents.

4.3 Two clocks and one explicit arithmetic regime

On an interval of physical prefixes from m_0 to n , write $m = m_0 2^v$ and $T = \log_2(n/m_0)$. Each doubling occupies one unit of the logarithmic clock. The final doubling has logarithmic fraction $1/T$, but asymptotic physical fraction $1/2$ when $m_0/n \rightarrow 0$. Thus averaging a holder's sign in the two clocks need not have the same concentration behaviour. If $Z(v)$ is that sign, the continuous-clock expressions are

$$\frac{1}{T} \int_0^T Z(v) dv, \quad \frac{\log 2}{2^T - 1} \int_0^T 2^v Z(v) dv.$$

The second identity is simply the change of variable $m = m_0 2^v$ in a physical average. Its weight concentrates near the endpoint instead of giving equal weight to all doublings.

To locate a genuine process theorem without introducing every cutoff parameter, here is an unconditional specialization of *Records and scale flows*. In the binary model (1.1), put, for large M ,

$$\begin{aligned} w_M &= \sqrt{\log M \log \log M}, & L_M &= \lfloor \log_2 M \rfloor - \left\lfloor \frac{w_M}{4 \log 2} \right\rfloor, \\ n &= M - L_M, & m_0 &= \lfloor n e^{-w_M/8} \rfloor, & T &= \log_2(n/m_0). \end{aligned} \quad (4.2)$$

Observe the prefixes $\lfloor m_0 2^{Tt} \rfloor + 1$, $0 \leq t \leq 1$. Let $N_M^{\text{rec}}(Tt)$ count the strict global record episodes recognized after the initial prefix and by this prefix. The historical maximum is not reset at $m_0 + 1$; only the counter is reset. Then [11, Theorem 2.4] gives

$$\left(\frac{N_M^{\text{rec}}(Tt) - Tt/2}{\sqrt{T}} \right)_{0 \leq t \leq 1} \implies \left(\frac{1}{2} B(t) \right)_{0 \leq t \leq 1} \quad \text{in } D([0, 1]), \quad J_1, \quad (4.3)$$

where B is standard Brownian motion. There are about $T/2$ recognized record episodes over the whole observation window, with fluctuations of order $\sqrt{T}$ and Brownian coefficient $1/2$. The source theorem also gives the joint signed imbalance and logarithmic leadership limit; (4.3) is only its count projection.

Parameter check (optional on a first reading). Its intensity is $\Lambda = n 2^{-L_M} = \exp(w_M/4 + O(1))$, $T \sim w_M/(8 \log 2)$, and $\Lambda 2^{-T} = \exp(w_M/8 + O(1)) \rightarrow \infty$. The hard scale in *Long runs* satisfies $V_M \sim w_M/\sqrt{2}$ and $\nu_M = \log M/V_M$, so the unconditional budget $\log \Lambda \leq V_M - c\nu_M$ holds for every fixed $c > 0$, eventually. For the shell reconstruction choose $J = \lceil T(1 + T)^2 \log M \rceil$ and $h = T/J$. Then $Th \rightarrow 0$, $m_0 h \rightarrow \infty$, $T = O(V_M)$ and $\log(2 + T + h^{-1}) = O(\log \log M) = o(\nu_M)$. These are exactly conditions (2.7), (2.11)–(2.12) of *Records and scale flows*. Thus (4.3) follows from the cited theorem, not from the bounded-intensity toy model [8, Proposition 4.2 and Theorem 7.7] [11, Theorems 2.3–2.4].

Returning to the two clocks. The same joint history result gives logarithmic positive leadership fraction tending to $1/2$ in probability. By contrast, along a subsequence with $\log_2 n - \lfloor \log_2 n \rfloor \rightarrow \theta$, the physical occupation fraction on the terminal prefix has a nondegenerate limit [11, Theorem 6.1]. That endpoint observation uses a different time interval as well as a different weight. Its law and its separation from the long Brownian history require the terminal-field argument in [11, Section 5]. Counts, record order and occupation are therefore three distinct uses of a common configuration, not three names for one Poisson limit.

5 A short concordance and a choice of route

5.1 Translate the object before translating the letter

These translations concern only the objects encountered above. The four editions are identified in the bibliography; theorem numbers below belong to those editions.

Object	Translation in the source articles
A run start	In [8], $J_{x,L}$ starts at x and uses $x - 1$; an absolute word $I_{x,w}$ starts at $x - 1$. The run label in [11] is $j = x - 1$.
Site and position	The dyadic limit in [8] uses x/N ; [11] uses j/n , $n = M - L$. The source theorem in [10] keeps integer sites. In [9], site sd has position $sd/N = sx$, where $x = d/N$ is the core coordinate. The core is not the site.
Hit probability	2^{-L} is the run-start reference; 2^{-L-e-2} is the exact signed-type target mean. <i>Arithmetic clusters</i> separates nominal mass a_n from true prime probability q_p . Corrected sites in [9] have probability τ/N .
Point intensity	The dyadic target in [8] has mean $N2^{-L}$; the common prefix target of [8,11] has $\Lambda = (M - L)2^{-L}$. The corrected intensity from [10] is given in equation (3.8) of the present article. In [9], $m_N = \tau(1 - B_{\mathcal{E}}(N)/N)$, with $B_{\mathcal{E}}(N) = \mathcal{D} \cap [1, N] $.
Germ intensity	In the toy it is $2a - b$, not $2a$. In [9], Λ_N counts active germs and $D_N = m_N - \Lambda_N$ is the extra-descendant deficit. By contrast, Λ in [11] is a target point intensity.
Marks and clocks	<i>Long runs</i> and <i>Records and scale flows</i> retain full excess and actual sign; the latter distinguishes physical prefix, logarithmic time and excess threshold. [10] keeps phases or descendant configurations; [9] keeps positions and angles enlarged by N .

For the source conventions, see [8, Sections 1.1 and 7.3–7.4], [11, Sections 2.1–2.4], [10, Sections 1.1–1.2] and [9, Section 2].

Quantifiers travel with the object. The critical law (3.4) has bounded positive intensity; the full-prefix coupling and long histories use diverging intensity. The rare boundary result needs relative precision at vanishing intensity. *Arithmetic clusters* averages conditional TV distances; the event-conditioned results of *Long runs* prescribe an information budget and a prime sigma-field. In *Moving prime environments*, phases may move but the exceptional prime set stays fixed.

The empirical count theorem in *Long runs* fixes almost every complete realization, for each fixed parameter pair and its prescribed dyadic scales [8, Corollary 7.8a]. It is not a statement about the particular Liouville function, nor a product-Poisson approximation of the entire local vector after fixing the function; Remark 7.8b proves a support obstruction. These are differences of observable and quantifier, not merely of letters.

5.2 Where to enter a proof

Choose a question rather than read all four articles in order. Links open the version-specific Zenodo records; printed pages locate the results. *Long runs* and *Records and scale flows* include technical companions.

Question	Entry and proof location
Why Poisson for logarithmic binary runs?	<i>Long runs</i> [8]: Theorem 1.1 (p. 3); Section 2: affine events; Theorem 3.1 (p. 13) and Section 3.5: weighted profile and energy; Sections 4–5: comparison.
How can one forced word control the field?	<i>Long runs</i> [8]: Theorem 7.7 (p. 53); companion F.4, Proposition F.7 (p. 33) for the conditional law, then F.5 for its finite budget. The scalar tail has an earlier independent proof.
What about the boundary or a fixed realization?	<i>Long runs</i> [8]: Theorem 1.2 (p. 4) and Section 7.5, companion E: boundary. Corollary 7.8a (p. 56) and Remark 7.8b: empirical count and local-field limitation.
Which sources and relations must be retained?	<i>Arithmetic clusters</i> [10]: Theorem 3.1 (p. 12) and Section 3.3: corrected sources and joint smoothing; Theorems 4.1 (p. 19) and 5.1 (p. 22): environment kernel and fixed-phase count law.
What survives moving phases or restriction?	<i>Moving prime environments</i> [9]: Theorem 3.1 (p. 12): criterion; Proposition 5.2 and Theorem 5.4 (p. 20): profiles; Theorems 8.2–8.4 (pp. 28–30): reconstruction; Sections 10 and 12: windows and full marks.
How does a field produce a history?	<i>Records and scale flows</i> [11]: Section 2 (p. 7): interface; Sections 3–4 and companion B (p. 6): records and recognition; Sections 5–6: history–endpoint law and occupation clocks.

For prescribed-word dictionaries, *Long runs* distinguishes uniform and typical quantifiers in Theorems 5.1 and 5.3. The long-pattern branch of *Arithmetic clusters*, Sections 8–9, reweights and caps those source populations rather than using just the final binary theorem. The spatial Gaussian hierarchy of *Records and scale flows*, Section 7.6 and companion F, has separate tail and joint-limit arguments. The finite-colour bridge identifies a target, not an arithmetic record theorem for every alphabet.

Long runs also supplies the binary input to *Records and scale flows*. The source and environment theorems of *Arithmetic clusters* have local proofs; *Moving prime environments* proves its Haar comparison directly, with *Arithmetic clusters* as the kernel and fixed-environment antecedent. None needs this exposition for a missing definition or proof.

What remains after the calculation. Ask what one active source unit produces, which dependence survives at the target’s resolution, and what the observation records. The toy model answers all three questions explicitly. In the arithmetic models, answering them chooses the target; the comparison and the analysis of its observations are then separate proofs. The surviving dependence is not always an error to eliminate. Sometimes it is exactly the structure the limit should retain.

Sources and assistance

The source archive contains this exposition, its bibliography, finite diagnostics and proof locators. Research results are cited in the public editions below. *Long runs* documents a companion Lean development relative to explicit literature inputs [8, pp. 64–65]; it does not formally verify this exposition or the other three articles.

The author organized and directed this work with substantial assistance from GPT-6 Astra (OpenAI), the only language-model system used for this article, for exposition, elementary calculations, source checks, diagnostics and file preparation. The author determines the claims and accepts responsibility for the text, references and remaining errors. AI-assisted checks and finite tests are not independent human peer review and do not replace the proofs.

References

- [1] Richard Arratia, Larry Goldstein and Louis Gordon. *Poisson approximation and the Chen–Stein method*. Statistical Science 5(4): 403–424. 1990. doi: [10.1214/ss/1177012015](https://doi.org/10.1214/ss/1177012015).
- [2] J. Cilleruelo and M. Z. Garaev. *Concentration of points on two and three dimensional modular hyperbolas and applications*. Geometric and Functional Analysis 21: 892–904. 2011. doi: [10.1007/s00039-011-0127-6](https://doi.org/10.1007/s00039-011-0127-6). arXiv: [1007.1526](https://arxiv.org/abs/1007.1526).
- [3] P. Erdős, A. Ivić and C. Pomerance. *On sums involving reciprocals of the largest prime factor of an integer*. Glasnik Matematički, Series III 21(41): 283–300. 1986. <https://math.dartmouth.edu/~carlp/PDF/54.pdf>.
- [4] L. Gordon, M. F. Schilling and M. S. Waterman. *An extreme value theory for long head runs*. Probability Theory and Related Fields 72(2): 279–287. 1986. doi: [10.1007/BF00699107](https://doi.org/10.1007/BF00699107).
- [5] C. D. Kemp and A. W. Kemp. *Some properties of the ‘Hermite’ distribution*. Biometrika 52(3–4): 381–394. 1965. doi: [10.1093/biomet/52.3-4.381](https://doi.org/10.1093/biomet/52.3-4.381).
- [6] Günter Last and Mathew Penrose. *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press. 2017. doi: [10.1017/9781316104477](https://doi.org/10.1017/9781316104477). <https://stoch.math.kit.edu/img/Last/lastpenrose2017.pdf>.
- [7] Lucien Le Cam. *An approximation theorem for the Poisson binomial distribution*. Pacific Journal of Mathematics 10(4): 1181–1197. 1960. doi: [10.2140/pjm.1960.10.1181](https://doi.org/10.2140/pjm.1960.10.1181).
- [8] Brice Pouly. *Long runs and rare patterns of a random completely multiplicative function*. 2026-09-21. Preprint, Version 3. Zenodo. With technical companion. doi: [10.5281/zenodo.22872154](https://doi.org/10.5281/zenodo.22872154).
- [9] Brice Pouly. *Moving prime environments and arithmetic cluster limits*. 2026-09-21. Preprint, Version 1. Zenodo. doi: [10.5281/zenodo.22878806](https://doi.org/10.5281/zenodo.22878806).
- [10] Brice Pouly. *Poisson laws and arithmetic clusters for rare multiplicative patterns*. 2026-09-21. Preprint, Version 1. Zenodo. doi: [10.5281/zenodo.22877400](https://doi.org/10.5281/zenodo.22877400).
- [11] Brice Pouly. *Records, scale flows, and Gaussian limits from lattice Poisson fields*. 2026-09-21. Preprint, Version 2. Zenodo. With technical companion. doi: [10.5281/zenodo.22876520](https://doi.org/10.5281/zenodo.22876520).