

# Poisson laws and arithmetic clusters for rare multiplicative patterns

Brice Pouly

Version 1.1

*This research received substantial assistance from AI systems;  
see the declaration at the end of the article.*

## Abstract

We determine when rare observations of completely multiplicative functions have a Poisson law and when arithmetic relations force clusters. For independent prime values on the circle with uniformly bounded  $L^2$  densities, we approximate the field of hits of site-dependent targets of Haar mass  $O(N^{-1})$  on deterministic masks in  $[2, CN]$ . The target retains the true prime-site intensities and uses Haar at composite sites, even when the total prime contribution diverges. The exponent 2 is sharp in the uniform  $L^q$  scale for this target and class of masks. Below this threshold, even correct prime intensities can fail at site resolution while the total count remains asymptotically Poisson. A fixed finite exceptional prime environment, possibly atomic and internally dependent, instead produces a conditional cluster law together with an exact environment-determined configuration. With Haar outside that environment, its integer relation lattice determines the compound-Poisson limit of the corrected prefix count and characterizes simple Poisson convergence. Further applications treat rational observations on sparse moving supports, permanent bounded prime biases, and—using a binary arithmetic relation profile—long Borel patterns in a finite abelian group times a torus. Classical marked Chen–Stein comparison provides a common interface. A separate deterministic-sequence interface gives correlation identities, transfer criteria and a resolution obstruction, without proving cancellation for Liouville.

**Keywords.** Random multiplicative functions; shrinking targets; Poisson approximation; rare patterns; marked fields; prime-value densities; relation lattices; compound Poisson laws.

**MSC 2020.** 11N64; 60F05; 60C05; 11D45; 13P10.

## Contents

|    |                                                                 |    |
|----|-----------------------------------------------------------------|----|
| 1  | Rare hits, prime sources and arithmetic clusters                | 3  |
| 2  | Conditional comparison and local smoothing                      | 7  |
| 3  | Shrinking targets under $L^2$ : retaining the prime source      | 11 |
| 4  | Finite prime environments and multiplicative clusters           | 18 |
| 5  | The exceptional-phase lattice determines the limiting count law | 21 |
| 6  | Shrinking targets of rational observations                      | 27 |
| 7  | Rational observations under permanent prime bias                | 31 |
| 8  | Borel patterns and coordinate-erasure cost                      | 34 |
| 9  | Dictionaries, quotients, cells and filtered relations           | 38 |
| 10 | Deterministic sequences: visible packets and spatial resolution | 42 |
| 11 | Scope, consequences and further questions                       | 46 |
| A  | One-window defects and deletion at a moving cutoff              | 48 |
| B  | Growing-order Bonferroni and abundant private primes            | 51 |
| C  | Arithmetic dependencies and the variable-base ledger            | 53 |
| D  | Finite identities for visible packets and aggregation           | 56 |
|    | Availability and declarations                                   | 61 |

# 1 Rare hits, prime sources and arithmetic clusters

A Poisson law for rare occurrences does not require the entire underlying sequence to resemble independent observations. It requires control of the information that those occurrences can detect. For a completely multiplicative function,  $F(ab) = F(a)F(b)$  persists at every scale. Some of its consequences disappear from rare-event statistics; others must remain in the target law. This paper distinguishes these situations quantitatively.

Two contrasts organize the main results. A population of integers of density zero may carry a nonvanishing, or even diverging, expected number of rare hits: the prime sites must then retain their own intensity. Conversely, perfectly uniform one-site marginals may hide exact relations that create clusters: the target must then retain more than an intensity correction. The theorems below treat these two phenomena before we turn to rational observations and long patterns. Their proofs are given locally in Sections 3–5; the external binary population estimates enter only the long-pattern branch.

## 1.1 A sparse population can carry a substantial rare-event mass

Let  $m$  be Haar probability on  $\mathbb{T}$ . Choose the prime values independently, with densities  $g_p$ , and extend  $F$  completely multiplicatively. Fix  $C \geq 1$ ,  $A > 0$ , a deterministic mask  $S_N \subseteq \{2, \dots, \lfloor CN \rfloor\}$ , and deterministic Borel targets  $J_{N,n}$  of masses  $a_n = m(J_{N,n}) \leq A/N$ . Record both positions and exact phases:

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

The comparison keeps the true prime-site law but uses Haar at composite sites:

$$\mu_N = \sum_{\substack{n \in S_N \\ n \text{ composite}}} \delta_n \otimes \mathbf{1}_{J_{N,n}} m + \sum_{\substack{p \in S_N \\ p \text{ prime}}} \delta_p \otimes \mathbf{1}_{J_{N,p}} g_p m. \quad (1.1)$$

**Theorem A** (Shrinking targets with a prime-source correction). *Suppose  $g_p \geq 0$ ,  $\int g_p dm = 1$  and  $\sup_p \int g_p^2 dm \leq K_2 < \infty$ , with  $K_2$  fixed. There exists  $\kappa = \kappa(K_2) > 0$  such that, with  $Y = (\log N)^4$  and  $\mathcal{F}_Y = \sigma(F(p) : p \leq Y)$ ,*

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

*The bound is uniform over the masks, targets and prime densities satisfying these hypotheses. No boundedness assumption is made on the total prime-site intensity. An explicit range of exponents and the proof are given in Theorem 3.1.*

Write  $Q_N = \sum_{p \in S_N} \int_{J_{N,p}} g_p dm$ , with the sum here and below restricted to primes, and put  $\Lambda_N = \sum_{n \in S_N} a_n$ ,  $\Lambda_N^p = \sum_{p \in S_N} a_p$ . The count has moving Poisson parameter

$$\lambda_N = \Lambda_N - \Lambda_N^p + Q_N.$$

The entirely Haar target is recovered if and only if  $Q_N \rightarrow 0$ . This is not merely a first-moment correction: the theorem controls the joint field, including prime–composite interactions and exceptional pairs of composite sites. The parameter may diverge and is retained exactly in total variation.

For a concrete contrast, take at every prime the density  $g(u) = \log(1/(2|u|))$  in angular turns,  $0 < |u| < 1/2$ , and an arc about 1 of mass  $\tau/N$ . At a prime its probability is  $\tau N^{-1}(1 + \log(N/\tau))$ . On  $[N, 2N)$ , the primes alone contribute asymptotic intensity  $\tau$ ; the composites contribute another  $\tau$ . The resulting count tends to  $\text{Pois}(2\tau)$ , not  $\text{Pois}(\tau)$ . A quadratic logarithmic cusp instead gives a diverging prime contribution; both examples are proved in Section 3. The

exponent 2 is sharp in the uniform  $L^q$  scale for Theorem A's particular target and arbitrary-mask conclusion: Proposition 3.5 gives the prime-square obstruction for each  $1 \leq q < 2$ . A second obstruction, Proposition 3.6, concerns resolution rather than the choice of intensities: on a prime-only mask, the site-labelled field can be at total-variation distance tending to one from the Poisson field with its true intensities, while its total count is close to Poisson with the exact same total mean. Both counterexamples lie below the  $L^2$  hypothesis, not within Theorem A.

## 1.2 Uniform marginals need not give Poisson occurrences

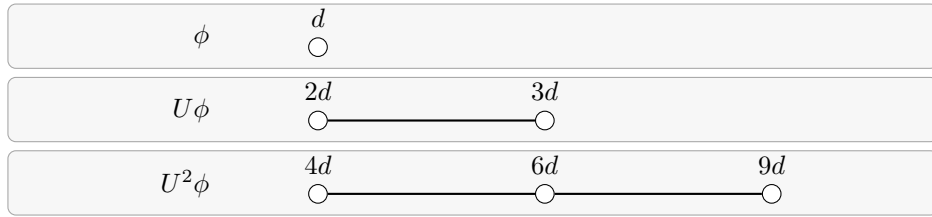
Let  $U$  be Haar on  $\mathbb{T}$ , put  $F(2) = F(3) = U$ , and choose every other prime value independently with Haar law and independently of  $U$ . Each  $F(n)$ ,  $n \geq 2$ , is individually Haar. Nevertheless, for the same shrinking arc,

$$\sum_{n=2}^N \mathbf{1}_{\{F(n) \in J_N\}} \xrightarrow{d_{\text{TV}}} \sum_{i=1}^P G_i, \quad P \sim \text{Pois}(2\tau/3), \quad \mathbb{P}(G_i = j) = \frac{2}{3} \left(\frac{1}{3}\right)^{j-1}, \quad (1.2)$$

where  $P, G_1, G_2, \dots$  are independent. The identities  $F(2d) = F(3d)$  and  $F(4d) = F(6d) = F(9d)$  create clusters of hits, as illustrated in Figure 1. Their geometric compound-Poisson distribution is the classical Pólya–Aeppli law [10]; the arithmetic relation determines its parameter. Exceptional-prime-only sites have vanishing activation mass in this particular random example, so no correction remains in (1.2).

One core  $d$ , with  $(d, 6) = 1$

$F(2) = F(3) = U$ ,  $\phi = F(d)$



Equal phases within a row: one hit may activate several sites.

Figure 1: An arithmetic cluster mechanism. The displayed sites in each row have exactly the same phase; only sites inside the observed prefix are retained. Different rows are not asserted to be independent in the original model. The diagram shows identities, not simulated data; the cluster law follows from the conditional germ comparison and the lattice calculation.

More generally, fix a finite exceptional set  $\mathcal{E}$  and phases  $\xi = (\xi_p)_{p \in \mathcal{E}}$ . Let  $\mathcal{D}_{\mathcal{E}}$  be the integers supported on  $\mathcal{E}$ , and write

$$n = sd, \quad s \in \mathcal{D}_{\mathcal{E}}, \quad (d, \prod_{p \in \mathcal{E}} p) = 1, \quad F(sd) = \chi_{\xi}(s)F(d), \quad \chi_{\xi}(s) = \prod_{p \in \mathcal{E}} \xi_p^{v_p(s)}.$$

The *core*  $d$  carries the remaining randomness. For  $d = 1$ , all values are determined by  $\xi$  and must be retained exactly. Under independent  $L^2$ -bounded densities outside  $\mathcal{E}$ , Theorem 4.1 constructs a conditional target of Poisson germs whose marks are all the hit sites  $sd$  and their phases. Prime cores retain their true density even at composite sites  $sd$ . The comparison is uniform in fixed  $\xi$  and therefore also permits an arbitrary joint exceptional law independent of the regular prime source.

To calculate a fixed limiting count law, assume from now on that the prime values outside  $\mathcal{E}$  are Haar. The integer lattice

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

partitions  $\mathcal{D}_{\mathcal{E}}$  into classes of equal phase. In each class  $\mathcal{A}$ , list the multipliers as  $s_{\mathcal{A},1} < s_{\mathcal{A},2} < \dots$ , padding finite lists with  $\infty$  and using  $1/\infty = 0$ . Put  $\delta_{\mathcal{E}} = \prod_{p \in \mathcal{E}} (1 - p^{-1})$  and

$$\lambda_j(\xi) = \tau \delta_{\mathcal{E}} \sum_{\mathcal{A}} \left( \frac{1}{s_{\mathcal{A},j}} - \frac{1}{s_{\mathcal{A},j+1}} \right), \quad j \geq 1. \quad (1.3)$$

**Theorem B** (A finite environment determines the cluster law). *Fix  $\mathcal{E}$  and  $\xi$ , take independent Haar prime values outside  $\mathcal{E}$ , and let  $J_N$  be the arc about 1 of mass  $\tau/N$ ,  $\tau > 0$ . Define*

$$B_N(\xi) = \sum_{s \in \mathcal{D}_{\mathcal{E}} \cap [2, N]} \mathbf{1}_{\{\chi_{\xi}(s) \in J_N\}}.$$

*Then the corrected prefix count satisfies*

$$\sum_{n=2}^N \mathbf{1}_{\{F(n) \in J_N\}} - B_N(\xi) \xrightarrow{d_{\text{TV}}} \sum_{j \geq 1} j P_j, \quad P_j \sim \text{Pois}(\lambda_j(\xi)) \text{ independently.}$$

*The jump intensities obey  $\sum_j j \lambda_j(\xi) = \tau$ . The limit is  $\text{Pois}(\tau)$  if and only if  $L_{\xi} = \{0\}$ . The prelimit environment kernel is proved in Theorem 4.1; its fixed-environment classification is Theorem 5.1.*

The subtraction is an exact configuration count, not centering by an expectation. The lattice is an integer subgroup, possibly nonsaturated, not merely a rational vector space. It determines the count law, but the converse identification is not asserted. The limit fixes  $\xi$  before  $N \rightarrow \infty$ ; near-relations in an  $N$ -dependent environment remain part of the prelimit cluster kernel and do not carry a uniform limit rate from this classification.

### 1.3 The observation decides which relations matter

The preceding theorems concern one value  $F(n)$  in a shrinking set. A rational observation  $\prod_{j=1}^{k_N} F(x + s_{N,j})^{e_{N,j}}$  can likewise be rare because its product enters one target of mass  $O(N^{-1})$ , while its individual coordinates are unconstrained. Theorem 6.1 handles sparse moving supports of diameter  $O(N)$  and complete sampled-value marks. A deficit measuring repeated activations within equality/inversion classes characterizes the return to simple Poisson. Theorem 7.2 treats permanent independent prime biases by local relative mixing. Neither proof uses the deep binary population estimates.

A long pattern has a different source of rarity: it accumulates many constraints of fixed probability. For Haar values in a fixed finite abelian group times a torus, Theorem 8.4 bounds the cost of erasing coordinates of a Borel event. This branch uses the binary profile in *Long runs and rare patterns* [26]: changing the base of its relation weight is distinct from changing the characteristic, and integer minors provide the second step. Quotients and character filtering may ignore relations that the event cannot see, while its full ambient values remain in the marks.

These applications share classical marked Chen–Stein comparison, with explicit deletion and conditional-marginal errors. Their arithmetic inputs are different. For deterministic signs observed from a random origin, Section 10 instead identifies signed correlation packets at scalar or spatially aggregated resolution and proves a finite-source support obstruction. It does not estimate the remaining Liouville packets.

### 1.4 Relation to earlier work and to process observables

Marked Poisson-process approximation and the folding of error terms over types are classical, especially in Arratia–Goldstein–Gordon [3,4]. The contribution here is the arithmetic verification of that comparison and the identification of the source-corrected or clustered target.

Najnudel [24] studies the character criterion, quantitative independence inside a block and empirical measures for general i.i.d. prime laws. Turk’s stronger threshold [36], in the formulation recalled in [24, arXiv version 2, Remark 2.3], is

$$n \geq \exp\left(ck \frac{\log \log(k+2)}{\log(k+1)}\right).$$

It already gives independence of a Steinhaus block for  $k \leq D \log n$ , with fixed  $D$  and large  $n$ . The comparison here concerns many moving occurrences, their marks and interactions, not independence inside one logarithmic block. We use Turk’s historical formulation as reported by Najnudel. Empirical convergence against fixed tests does not control a target of mass  $N^{-1}$ .

There is also an arithmetic antecedent for consecutive sign chains: Basak–Nath–Zaharescu [8] prove Poisson behavior for chains of quadratic non-residues in initial ranges of size  $(\log p)^A$ ,  $A > 2$ , for almost all primes  $p$  in short intervals. This concerns a family of quadratic characters as the modulus varies, rather than the independent prime-value models studied here. The comparison recorded here uses their published abstract; it does not assert an equivalence or a separation of the full growth regimes of chain lengths and the long-pattern theorem below.

Poisson and compound-Poisson approximations for several rare words in Markov chains were developed by Reinert–Schbath [29] and Roquain–Schbath [30]. Word labels, local overlaps and clustering are not new probabilistic principles. Here local overlap is separated from long-range multiplicative dependence; the marker construction is attributed directly to cross-bifix-free coding [9].

Change-of-measure stability belongs to the mixing-limit principles of Eagleson [12,15]; Kakutani [19] supplies the context of infinite-product equivalence and singularity. Our direct conditional argument states its full-prime-sigma-field measurability assumption. Permanent biases may be globally singular relative to Haar and require the separate local reconstruction.

Hochman–Paviato [18, Theorem 1.1] study a different form of robustness: Poisson genericity for independent letters with nonstationary biases  $\gamma_n = O((\log n)^{-c})$ . For  $c > 1/2$  almost every sequence is simply Poisson generic, while slower-decaying biases give counterexamples. This includes product measures singular relative to the uniform product. In their basic statistic a sequence is fixed and the searched word is sampled uniformly. Their decay threshold is not the  $L^2$  threshold of Theorem 3.1: our source correction and exceptional relation lattices concern multiplicative values, not independent letters.

Compound-Poisson return laws and spatio-temporal small-target processes also occur in dynamics [17,20,28,25]; they are conceptual antecedents, not proof inputs for our arithmetic estimates. Likewise, monomial and lattice ideal theory is classical [5,13,23]; Section 5.4 identifies the initial ideal whose evaluation gives the cluster parameters.

The long-pattern branch uses the population estimates of *Long runs and rare patterns* [26]; Section C specifies each estimate before reweighting and capping. Its finite-colour geometric target also connects with the process model in *Records, scale flows, and Gaussian limits* [27]. The target identification is given in Section 9.3, and its process-level limitations are discussed in Section 11.4.

Adjacent work on polynomial square-product counts [11], normalized short-interval sums [16], unusual factorization anatomy [34] and dominant-chaos Poisson criteria [37] concerns different observables or spectral hypotheses. None is used as a black-box proof of the shrinking-target or environment theorems. For deterministic signs, [22,33] gives context but not the growing even correlation packets required here.

## 1.5 Reading guide and proof responsibilities

The central results are the source theorem and the environment theorems, not the general transfer inequality. The following paths separate their conclusions from the extensions and from the deterministic interface.

| Reading path                   | Sections and role                                                                                                                                                |
|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sources and environments       | Sections 3 to 5: prime correction, conditional cluster kernels and fixed-environment count laws. These are the main results.                                     |
| Extensions by observation      | Sections 6 and 7: sparse rational probes, without the deep binary profile.                                                                                       |
| Deterministic interface        | Sections 8 and 9: long Borel patterns, using that profile.                                                                                                       |
| Common tools and proof records | Section 10: exact packets, sufficient transfer criteria and a resolution obstruction. No Liouville cancellation is imported from the random models.              |
|                                | Section 2: conditional comparison and smoothing. Appendices A–D: defect estimates, prime abundance, the imported arithmetic ledger and finite packet identities. |

Readers interested in the two main contrasts may use the tools in Section 2 and proceed directly through Sections 3 to 5; the later branches are not prerequisites for those results.

The source register in Section C.1 identifies the precise arithmetic inputs to the pattern branch. All logarithms without a base are natural. Total variation compares a common prelimit space; a diffuse spatial limit is a subsequent weak limit, not a total-variation comparison with a lattice.

## 2 Conditional comparison and local smoothing

The two main results use the same three inputs: the mass of discarded hits, a relative approximation of conditional marked marginals, and pair bounds on a genuine conditional dependency graph. We state the common comparison first, then the smoothing and prime-factor estimates needed for the prime-source theorem. Throughout,  $v_p(n)$  is the exponent of the prime  $p$  in  $n$ , and  $\omega(n)$  is the number of distinct prime factors. On  $\mathbb{T}$ ,  $m$  denotes Haar probability, and  $\hat{g}(j) = \int z^{-j} g(z) dm(z)$ .

### 2.1 Measure conventions

All mark spaces are standard Borel. For a finite measure  $\nu$ ,  $\text{PPP}(\nu)$  denotes the *law* of the Poisson point measure of intensity  $\nu$ . For random  $\nu$  it denotes a conditional kernel, not a Poisson law with intensity  $\mathbb{E}\nu$ . We use

$$d_{\text{TV}}(P, Q) = \sup_A |P(A) - Q(A)| = \frac{1}{2} \|P - Q\|_{\text{var}}.$$

For subprobabilities of different masses, the latter norm is the total variation norm of a signed measure. A measurable map contracts  $d_{\text{TV}}$ ; recording less information costs no new estimate if the same map is applied to the target.

**Lemma 2.1** (Comparing intensities). *For finite measures on the same space,*

$$d_{\text{TV}}(\text{PPP}(\nu), \text{PPP}(\tilde{\nu})) \leq 1 - e^{-\|\nu - \tilde{\nu}\|_{\text{var}}} \leq \|\nu - \tilde{\nu}\|_{\text{var}}.$$

*Proof.* Use one Poisson process of intensity  $\nu \wedge \tilde{\nu}$ , taking the minimum of the densities relative to  $\nu + \tilde{\nu}$ , and independent processes for the remaining measures. Both configurations agree if the remainders are empty. The sum of their masses is the stated norm.  $\square$

## 2.2 The common inequality

On a finite index set  $I$ , an event  $A_x$  carries a mark  $M_x$  and defines  $\Xi = \sum_{x \in I} \mathbf{1}_{A_x} \delta_{(x, M_x)}$ . Fix deterministic subprobabilities  $\nu_x$  of masses  $a_x$ , and the target PPP( $\mu$ ) for  $\mu = \sum_x \delta_x \otimes \nu_x$ . On a deterministic  $G \subseteq I$ , assume a deterministic simple undirected graph  $\mathcal{G} = (G, E(\mathcal{G}))$  is an exact dependency graph for the *complete marked variables*, conditionally on a sigma-field  $\mathcal{F}$ : collections with no edge between them are conditionally independent. Use its closed neighborhoods

$$B_x = \{x\} \cup \{y \in G : \{x, y\} \in E(\mathcal{G})\}, \quad x \in G.$$

Thus every neighbor is included, not only the site  $x$  itself.

Write

$$\nu_x^{\mathcal{F}}(E) = \mathbb{P}(A_x, M_x \in E \mid \mathcal{F}), \quad p_x^{\mathcal{F}} = \nu_x^{\mathcal{F}}(\mathbf{M}),$$

and define

$$\mathcal{B} = \sum_{x \notin G} (\mathbb{P}(A_x) + a_x), \quad (2.1)$$

$$\mathcal{M} = \sum_{x \in G} \mathbb{E} \|\nu_x^{\mathcal{F}} - \nu_x\|_{\text{var}}, \quad (2.2)$$

$$b_1(\mathcal{F}) = \sum_{x \in G} \sum_{y \in B_x} p_x^{\mathcal{F}} p_y^{\mathcal{F}}, \quad b_2(\mathcal{F}) = \sum_{x \in G} \sum_{y \in B_x \setminus \{x\}} \mathbb{P}(A_x \cap A_y \mid \mathcal{F}). \quad (2.3)$$

**Theorem 2.2** (Marked interface with approximate marginals). *Under these assumptions, with the closed neighborhoods  $B_x$  defined above,*

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\Xi \mid \mathcal{F}), \text{PPP}(\mu)) \leq \mathcal{B} + \mathcal{M} + 2\mathbb{E}b_1(\mathcal{F}) + 2\mathbb{E}b_2(\mathcal{F}). \quad (2.4)$$

Moreover  $\mathbb{E}b_2(\mathcal{F}) = \sum_{x \in G} \sum_{y \in B_x \setminus \{x\}} \mathbb{P}(A_x \cap A_y)$ . The constants do not depend on the number or resolution of marks.

*Proof.* Deleting real and target sites outside  $G$  costs  $\mathcal{B}$  after averaging. On each conditional fibre, partition the mark space into finitely many cells  $E_i$  and apply [4, Theorem 2] to

$$J_{x,i} = \mathbf{1}_{A_x} \mathbf{1}_{\{M_x \in E_i\}}, \quad B_{x,i} = B_x \times \{\text{partition cells}\}.$$

Marks at one site are exclusive. Summing over cells recovers exactly  $b_1, b_2$ . The third Chen–Stein term vanishes because each  $B_x$  contains all graph neighbors and the graph controls the complete marks. With our convention, half the norm in that reference, the process bound is  $2(b_1 + b_2)$ .

The first target has intensity  $\sum_{x \in G} \delta_x \otimes \nu_x^{\mathcal{F}}$ . Lemma 2.1 changes it to the deterministic reference at cost at most  $\sum_x \|\nu_x^{\mathcal{F}} - \nu_x\|_{\text{var}}$ .

For standard Borel marks take increasing finite partitions generating the sigma-field. Their counts generate the sigma-field of finite configurations. Total variation on increasing sigma-fields converges to full total variation, by convergence of conditional expectations of the two densities relative to their sum. This holds for the conditional kernels; monotone convergence permits averaging. The index set and graph are deterministic, so the tower property gives the final identity.  $\square$

The mark-folding argument is already explicit in [4, Section 3.1]; see also [3]. The theorem organizes its conditional use, deletions and intensity replacement. The arithmetic tasks are to produce the good set, graph, marginal subprobabilities and pair estimates.

### 2.3 Exact marginals, caps and signed sums

If  $\nu_x^{\mathcal{F}} = \nu_x$ , then  $\mathcal{M} = 0$ . Use the distinct ordered edges  $\mathcal{E}_G = \{(x, y) : x \in G, y \in B_x \setminus \{x\}\}$ , so that each undirected edge contributes both orientations. Put

$$\mathcal{V} = \sum_G a_x^2, \quad \mathcal{Q} = \sum_{(x,y) \in \mathcal{E}_G} a_x a_y, \quad \mathcal{C} = \sum_{(x,y) \in \mathcal{E}_G} (\mathbb{P}(A_x \cap A_y) - a_x a_y).$$

Then  $\mathbb{E}b_1 = \mathcal{V} + \mathcal{Q}$  and  $\mathbb{E}b_2 = \mathcal{Q} + \mathcal{C}$ , hence

$$\mathbb{E}d_{\text{TV}} \leq \mathcal{B} + 2\mathcal{V} + 4\mathcal{Q} + 2\mathcal{C}. \quad (2.5)$$

The signed sum  $\mathcal{C}$  may be retained until a positive bound is needed. It does not allow dropping  $\mathcal{Q}$  when pairwise independence holds after averaging.

For an indicator of activation probability  $a$ , if  $\mathbb{P}(A_x \cap A_y) \leq a^2(1 + K(x, y))$ ,  $K(x, y) \geq 0$ , the marginal ceiling gives

$$\mathbb{P}(A_x \cap A_y) \leq a^2[1 + \min\{a^{-1} - 1, K(x, y)\}]. \quad (2.6)$$

Apply this before summing a dictionary of mutually exclusive words. It retains all labels without paying an uncapped relation weight.

Relative local approximation serves another purpose. If a density satisfies  $\|h_x^{\mathcal{F}} - 1\|_{\infty} \leq \delta_N$ , restricting to an event of reference mass  $a_x$  gives

$$\|\nu_x^{\mathcal{F}} - \nu_x\|_{\text{var}} \leq a_x \delta_N.$$

The total is  $\delta_N \sum_x a_x$ , not  $|G|\delta_N$ . A per-block absolute  $o(1)$  error is insufficient at the rare-event scale.

### 2.4 Clusters and retained environments

A unit may produce a whole configuration. If a germ  $u$  at index  $d$  creates a measurable cluster  $\mathcal{G}_d(u)$ , transfer the germ process and then apply

$$\eta \mapsto \sum_d \int \mathcal{G}_d(u) \eta_d(du).$$

Contraction does not multiply the error by cluster size. Internal cluster dependence remains in the target.

For an environment  $\xi$  fixing targets, rotations or an exact configuration, work for each  $\xi$  with a target kernel  $Q_N(\xi)$ . A uniform bound integrates to a joint comparison retaining that environment. Mixtures of Poisson or cluster laws cannot generally be replaced by a Poisson law of averaged intensity.

The mark assumption matters: independent activations of probability  $1/N$ , all given one random colour chosen once, have a count close to  $\text{Pois}(1)$  but cannot approximate independently coloured Poisson points. The latter displays both equiprobable colours with probability  $(1 - e^{-1/2})^2 > 0$ , whereas the former never does.

### 2.5 Private primes and the conditional graph

In a window of diameter less than  $B$ , a prime  $p > B$  divides at most one integer, so its coordinate is private. Valuation one supplies a Haar coordinate in every compact abelian group. On a torus any nonzero valuation suffices: every nonzero integer power pushes Haar to Haar.

When the prime coordinates are independent, after exposing  $\mathcal{F}_Y = \sigma(F(p) : p \leq Y)$  a site only depends on unexposed coordinates  $p > Y$  dividing its observed integers. Collections with

disjoint such supports are conditionally independent, giving an exact graph for complete values and then their observations. Pairwise independence after averaging would not suffice.

For  $k$  positions  $0 \leq s_1 < \dots < s_k \leq CN$ , a prime meets at most  $k(N/p + 1)$  sites  $x \in [N, 2N)$ . Squared incidences give the distinct ordered edge count

$$E_Y \ll_C k^2 \left( \frac{N^2}{Y \log Y} + N \log \log(3N) + \frac{N}{\log(3N)} \right). \quad (2.7)$$

For one-coordinate observations in a prefix, floor functions give

$$E_Y \leq \sum_{p > Y} \lfloor M/p \rfloor^2 \ll \frac{M^2}{Y \log Y}. \quad (2.8)$$

The bounds  $\sum_{p > Y} p^{-2} \ll (Y \log Y)^{-1}$  and  $\sum_{p \leq M} p^{-1} \ll \log \log(3M)$  follow from classical prime estimates and partial summation; see [35, 31].

**Two ways to control dependence.** The proofs below use an exact graph for the retained marked collections, together with deletion and conditional-marginal bounds. A different route can be available when one can couple the full configuration to its law conditioned on one prescribed occurrence. The binary-run construction in [26, technical companion, Appendix F] forces selected private prime coordinates and charges only the coordinates that this forcing may change; this directed set can be smaller than the graph of all shared primes. The common probabilistic mechanism is classical Palm–Stein comparison, not a universal arithmetic pivot theorem. Its exact conditional-law construction and its modified-coordinate estimates are model-specific. We neither import that binary coupling into the  $L^2$  or finite-environment proofs nor infer a deterministic-sequence comparison from it.

## 2.6 Prime densities under an $L^2$ bound

Let  $K_2$  bound the squared norm:

$$\sup_p \int g_p^2 dm \leq K_2 < \infty. \quad (2.9)$$

Every nonzero integer power map contracts  $L^2$  by Jensen. More generally a full-rank integer map  $\mathbb{T}^d \rightarrow \mathbb{T}^d$  is surjective; its transfer operator averages over fibres and contracts  $L^2$ .

Two distinct prime factors smooth a marginal, since  $\|h_1 * h_2\|_\infty \leq \|h_1\|_2 \|h_2\|_2 \leq K_2$ . Thus

$$\mathbb{P}(F(n) \in J) \leq K_2 m(J) \quad \text{if } \omega(n) \geq 2. \quad (2.10)$$

For a prime power the available bound is

$$\mathbb{P}(F(p)^r \in J) \leq \sqrt{K_2 m(J)}. \quad (2.11)$$

A density of  $L^2$  norm at most  $\sqrt{K_2}$  puts at most half its mass on a set of measure  $1/(4K_2)$ . Push by the nonzero power in question and rotate so that the first Fourier coefficient is real and nonnegative. The real moment is at most  $(1 + \cos(\pi/(4K_2)))/2 \leq 1 - 1/(16K_2^2)$ . Thus

$$|\widehat{g}_p(n)| \leq \varrho_{K_2} := 1 - \frac{1}{16K_2^2}, \quad n \neq 0.$$

For a product of  $t$  independent prime variables raised to nonzero integer powers  $e_i$ , its nontrivial Fourier coefficients have sum at most

$$\sum_{j \neq 0} \prod_{i=1}^t |\widehat{g}_i(e_i j)| \leq \varrho_{K_2}^{t-2} \left( \sum_{j \neq 0} |\widehat{g}_1(e_1 j)|^2 \right)^{1/2} \left( \sum_{j \neq 0} |\widehat{g}_2(e_2 j)|^2 \right)^{1/2} \leq (K_2 - 1) \varrho_{K_2}^{t-2}.$$

Here Parseval bounds each sum of squares by  $K_2 - 1$ . The absolutely convergent Fourier series represents the product density and gives

$$\|h_t - 1\|_\infty \leq \delta_2(t) := (K_2 - 1)\varrho_{K_2}^{t-2}, \quad t \geq 2. \quad (2.12)$$

The laws may differ at every prime and the bound is uniform in nonzero power exponents. The symbol  $\varrho_{K_2}$  distinguishes this spectral constant from binary nullities.

## 2.7 Two disjoint bases smooth a pair

One basis of two prime coordinates gives only an  $L^2$  density on  $\mathbb{T}^2$ . If the columns  $(v_p(n), v_p(m))$  possess two bases disjoint in prime indices, each pushes forward a density of  $L^2$  norm at most  $K_2$ . Their independent convolution has  $L^\infty$  norm at most  $K_2^2$ , hence

$$\mathcal{L}(F(n), F(m)) \leq K_2^2 m^{\otimes 2}. \quad (2.13)$$

Section 3 identifies and pays for the pairs lacking these bases. These pair bounds are used after averaging over small prime values, not asserted on every conditional fibre.

## 2.8 Abundant unexposed prime factors

Put  $U = \log N$ ,  $V = \log \log N$ , and choose

$$Y = U^4, \quad \ell_N = 2\lceil 4V \rceil, \quad Z = N^{1/(4\ell_N)}, \quad A_N = \sum_{Y < p \leq Z} \frac{1}{p}. \quad (2.14)$$

Mertens' formula gives

$$A_N = V - 2 \log V - \log 128 + o(1) \sim V. \quad (2.15)$$

For fixed  $\theta \in (0, 1)$  let  $I(\theta) = 1 - \theta + \theta \log \theta$  and  $t_N = \lfloor \theta A_N \rfloor$ .

Lemma B.2 in Section B proves uniformly on translated intervals of  $N$  integers that

$$\frac{1}{N} \#\{n : \omega_{Y,Z}(n) < \theta A_N\} \ll_\theta e^{-I(\theta)A_N}. \quad (2.16)$$

This computes factorial moments of divisibility with their congruence errors; it assumes no independence of divisibility events. Moving the upper cutoff is essential:  $Z^{\ell_N} = N^{1/4}$  pays for the growing-order truncation.

# 3 Shrinking targets under $L^2$ : retaining the prime source

This section proves Theorem A. A set of sites of vanishing arithmetic density may still carry a persistent mass of hits. The correct target retains the prime-site law and reconstructs Haar on composite sites. Both populations are handled in one graph, so their asymptotic independence is a conclusion, not an assumption.

## 3.1 A prime source that cannot be deleted

Mixing at typical composite sites alone does not determine the correct target. On the circle in turns, take the same prime density at every prime:

$$g_2^{\log}(u) = \frac{1}{2} \left( \log \frac{1}{2|u|} \right)^2, \quad 0 < |u| < 1/2.$$

It is a density, with  $\int (g_2^{\log})^q dm = \Gamma(2q+1)/2^q$  and squared  $L^2$  norm 6. Products mix exponentially by (2.12). Yet a centred arc of mass  $a$  has probability

$$q_2(a) = a \left( 1 + \log(1/a) + \frac{1}{2} \log^2(1/a) \right). \quad (3.1)$$

At  $a = \tau/N$ , prime sites in  $[N, 2N]$  alone give independent indicators of total mean asymptotic to  $(\tau/2) \log N$ , by the prime number theorem [35]. Their variance is at most their mean, so the count diverges in probability and its distance to every fixed-parameter Poisson tends to one.

This does not contradict empirical equidistribution against fixed tests, as in [24, arXiv version 2, Proposition 7.3]. Few sites can carry substantial microscopic mass. This section retains that mass instead of deleting it.

### 3.2 A theorem on arbitrary masks

Fix  $C \geq 1$ ,  $A > 0$ ,  $K_2 \geq 1$ . Let  $M = \lfloor CN \rfloor$  and choose deterministic  $S_N \subseteq \{2, \dots, M\}$ . Each site has a deterministic Borel target  $J_{N,n}$  of mass  $a_n \leq A/N$ . Write  $a_* = \max_{n \in S_N} a_n$ , zero for an empty mask, and

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

Let  $S_N^p, S_N^c$  be prime and composite sites, and  $q_p = \int_{J_{N,p}} g_p dm$ . The target intensity is

$$\mu_N = \sum_{n \in S_N^c} \delta_n \otimes \mathbf{1}_{J_{N,n}} m + \sum_{p \in S_N^p} \delta_p \otimes \mathbf{1}_{J_{N,p}} g_p m. \quad (3.2)$$

Prime-site marks retain their true restricted law. No bound on  $\sum_p q_p$  is assumed.

**Theorem 3.1** (Shrinking-target field on a mask). *Assume independent prime values and (2.9). Fix  $0 < \theta < 1$  and*

$$0 < \kappa < \beta_2(\theta) := \min\{1 - \theta + \theta \log \theta, -\theta \log \varrho_{K_2}\}, \quad \varrho_{K_2} = 1 - \frac{1}{16K_2^2}.$$

For  $Y = (\log N)^4$ ,

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\Xi_N \mid \mathcal{F}_Y), \text{PPP}(\mu_N)) \ll_{C,A,K_2,\theta,\kappa} (\log N)^{-\kappa}. \quad (3.3)$$

*Constants and thresholds are uniform in all masks, targets and densities satisfying the hypotheses. The lattice of sites and exact occurrence phases are retained.*

Set  $\Lambda_N = \sum_{n \in S_N} a_n$ ,  $A_N^p = \sum_{p \in S_N^p} a_p$  and  $Q_N = \sum_{p \in S_N^p} q_p$ . The total count has target  $\text{Pois}(\lambda_N)$  with

$$\lambda_N = \Lambda_N - A_N^p + Q_N. \quad (3.4)$$

Here  $\Lambda_N \leq CA$  and  $A_N^p = O(1/\log N)$ , whereas  $Q_N$  may diverge. The exact moving parameter, not only a leading equivalent, is used in total variation.

### 3.3 Two disjoint bases and their precise defect

The intended bound for an ordinary pair is proportional to  $a_n a_m$ , the product of two small target masses. This is why rank two alone is not the resource we seek. One basis of prime columns gives a nonsingular integer map of  $\mathbb{T}^2$  and hence a pushed-forward density of  $L^2$  norm at most  $K_2$ ; that density need not be bounded. Two bases disjoint in prime indices supply two independent  $\mathbb{T}^2$ -valued factors. Cauchy–Schwarz bounds the density of their convolution by  $K_2^2$ , and the remaining independent prime coordinates preserve this bound by further convolution. This is the joint smoothing in (2.13), used after averaging the small-prime environment.

The next lemma identifies exactly when this resource fails under the retained-site hypotheses. The failure leaves a prime-power ratio, rather than an unrestricted family of rank-two pairs. Lemma 3.3 then pays for these pairs with a weaker individual estimate and a sufficiently small arithmetic count. The proof needs both parts of this division, not a quadratic bound imposed on every pair.

**Lemma 3.2** (Failure of two disjoint bases). *Suppose distinct  $n, m$  each have at least three distinct prime factors  $> Y$ , and neither is divisible by  $p^2$  for  $p > Y$ . The columns  $(v_p(n), v_p(m))$ ,  $p \mid nm$ , fail to have two bases of  $\mathbb{Q}^2$  disjoint in indices if and only if*

$$n/m = p^r \quad \text{for a prime } p \text{ and } r \in \mathbb{Z} \setminus \{0\}. \quad (3.5)$$

*Proof.* A rank-two family of at least four nonzero vectors without two disjoint bases has all but at most one vector on a single line. Remove one basis: if the remainder has rank two it supplies the other. Otherwise at least two remaining vectors lie on a line  $L$ . If both removed vectors lie outside  $L$ , pair each with a different remaining vector; otherwise only one vector of the original family is outside  $L$ .

The arithmetic family has at least four columns: with at most three, all three required large factors would be shared, simple and the only factors, implying  $n = m$ . Its rank is two: a one-dimensional family containing a large-prime column  $(1, 0)$  implies  $m = 1$ , and one containing  $(1, 1)$  implies  $n = m$ .

Under failure of two bases, at least two of the three large simple factors of  $n$  are not the exceptional prime. Their columns are  $(1, 0)$  or  $(1, 1)$ , forcing  $L$  horizontal or diagonal. Horizontal would leave  $m$  with at most one prime factor. Diagonal means all valuations agree except at one prime, giving (3.5). Conversely every basis for such a pair must use that prime.  $\square$

The exceptional ratio is impossible within  $[N, 2N)$ , but must be paid on a prefix.

**Lemma 3.3** (Exceptional pair cost). *Under the preceding hypotheses and (2.9), a pair satisfying (3.5) obeys*

$$\mathbb{P}(F(n) \in J_n, F(m) \in J_m) \leq K_2^{3/2} \min\{a_n \sqrt{a_m}, a_m \sqrt{a_n}\}. \quad (3.6)$$

*There are  $O(M \log \log(3M))$  ordered pairs of this form in  $\{2, \dots, M\}$ .*

*Proof.* Write  $n = p^u d$ ,  $m = p^v d$ ,  $p \nmid d$ ,  $u \neq v$ . Removing  $p$  leaves  $d$  with at least two distinct prime factors. Thus  $W = F(d)$  has density bounded by  $K_2$ , independently of  $Z = F(p)$  whose density has  $L^2$  norm at most  $\sqrt{K_2}$ . Conditioning on  $Z$ ,

$$\mathbb{P}(WZ^u \in J_n, WZ^v \in J_m) \leq K_2 \int g_p(z) \int_{t \in J_n} \mathbf{1}_{J_m}(tz^{v-u}) dm(t) dm(z).$$

The nonzero power  $v - u$  contracts  $L^2$ , so the inner probability in  $z$  is at most  $\sqrt{K_2 a_m}$ . Exchange  $n, m$  for the other bound.

The exact ordered count is

$$2 \sum_{\substack{p^r \leq M \\ r \geq 1}} (\lfloor M/p^r \rfloor - 1)_+ \leq 2M \sum_{p^r \leq M} p^{-r} \ll M \log \log(3M).$$

The terms  $r \geq 2$  are summable;  $r = 1$  uses the reciprocal-prime sum.  $\square$

At  $a_* \leq A/N$ , all exceptional pairs cost  $O(\log \log N / \sqrt{N})$ . A weaker individual bound is sufficient on a sparse family.

For  $385 = 5 \cdot 7 \cdot 11$  and  $770 = 2 \cdot 385$ , the columns at 5, 7, 11 are diagonal, while the column at 2 is the only one off that line. Every basis must use prime 2. The following example shows how this shared coordinate retains its density singularity.

*Example 3.4* (A quadratic bound really can fail). Take Haar at all odd primes and density  $g_2^{\log}$  from (3.1) at prime 2. The pair 385, 770 has unbounded joint density  $g_2^{\log}(z_2/z_1)$ . For the same centred arc of mass  $a < 1/2$ ,

$$\mathbb{P}(F(385) \in J_a, F(770) \in J_a) = a^2 \left( \frac{L^2}{2} + \frac{3L}{2} + \frac{7}{4} \right), \quad L = \log \frac{1}{2a}.$$

Integrating  $(a - |u|)g_2^{\log}(u)$  over  $[-a, a]$  proves the formula. Its ratio to  $a^2$  diverges, so (3.6) cannot be replaced by a uniform quadratic bound on all pairs.

### 3.4 Mixed edges and deletion costs

If a retained prime  $p > Y$  shares a coordinate with a good composite  $n$ , then  $n = pd$ ,  $p \nmid d$ , and  $d$  has at least two distinct prime factors. Conditional on  $F(p)$ ,

$$\mathbb{P}(F(p) \in J_p, F(n) \in J_n) \leq K_2 q_p a_n. \quad (3.7)$$

With  $W_N = \sum_{p \in S_N^p, p > Y} q_p \sum_{n \in G_N, p|n} a_n$ ,

$$W_N \leq M a_* \sum_p \frac{q_p}{p} \ll K_2^{1/2} M a_*^{3/2} \log \log(3M), \quad (3.8)$$

since  $q_p \leq \sqrt{K_2 a_*}$ . This weighted sum can vanish while  $Q_N$  diverges.

Use (2.14) and  $t_N = \lfloor \theta A_N \rfloor \geq 3$ . Good composites  $G_N$  have at least  $t_N$  factors in  $(Y, Z]$  and no prime square  $> Y$ . Lemma B.2 on the prefix of length  $M \asymp N$  gives, for  $D_N = S_N^c \setminus G_N$ ,

$$|D_N| \ll N \left( e^{-I(\theta)A_N} + \frac{1}{Y \log Y} \right). \quad (3.9)$$

Deleted composites with two distinct prime factors cost  $K_2 a_*$  each. Single-prime powers  $p^r$ ,  $r \geq 2$ , only satisfy (2.11), but their number  $H_M$  obeys

$$H_M \ll \frac{\sqrt{M}}{\log M}. \quad (3.10)$$

Prime squares give this bound by Chebyshev; exponents  $r \geq 3$  contribute at most  $M^{1/3} \lfloor \log_2 M \rfloor$ , which is smaller. Prime sites  $p \leq Y$  are exposed and deleted from real and target fields at cost  $2q_p$ . A safe budget is

$$B_N = (K_2 + 1) a_* |D_N| + \sqrt{K_2 a_*} H_M + 2 \sum_{p \in S_N^p, p \leq Y} q_p. \quad (3.11)$$

The harmless first-term overcount includes target masses of the prime powers.

### 3.5 Proof of the joint transfer

On  $G_N$ , the conditional density of  $F(n)$  has uniform error  $\delta_N = (K_2 - 1) \varrho_{K_2}^{t_N - 2}$  from one. At prime sites  $p > Y$  the conditional subprobability is exactly  $\mathbf{1}_{J_p} g_p m$ . Join both populations when they share a prime  $> Y$ . Collections without edges use disjoint independent unexposed coordinates; the graph is exact.

Let  $E_N$  count ordered good-composite edges and  $R_N^G$  the exceptional ordered pairs, possibly overcounted beyond the graph. Then  $E_N \ll M^2/(Y \log Y)$  and  $R_N^G \ll M \log \log(3M)$ . The interface gives

$$\begin{aligned} \mathcal{E}_N &\leq B_N + \delta_N \sum_{n \in G_N} a_n + 2(1 + \delta_N)^2 \sum_{n \in G_N} a_n^2 + 2 \sum_{\substack{p > Y \\ p \in S_N^p}} q_p^2 \\ &\quad + 2((1 + \delta_N)^2 + K_2^2) a_*^2 E_N + 2K_2^{3/2} a_*^{3/2} R_N^G + 4(1 + \delta_N + K_2) W_N, \end{aligned} \quad (3.12)$$

where  $\mathcal{E}_N$  is the left side of (3.3). The two mixed orientations contribute  $2(1 + \delta_N)W_N$  to  $b_1$  and  $2K_2W_N$  to  $\mathbb{E}b_2$ . There are no edges between distinct primes.

Choose  $\kappa'$  strictly between  $\kappa$  and  $\beta_2(\theta)$ . Bad-site and mixing terms are  $O((\log N)^{-\kappa'})$ . The remaining bounds are

$$\sqrt{a_*}H_M \ll (\log N)^{-1}, \quad \sum_{p \leq Y} q_p \ll \pi(Y)/\sqrt{N}, \quad \sum_{p \leq M} q_p^2 \leq K_2 a_* \pi(M) \ll (\log N)^{-1},$$

$$a_*^2 E_N \ll (Y \log Y)^{-1}, \quad a_*^{3/2} R_N^G + W_N \ll \log \log N / \sqrt{N}.$$

The composite diagonal is  $O(N^{-1})$ . All are absorbed by  $(\log N)^{-\kappa}$ , proving Theorem 3.1. Pair bounds are used only after averaging the environment; no uniform conditional domination of all pairs was assumed.

### 3.6 The Haar criterion and corrected intensities

The theorem implies  $|\mathbb{P}(Z_N = 0) - e^{-\Lambda_N}| \leq \mathcal{E}_N$ . If  $Q_N \rightarrow 0$ , replacing prime intensities by Haar costs at most

$$\sum_{p \in S_N^p} \int_{J_p} |g_p - 1| dm \leq Q_N + A_N^p \rightarrow 0.$$

Conversely, closeness of the count to  $\text{Pois}(\Lambda_N)$  forces

$$e^{-\Lambda_N} |e^{-(Q_N - A_N^p)} - 1| \rightarrow 0.$$

Since  $\Lambda_N$  is bounded and  $A_N^p \rightarrow 0$ , this implies  $Q_N \rightarrow 0$ . Thus prime-source disappearance, the Haar count target and the Haar field target are equivalent under the theorem's hypotheses.

For the common prime density

$$g_\alpha^{\log}(u) = \frac{(\log(1/(2|u|)))^\alpha}{\Gamma(\alpha + 1)}, \quad 0 < |u| < 1/2, \quad \alpha \geq 0,$$

we have

$$\int (g_\alpha^{\log})^q dm = \frac{\Gamma(\alpha q + 1)}{\Gamma(\alpha + 1)^q}, \quad \nu_\alpha(J_a) = \frac{1}{\Gamma(\alpha + 1)} \int_{\log(1/a)}^\infty t^\alpha e^{-t} dt \sim \frac{a \log(1/a)^\alpha}{\Gamma(\alpha + 1)}.$$

On  $[N, 2N]$ , at  $a = \tau/N$ , the prime number theorem gives  $Q_N \sim \tau(\log N)^{\alpha-1}/\Gamma(\alpha + 1)$ . The limit is  $\text{Pois}(\tau)$  for  $\alpha < 1$ ,  $\text{Pois}(2\tau)$  for  $\alpha = 1$ , and a Poisson law of diverging moving parameter for  $\alpha > 1$ . For the quadratic density the exact comparison parameter is

$$\lambda_N = \tau + \frac{\tau P_N}{N} (L_N + \frac{1}{2} L_N^2), \quad L_N = \log(N/\tau), \quad P_N = \pi(2N - 1) - \pi(N - 1).$$

Its equivalent  $(\tau/2) \log N$  is not substituted without a parameter-error bound.

The theorem is joint in space. For  $g_1^{\log}$ , a common target of mass  $\tau/N$  on  $[N, 2N]$  and  $[2N, 4N]$  gives

$$(Z_N^{(1)}, Z_N^{(2)}) \xrightarrow{d_{\text{TV}}} \text{Pois}(2\tau) \otimes \text{Pois}(4\tau).$$

Independence comes from restrictions of the same target field, not two marginal convergence statements. The identity  $F(2n) = F(2)F(n)$  is paid on exceptional edges. On a single dyadic block, the prime and composite fields are actually independent: every prime factor of a composite below  $2N$  is  $< N$ . That simplification is not used on general masks.

### 3.7 Sharpness in the uniform $L^q$ scale

**Proposition 3.5** (Prime-square obstruction below  $L^2$ ). *For every  $1 \leq q < 2$ , a uniform  $L^q$  bound on independent prime densities does not suffice for Theorem 3.1 with target (3.2) on all deterministic masks. One may choose a fixed common density and a prime-square mask for which the count's distance from that target tends to one.*

*Proof.* Choose  $1/2 < \alpha < 1/q$  and use at every prime the density

$$g_\alpha^{\text{pow}}(u) = (1 - \alpha)2^{-\alpha}|u|^{-\alpha}, \quad 0 < |u| < 1/2.$$

It integrates to one and  $\int (g_\alpha^{\text{pow}})^q dm = (1 - \alpha)^q / (1 - \alpha q) < \infty$ . For fixed  $\tau > 0$ , let  $S_N = \{p^2 : p \leq \sqrt{N}, p \text{ prime}\}$  and use the centred arc of mass  $a_N = \tau/N$ . There are no prime sites, so (3.2) has total mass  $\Lambda_N = a_N \pi(\sqrt{N}) \rightarrow 0$ .

Write  $F(p) = e^{2\pi i U_p}$ . The event  $|U_p| \leq a_N/4$  forces  $F(p)^2 \in J_N$ , and has probability

$$\int_{-a_N/4}^{a_N/4} g_\alpha^{\text{pow}}(u) du = 2^{\alpha-1} a_N^{1-\alpha}.$$

The square-site indicators use distinct independent prime coordinates. Their sum satisfies

$$\mathbb{E}Z_N \gg \pi(\sqrt{N})N^{\alpha-1} \asymp \frac{N^{\alpha-1/2}}{\log N} \rightarrow \infty.$$

Thus  $\mathbb{P}(Z_N = 0) \leq e^{-\mathbb{E}Z_N} \rightarrow 0$ , while the proposed target is empty with probability tending to one. This already forces total variation to one for the counts.  $\square$

This proves sharpness of exponent 2 for the stated uniform  $L^q$  assumption, target and arbitrary-mask conclusion. It does not say every non- $L^2$  density destroys Poisson, or preclude a different target retaining additional prime-power sources.

### 3.8 Correct intensities at an inadmissible resolution

The prime-square obstruction concerns an incorrect intensity rule on composite sites. There is a second obstruction below  $L^2$ : keeping the correct intensities does not by itself make a site-labelled Poisson field admissible. This is the classical distinction between scalar and process Poisson approximation [4, Theorems 1–2], applied to the same family of prime densities.

**Proposition 3.6** (Correct prime intensities, but no sitewise Poisson approximation). *For every  $1 \leq q < 2$ , there is a fixed common  $L^q$  prime density for which the following holds. Let  $S_N$  consist only of the primes in  $[N, 2N]$  and let  $J_N$  be the arc about 1 of Haar mass  $a_N = \tau/N$ , with  $\tau > 0$  fixed. Write*

$$k_N = |S_N|, \quad q_N = \mathbb{P}(F(p) \in J_N), \quad Z_N = \sum_{p \in S_N} \mathbf{1}_{\{F(p) \in J_N\}}.$$

*Then  $q_N \rightarrow 0$  and  $k_N q_N \rightarrow \infty$ , but*

$$d_{\text{TV}} \left( \bigotimes_{p \in S_N} \text{Bern}(q_N), \bigotimes_{p \in S_N} \text{Pois}(q_N) \right) \rightarrow 1, \quad (3.13)$$

$$d_{\text{TV}}(\mathcal{L}(Z_N), \text{Pois}(k_N q_N)) \leq q_N \rightarrow 0. \quad (3.14)$$

*The obstruction in (3.13) also holds when each occurrence retains its exact phase, even if the target uses the correct restricted phase law at every prime.*

*Proof.* First consider independent Bernoulli variables with parameters  $p_i \in [0, 1]$  and independent Poisson variables with the same means. The event that some site contains at least two points is impossible for the Bernoulli field. Its Poisson probability gives

$$d_{\text{TV}} \left( \bigotimes_i \text{Bern}(p_i), \bigotimes_i \text{Pois}(p_i) \right) \geq 1 - \prod_i e^{-p_i} (1 + p_i) = 1 - \exp \left\{ - \sum_i (p_i - \log(1 + p_i)) \right\}. \quad (3.15)$$

Moreover

$$\frac{p^2}{2(1+p)} \leq p - \log(1+p) = \int_0^p \frac{t}{1+t} dt \leq \frac{p^2}{2}. \quad (3.16)$$

Thus the lower bound tends to one when  $\sum_i p_i^2 \rightarrow \infty$ .

For the arithmetic example, choose  $1/2 < \alpha < 1/q$  and take  $g_\alpha^{\text{pow}}$  from the preceding proof at every prime. The normalization and  $L^q$  norm are as computed there, and  $g_\alpha^{\text{pow}} \notin L^2$  since  $2\alpha > 1$ . Angular coordinates are measured in turns:  $J_N$  corresponds to  $|u| < a_N/2$ . Consequently its exact prime probability is

$$q_N = 2(1-\alpha)2^{-\alpha} \int_0^{a_N/2} u^{-\alpha} du = a_N^{1-\alpha}. \quad (3.17)$$

The prime number theorem gives  $k_N \sim N/\log N$ , and therefore

$$k_N q_N^2 \sim \tau^{2(1-\alpha)} \frac{N^{2\alpha-1}}{\log N} \rightarrow \infty. \quad (3.18)$$

Equations (3.15)–(3.16) prove the field assertion. Here  $Z_N$  is exactly binomial because distinct prime sites use independent coordinates.

For completeness, let  $X_i$  be independent Bernoulli variables,  $W = \sum_i X_i$ ,  $W_i = W - X_i$  and  $\lambda = \sum_i p_i > 0$ . For an indicator test, the scalar Poisson Stein equation has a solution with  $\|\Delta f\|_\infty \leq (1 - e^{-\lambda})/\lambda$ ; this is the scalar factor in [4, Theorem 1], also recorded in [26, technical companion, Appendix B.2]. Independence gives the leave-one-out identity

$$\mathbb{E}[\lambda f(W+1) - W f(W)] = \sum_i p_i^2 \mathbb{E}[\Delta f(W_i + 1)].$$

Taking the supremum over indicator tests yields

$$d_{\text{TV}}(\mathcal{L}(W), \text{Pois}(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda} \sum_i p_i^2. \quad (3.19)$$

For  $p_i = q_N$  the right side is  $(1 - e^{-k_N q_N})q_N \leq q_N$ . Finally the projection of any phase-marked field to its counts by site contracts total variation. The event used in (3.15) is still impossible in the real marked field, proving the last assertion.  $\square$

**Collision mass and spatial aggregation.** The classical independent-source approximation of Le Cam [21] and the scalar refinements of Barbour–Hall [7] distinguish sitewise coupling from count approximation. The coordinatewise Bernoulli–Poisson coupling costs at most  $\sum_i p_i^2$ . Together with (3.15) and  $p_i \leq 1$ , this shows that for independent sources sitewise Poisson approximation holds if and only if  $\sum_i p_i^2 \rightarrow 0$ . Under (2.9) and the target ceiling of Theorem 3.1, this mass is  $O(1/\log N)$  at prime sites; the proposition does not contradict that theorem.

In the prime-only example, partition  $S_N$  deterministically into  $r_N$  nonempty cells of sizes  $k_{N,c}$ . The cell counts are independent binomials, so (3.19) and product coupling give

$$d_{\text{TV}}\left(\mathcal{L}((Z_{N,c})_c), \bigotimes_c \text{Pois}(k_{N,c}q_N)\right) \leq q_N \sum_c (1 - e^{-k_{N,c}q_N}) \leq r_N q_N. \quad (3.20)$$

Thus  $r_N = o(q_N^{-1})$  is a sufficient growing-resolution regime, while the complete prime-labelled field fails. This deduction concerns independent prime sites only; it does not settle interactions with composites or with powers of the same prime. The exact moving parameter  $k_N q_N$  is retained throughout. Its divergence does not invalidate the scalar approximation, but an asymptotic equivalent cannot replace it in total variation without a parameter-error estimate.

*Remark 3.7* (Why an irregular environment must be retained). If odd prime values are Haar but  $F(2) = 1$ , hits at odd  $n \in [N, 2N)$  and at  $2n$  coincide exactly. Their combined count tends to  $2\text{Pois}(\tau/2)$ , not  $\text{Pois}(\tau)$ . One atomic prime violates (2.9) and remains visible. The next section builds the target retaining this relation.

## 4 Finite prime environments and multiplicative clusters

A nonnegligible exceptional source need not be deleted. We retain finitely many prime coordinates and group sites which become known rotations of one value after those coordinates are fixed. The target is a conditional cluster law, generally not Poisson with an averaged intensity.

### 4.1 Decomposition and conditional target

Fix a finite set  $\mathcal{E}$  of primes, possibly empty, and let

$$q_{\mathcal{E}} = \prod_{p \in \mathcal{E}} p, \quad b = |\mathcal{E}|, \quad \mathcal{D}_{\mathcal{E}} = \{s \geq 1 : p \mid s \Rightarrow p \in \mathcal{E}\}.$$

The vector  $\xi = (F(p))_{p \in \mathcal{E}}$  has arbitrary joint law: atoms and internal dependence are allowed. It is independent of the remaining prime values, which are mutually independent and satisfy

$$F(p) \sim g_p m, \quad \sup_{p \notin \mathcal{E}} \int g_p^2 dm \leq K_2 < \infty. \quad (4.1)$$

Both  $\mathcal{E}$  and  $K_2$  are fixed before  $N$ . Every integer decomposes uniquely as

$$n = sd, \quad s \in \mathcal{D}_{\mathcal{E}}, \quad (d, q_{\mathcal{E}}) = 1, \quad F(sd) = \xi(s)F(d), \quad (4.2)$$

where  $\xi(s) = \prod_{p \in \mathcal{E}} \xi_p^{v_p(s)}$ . We call  $d$  the core; it is not a squarefree kernel.

Take a mask  $S_N \subseteq \{2, \dots, M\}$ ,  $M = \lfloor CN \rfloor$ , and targets of masses at most  $a_* = A/N$ . For fixed  $\xi$ , a core  $d \geq 2$  coprime to  $q_{\mathcal{E}}$  has phase target

$$K_{N,d}^{\xi} = \bigcup_{\substack{s \in \mathcal{D}_{\mathcal{E}} \\ sd \in S_N}} \xi(s)^{-1} J_{N, sd}. \quad (4.3)$$

A phase germ  $u$  creates

$$\mathcal{G}_{N,d}^\xi(u) = \sum_{\substack{s \in \mathcal{D}_\mathcal{E} \\ sd \in S_N}} \mathbf{1}_{\{\xi(s)u \in J_{N,sd}\}} \delta_{(sd, \xi(s)u)}. \quad (4.4)$$

Keep all multiplicities. Let  $h_d = g_d$  for prime  $d$ , and  $h_d = 1$  for composite  $d$ . Given  $\xi$ , take independent Poisson measures  $\eta_{N,d}^\xi$  of intensity  $\mathbf{1}_{K_{N,d}^\xi} h_d m$ .

Core  $d = 1$  has no regular randomness and contributes exactly

$$\Xi_{N,0}^\xi = \sum_{s \in \mathcal{D}_\mathcal{E} \cap S_N} \mathbf{1}_{\{\xi(s) \in J_{N,s}\}} \delta_{(s, \xi(s))}. \quad (4.5)$$

The full target kernel  $\mathcal{Q}_N(\xi)$  is the law of

$$\Xi_N^{\text{tar}, \xi} = \Xi_{N,0}^\xi + \sum_{\substack{2 \leq d \leq M \\ (d, q_\mathcal{E})=1}} \int \mathcal{G}_{N,d}^\xi(u) \eta_{N,d}^\xi(du). \quad (4.6)$$

The exact part may have diverging cardinality; it is not Poissonized. Unconditionally, draw  $\xi$  from its actual law and then use this kernel.

**Theorem 4.1** (Retaining a finite prime environment). *Under (4.1) and the preceding assumptions, fix  $0 < \theta < 1$ ,  $0 < \kappa < \beta_2(\theta)$  as in Theorem 3.1. Let  $Y = (\log N)^4$  and  $\mathcal{F}_Y^{\text{reg}} = \sigma(F(p) : p \leq Y, p \notin \mathcal{E})$ . Then*

$$\sup_{\xi \in \mathbb{T}^\mathcal{E}} \mathbb{E}_{\text{reg}} d_{\text{TV}}(\mathcal{L}(\Xi_N^\xi \mid \mathcal{F}_Y^{\text{reg}}), \mathcal{Q}_N(\xi)) \ll_{\mathcal{E}, C, A, K_2, \theta, \kappa} (\log N)^{-\kappa}. \quad (4.7)$$

Here the real field  $\Xi_N^\xi$  is constructed by fixing the exceptional coordinates. Thus the kernel is defined at every  $\xi$ , not only at positive-probability conditioning atoms. Integrating yields a joint comparison of environment and field.

The reference rule concerns the core: a composite site  $sp$  can retain the density of its prime core  $p$ . Regular  $L^2$  bounds do not justify replacing that density by Haar.

## 4.2 Layer sums and two different ceilings

For every  $\alpha > 0$ ,

$$Z_\mathcal{E}(\alpha) := \sum_{s \in \mathcal{D}_\mathcal{E}} s^{-\alpha} = \prod_{p \in \mathcal{E}} (1 - p^{-\alpha})^{-1} < \infty. \quad (4.8)$$

Also  $B_\mathcal{E}(x) = |\mathcal{D}_\mathcal{E} \cap [1, x]| = O_\mathcal{E}((1 + \log x)^b)$  and

$$\sum_{\substack{s > H \\ s \in \mathcal{D}_\mathcal{E}}} s^{-1} \leq H^{-1/2} Z_\mathcal{E}(1/2). \quad (4.9)$$

These are finite products of geometric series, not uniform smooth-number results with a growing prime cutoff.

Two layer counts are useful:

$$\sum_{s \in \mathcal{D}_\mathcal{E}} \pi(M/s) \ll_\mathcal{E} \frac{M}{\log M}, \quad \sum_{s \in \mathcal{D}_\mathcal{E}} H_0(M/s) \ll_\mathcal{E} \frac{\sqrt{M}}{\log M}, \quad (4.10)$$

where  $H_0(x)$  counts prime powers of exponent at least two. For the first,  $s \leq \sqrt{M}$  costs  $O(M/(s \log M))$  by Chebyshev; the tail is at most  $M \sum_{s > \sqrt{M}} s^{-1} = O(M^{3/4})$ . For the second use  $H_0(x) \ll \sqrt{x}/\log(2x)$  below  $\sqrt{M}$ ; its tail is bounded by

$$\sqrt{M} \sum_{s > \sqrt{M}} s^{-1/2} \leq M^{3/8} Z_\mathcal{E}(1/4).$$

Put  $U = \log N$ ,  $V = \log \log N$  and  $H = U^4$ . For composite cores, temporarily retain only multipliers  $s \leq H$ ; for prime cores retain all multipliers. If a core has two distinct prime factors its density is bounded by  $K_2$ , so real and target mass of discarded layers costs

$$(K_2 + 1)a_*M \sum_{s>H} s^{-1} = O(U^{-2}).$$

For cores  $d = p^r$ ,  $r \geq 2$ , even all layers can be bounded by

$$(\sqrt{K_2 a_*} + a_*) \sum_s H_0(M/s) = O(U^{-1}).$$

These estimates justify final restoration of composite layers.

Write  $K_d^H$  for a truncated composite-core target,  $b_d = m(K_d^H)$ . For a prime core retain  $K_p$  in full and write  $b_p = m(K_p)$ ,  $q_p = \int_{K_p} g_p dm$ . The two ceilings are

$$b_c^* := \sup_{d \text{ composite}} b_d \ll_{\mathcal{E}, A} \frac{V^b}{N}, \quad b_p \ll_{\mathcal{E}, A} \frac{U^b}{N}, \quad q_p \ll_{\mathcal{E}, A, K_2} N^{-1/2} U^{b/2}. \quad (4.11)$$

The total composite-target mass is at most  $CAZ_{\mathcal{E}}(1)$ . At prime cores,

$$\sum_{p \notin \mathcal{E}} q_p^2 \leq K_2 \sum_p b_p \leq K_2 a_* \sum_s \pi(M/s) \ll (\log N)^{-1}. \quad (4.12)$$

Their mean germ count  $\sum q_p$  may diverge. Truncating prime cores on cardinality alone would lose the needed control.

### 4.3 Proof of the core transfer

Once  $Y > \max \mathcal{E}$ , the band  $(Y, Z]$  in (2.14) contains only regular primes. Good composite cores are coprime to  $q_{\mathcal{E}}$ , have at least  $t_N = \lfloor \theta A_N \rfloor$  factors in that band, and no square  $p^2$  with  $p > Y$ . Their bad complement has

$$|D| \ll N(e^{-I(\theta)A_N} + (Y \log Y)^{-1}).$$

A good core has conditional density within  $\delta_N = (K_2 - 1)\varrho_{K_2}^{t_N-2}$  of one, so restriction costs  $\delta_N b_d$ . Prime cores  $p > Y$  have exact subprobability  $\mathbf{1}_{K_p} g_p m$ .

A deletion budget is

$$B_{\text{core}} = (K_2 + 1)b_c^*|D| + \sqrt{K_2 b_c^*} H_0(M) + 2 \sum_{p \leq Y, p \notin \mathcal{E}} q_p. \quad (4.13)$$

The last term is  $O(N^{-1/2} U^{4+b/2})$ .

Join cores sharing a regular prime  $> Y$ . Conditionally this is the exact graph of core phases and hence cluster marks. The preceding pair arguments apply since every core prime lies outside  $\mathcal{E}$ . In particular an exceptional ratio  $p^r$  cannot have  $p \in \mathcal{E}$ : those irregular directions have already been grouped.

Let  $E_G \ll M^2/(Y \log Y)$ ,  $R_M \ll M \log \log(3M)$ , and

$$W = \sum_{p>Y, p \notin \mathcal{E}} q_p \sum_{d \in G, p|d} b_d \ll N^{-1/2} V^{b+1} U^{b/2}.$$

The finite ledger is

$$\begin{aligned} \mathcal{E}_{\text{core}} &\leq B_{\text{core}} + \delta_N \sum_{d \in G} b_d + 2(1 + \delta_N)^2 \sum_{d \in G} b_d^2 + 2 \sum_{p>Y} q_p^2 \\ &\quad + 2((1 + \delta_N)^2 + K_2^2)(b_c^*)^2 E_G + 2K_2^{3/2} (b_c^*)^{3/2} R_M + 4(1 + \delta_N + K_2)W. \end{aligned} \quad (4.14)$$

Pair probabilities have been averaged over small regular coordinates; there is no fibrewise pair assumption.

Choose  $\kappa'$  strictly between  $\kappa$  and  $\beta_2(\theta)$ . Factor deficits, prime powers, small prime cores, marginal errors and prime diagonals have respective orders

$$V^b U^{-\kappa'}, \quad V^{b/2}/U, \quad N^{-1/2} U^{4+b/2}, \quad U^{-\kappa'}, \quad U^{-1}.$$

The costs of ordinary edges, exceptional pairs and mixed edges are, respectively,

$$O\left(\frac{V^{2b}}{U^4 V}\right), \quad O(N^{-1/2} V^{1+3b/2}), \quad O(N^{-1/2} V^{b+1} U^{b/2}).$$

As  $b$  is fixed and  $\kappa < \beta_2(\theta) < 1$ , each term is  $O(U^{-\kappa})$ .

The measurable germ-to-cluster map (4.4) contracts total variation without multiplying by cluster size. Restoring discarded composite layers adds  $O(U^{-1})$ . Adding (4.5) to both sides is an identical deterministic transformation at fixed  $\xi$ . This proves (4.7), uniformly in  $\xi$ .

#### 4.4 What conditioning permits

The same proof allows targets Borel measurable in  $\xi$  alone, with the pointwise ceiling  $A/N$  for every environment. The target kernel must retain that dependence. For instance  $F(2)J_N$  at site  $2d$  is permitted when  $2 \in \mathcal{E}$ , but creates the duplication encoded in (4.4); it does not preserve an earlier independent target.

For  $r_{N,d}^\xi(u) = |\mathcal{G}_{N,d}^\xi(u)|$  and  $B_N(\xi) = |\Xi_{N,0}^\xi|$ , the conditional count generating function is

$$z^{B_N(\xi)} \exp \left\{ \sum_{\substack{d \geq 2 \\ (d,q\mathcal{E})=1}} \int_{\mathbb{T}} (z^{r_{N,d}^\xi(u)} - 1) h_d(u) dm(u) \right\}, \quad 0 < z \leq 1. \quad (4.15)$$

All sums are finite at fixed size. This retains a whole conditional law rather than its mean. Mixing over  $\xi$  can remain non-Poisson even if all conditional means coincide.

The exceptional set is fixed and finite and independent of the regular source. Arbitrary dependence across those two parts is not covered. The next classification assumes Haar outside  $\mathcal{E}$ , separating exact environmental relations from additional microscopic prime-core densities.

### 5 The exceptional-phase lattice determines the limiting count law

We now calculate the target of Theorem 4.1 on the prefix  $\{2, \dots, N\}$ , for a common arc  $J_N$  about 1 of Haar mass  $\tau/N$ , with fixed  $\tau > 0$ . Throughout this section, prime values outside  $\mathcal{E}$  are *Haar*. The reference for every core  $d \geq 2$  is therefore Haar; no additional prime-source correction is suppressed.

Compound-Poisson and geometric-cluster limits are classical in return-time problems, including [17,20,28]. The specific assertion below identifies the integer phase-relation lattice and the arithmetic jump intensities; it does not claim the general occurrence of compound-Poisson laws as new.

#### 5.1 Lattice, phase classes and jump intensities

First fix  $\xi \in \mathbb{T}^\mathcal{E}$  and write

$$\delta_\mathcal{E} = \prod_{p \in \mathcal{E}} (1 - p^{-1}), \quad \chi_\xi(s) = \prod_{p \in \mathcal{E}} \xi_p^{v_p(s)}.$$

The relation lattice is

$$L_\xi = \{v \in \mathbb{Z}^\mathcal{E} : \xi^v = 1\}. \quad (5.1)$$

This subgroup of  $\mathbb{Z}^\mathcal{E}$  need not be saturated or full rank. Negative exponents are allowed: phases of infinite individual order may still satisfy a common relation.

Partition  $\mathcal{D}_\mathcal{E}$  by  $s \sim t$  when  $\chi_\xi(s) = \chi_\xi(t)$ . In each class  $\mathcal{A}$ , list the multipliers in increasing order,

$$s_{\mathcal{A},1} < s_{\mathcal{A},2} < \cdots.$$

Pad a finite class with  $s_{\mathcal{A},j} = \infty$ , using  $1/\infty = 0$ . The partition depends only on  $L_\xi$ .

Remove the exact contribution of integers determined entirely by the environment:

$$\begin{aligned} B_N(\xi) &= \sum_{s \in \mathcal{D}_\mathcal{E} \cap [2,N]} \mathbf{1}_{\{\chi_\xi(s) \in J_N\}}, \\ W_N^\xi &= \sum_{n=2}^N \mathbf{1}_{\{F(n) \in J_N\}} - B_N(\xi). \end{aligned} \quad (5.2)$$

This may be a non-negligible correction. It subtracts a known subconfiguration, not an expectation.

For  $j \geq 1$ , define

$$\lambda_j(\xi) = \tau \delta_\mathcal{E} \sum_{\mathcal{A}} \left( \frac{1}{s_{\mathcal{A},j}} - \frac{1}{s_{\mathcal{A},j+1}} \right). \quad (5.3)$$

All terms are nonnegative and the series converge, since  $\sum_{s \in \mathcal{D}_\mathcal{E}} s^{-1} = \delta_\mathcal{E}^{-1}$ .

**Theorem 5.1** (Classification for a fixed environment). *Under the assumptions of this section,*

$$W_N^\xi \xrightarrow{d_{\text{TV}}} W^\xi := \sum_{j \geq 1} j P_j, \quad P_j \sim \text{Pois}(\lambda_j(\xi)) \text{ independently.} \quad (5.4)$$

*The sum is almost surely finite, and*

$$\begin{aligned} \sum_{j \geq 1} j \lambda_j(\xi) &= \tau, \quad \sum_{j \geq 1} \lambda_j(\xi) = \tau \Theta(\xi), \\ \Theta(\xi) &= \delta_\mathcal{E} \sum_{\mathcal{A}} s_{\mathcal{A},1}^{-1} \in [\delta_\mathcal{E}, 1]. \end{aligned} \quad (5.5)$$

Moreover,

$$L_\xi = \{0\} \iff \Theta(\xi) = 1 \iff W^\xi \sim \text{Pois}(\tau). \quad (5.6)$$

*These conditions are equivalent to  $d_{\text{TV}}(\mathcal{L}(W_N^\xi), \text{Pois}(\tau)) \rightarrow 0$ .*

The number of germs has law  $\text{Pois}(\tau \Theta(\xi))$ . Given a germ, its size satisfies  $\mathbb{P}(G = j) = \lambda_j(\xi) / (\tau \Theta(\xi))$ . Thus

$$\mathbb{E}[z^{W^\xi}] = \exp \left\{ \sum_{j \geq 1} \lambda_j(\xi) (z^j - 1) \right\}, \quad \mathbb{P}(W^\xi = 0) = e^{-\tau \Theta(\xi)}. \quad (5.7)$$

The ratio  $\Theta$  of germ intensity to point intensity is analogous to an extremal index in rare-return theory [17,28]. The result here is its arithmetic expression, not a new abstract notion of clustering. Nor does the count law determine the lattice: the rank-one calculation below shows that part of the spatial geometry may be lost.

## 5.2 Proof: finite separation followed by tail control

Fix  $H \geq 1$ , and let  $\mathcal{D}_H = \mathcal{D}_{\mathcal{E}} \cap [1, H]$ ,  $L_H = |\mathcal{D}_H|$ . By Theorem 4.1, the field is within  $\epsilon_N \rightarrow 0$  of its prelimit cluster target. After removing the exact configuration (4.5), delete multipliers  $s > H$  in that target. Their cost is at most

$$\frac{\tau}{N} \sum_{s>H} \left\lfloor \frac{N}{s} \right\rfloor \leq \tau \sum_{s>H} s^{-1}. \quad (5.8)$$

Use the same germs before and after deletion; the expected number of removed points bounds the probability of a changed configuration.

The finitely many distinct values  $\chi_{\xi}(s)$  for  $s \in \mathcal{D}_H$  have a positive minimum circular separation  $\Delta_H(\xi)$ . Put  $\Delta_H = 1$  if there is only one value. When  $\tau/N < \Delta_H$ , the arcs  $\chi_{\xi}(s)^{-1}J_N$  are disjoint for different classes and identical within one class. Restrictions of Poisson germs to these arcs are independent.

A truncated class  $\{s_1 < \dots < s_l\}$  produces exactly  $j$  points for cores satisfying

$$N/s_{j+1} < d \leq N/s_j,$$

where  $s_{l+1} = \infty$ . Finite inclusion–exclusion gives

$$\#\{2 \leq d \leq x : (d, q_{\mathcal{E}}) = 1\} = \delta_{\mathcal{E}} x + O(2^b + 1), \quad b = |\mathcal{E}|.$$

The sum of errors between the prelimit intensities and (5.3) truncated at  $H$  is consequently at most  $2(2^b + 1)\tau L_H/N$ . Coupling Poisson processes by their common intensity gives the same bound for the sums weighted by germ sizes.

A common limiting construction makes the final truncation transparent. For every class  $\mathcal{A}$ , take independent germs  $t \in (0, 1/s_{\mathcal{A},1}]$  with intensity  $\tau \delta_{\mathcal{E}} dt$ . A germ produces one point for every  $s \in \mathcal{A}$  with  $st \leq 1$ . Its size is finite for  $t > 0$ , and the jump intensities are (5.3). Truncating  $s \leq H$  costs at most  $\tau \delta_{\mathcal{E}} \sum_{s>H} s^{-1}$ . Hence

$$d_{\text{TV}}(\mathcal{L}(W_N^{\xi}), \mathcal{L}(W^{\xi})) \leq \epsilon_N + 2\tau \sum_{s>H} s^{-1} + \frac{2(2^b + 1)\tau L_H}{N}, \quad (5.9)$$

$$\tau/N < \Delta_H(\xi).$$

First let  $N \rightarrow \infty$  with  $H$  fixed, and then  $H \rightarrow \infty$ .

The mass identities follow by positive telescoping. For a class,

$$\sum_{j \geq 1} (s_{\mathcal{A},j}^{-1} - s_{\mathcal{A},j+1}^{-1}) = s_{\mathcal{A},1}^{-1}, \quad \sum_{j \geq 1} j(s_{\mathcal{A},j}^{-1} - s_{\mathcal{A},j+1}^{-1}) = \sum_{j \geq 1} s_{\mathcal{A},j}^{-1}.$$

The total germ intensity is at most  $\tau$ , so there are almost surely finitely many germs. The lower bound  $\Theta \geq \delta_{\mathcal{E}}$  follows from the class containing 1.

If  $L_{\xi} = 0$ , every class is a singleton, so  $\lambda_1 = \tau$  and  $\lambda_j = 0$  for  $j \geq 2$ . A nonzero  $v \in L_{\xi}$  has positive and negative parts giving two distinct multipliers of the same phase. Some class then has at least two elements and

$$1 - \Theta(\xi) = \delta_{\mathcal{E}} \sum_{\mathcal{A}} \sum_{j \geq 2} s_{\mathcal{A},j}^{-1} > 0.$$

The void probability yields

$$\liminf_N d_{\text{TV}}(\mathcal{L}(W_N^{\xi}), \text{Pois}(\tau)) \geq e^{-\tau\Theta(\xi)} - e^{-\tau} > 0. \quad (5.10)$$

This proves the theorem without passing from convergence in law to convergence of unbounded moments.

*Remark 5.2* (Order of limits). The prelimit comparison (4.7) is uniform in  $\xi$ . Its reduction to (5.4) has no uniform rate over irrational environments here:  $\Delta_H(\xi)$  can be extremely small. A sequence  $\xi_N$  approaching a resonance at scale  $N^{-1}$  is governed by the prelimit target, not automatically by the fixed-environment classification. Identifying limits for size-dependent environments requires retaining the rescaled angular separations (in turns) in the intersections defining  $K_{N,d}^{\xi_N}$  and proving convergence of the resulting cluster marks and their tails. The fixed-environment classification does not identify all limits when the exceptional phases vary at the target scale.

### 5.3 Finite divisibility certificates and rational jump intensities

A multiplier  $s$  is *nonminimal* when a smaller  $t$  has the same phase. Then  $us$  is nonminimal for every  $u \in \mathcal{D}_{\mathcal{E}}$ . Its exponent vectors form an upward-closed subset of  $\mathbb{N}_0^b$ . Dickson's lemma [5] gives finitely many minimal generators:

$$\{s : s \text{ is nonminimal}\} = u_1 \mathcal{D}_{\mathcal{E}} \cup \cdots \cup u_m \mathcal{D}_{\mathcal{E}}. \quad (5.11)$$

For completeness, every infinite sequence of nonnegative integers has an infinite nondecreasing subsequence. Successive extraction in the  $b$  coordinates rules out an infinite antichain of minimal vectors. Every element lies above a minimal one, by descending among its finitely many divisors.

Since  $\delta_{\mathcal{E}} \sum_{s \in u \mathcal{D}_{\mathcal{E}}} s^{-1} = 1/u$ , inclusion–exclusion gives

$$\Theta(\xi) = \sum_{J \subseteq \{1, \dots, m\}} \frac{(-1)^{|J|}}{\text{lcm}(u_j : j \in J)}, \quad \text{lcm}(\emptyset) = 1. \quad (5.12)$$

Thus  $\Theta(\xi)$  is rational once the complete generator list is certified. A single relation  $\chi_{\xi}(A) = \chi_{\xi}(B)$  with coprime  $A > B$  already gives  $1 - \Theta(\xi) \geq 1/A$ , since all exceptional-prime multiples of  $A$  are nonminimal.

The same argument applies to every fixed jump size. If  $r(s)$  is the rank of  $s$  in its class, put

$$\mathcal{I}_j = \{s : r(s) \geq j\}, \quad w_j = \delta_{\mathcal{E}} \sum_{s \in \mathcal{I}_j} s^{-1}.$$

Multiplying  $j$  distinct predecessors or the element itself by a common multiplier proves that  $\mathcal{I}_j$  is again a divisibility ideal. Its mass is rational. Since elements of rank exactly  $j$  are the  $s_{\mathcal{A},j}$ ,

$$\lambda_j(\xi)/\tau = w_j - 2w_{j+1} + w_{j+2} \in \mathbb{Q}. \quad (5.13)$$

Nonnegativity follows from (5.3). Finiteness is asserted for each fixed  $j$ , not for one certificate necessarily encoding all jump sizes.

### 5.4 The initial ideal behind the certificates

The preceding rationality has a standard commutative-algebra interpretation. We spell out the identification to distinguish it from the arithmetic probability statement. Let  $\mathcal{E} = \{p_1, \dots, p_b\}$  and form the lattice ideal

$$I_L = \langle X^{\alpha} - X^{\beta} : \alpha, \beta \in \mathbb{N}_0^b, \alpha - \beta \in L_{\xi} \rangle \subseteq \mathbb{Q}[X_1, \dots, X_b].$$

This is equivalently generated by  $X^{v^+} - X^{v^-}$ ,  $v \in L_{\xi}$ . Order monomials by the corresponding integer  $p^{\alpha} = \prod_i p_i^{\alpha_i}$ , or by its logarithm. Unique factorization makes the order total; it is multiplicative and well founded, with finitely many monomials below each fixed bound. It is therefore a monomial order.

The initial ideal for this order is exactly the ideal of nonminimal multipliers in (5.11). One inclusion follows because a nonminimal  $p^\alpha$  has a smaller representative  $p^\beta$  in its phase class, making  $X^\alpha$  the leading monomial of  $X^\alpha - X^\beta$ . For the converse, map the polynomial ring to the group algebra  $\mathbb{Q}[\mathbb{Z}^b/L_\xi]$ . Its kernel is  $I_L$ : a polynomial is in the kernel exactly when its coefficients sum to zero in every exponent coset, and within a coset it is a sum of binomial differences. The leading term of a nonzero kernel element must therefore have another, smaller monomial in its coset. Its multiplier is nonminimal. This proof does not require  $L_\xi$  to be saturated; the quotient group may have torsion.

Standard monomials thus represent exactly the smallest multipliers in the phase classes. If  $H_{\mathbb{Q}[X]/\text{in}(I_L)}$  denotes their multigraded generating series, then

$$\Theta(\xi) = \delta_{\mathcal{E}} H_{\mathbb{Q}[X]/\text{in}(I_L)}(p_1^{-1}, \dots, p_b^{-1}).$$

The rationality of monomial-ideal generating series and the theory of lattice and binomial ideals are classical [13,23]. Their identification with the intensity ratio is the application here. Formula (5.12) is the elementary inclusion–exclusion evaluation of this series.

Neither this interpretation nor Dickson’s lemma recognizes exact phase relations from floating-point approximations. A list of proven relations may generate a proper sublattice of  $L_\xi$ ; a rational basis may miss its integer index. A finite search of multipliers does not certify that the generator list is complete. Exact computation becomes possible only after the relevant integer lattice or a complete reduction certificate has been supplied.

## 5.5 A single generating relation: the Pólya–Aeppli law

Suppose  $L_\xi$  has rank one and is nonzero. Choose a generator of the *actual integer lattice*, oriented so its positive and negative parts give

$$A = \prod_{p \in \mathcal{E}} p^{(v_p)^+}, \quad B = \prod_{p \in \mathcal{E}} p^{(-v_p)^+}, \quad A > B \geq 1.$$

The integers  $A, B$  are coprime. The generator need not be primitive in  $\mathbb{Z}^b$ : one phase of finite order  $r$  at prime  $p$  gives  $L_\xi = r\mathbb{Z}$ .

**Corollary 5.3** (One generating relation). *Put  $q_A = 1/A$  and  $p_A = 1 - q_A$ . Then*

$$\Theta(\xi) = p_A, \quad \lambda_j(\xi) = \tau p_A^2 q_A^{j-1}, \quad W^\xi \stackrel{d}{=} \sum_{i=1}^{P_A} G_{A,i}, \quad (5.14)$$

where  $P_A \sim \text{Pois}(\tau p_A)$  and the independent positive geometric marks satisfy  $\mathbb{P}(G_{A,i} = j) = p_A q_A^{j-1}$ .

*Proof.* Ratios of elements in the same class are  $(A/B)^l$ ,  $l \in \mathbb{Z}$ . The smaller predecessors of  $s$  are exactly  $s(B/A)^l$ ,  $1 \leq l \leq \nu_A(s)$ , where  $\nu_A(s) = \max\{l : A^l \mid s\}$ . Coprimality makes integrality equivalent to this divisibility. Hence  $r(s) = 1 + \nu_A(s)$ ,  $\mathcal{I}_j = A^{j-1} \mathcal{D}_{\mathcal{E}}$  and  $w_j = A^{-(j-1)}$ . Apply (5.13).  $\square$

This is the classical Pólya–Aeppli distribution [10]. Here the geometric parameter is determined by the larger integer height  $A$ , not only by  $A/B$ . The smaller integer  $B$  remains visible in the spatial geometry of clusters, but not in this prefix count law.

For explicit probabilities, set  $\lambda = \tau p_A$ ,  $p = p_A$ ,  $q = q_A$ . Conditioning on the number of germs gives, for  $n \geq 1$ ,

$$\mathbb{P}(W^\xi = n) = e^{-\lambda} \sum_{k=1}^n \frac{\lambda^k}{k!} \binom{n-1}{k-1} p^k q^{n-k}, \quad \mathbb{P}(W^\xi = 0) = e^{-\lambda}.$$

Equivalently, the standard Laguerre generating function [14, Eq. 18.12.13] gives

$$\mathbb{P}(W^\xi = n) = e^{-\lambda} q^n L_n^{(-1)}\left(-\frac{\lambda p}{q}\right), \quad n \geq 0.$$

This special-function representation is not an additional probability theorem; it is a coefficient formula for  $\exp\{-\lambda + \lambda pz/(1 - qz)\}$ .

For one exceptional prime  $p$  with a phase of order  $r$ , take  $A = p^r$ ,  $B = 1$ . In particular,  $F(2) = 1$  gives

$$Z_N - \lfloor \log_2 N \rfloor \xrightarrow{d_{\text{TV}}} \sum_{i=1}^P G_i, \quad P \sim \text{Pois}(\tau/2), \quad \mathbb{P}(G_i = j) = 2^{-j}.$$

An infinite-order phase at the sole prime 2 has zero relation lattice, so its corrected count tends to  $\text{Pois}(\tau)$ . In both cases  $B_N(\xi)$  is subtracted exactly; its disappearance does not follow merely from absence of torsion.

## 5.6 Uniform marginals may retain a nontrivial joint lattice

Let  $U$  be Haar, set  $F(2) = F(3) = U$ , and use independent Haar elsewhere. Almost surely  $U$  has infinite order and the lattice is generated by  $(1, -1)$ , with  $A = 3$ ,  $B = 2$ . Yet every  $F(n)$ ,  $n \geq 2$ , is Haar: a prime outside  $\{2, 3\}$  supplies an independent Haar factor, or else  $F(2^a 3^b) = U^{a+b}$  with  $a + b > 0$ .

There are  $O((\log N)^2)$  integers supported on  $\{2, 3\}$ , and

$$\mathbb{E}B_N(U) = \frac{\tau}{N}(|\mathcal{D}_{\{2,3\}} \cap [1, N]| - 1) \longrightarrow 0.$$

The uncorrected count therefore satisfies

$$Z_N \xrightarrow{d_{\text{TV}}} \sum_{i=1}^P G_i, \quad P \sim \text{Pois}(2\tau/3), \quad \mathbb{P}(G_i = j) = \frac{2}{3} \left(\frac{1}{3}\right)^{j-1}. \quad (5.15)$$

The convergence is joint with  $U$ , with an independent limit, because the limiting kernel is the same for almost every  $U$ . The law of  $(U, U)$  is singular relative to product Haar; Proposition 7.3 does not cover it.

For fixed  $\zeta$  of infinite order, the pairs  $(\zeta, \zeta^2)$  and  $(\zeta^2, \zeta^3)$  at primes 2, 3 give the generating relations  $F(4) = F(3)$  and  $F(9) = F(8)$ , hence geometric parameters  $3/4$  and  $8/9$ . The pair  $(\zeta, -\zeta)$  has generating relation  $F(9) = F(4)$ : replacing this by  $F(3) = F(2)$  would be false. The actual integer lattice matters.

## 5.7 A complete certificate at three primes

Fix  $\zeta$  of *infinite order*, let  $\mathcal{E} = \{2, 3, 5\}$  and set  $(\xi_2, \xi_3, \xi_5) = (\zeta, \zeta^2, \zeta^3)$ . Two multipliers  $2^a 3^b 5^c$  have the same phase exactly when their integer weights  $a + 2b + 3c$  coincide. The reductions

$$4 \mapsto 3, \quad 6 \mapsto 5, \quad 10 \mapsto 9, \quad 27 \mapsto 25$$

preserve the weight and strictly decrease the integer. The irreducibles are exactly

$$\{2\} \cup \{3^b 5^c : 0 \leq b \leq 2, c \geq 0\}.$$

Their weights are distinct: the residues  $2b \pmod{3}$  for  $b = 0, 1, 2$  are distinct, and the weight 1 of 2 is separate. Reduction terminates, so each phase class has a unique irreducible, necessarily its minimum. The complete nonminimal ideal generators are 4, 6, 10, 27, and

$$\Theta = \frac{4}{15} \left[ \frac{1}{2} + \frac{1 + 1/3 + 1/9}{1 - 1/5} \right] = \frac{83}{135}. \quad (5.16)$$

The corrected count has limiting void probability  $e^{-83\tau/135}$ . Geometric cluster sizes are not asserted. When both phases at 2, 3 equal 1, the unique class starts 1, 2, 3, 4, 6,  $\dots$ ; its first three size probabilities are  $1/2, 1/6, 1/12$ , another nongeometric example.

## 5.8 Random environments and scope of the classification

For any law of  $\xi$  independent of the regular Haar source, countably many character equalities and nonnegative sums make  $\xi \mapsto \mathcal{L}(W^\xi)$  a measurable kernel. Conditional distances tend to zero for every fixed  $\xi$  and are bounded by one. Dominated convergence gives

$$d_{\text{TV}}(\mathcal{L}(\xi, Z_N - B_N(\xi)), \mathcal{L}(\xi, W^\xi)) \longrightarrow 0. \quad (5.17)$$

In particular,

$$Z_N - B_N(\xi) \xrightarrow{d_{\text{TV}}} \text{Pois}(\tau) \iff \mathbb{P}(L_\xi = \{0\}) = 1. \quad (5.18)$$

If the lattice is nonzero with positive probability, the mixture has void probability  $\mathbb{E}e^{-\tau\Theta(\xi)} > e^{-\tau}$ , although every conditional limiting law has mean  $\tau$ .

A joint law of  $\xi$  absolutely continuous relative to Haar on  $\mathbb{T}^b$  gives  $L_\xi = 0$  almost surely: each nontrivial integer character equals one on a Haar-null set, and the potential relations are countable. Absolute continuity of individual marginals alone does not suffice, as (5.15) shows.

The classification concerns a prefix and one common shrinking arc. A mask may avoid all repeated phases, and other Borel targets may overlap their rotations differently. In those situations the prelimit target (4.6) remains the general object. Finally, (5.4) is a law of the corrected count; it does not assert independent spatial increments within a cluster.

## 6 Shrinking targets of rational observations

This branch uses prime witnesses, Haar characters and elementary counts, not (8.6) or the Runge–Pell analysis in *Long runs and rare patterns* [26]. Rarity comes from a target of mass  $N^{-1}$  rather than accumulated letter constraints.

### 6.1 Valuation relations under Haar

For  $r \in \mathbb{Q}_{>0}$  let  $v(r) = (v_p(r))_p$  be its finitely supported integer valuation vector. Extend  $F$  by  $F(a/b) = F(a)/F(b)$ . Under independent Haar prime values,

$$\mathbb{E} \prod_{j=1}^q F(r_j)^{u_j} = \mathbf{1}_{\{\sum_j u_j v(r_j) = 0\}}, \quad u_j \in \mathbb{Z}. \quad (6.1)$$

Thus the phases are independent and Haar exactly when the valuation vectors are independent over  $\mathbb{Q}$ , the classical character criterion of [24, arXiv version 2, Lemma 2.1].

### 6.2 Sparse supports and a primitive certificate

For  $x \in I_N = [N, 2N) \cap \mathbb{Z}$ , choose distinct integers  $0 \leq s_1 < \dots < s_k \leq CN$  and nonzero integer exponents  $e_j$ , and set

$$R(x) = \prod_{j=1}^k (x + s_j)^{e_j}, \quad T_x = F(R(x)). \quad (6.2)$$

Data may depend on  $N$ ; require  $R(x) \neq 1$  at every site. For a deterministic Borel  $J \subseteq \mathbb{T}$ , let  $a = m(J) > 0$ ,  $\Lambda = Na$ . Under Haar,  $\mathbb{P}(T_x \in J) = a$  at every site, including deleted ones. We retain

$$Z_N = \sum_x \mathbf{1}_{\{T_x \in J\}}, \quad \Xi_N^{\text{ph}} = \sum_x \mathbf{1}_{\{T_x \in J\}} \delta_{(x, T_x)}, \quad \Xi_N^{\text{full}} = \sum_x \mathbf{1}_{\{T_x \in J\}} \delta_{(x, (F(x+s_j))_{j \leq k})}.$$

The full mark records sampled positions, not intermediate values.

Put  $d_0 = \gcd_j |e_j|$ ,  $f_j = e_j/d_0$ ,  $Q(x) = \prod_j (x + s_j)^{f_j}$ . Choose a deterministic set  $J_*$  of  $r$  positions with  $\gcd(|f_j| : j \in J_*) = 1$ ; the full set always works. If each selected position has a private prime  $p_j > Y$  of valuation one,  $v_{p_j}(Q(x)) = f_j$ , so  $v(Q(x))$  is primitive. Two proportional primitive integer vectors are equal or opposite: their ratio's denominator divides the first vector and its numerator the second. By (6.1), two good phases are therefore independent except when

$$R(x) = R(y) \quad \text{or} \quad R(x)R(y) = 1. \quad (6.3)$$

The common power  $d_0$  changes neither rank nor Haar. No original coefficient need be  $\pm 1$ : (2, 3) has a two-position certificate, (6, 10, 15) a three-position one.

### 6.3 The cost follows the support rather than its diameter

A prime  $p > Y$  shared by  $x + s_i, x + s_j$  divides  $s_j - s_i$ . With  $M = \lceil (C + 2)N \rceil$ , a family  $\mathcal{P}$  of position pairs has collision set bounded by

$$|\mathcal{C}_Y(\mathcal{P})| \leq \sum_{(i,j) \in \mathcal{P}} \sum_{\substack{p > Y \\ p | s_j - s_i}} (N/p + 1) \leq |\mathcal{P}| \frac{\log M}{\log Y} (N/Y + 1). \quad (6.4)$$

The  $+1$  is retained; no condition  $Y > s_k - s_1$  is imposed.

Let  $R_Y = \sum_{p > Y} \lfloor M/p^2 \rfloor \ll_C N/(Y \log Y)$ . For phase marks keep sites with all certificate witnesses. For full marks also require no sampled  $Y$ -smooth integer and no large-prime collision. Their complements satisfy

$$|D^{\text{ph}}| \leq r\Psi(M, Y) + rR_Y + |\mathcal{C}_Y(\mathcal{P}_*)|, \quad |\mathcal{P}_*| \leq r(k - 1), \quad (6.5)$$

$$|D^{\text{full}}| \leq k\Psi(M, Y) + rR_Y + |\mathcal{C}_Y(\{(i, j) : i < j\})|. \quad (6.6)$$

Without collisions an integer lacking a simple witness is smooth or divisible by a large square. Positions outside the certificate only require a private prime of nonzero valuation, explaining the coefficient  $r$ , not  $k$ , of  $R_Y$ .

On retained sites phase marginals are conditionally Haar and full-value marginals are  $m^{\otimes k}$ . The large-prime graph obeys (2.7). A shared integer either has an unexposed prime and is detected by the graph, or has its value fixed by the environment. No additional geometric edges are needed.

### 6.4 Poissonizing resonance classes

Equality and inversion in (6.3) define classes  $\mathcal{C}$  on all of  $I_N$ . Write

$$R(x) = r_C^{\sigma_x}, \quad r_C > 1, \quad \sigma_x \in \{-1, +1\}, \quad U_C = F(r_C).$$

A class is active if  $U_C \in K_C := \bigcup_{x \in \mathcal{C}} J^{\sigma_x}$ . Its reference mass  $p_C$  is  $a$  for one orientation,  $2a - b$  for two, where  $b = m(J \cap J^{-1})$ . Thus  $p_C \leq 2a$  and  $p_C \leq |\mathcal{C}|a$ .

The target  $\mathcal{P}_N^{\text{cl}}$  uses independent Poisson germs of intensity  $\mathbf{1}_{K_C} m$  for each class. A germ  $u$  produces

$$\sum_{x \in \mathcal{C}} \mathbf{1}_{\{u^{\sigma_x} \in J\}} \delta_{(x, u^{\sigma_x})}.$$

Every site has intensity  $\mathbf{1}_J m$ . Define

$$\Lambda^{\text{cl}} = \sum_{\mathcal{C}} p_C, \quad \mathfrak{d} = \Lambda - \Lambda^{\text{cl}} = \sum_{\mathcal{C}} \int \left( \sum_{x \in \mathcal{C}} \mathbf{1}_{\{u^{\sigma_x} \in J\}} - 1 \right)_+ dm(u). \quad (6.7)$$

Under Haar,  $\mathfrak{d}$  is exactly the mean number of points removed by retaining only the first active site in each class. A class of size  $m$  contributes  $(m-1)a$  with one orientation and  $(m-2)a+b$  with two.

**Theorem 6.1** (Rational resonances under Haar). *For (6.2) with  $R(x) \neq 1$ , put*

$$\epsilon_N^\star = 2a|D^\star| + 4Na^2 + 16a^2E_Y, \quad \star \in \{\text{ph}, \text{full}\}.$$

*The phase field is within mean conditional total variation  $\epsilon_N^{\text{ph}}$  of  $\mathcal{P}_N^{\text{cl}}$ . The simple full-value target has intensity*

$$\mu_N^{\text{full}} = \sum_x \delta_x \otimes \left( \mathbf{1}_{\{\prod_j z_j^{e_j} \in J\}} m^{\otimes k}(\text{d}z) \right),$$

and

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\Xi_N^{\text{full}} \mid \mathcal{F}_Y), \text{PPP}(\mu_N^{\text{full}})) \leq \epsilon_N^{\text{full}} + 2\mathfrak{d}, \quad (6.8)$$

$$d_{\text{TV}}(\mathcal{L}(Z_N), \text{Pois}(\Lambda)) \geq e^{-\Lambda}(e^{\mathfrak{d}} - 1) - \epsilon_N^{\text{ph}}. \quad (6.9)$$

*Class sizes need not be bounded. If  $\Lambda_N$  is bounded and  $\epsilon_N^{\text{full}} \rightarrow 0$ , then  $\mathfrak{d}_N \rightarrow 0$  is equivalent to the simple Poisson target for both the count and complete marked field, including mean conditional total variation for the latter.*

The following bookkeeping lemma records both the graph contraction and the complete mark used in the proof. Its relative version will also be used under permanent bias.

**Lemma 6.2** (Class contraction and first-active marks). *Fix a deterministic retained set  $G$  and restrict the resonance classes to nonempty  $\mathcal{C}_G = \mathcal{C} \cap G$ . Assume the site graph is conditionally exact for the complete site variables. Join two distinct classes whenever a site edge joins their members. This graph is conditionally exact for any class variables measurable from their complete site variables, and its number of distinct ordered edges is at most that of the site graph.*

*Order each restricted class by its sites and define*

$$L_x = J^{\sigma_x} \setminus \bigcup_{\substack{y \in \mathcal{C}_G \\ y < x}} J^{\sigma_y}, \quad x \in \mathcal{C}_G.$$

*These sets partition  $K_{\mathcal{C}_G} = \bigcup_{x \in \mathcal{C}_G} J^{\sigma_x}$ . Put  $p_{\mathcal{C}_G} = m(K_{\mathcal{C}_G})$  and  $J'_x = \{u^{\sigma_x} : u \in L_x\} \subseteq J$ . If the complete values at site  $x$  have conditional law  $m^{\otimes k}$ , the subprobability of selecting  $x$  as the first active site and recording its values is exactly*

$$\mathbf{1}_{\{\prod_j z_j^{e_j} \in J'_x\}} m^{\otimes k}(\text{d}z).$$

*More generally, if each conditional complete-value density  $h_x^{\mathcal{F}}$  satisfies  $\|h_x^{\mathcal{F}} - 1\|_\infty \leq \Delta$ , let  $\eta_{\mathcal{C}_G}^{\mathcal{F}}$  be the marked subprobability of the selected site and its values, and let  $\eta_{\mathcal{C}_G}^0$  be its Haar reference. Then*

$$\|\eta_{\mathcal{C}_G}^{\mathcal{F}} - \eta_{\mathcal{C}_G}^0\|_{\text{var}} \leq \Delta \sum_{x \in \mathcal{C}_G} m(J'_x) = \Delta p_{\mathcal{C}_G}.$$

*No product law for the complete values of an entire class is required.*

*Proof.* Collections of classes without an edge have unions of sites without an edge; conditional independence of the complete site collections proves the graph assertion. Choose one ordered site-edge witness for each ordered class edge. Distinct class edges have disjoint sets of witnesses, so this choice is injective. For the prime-support graph this is equivalently the graph obtained from the unions of unexposed prime supports.

The identity  $T_x = U_C^{\sigma_x}$  shows that  $x$  is the first active site exactly when  $U_C \in L_x$ , or equivalently  $T_x \in J'_x$ . Selection is thus a restriction of that site's own phase. Its joint law with the full local values is the restriction of their single-site conditional law. The nontrivial character  $z \mapsto \prod_j z_j^{e_j}$  sends  $m^{\otimes k}$  to  $m$ , so the reference restriction has mass  $m(J'_x)$ . The selected-site labels are disjoint; summing their variation norms gives the stated bound. Finally  $u \mapsto u^{\sigma_x}$  preserves Haar, and the  $L_x$  partition the active domain, giving  $\sum_x m(J'_x) = p_{C_G}$ .  $\square$

*Proof of Theorem 6.1.* On good sites each class phase is conditionally Haar. Two distinct class phases are independent after averaging, by the primitive certificate. By Lemma 6.2, the contracted graph is conditionally exact and has at most  $E_Y$  ordered edges. Since  $\sum p_C \leq Na$  and  $p_C \leq 2a$ , Theorem 2.2 gives

$$2 \sum p_C^2 + 4 \sum_{C \sim C'} p_C p_{C'} \leq 4Na^2 + 16a^2 E_Y.$$

Deleting bad sites from real and target fields costs  $2a|D|$ . Restricting the cluster target to good sites produces the directly constructed restricted target: germs producing no remaining point are ignored.

For full marks, Lemma 6.2 identifies the correct first-active marked subprobabilities from the single-site product marginals. Transfer classes with the mark “selected site and its values”. The real loss is  $\mathfrak{d}_G \leq \mathfrak{d}$ . The resulting intensity is a submeasure of  $\mu_N^{\text{full}}$ ; independent Poisson filling costs another  $\mathfrak{d}_G$ , besides the already paid bad sites. This proves (6.8).

The cluster target is empty with probability  $e^{-\Lambda^{\text{cl}}}$ . Compare with  $e^{-\Lambda}$  to obtain (6.9). For  $\Lambda \leq A_0$ , the lower bound is at least  $e^{-A_0} \mathfrak{d} - \epsilon_N^{\text{ph}}$ . The equivalence follows by contraction and these bounds, without using unbounded-moment convergence.  $\square$

## 6.5 Budgets and applications

With  $Y = \exp(\sqrt{\log N \log \log N}/\sqrt{2})$ , the Rankin calculation in Section A gives  $\Psi(M, Y)/N \leq \exp[-(1/\sqrt{2} + o(1))\sqrt{\log N \log \log N}]$ . Thus for  $\log k_N = o(\sqrt{\log N \log \log N})$  and bounded  $\Lambda_N$ ,

$$\epsilon_N^* \ll_{C,c} (\Lambda_N + \Lambda_N^2) e^{-c\sqrt{\log N \log \log N}}, \quad 0 < c < 1/\sqrt{2}. \quad (6.10)$$

The bound is independent of exponent size.

For phase alone a wider range is available with a small certificate. Summing  $\log n$  over smooth integers gives  $\sum \log n \leq M \sum_{p \leq Y} \log p / (p-1) \ll M \log Y$ , hence

$$\Psi(M, Y) \ll \sqrt{M} + M \frac{\log Y}{\log M}.$$

For  $k_N = N^{o(1)}$ , choose  $\log Y = 4 \log(2k_N) + \sqrt{\log N}$ . Other terms vanish and, at bounded intensity,

$$\epsilon_N^{\text{ph}} \ll_{C,A_0} r_N \left( \frac{\log(2k_N)}{\log N} + \frac{1}{\sqrt{\log N}} \right). \quad (6.11)$$

A bounded certificate size gives convergence throughout this range. Full marks retain a factor  $k_N$  in the smooth-number cost and do not inherit this extension for free.

For  $R(x) = x/(x + h_N)$ ,  $1 \leq h_N \leq CN$ , strict increase and values in  $(0, 1)$  imply  $\mathfrak{d}_N = 0$ . Quasi-alignments in a target of mass  $\tau/N$  therefore give  $\text{Pois}(\tau)$  with (6.10), including  $h_N = N$ . Positive-exponent products are similarly injective and exceed one. Multiplicative differences

$$R_{r,h}(x) = \prod_{j=0}^r (x + jh)^{(-1)^{r-j} \binom{r}{j}}, \quad rh \leq CN,$$

are admissible because

$$\log R_{r,h}(x) = (-1)^{r-1}(r-1)! \int_{[0,h]^r} (x + t_1 + \cdots + t_r)^{-r} dt_1 \cdots dt_r$$

has strict sign and derivative of the opposite sign. An endpoint coefficient gives a one-position certificate.

Non-monotonicity can be paid rather than excluded. For  $R_N(x) = x(x+5N)/(x+N)^2 > 1$ , distinct  $x, y$  satisfy

$$R_N(x) = R_N(y) \iff (3x - N)(3y - N) = 16N^2.$$

There are at most  $\tau_0(16N^2) = N^{o(1)}$  ordered coincidences. At  $a = \tau/N$ ,  $\mathfrak{d}_N \leq (a/2)\tau_0(16N^2) = N^{-1+o(1)}$ , so simple Poisson survives. Coincidences genuinely occur: at  $N = 14$ , sites 21, 26 both give 39/25.

In contrast,  $x(x+7N)/(x+2N)^2 = 1$  at  $x = 4N/3$  when  $3 \mid N$ . If  $1 \in J_N$ , that site is certain and the count's distance to  $\text{Pois}(\tau)$  is at least  $e^{-\tau}$ . A single site of proportion  $N^{-1}$  cannot be deleted at its fictitious Haar mass.

## 7 Rational observations under permanent prime bias

Retain (6.2), but use independent prime laws  $g_p m$  with  $0 \leq g_p \leq K$  and  $\int g_p dm = 1$ , for fixed  $K$ . Equality and inversion classes of  $R_N(x)$  remain exact. Their phases are no longer exactly Haar, and rational rank does not imply statistical independence. The two different bounds (7.2) and (7.3) replace these properties.

### 7.1 Uniformly bounded prime densities

Suppose  $g_p \geq 0$ ,  $\int g_p dm = 1$ ,  $g_p \leq K$ , with fixed  $K$ . The image under  $z \mapsto z^e$ ,  $e \neq 0$ , averages over  $|e|$  preimages and remains bounded by  $K$ . For every nonzero integer  $n$ ,

$$|\widehat{g}_p(n)| \leq \rho_K := 1 - \frac{1}{4K^2} < 1, \quad \sum_{n \neq 0} |\widehat{g}_p(n)|^2 \leq K - 1. \quad (7.1)$$

For the first bound push by the  $n$ th power, then rotate to make the first coefficient nonnegative real. An arc about 1 of normalized length  $1/(2K)$  has mass at most  $1/2$ . The real moment is at most  $(1 + \cos(\pi/(2K)))/2 \leq 1 - 1/(4K^2)$ . Parseval and  $\int g_p^2 \leq K$  give the second bound.

**Lemma 7.1** (Uniform convolution bound). *For  $t \geq 2$  independent prime variables of densities bounded by  $K$ , raised to arbitrary nonzero integer powers, the product density satisfies*

$$\|h_t - 1\|_\infty \leq (K - 1)\rho_K^{t-2}. \quad (7.2)$$

*A rotation or an additional independent convolution preserves this bound.*

*Proof.* The Fourier coefficient at  $n$  is  $\prod_i \widehat{g}_i(e_i n)$ . Use the gap on  $t-2$  factors and Cauchy–Schwarz on the last two:

$$\sum_{n \neq 0} \prod_i |\widehat{g}_i(e_i n)| \leq \rho_K^{t-2} \left( \sum_{n \neq 0} |\widehat{g}_1(e_1 n)|^2 \right)^{1/2} \left( \sum_{n \neq 0} |\widehat{g}_2(e_2 n)|^2 \right)^{1/2} \leq (K - 1)\rho_K^{t-2}.$$

The Fourier series is absolutely convergent and represents the density. Convolution with a probability contracts uniform distance to Haar.  $\square$

For  $r \neq 1$ , fixing all but one prime with nonzero valuation gives  $\mathcal{L}(F(r)) \leq Km$ . If  $v(r), v(s)$  have rank two, extract two primes giving a nonsingular integer matrix. Its map on  $\mathbb{T}^2$  is surjective and preserves Haar. Pushing forward the input domination and integrating other coordinates gives

$$\mathcal{L}(F(r), F(s)) \leq K^2 m^{\otimes 2}. \quad (7.3)$$

This replaces exact independence. No unimodular determinant is required on the torus, but the same step is not automatic on a finite group.

## 7.2 Prime factors and relative precision

Use the common cutoff band (2.14)–(2.15), the function  $I(\theta) = 1 - \theta + \theta \log \theta$ , and  $t_N = \lfloor \theta A_N \rfloor$ . The abundance estimate (2.16) applies uniformly to every translated interval used below. Its proof, including the growing-order congruence budget, is in Section B.

A good site has at least  $t_N$  factors in  $(Y, Z]$  in each sampled integer, no prime square  $p^2$  with  $p > Y$  in those integers, and no large-prime collision between positions. Each coordinate then has  $t_N$  private simple prime factors. By (6.4) and (2.16), the complement obeys

$$\frac{|D_N|}{N} \ll_{C,\theta} k e^{-I(\theta)A_N} + \frac{k}{Y \log Y} + k^2 \frac{\log M}{\log Y} \left( \frac{1}{Y} + \frac{1}{N} \right). \quad (7.4)$$

These witnesses also make the normalized valuation vector primitive without a coefficient- $\pm 1$  assumption.

Let  $\delta_N = (K - 1)\rho_K^{t_N - 2}$  and  $\Delta_N = (1 + \delta_N)^k - 1$ . After fixing all other prime coordinates, each component density lies in  $[1 - \delta_N, 1 + \delta_N]$ , for  $\delta_N \leq 1$ . The private coordinates are disjoint, so integration preserves

$$\|h_x^{\mathcal{F}_Y} - 1\|_\infty \leq \Delta_N. \quad (7.5)$$

Restriction to the rare event costs  $a\Delta_N$ . One component already mixes the monomial phase itself, whose conditional activation probability is at most  $(1 + \delta_N)a$ , without a factor  $k$ .

## 7.3 The quantitative theorem and its resonances

The Haar targets and deficit  $\mathfrak{d}_N$  remain those of Section 6. Under bias the deficit is a reference quantity, no longer the exact expected real first-active loss.

**Theorem 7.2** (Robustness under bounded densities). *Fix  $C, K, A_0$ ,  $0 < \theta < 1$  and*

$$0 < \kappa < \beta_K(\theta) := \min\{I(\theta), -\theta \log \rho_K\}.$$

*Assume  $R_N(x) \neq 1$ ,  $\Lambda_N \leq A_0$  and  $1 \leq k_N \leq (\log N)^\kappa$ . A common envelope*

$$\epsilon_N \ll_{C,K,A_0,\theta,\kappa} (\Lambda_N + \Lambda_N^2) k_N (\log N)^{-\kappa} \quad (7.6)$$

*controls mean conditional total variation of the phase field from its Haar cluster target, and*

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\Xi_N^{\text{full}} \mid \mathcal{F}_Y), \text{PPP}(\mu_N^{\text{full}})) \leq \epsilon_N + (2 + \delta_N) \mathfrak{d}_N \leq \epsilon_N + 3\mathfrak{d}_N, \quad (7.7)$$

$$d_{\text{TV}}(\mathcal{L}(Z_N), \text{Pois}(\Lambda_N)) \geq e^{-\Lambda_N} (e^{\mathfrak{d}_N} - 1) - \epsilon_N. \quad (7.8)$$

*If  $k_N = o((\log N)^\kappa)$ , vanishing  $\mathfrak{d}_N$  is equivalent to the simple target for the count and full marked field.*

*Proof.* At every site  $R_N(x) \neq 1$ , so the scalar density domination gives  $\mathbb{P}(A_x) \leq Ka$ . Real and target deletions cost  $(K + 1)a|D_N|$ . On good sites complete subprobabilities differ from Haar

by at most  $a\Delta_N$ . Two distinct classes have a pair density bounded by  $K^2m^{\otimes 2}$ , by primitivity and (7.3).

Apply the interface to classes with their true conditional marginals, then change to their references. The bounds  $p_C \leq 2a$  and  $\sum p_C \leq Na$  give

$$\begin{aligned} \epsilon_N &= (K+1)a|D_N| + \Lambda\Delta_N + 4(1+\delta_N)^2Na^2 \\ &\quad + 8((1+\delta_N)^2 + K^2)a^2E_Y. \end{aligned} \tag{7.9}$$

Pair probabilities are averaged after conditioning. A constant times  $a^2$  suffices because the graph is sparse; no pointwise  $o(a^2)$  excess is needed.

By Lemma 6.2, first-active selection remains a restriction  $J'_x$  of the selected site's own phase. Its marked subprobability error is at most  $\Delta_N m(J'_x)$ , summing to  $\Delta_N p_C$  over a class. Real selection loss is at most  $(1+\delta_N)\mathfrak{d}_G$ ; filling the missing reference intensity costs  $\mathfrak{d}_G$ . This proves (7.7); the cluster target's void gives (7.8).

Choose  $\kappa'$  strictly between  $\kappa$  and  $\beta_K(\theta)$ . Since  $A_N/V \rightarrow 1$ ,  $e^{-I(\theta)A_N} + \delta_N \ll U^{-\kappa'}$ . Thus  $k\delta_N \rightarrow 0$  uniformly and  $\Delta_N \leq 2k\delta_N$ . The remaining collision term is  $O(k^2/(U^3V))$  and the graph term is smaller. Since  $\kappa < \log(4/3) < 0.288$ , these and the  $N^{-1}$  terms fit  $kU^{-\kappa}$ . This proves (7.6).  $\square$

The exponent is sufficient, not optimal. With  $\theta = 1/2$  one may use

$$\kappa < \min \left\{ \frac{1 - \log 2}{2}, -\frac{1}{2} \log \left( 1 - \frac{1}{4K^2} \right) \right\}.$$

Within this cutoff argument the best threshold balances  $I(\theta) = -\theta \log \rho_K$ . Under Haar, (6.10) is stronger.

## 7.4 Permanent bias and global singularity

For the same law  $g(e^{it}) = 1 + \frac{1}{2} \cos t$  at all primes,  $K = 3/2$ ,  $\rho_K = 8/9$ , and  $\kappa = 1/20$  is valid at  $\theta = 1/2$ . For  $1 \leq h_N \leq CN$  and  $m(J_N) = \tau/N$ , hits of  $F(x)/F(x+h_N)$  have error  $O((\log N)^{-1/20})$  from  $\text{Pois}(\tau)$ , also retaining both values. Positive-exponent products at  $k_N \leq (\log N)^{1/100}$  positions have error  $O((\log N)^{-1/25})$  with full marks.

This is not a global change-of-measure corollary. For  $g_b(e^{it}) = 1 + b \cos t$ ,  $b \neq 0$ , the average of the enumerated prime values converges almost surely to  $b/2$ , and to zero under Haar. The two infinite-sequence laws are singular. The proof reconstructs local uniformity directly under bias. It also permits densities vanishing on part of the circle, such as the uniform law on a semicircle, and densities varying with the prime.

Global singularity also occurs in the Poisson-generic product measures of Hochman–Paviato [18], but there the independent letter biases decay. Here the prime bias may persist: products of many private prime coordinates reconstruct local Haar marginals. Neither comparison supplies the other's hypotheses or observable.

## 7.5 The target can expose an exact relation

For  $U_x = F(x)/F(x+h_N)$ , record both  $U_x$  and  $U_x^{-1}$ . Apply the phase theorem to  $J_N \cup J_N^{-1}$  and retain channel labels. With  $b_N = m(J_N \cap J_N^{-1})$ ,

$$(Z_N^+, Z_N^-) \approx (P_{10} + P_{11}, P_{01} + P_{11}),$$

where the three target variables are independent and

$$P_{10}, P_{01} \sim \text{Pois}(N(a_N - b_N)), \quad P_{11} \sim \text{Pois}(Nb_N).$$

If  $a_N = \tau/N$  and  $Nb_N \rightarrow \vartheta$ , the total tends to  $P_1 + 2P_2$  with independent parameters  $2(\tau - \vartheta)$  and  $\vartheta$ . An arc of length  $\tau/N$  centred, in turns, at  $c/N$  has  $Nb_N = (\tau - 2|c|)_+$ . The same inversion relation yields duplication, a common component or asymptotically independent channels according to the target. Under bias these are reference parameters, not exact finite-size probabilities.

## 7.6 A fixed global change of measure

**Proposition 7.3** (Prime-coordinate density). *Work on the canonical prime-value space, with completed sigma-field  $\mathcal{F}_\infty = \sigma(F(p) : p \text{ prime})$ . Suppose  $Y_N \rightarrow \infty$  and, for deterministic laws  $Q_N$ ,*

$$e_N = d_{\text{TV}}(\mathcal{L}_{\mathbb{P}}(X_N | \mathcal{F}_{Y_N}), Q_N), \quad \mathbb{E}e_N \rightarrow 0.$$

*For a fixed density  $W \geq 0$ ,  $\mathbb{E}W = 1$ , measurable with respect to the completed  $\mathcal{F}_\infty$ , the law  $\tilde{\mathbb{P}} = W\mathbb{P}$  satisfies  $d_{\text{TV}}(\mathcal{L}_{\tilde{\mathbb{P}}}(X_N), Q_N) \rightarrow 0$ . On an enlarged space the same measurability assumption on  $W$  is required.*

*Proof.* Let  $W_Y = \mathbb{E}(W | \mathcal{F}_Y)$ . The law  $W_Y\mathbb{P}$  preserves conditional kernels and

$$d_{\text{TV}}(\mathcal{L}_{\tilde{\mathbb{P}}}(X_N), Q_N) \leq \frac{1}{2}\mathbb{E}|W - W_{Y_N}| + \mathbb{E}(W_{Y_N}e_N).$$

The first term vanishes because conditional expectations converge in  $L^1$  to  $\mathbb{E}(W | \mathcal{F}_\infty) = W$ . The family  $(W_Y)$  is uniformly integrable; truncate at a fixed height and use  $0 \leq e_N \leq 1$  for the second term. If  $W \in L^r$ ,  $r > 1$ , it is at most  $\|W\|_r(\mathbb{E}e_N)^{1-1/r}$ .  $\square$

Without measurability the abstract statement is false: adjoin an independent  $B \sim \text{Bernoulli}(1/2)$ , put  $X_N = B$ ,  $Q_N = \text{Bernoulli}(1/2)$  and  $W = 2\mathbf{1}_{\{B=1\}}$ . Then  $e_N = 0$ , but under  $W\mathbb{P}$  the variable is identically one, at distance  $1/2$  from  $Q_N$ . If all  $X_N$  are prime-coordinate measurable one may instead project  $W$  onto  $\mathcal{F}_\infty$  at the start.

The density concerns the entire prime sequence and is fixed, not an arbitrary  $W_N$  concentrated on rare events. Finite-dimensional densities alone are insufficient. This conditional proof is in the spirit of Eagleson [12,15]; the infinite-product distinction is classical [19]. Permanent independent biases may be globally singular and are handled by the different local reconstruction in Section 7.

## 8 Borel patterns and coordinate-erasure cost

This branch uses the arithmetic estimates of [26]. The weight base varies but the relation space stays binary. Projection, integer comparison and deletion are proved here; the imported Diophantine population estimates are recorded precisely in Section C.

### 8.1 Value groups and integer relations

For a finite abelian group  $H$  of order  $h$ , define

$$\mathcal{K}_H(n_1, \dots, n_b) = \{(\chi_j) \in \hat{H}^b : \prod_j \chi_j^{v_p(n_j)} = 1 \text{ for every } p\}.$$

The image of the prime-value map is a subgroup of  $H^b$  with annihilator  $\mathcal{K}_H$ . Hence

$$\mathbb{P}((F(n_j))_j = w) = h^{-b} |\mathcal{K}_H| \mathbf{1}_{\{w \text{ compatible}\}}. \quad (8.1)$$

For cyclic prime order  $\ell$ ,  $|\mathcal{K}_H| = \ell^{\rho_\ell}$ ; in general we retain its cardinality rather than a dimension over a field.

An integer minor nonzero modulo 2 is nonzero over  $\mathbb{Q}$ , so

$$b - \text{rank}_{\mathbb{Q}} M \leq b - \text{rank}_{\mathbb{F}_2}(M \bmod 2). \quad (8.2)$$

This does not compare arbitrary finite characteristics. The valuations of 12, 18 give  $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ , invertible modulo 2 and singular modulo 3.

## 8.2 Events, complete marks and projections

Fix  $K = H \times \mathbb{T}^d$ , with  $H$  finite abelian (possibly trivial) and  $d \geq 0$  fixed. Prime values are independent of Haar law  $\mu$  on  $K$ . For  $I_N = [N, 2N) \cap \mathbb{Z}$  put

$$V_x = \{x - 1, \dots, x + B - 2\}, \quad \mathbf{F}_x = (F(x - 1), \dots, F(x + B - 2)),$$

where  $\beta_- \log N \leq B \leq \beta_+ \log N$ ,  $0 < \beta_- < \beta_+$  fixed. Let  $A = A_N \subseteq K^B$  be Borel with  $a = \mu^{\otimes B}(A) > 0$ ,  $\Lambda = Na$ , and set

$$\Xi_N = \sum_{x \in I_N} \mathbf{1}_{\{\mathbf{F}_x \in A\}} \delta_{(x, \mathbf{F}_x)}, \quad \nu_A = \mathbf{1}_A \mu^{\otimes B}.$$

The target is  $\text{PPP}(\sum_x \delta_x \otimes \nu_A)$ .

**Definition 8.1** (Erasure cost). For  $0 \leq r \leq B$ , define

$$D_r(A) = \frac{1}{a} \max_{\substack{J \subseteq \{0, \dots, B-1\} \\ |J|=B-r}} (\mu^{\otimes J})^*(\text{pr}_J A). \quad (8.3)$$

An erasure envelope  $(C_0, z)$  means fixed  $C_0 \geq 1$ ,  $z > 1$  with

$$D_r(A_N) \leq C_0 z^r \quad (0 \leq r \leq B_N). \quad (8.4)$$

Outer measure avoids assuming a Borel projection is Borel. The examples have measurable projections. The actual sets matter, not only their classes modulo Haar-null sets, since a singular arithmetic law can detect such a set. We have  $D_0 = 1$ ,  $D_B = 1/a$  and  $az^B \geq C_0^{-1}$ .

In the independent reference put

$$o_s(A) = \mu^{\otimes(B+s)} \{z : (z_0, \dots, z_{B-1}) \in A, (z_s, \dots, z_{s+B-1}) \in A\},$$

$$\Omega(A) = a^{-1} \sum_{s=1}^{B-1} o_s(A), \quad \mathfrak{m} = \max\{1, az^B\}. \quad (8.5)$$

The effective size  $\mathfrak{m}$  belongs to the proved envelope, not a claimed canonical optimum.

## 8.3 The binary profile at a variable weight base

For separated windows let  $M_{x,y}$  have prime rows and the  $2B$  occurrences as columns, and write  $\rho_2(x, y) = 2B - \text{rank}_{\mathbb{F}_2}(M_{x,y} \bmod 2)$ .

**Proposition 8.2** (Variable-base arithmetic profile). *In the fixed logarithmic band, for fixed  $z > 1$ , any  $\varepsilon > 0$ , uniformly in  $T \geq 1$ ,*

$$\mathcal{W}_{2,z}^{[T]} := \sum_{|x-y| \geq B} \min\{T, z^{\rho_2(x,y)} - 1\} \ll N^\varepsilon \left( N^{3/2} z^{B/6} + N z^{2B/3} + N^{2/3} \min\{T, z^B\} \right). \quad (8.6)$$

Constants depend on  $z$ , the band and  $\varepsilon$ , not on  $T$ .

*Derivation from the source population estimates.* Retain the rational subcode, its dimension  $\sigma$ , residual dimension  $\tau$  and eight-sector partition of [26, Section 3]. The identity is

$$z^\rho - 1 = (z^\sigma - 1) + z^\sigma(z^\tau - 1).$$

The population ledger in Section C.2 separates host counts, dimensions and the placement of the cap, with exact locations in [26]. Its bounds retain  $N^{o(1)}$  factors and the stated  $N^\varepsilon$  at the energy step. The three surviving envelopes are  $N^{3/2}z^{B/6}$ ,  $Nz^{2B/3}$  and  $N^{2/3}\min(T, z^B)$ . The following argument records why the reweighting and cap preserve their uniformity. For fixed  $z$ ,  $z^{o(B)} = N^{o(1)}$ . In sectors 1 and 2,  $z^\sigma \leq z(z^\sigma - 1)/(z - 1)$  when  $\sigma > 0$ , so  $z > 1$  suffices. No support count changes characteristic.

The terms carrying  $z^B$  have host bounds fixed before the cap. Thus  $\min(T, Ez^B) \leq E\min(T, z^B)$ , with  $E \geq 1$  independent of  $T$ , preserves uniformity. The two terminal pointwise bounds are  $z^{B+2}$  and  $z^{2B/3+8/3}$ . Cap the first; leave the second and other populations uncapped. Their masses fit (8.6).

The two block-parity conditions remove at most two dimensions. If  $\rho_{\text{rel}}$  includes them,

$$z^{\rho_2} - 1 \leq z^2(z^{\rho_{\text{rel}}} - 1) + (z^2 - 1)\mathbf{1}_{\{\rho_2 > 0\}}.$$

The same cap inequality applies to the first term. The last indicator totals  $N^{3/2+o(1)}$  by [26, Proposition 3.7], and is absorbed by the first monomial. This is the prescribed-value step of [26, Propositions 3.24–3.25]. The proof consumes the population-level arguments, not a substitution into the final binary numerical bound.  $\square$

## 8.4 An integer defect common to all value groups

Set  $H_{\mathbb{Z}}(x, y) = \mathbf{1}_{\{M_{x,y}^T \mathbb{Z}^P \neq \mathbb{Z}^{2B}\}}$ . This includes rational rank deficiency and a nontrivial integer index at full rank. If  $H_{\mathbb{Z}} = 0$ , an integer right inverse makes the value map surjective on  $K^{2B}$  for every compact abelian  $K$ . The values are then unconditionally independent of Haar law.

**Lemma 8.3.** *In the logarithmic band,*

$$\sum_{|x-y| \geq B} H_{\mathbb{Z}}(x, y) \leq N^{3/2+o(1)}. \quad (8.7)$$

*Proof.* A proper integer image has a nonzero kernel modulo some prime  $\ell$ : use a prime divisor of its index at full rank, or reduction modulo 2 at deficient rational rank. Choose an occurrence  $n$  with nonzero coefficient and let

$$J_B(n) = \prod_{\substack{p > B \\ v_p(n)=1}} p.$$

Every such prime must divide an occurrence in the other window. There are at most  $B^{\omega(J_B(n))}$  assignments to its positions; the Chinese remainder theorem then fixes the partner start modulo  $J_B(n)$ , with at most  $N/J_B(n) + 1 \leq 4N/J_B(n)$  choices. Orientation and anchor position cost  $O(B)$ , hence

$$\sum H_{\mathbb{Z}} \ll BN \sum_{n \leq 3N} \frac{B^{\omega(J_B(n))}}{J_B(n)}.$$

Write uniquely  $n = urs$ , where  $u$  contains small primes,  $r = J_B(n)$  and  $s$  is powerful, supported on primes  $> B$  with exponents at least two. Powerful integers up to  $X$  number  $O(\sqrt{X})$ , by writing them as  $a^2b^3$  with  $b$  squarefree. Dropping coprimality restrictions,

$$\sum_{n \leq 3N} \frac{B^{\omega(J_B(n))}}{J_B(n)} \ll \sqrt{N} \prod_{p \leq B} (1 - p^{-1/2})^{-1} \prod_{p > B} (1 + Bp^{-3/2}) \leq \sqrt{N} e^{O(\sqrt{B})}.$$

Now  $Be^{O(\sqrt{B})} = N^{o(1)}$ . No sum over characteristics was inserted: one divisibility condition captures all certificates.  $\square$

### 8.5 A unimodular minor after controlled erasure

Assume every integer in the two windows has a prime factor  $> B$ . Remove the  $t$  columns containing a square of such a prime. A remaining large-prime row has one or two entries, all equal to one, at most one per window. Two-entry rows form a bipartite graph; one-entry rows pin a vertex. No retained vertex lacks a constraint.

Let  $c$  be the number of unpinned components. In a pinned component select a pinning row and a spanning tree. In a free component remove one vertex and select a spanning tree. Reverse the signs of second-window columns; leaf elimination gives a minor of determinant  $\pm 1$  on  $2B - t - c$  coordinates. These values are independent Haar in every compact abelian group, even after fixing the other prime coordinates.

For clarity, let  $M'$  be the restriction to the retained columns. The kernel of its large-prime equations has dimension  $c$  over  $\mathbb{F}_2$ . Adding small-prime rows removes at most  $\pi(B)$  dimensions, and extending a vector in  $\ker M'$  by zeros on the deleted columns injects it into the full kernel. Consequently

$$c - \pi(B) \leq \dim \ker M' \leq \rho_2(x, y), \quad c \leq \rho_2(x, y) + \pi(B).$$

There is no additional  $t$  in this inequality;  $t$  is already paid by erasure. Projecting the two events onto the retained independent coordinates gives

$$\mathbb{P}(\mathbf{F}_x \in A, \mathbf{F}_y \in A) \leq C_0^2 a^2 z^{t+c} \leq C_0^2 a^2 z^{t+\pi(B)+\rho_2(x,y)}. \quad (8.8)$$

For non-Borel projections take Borel envelopes and then the infimum of their masses. This erases observed constraints; it does not condition deleted values to the identity.

### 8.6 The pattern theorem

**Theorem 8.4** (Controlled-erasure patterns). *In the preceding framework, under (8.4) and Proposition 8.2, for  $0 < c < 1/\sqrt{2}$  and  $0 < \varepsilon < 1/3$ ,*

$$\begin{aligned} & \mathbb{E} d_{\text{TV}} \left( \mathcal{L}(\Xi_N \mid \mathcal{F}_Y), \text{PPP} \left( \sum_{x \in I_N} \delta_x \otimes \nu_A \right) \right) \\ & \ll \Lambda \Omega(A) + \Lambda(1 + \Lambda) e^{-c\sqrt{\log N \log \log N}} + N^{-1/3+\varepsilon} \left( \Lambda + \Lambda^{11/6} \mathbf{m}^{1/6} + \Lambda^{4/3} \mathbf{m}^{2/3} \right), \end{aligned} \quad (8.9)$$

where  $Y = \exp(\sqrt{\log N \log \log N}/\sqrt{2})$ . Constants may depend on  $H, d, C_0, z, \beta_-, \beta_+, c, \varepsilon$ , not on the Borel description of  $A_N$ . If  $\Lambda_N$  stays in a compact subset of  $(0, \infty)$ ,  $\mathbf{m}_N \leq N^{1/2-\eta}$ ,  $0 < \eta < 1/2$ , and  $\Omega(A_N) \rightarrow 0$ , the error is

$$O(\Omega(A_N) + e^{-c\sqrt{\log N \log \log N}} + N^{-\eta/2}). \quad (8.10)$$

*Proof.* Section A proves the needed selection bounds. Put  $r_N = \lceil \sqrt{2 \log N / \log \log N} \rceil$ . Retain windows having fewer than  $r_N$  vertices carrying a large-prime square, and a valuation-one prime  $> Y$  at every vertex. Their complement  $D$  satisfies  $|D|/N \ll e^{-c\sqrt{\log N \log \log N}}$ . The weighted one-window defect satisfies

$$\mathcal{U} := \sum_x (z^{u_B(x)} - 1) \leq N^{1/2+o(1)},$$

and real and target deletions cost  $O(a\mathcal{U} + 2a|D|)$ . The defect  $u_B(x)$  is defined there using only 2 and the prime divisors of  $|H|$ .

On a retained window, private valuation-one primes give the full conditional law  $\mu^{\otimes B}$ . Overlapping unions have the same property since  $Y > 2B$ . Join overlapping windows or windows sharing a prime  $> Y$ . This is the exact marked graph. Local pairs cost  $2\Lambda\Omega(A)$ ; separated edges obey (2.7) with  $k = B$ .

Let  $\Gamma_N = C_0^2 z^{\pi(B)+2r_N} = N^{o(1)}$ ,  $T = 1/a$ . Integer surjectivity, (8.8) and the marginal cap give, after averaging the environment,

$$\mathbb{P}(\mathbf{F}_x \in A, \mathbf{F}_y \in A) \leq a^2[1 + \Gamma_N \min\{T, z^{\rho_2(x,y)} - 1\} + (\Gamma_N - 1)H_{\mathbb{Z}}(x, y)]. \quad (8.11)$$

Theorem 2.2 yields the finite ledger

$$\begin{aligned} \mathbb{E}d_{TV} &\ll a\mathcal{U} + a|D| + \Lambda\Omega(A) \\ &\quad + a^2 \left( NB + E_Y + \Gamma_N \mathcal{W}_{2,z}^{[1/a]} + (\Gamma_N - 1) \sum H_{\mathbb{Z}} \right). \end{aligned}$$

Crucially  $\Gamma_N$  does not multiply the baseline  $a^2 E_Y$ . Its logarithm can exceed  $\sqrt{\log N \log \log N}$ ; attaching it to the whole graph would destroy the desired subexponential error. It multiplies only polynomially saving arithmetic terms.

Write  $Q = z^B$ ,  $m_0 = aQ \leq \mathbf{m}$ . After multiplying the first two monomials by  $a^2$ , they equal

$$N^{-1/3} \Lambda^{11/6} m_0^{1/6}, \quad N^{-1/3} \Lambda^{4/3} m_0^{2/3}.$$

The terminal term is at most  $N^{-1/3} \Lambda$ , whether or not  $Q \geq T$ . The host sum (8.7) fits the first monomial. Start with an arithmetic exponent smaller than  $\varepsilon$  to absorb  $\Gamma_N$  and other subpolynomial factors. Deletions and edges give  $\Lambda(1 + \Lambda)e^{-c\sqrt{\log N \log \log N}}$ ; one-window defects cost  $\Lambda N^{-1/2+o(1)}$ . This proves (8.9). At bounded intensity,  $\varepsilon = \eta/6$  gives exponents at most  $-1/4, -\eta/2, -1/3 + \eta/6$ , proving (8.10).  $\square$

For  $1 < z < 2$  we never use the false symmetric estimate  $|P/a - 1| \leq z^u - 1$  when  $P$  might vanish. Projection gives an upper bound and exceptional mass is paid separately. The fixed group  $H \times \mathbb{T}^d$  is used in the weighted one-window argument; the universal integer host count alone does not extend the theorem to arbitrary profinite groups.

## 9 Dictionaries, quotients, cells and filtered relations

### 9.1 Finite dictionaries and cell words

Let  $H$  be a nontrivial finite abelian group,  $h = |H|$ , and  $W \subseteq H^B$  a nonempty dictionary of size  $m_W$ . The event  $A = W$  has mass  $a = m_W h^{-B}$ . Erasing  $r$  positions leaves at most  $m_W$  words of length  $B - r$ , so  $D_r \leq h^r$  and  $\mathbf{m} = m_W$ .

For  $1 \leq s < B$ , write  $w \sim_s v$  when the suffix of  $w$  of length  $B - s$  equals the corresponding prefix of  $v$ . Then

$$\Omega_H(W) = \frac{1}{m_W} \sum_{w,v \in W} \sum_{s=1}^{B-1} h^{-s} \mathbf{1}_{\{w \sim_s v\}}. \quad (9.1)$$

This includes self-overlaps and overlaps between distinct words.

**Corollary 9.1** (Complete dictionary field). *Fix  $H$  and  $0 < \eta < 1/2$ . If*

$$m_N \leq N^{1/2-\eta}, \quad B_N = \log_h(Nm_N) + O(1), \quad \Omega_H(W_N) \rightarrow 0,$$

*the field  $I_{x,w} = \mathbf{1}_{\{\mathbf{F}_x=w\}}$  is within  $O(\Omega_H(W_N) + e^{-c\sqrt{\log N \log \log N}} + N^{-\eta/2})$  of independent  $\text{Pois}(h^{-B_N})$  coordinates, for  $x \in I_N, w \in W_N$ . The same bound holds in mean conditional total variation at the cutoff of Theorem 8.4.*

*Proof.* Apply Theorem 8.4 to  $A = W$  and identify the complete mark with its word label. All hypotheses follow from the projection count and (9.1).  $\square$

Only the error sums are aggregated; positions and identities remain. The independent-letter field has only local dependencies, with error  $O(\Lambda\Omega_H + \Lambda^2 B/N)$ , so the two rare-word fields are also close to each other.

The same proof applies to a partition into  $q$  Borel cells of measure  $1/q$  on a torus or on  $H \times \mathbb{T}^d$ . For a dictionary of cell labels use  $z = q$ ,  $a = |W|q^{-B}$  and  $\mathfrak{m} = |W|$ . The mark may retain all exact phases. The label sequence need not itself be multiplicative. On the circle exact equalities of two distinct values have probability zero under Haar, because  $F(n)/F(m)$  is Haar for  $n \neq m$ ; cells provide a nondegenerate notion of a constant pattern.

## 9.2 The cubic example and excess lengths

For  $H = \{1, \omega, \omega^2\}$ ,  $\omega = e^{2\pi i/3}$ , the corollary gives ternary word fields. An explicit overlap-free family uses words  $1^k \omega u \omega$ , where  $k = \lceil \log_3(2B) \rceil$ ,  $|u| = B - k - 2$ , and  $u$  contains no  $1^k$ . The marker occurs only at the beginning. A short overlap contradicts the final letter; a long one creates an interior marker. A union bound retains at least half the interiors, giving at least  $3^B/(108B)$  words for large  $B$ . This is a ternary specialization of the cross-bifix-free marker construction of [9], not a new coding principle.

For starts of constant runs of length at least  $L$ , use

$$W_L = \{(a, b, \dots, b) : a, b \in H, a \neq b\}, \quad B = L + 1.$$

Its mass and overlap weight both equal  $(h-1)h^{-L}$ . Shifts below  $L$  are incompatible; at shift  $L$  each word has  $h-1$  continuations. Consequently

$$Z_{N,L}^{(H)} \approx \text{Pois}((h-1)Nh^{-L}), \quad L = \log_h N + O(1). \quad (9.2)$$

The factor 2 for the cubic case is essential.

One may retain the run value  $b$  and exact excess  $e \geq 0$ . The event with a left mismatch,  $L+e$  consecutive values equal to  $b$ , and a right mismatch has elementary target intensity

$$(h-1)^2 h^{-(L+e+2)}. \quad (9.3)$$

For all excess marks, take  $E = \lceil \sqrt{\log N \log \log N} / \log h \rceil$  and a dictionary of length  $L + E + 2$  imposing only the initial start. It has  $(h-1)h^{E+2} = N^{o(1)}$  words. Their full labels determine excess up to  $E$  or a censored mark; the overlap is  $(E+2)(h-1)h^{-L}$ . Apply Corollary 9.1. The real and target excess tails cost  $O(Nh^{-L}h^{-E})$  by the one-window estimates of Section A. The final error is  $O(e^{-c\sqrt{\log N \log \log N}})$ .

In the target the run value is uniform on  $H$  and independent of excess, with  $\mathbb{P}(e = j) = (h-1)h^{-(j+1)}$ . Counting all constant windows rather than starts gives compound-Poisson clusters of size  $e+1$ . Runs crossing an endpoint cost  $N^{-1+o(1)}$  by the same marginal bound. This is not a classification of arbitrary periodic-word clusters; general rare-word clustering has its own classical theory [29,30].

*Remark 9.2* (Recovery of the binary case). For  $h = 2$ , the start probability, the intensity of a fixed colour and exact excess, and the excess law are respectively

$$2^{-L}, \quad 2^{-L-e-2}, \quad \mathbb{P}(E = e) = 2^{-e-1}.$$

These are the normalizations of [26, Theorem 1.1]. A binary dictionary has  $z = 2$ ,  $C_0 = 1$  and  $\mathfrak{m} = |W|$ ; the powers of intensity and dictionary size in (8.9) agree with [26, Theorem 5.1]. This recovers the binary specialization, not an independent proof of its arithmetic input. The binary source retains its sharper cutoff analyses, distinctions between aggregated and site-labelled comparisons, and relative boundary results.

These method-specific cutoff comparisons are not intrinsic lower bounds on the cost of retaining sites; the stronger binary input is discussed next.

### 9.3 The finite-colour geometric target and process observables

The finite-colour target in [27, Section 8, equation (8.1)] has a geometric base  $b_* > 1$ ,  $p_* = 1 - b_*^{-1}$  and colour law  $\pi$ . Our run field has precisely this target with

$$b_* = h, \quad p_* = 1 - 1/h, \quad \pi(c) = 1/h \quad (c \in H).$$

To compare centered integer lengths, define

$$b_N = \lfloor \log_h((h-1)N) \rfloor, \quad \theta_N = \log_h((h-1)N) - b_N. \quad (9.4)$$

At exact length  $b_N + r$ , the target intensity per unit of rescaled position and per colour equals

$$N(h-1)^2 h^{-(b_N+r+2)} = \frac{p_*}{h} h^{\theta_N-r}. \quad (9.5)$$

Thus a fixed collection of critical levels matches the corresponding restriction of the finite-colour geometric target. The factor  $h-1$  in the start intensity is essential to the centering (9.4).

This identifies the targets and provides a local arithmetic comparison. It does not by itself transfer the record, occupation or scale-flow limits of [27] to every finite alphabet. Such conclusions require a suitable arithmetic comparison on growing logarithmic windows, retaining the joint colour-excess marks under declared support, information and intensity budgets, followed by tail, localization and ordering estimates. Theorem 8.1 of that article is a result within the finite-colour target model. Matching (9.5) alone does not supply these further inputs.

For binary runs, the stronger same-grid input and the distinct direct-shell route are summarized in Section 11.4; their hypotheses must still be checked before taking a readout. Multiplicative clusters from Section 5 cannot be substituted for independent geometric levels: descendants of a shared germ remain linked, even in different spatial windows.

### 9.4 Detection in a quotient, marks in the ambient group

Let  $\pi : K \rightarrow Q$  be a continuous surjective homomorphism with  $|Q| = q < \infty$ , and  $W_0 \subseteq Q^B$ . For  $K = H \times \mathbb{T}^d$ ,

$$A = (\pi^{\otimes B})^{-1}(W_0), \quad a = |W_0|q^{-B}, \quad D_r(A) \leq q^r.$$

The cost depends on  $|W_0|$ , while the mark retains  $K^B$ .

This lift also holds for an arbitrary compact metrizable abelian  $K$  with the given finite quotient. Indeed  $\pi \circ F$  is the uniform multiplicative model on  $Q$ . On the retained windows of Section A, valuation-one private primes give the correct complete ambient marginals for every  $K$ . Activation and deletion costs are computed in  $Q$ , while the complete graph still satisfies (2.7). Theorem 2.2 therefore gives (8.9) with  $\mathbf{m} = |W_0|$  and  $\Omega_Q(W_0)$ .

If  $K$  is finite with fibre size  $d_0$ , the lifted dictionary has  $d_0^B |W_0|$  labels. For  $\mu_4 \rightarrow \mu_2$ ,  $z \mapsto z^2$ , the quotient word  $(-1, +1, \dots, +1)$  has no proper overlap. At  $B = \log_2 N + O(1)$ , its lift retains  $2^B \asymp N$  distinct words. This is not a uniform result for arbitrary dictionaries of that size; detection remains described by the small quotient.

### 9.5 Filtering by characters actually seen by a pattern

Erasure is robust, but algebraic cancellation can be finer. Adopt the Fourier convention

$$b_W(\chi) = \frac{1}{|W|} \sum_{w \in W} \prod_{j=0}^{B-1} \chi_j(w_j), \quad |b_W| \leq 1.$$

On separated windows define

$$\Phi_W(x, y) = \sum_{(\chi, \psi) \in \mathcal{K}_H(x, y) \setminus \{1\}} |b_W(\chi)| |b_W(\psi)|.$$

Character orthogonality gives

$$\mathbb{P}(A_x \cap A_y) = a^2 \sum_{(\chi, \psi) \in \mathcal{K}_H(x, y)} b_W(\chi) b_W(\psi).$$

The trivial term is  $a^2$ . On good windows, the marginal cap yields

$$\mathbb{P}(A_x \cap A_y) \leq a^2 [1 + \min\{a^{-1} - 1, \Phi_W(x, y)\}]. \quad (9.6)$$

The transfer can therefore use this filtered profile instead of the full annihilator weight. Nontrivial characters of a free coordinate average to zero. For a coset  $W = w_0 V$ ,  $|b_W(\chi)| = \mathbf{1}_{V^\perp}(\chi)$ , so only relations in  $V^\perp \times V^\perp$  remain.

For a finite quotient lift, averaging along fibres kills exactly the characters not descending from the quotient. Thus

$$\Phi_{K, W}(x, y) = \Phi_{Q, W_0}(x, y), \quad \Omega_K(W) = \Omega_Q(W_0).$$

For the overlap identity a compatible quotient pair at shift  $s$  has  $d_0^{B+s}$  compatible lifts; normalization cancels that factor. Filtering and erasure are two ways to verify the same probabilistic input, not a universal hierarchy. The coset identity is an exact algebraic reduction; by itself it supplies no small-overlap certificate for a chosen coset. A separate typical result is proved in [26, Theorem 5.3 and Corollary 5.4]: for binary words in a fixed logarithmic length band and with total intensity bounded above and away from zero, the distance is small for a high proportion of uniformly sampled affine dictionaries. The matrix is uniform among full-row-rank matrices and its right-hand side is independently uniform; the dictionary is selected independently of the prime signs and then held fixed. The proof averages the conditional distance after fixing the dictionary and uses one- and two-point inclusion probabilities with full-word collision bounds. It neither follows from the annihilator identity alone nor certifies a prescribed matrix. A full lift of a small quotient has a different mechanism. This comparison is contextual: it is not used in the proofs of the present dictionary corollary or the autonomous shrinking-target and fixed-environment theorems.

## 9.6 Unequal cell masses and nonrectangular restrictions

For a fixed cell  $A_0$  of measure  $p \in (0, 1)$ , the start pattern  $A_0^c \times A_0^L$  has mass  $(1 - p)p^L$ . Erasing  $r$  coordinates costs  $p^{-r}$  if the marker is retained and  $p(1 - p)^{-1}p^{-r}$  otherwise. Take

$$z = p^{-1}, \quad C_0 = \max\{1, p/(1 - p)\}, \quad \mathbf{m} = \max\{1, (1 - p)/p\}.$$

Starts at distances  $1, \dots, L$  are incompatible. Hence at  $L = \log_{1/p} N + O(1)$  the count is close to  $\text{Pois}(N(1 - p)p^L)$ , with error  $O(e^{-c\sqrt{\log N \log \log N}})$  and all block values retained. The choice  $p = 3/4$  genuinely uses  $z = 4/3 < 2$ .

For a partition  $K = A_1 \sqcup \dots \sqcup A_q$ ,  $\mu(A_j) = p_j > 0$ , let  $p_* = \max p_j$  and

$$A_L = \bigcup_j A_j^c \times A_j^L, \quad a_L = \sum_j (1 - p_j) p_j^L.$$

If  $s_*$  cells attain  $p_*$ , then  $a_L \geq s_*(1 - p_*)p_*^L$ . Projected contributions remain disjoint when any run coordinate is retained, giving

$$D_r(A_L) \leq \frac{q}{s_*(1 - p_*)} p_*^{-r}.$$

The effective size is bounded. Shifts below  $L$  are impossible, and

$$o_L(A_L) = \sum_{j \neq k} (1 - p_j) p_j^L p_k^L, \quad \Omega(A_L) \leq q p_*^L.$$

At  $L = \log_{1/p_*} N + O(1)$  the target has independent types with intensities

$$\lambda_{N,j} = N(1 - p_j) p_j^L. \quad (9.7)$$

Cells below maximal mass disappear in the limit, and maximal cells have equal weights. Absolute error does not give free relative accuracy for minority types.

If  $A' \subseteq A$  retains a fixed fraction  $\delta > 0$  of its mass, then  $D_r(A') \leq \delta^{-1} D_r(A)$  and  $\Omega(A') \leq \delta^{-1} \Omega(A)$ . One can also simply restrict the complete marked field of  $A$ . For a symmetric arc  $\{e^{it} : |t| < \theta\}$ , adding  $\sum_{j < L} \arg F(x + j) > 0$  to a run retains exactly half the independent mass. The event is not rectangular; the target retains its actual joint mark law.

Rarity alone does not imply small erasure cost. For  $A = \{z_1 \cdots z_B \in J_N\}$  every proper projection has full measure, so  $D_1 = 1/m(J_N)$ . Section 6 treats this observation by primitivity and resonance classes instead.

## 10 Deterministic sequences: visible packets and spatial resolution

A transfer based on independent prime values is not a theorem for a deterministic function observed from a random origin. The observables and their algebraic expansions remain useful, but the probabilistic input must change. This section gives exact correlation identities and finite factorial-moment budgets. It supplies no new cancellation estimate for Liouville. The spatially labelled formulation makes explicit how the required arithmetic input changes when the observation is refined.

Poisson genericity supplies a related but different empirical viewpoint: see Álvarez–Becher–Mereb [1] and its extension to exponentially mixing probabilities [2]. It concerns the distribution of occurrence counts as the searched word varies in a fixed typical sequence. Here the pattern is prescribed and the origin is sampled; a full field further retains positions and marks. These are different observables, so the genericity results do not supply the growing arithmetic correlation estimates required below.

### 10.1 A random origin and the transition identity

Let  $\varepsilon(n) \in \{-1, 1\}$  be deterministic. Choose a random origin  $U$  with a specified law, supported where the required values are defined. A reference example is uniform  $U \in \{X, \dots, 2X - 1\}$  with  $1 \leq H \leq X$ . Here  $X$  is the averaging scale,  $H$  the observation horizon and  $L$  the minimum run length.

Set

$$R_L(n) = \mathbf{1}_{\{\varepsilon(n-1) \neq \varepsilon(n) = \dots = \varepsilon(n+L-1)\}}, \quad Z = \sum_{j=1}^H R_L(U + j), \quad a = 2^{-L}, \quad \tau_H = Ha.$$

All randomness comes from  $U$ . For  $D(n) = \varepsilon(n-1)\varepsilon(n)$ ,

$$R_L(n) = 2^{-L} (1 - D(n)) \prod_{t=1}^{L-1} (1 + D(n+t)). \quad (10.1)$$

Two starts at distance less than  $L$  are incompatible. Distance exactly  $L$  is allowed for the unsigned start count.

For finite  $S$ , let  $C_U(S) = \mathbb{E} \prod_{s \in S} \varepsilon(U + s)$ , with  $C_U(\emptyset) = 1$ . Let  $\mathcal{J}_k$  be the ordered  $k$ -tuples of starts in  $\{1, \dots, H\}$  at pairwise distances at least  $L$ . For  $\mathbf{j} \in \mathcal{J}_k$ , define

$$E_{\mathbf{j}} = \bigcup_{i=1}^k \{j_i, \dots, j_i + L - 1\}, \quad J_{\mathbf{j}} = \{j_1, \dots, j_k\}.$$

The transition intervals are disjoint even at separation  $L$ . For  $A \subseteq E_{\mathbf{j}}$ , put

$$\partial A = A \triangle (A - 1), \quad \sigma_{\mathbf{j}}(A) = (-1)^{|A \cap J_{\mathbf{j}}|}.$$

Interior signs cancel in  $\prod_{t \in A} D(U + t)$ , leaving precisely the sign product on  $\partial A$ . For finite nonempty  $A$ , this boundary is nonempty and has even cardinality. Expansion gives the exact identity

$$\mathbb{E} \prod_{i=1}^k R_L(U + j_i) = a^k \left( 1 + \sum_{\emptyset \neq A \subseteq E_{\mathbf{j}}} \sigma_{\mathbf{j}}(A) C_U(\partial A) \right). \quad (10.2)$$

## 10.2 The signed packet prescribed by the count

Define

$$\mathcal{A}_k(U, H, L) = H^{-k} \left| \sum_{\mathbf{j} \in \mathcal{J}_k} \sum_{\emptyset \neq A \subseteq E_{\mathbf{j}}} \sigma_{\mathbf{j}}(A) C_U(\partial A) \right|. \quad (10.3)$$

The absolute value is taken after the prescribed sum over placements. It is not replaced by a supremum over freely chosen spectral masks. With  $(z)_k = z(z-1)\cdots(z-k+1)$ , incompatible tuples contribute zero, so

$$|\mathbb{E}(Z)_k - a^k |\mathcal{J}_k|| = \tau_H^k \mathcal{A}_k(U, H, L). \quad (10.4)$$

Ordering the starts and removing  $L-1$  spaces after each of the first  $k-1$  gives

$$|\mathcal{J}_k| = k! \binom{H - (L-1)(k-1)}{k}, \quad (10.5)$$

with value zero when the placement is impossible. A union bound over pairs that are too close also gives

$$0 \leq 1 - |\mathcal{J}_k|/H^k \leq k(k-1)L/H.$$

Consequently,

$$|\mathbb{E}(Z)_k - \tau_H^k| \leq \tau_H^k \left( \mathcal{A}_k(U, H, L) + \frac{k(k-1)L}{H} \right). \quad (10.6)$$

This is an arithmetic representation of a factorial-moment discrepancy, not its cancellation. The parameter  $k$  counts occurrences, not signs in a correlation: even  $k=1$  contains degrees up to  $L+1$ . For several occurrences, support diameters may reach  $H+L-1$ . Fixed-shift, fixed-order estimates do not automatically provide these growing packets.

To combine terms with the same support, define the coefficients of the *unnormalized factorial-moment discrepancy* by

$$b_{k,H,L}(S) := a^k \sum_{\mathbf{j} \in \mathcal{J}_k} \sum_{\substack{\emptyset \neq A \subseteq E_{\mathbf{j}} \\ \partial A = S}} \sigma_{\mathbf{j}}(A), \quad S \neq \emptyset. \quad (10.7)$$

Then the three normalizations are related exactly by

$$\mathbb{E}(Z)_k - a^k |\mathcal{J}_k| = \sum_{S \neq \emptyset} b_{k,H,L}(S) C_U(S), \quad \mathcal{A}_k = \tau_H^{-k} \left| \sum_{S \neq \emptyset} b_{k,H,L}(S) C_U(S) \right|. \quad (10.8)$$

In particular,

$$\sum_{S \neq \emptyset} |b_{k,H,L}(S)| \leq a^k |\mathcal{J}_k| (2^{kL} - 1) = (1 - a^k) |\mathcal{J}_k| \leq |\mathcal{J}_k| \leq H^k. \quad (10.9)$$

The factor  $a^k$  belongs to these coefficients; it is not present inside the raw sum in (10.3). Without it, the crude absolute coefficient mass is  $|\mathcal{J}_k| (2^{kL} - 1)$ . For  $H = k = 1, L = 3$ , the raw expansion has seven nontrivial characters, of total absolute mass 7, whereas the moment coefficients have mass  $7/8$ . The coefficients are fixed by the observable, so an arithmetic decomposition must be summed against them before replacing any part by an absolute bound.

### 10.3 A finite scalar transfer

The use of factorial moments for Poisson convergence belongs to the classical Brun method; see, for example, Barbour [6]. We give the precise finite inversion and its remainder here, rather than attributing the following inequality verbatim to that reference.

**Lemma 10.1** (Factorial-moment budget). *For an  $\mathbb{N}_0$ -valued variable  $Z$ ,  $\tau > 0$  and  $K \geq 1$  with  $\mathbb{E}(Z)_K < \infty$ ,*

$$d_{\text{TV}}(\mathcal{L}(Z), \text{Pois}(\tau)) \leq \frac{1}{2} \sum_{k=1}^K \frac{2^k}{k!} |\mathbb{E}(Z)_k - \tau^k| + \frac{(2\tau)^K}{K!}. \quad (10.10)$$

*Proof.* The finite binomial-inversion argument is given in Section D.4.  $\square$

Combining the lemma with (10.6) gives

$$\begin{aligned} d_{\text{TV}}(\mathcal{L}(Z), \text{Pois}(\tau_H)) &\leq \frac{1}{2} \sum_{k=1}^K \frac{(2\tau_H)^k}{k!} \mathcal{A}_k(U, H, L) \\ &\quad + 2\tau_H^2 e^{2\tau_H} \frac{L}{H} + \frac{(2\tau_H)^K}{K!}. \end{aligned} \quad (10.11)$$

Here  $\sum_{k \geq 0} k(k-1)(2\tau_H)^k/k! = 4\tau_H^2 e^{2\tau_H}$ .

**Corollary 10.2** (Conditional criterion for a sign sequence). *If  $H \rightarrow \infty$ ,  $L/H \rightarrow 0$ ,  $\tau_H$  remains bounded and  $\mathcal{A}_k(U, H, L) \rightarrow 0$  for every fixed  $k \geq 1$ , then  $d_{\text{TV}}(\mathcal{L}(Z), \text{Pois}(\tau_H)) \rightarrow 0$ . If also  $\tau_H \rightarrow \tau > 0$ , the limiting law is  $\text{Pois}(\tau)$ .*

*Proof.* First choose  $K$  so the last term of (10.11) is uniformly small at bounded intensity, then pass to the size limit with  $K$  fixed. The moving parameter avoids an unjustified convergence assumption on the integer-rounded length.  $\square$

The first packet admits further exact simplifications, detailed in Section D: its spectrum is even, Krawtchouk polynomials compress all degrees into one Hamming statistic, and the full diameter- $L - 1$  band is the endpoint difference in (D.3). These identities preserve signed cancellation within the first moment; they are not estimates for the remaining even correlations or for higher occurrence packets.

### 10.4 Transfer at a chosen spatial resolution

Let  $\mathcal{P} = \{B_1, \dots, B_m\}$  partition the starts  $\{1, \dots, H\}$ . Put  $Z_c = \sum_{j \in B_c} R_L(U + j)$  and  $\lambda_c = |B_c|a$ . For a cell tuple  $\mathbf{c} = (c_1, \dots, c_k)$ , let  $\mathcal{J}_k(\mathbf{c})$  contain compatible ordered starts with  $j_i \in B_{c_i}$ . Define

$$\mathcal{A}_k^{\mathcal{P}} = H^{-k} \sum_{\mathbf{c} \in \{1, \dots, m\}^k} \left| \sum_{\mathbf{j} \in \mathcal{J}_k(\mathbf{c})} \sum_{\emptyset \neq A \subseteq E_{\mathbf{j}}} \sigma_{\mathbf{j}}(A) C_U(\partial A) \right|. \quad (10.12)$$

Absolute values are taken after summing placements with the same cell labels, not before and not after all labels have been forgotten.

**Proposition 10.3** (Resolution-adapted factorial transfer). *For every  $K \geq 1$ ,*

$$d_{\text{TV}} \left( \mathcal{L}((Z_c)_{c=1}^m), \bigotimes_{c=1}^m \text{Pois}(\lambda_c) \right) \leq \frac{1}{2} \sum_{k=1}^K \frac{(2\tau_H)^k}{k!} \mathcal{A}_k^{\mathcal{P}} + 2\tau_H^2 e^{2\tau_H} \frac{L}{H} + \frac{(2\tau_H)^K}{K!}. \quad (10.13)$$

If  $\mathcal{Q}$  is coarser than  $\mathcal{P}$ , then  $\mathcal{A}_k^{\mathcal{Q}} \leq \mathcal{A}_k^{\mathcal{P}}$ . There is no hidden factor depending on the number of cells in the bound; its input norm may nevertheless grow under refinement.

*Proof.* The multivariate finite inversion in Lemma D.2 applies to the binomial moments of  $(Z_c)$ . Summing ordered cell labels with fixed multiplicities turns its coefficient into  $2^{k-1}/k!$  times the corresponding ordered factorial-moment discrepancy. Formula (10.2) identifies the signed correlation part. The sum over all cell labels of the excluded independent placements is at most  $H^k k(k-1)L/H$ , by the same pair union bound as before. Summing this error proves (10.13). Coarsening combines signed packets before taking absolute values, so the triangle inequality proves the contraction.  $\square$

One cell recovers (10.11). Singleton cells demand the full collection of labelled inclusion probabilities. A scalar packet estimate therefore cannot simply be reused for a refined process.

Unlike the random-model readout discussed in Section 11.4, this coarsening contracts an arithmetic norm of signed packets whose smallness must still be proved. No random-prime forcing is available for a deterministic sequence observed from a random origin.

The prescribed signed packets in (10.3) and (10.13) must also be distinguished from a sum of the absolute values of all factorial cumulants. The binary example in [26, technical companion, Theorem G.9] has small total-variation distance and large absolute cumulant activity. It does not disprove these resolution-adapted sufficient criteria or estimate their deterministic arithmetic input. The Palm deficit in that companion is a downstream consequence of the random microscopic comparison, not a cancellation estimate for Liouville.

## 10.5 A support obstruction before arithmetic

**Proposition 10.4** (Finite-source resolution obstruction). *Suppose a random vector  $V \in \mathbb{N}_0^m$  is a deterministic function of an origin with at most  $N_{\text{orig}}$  possible values. Let  $\lambda = \sum_c \lambda_c$  and  $\lambda_{\max} = \max_c \lambda_c \leq 1$ . For every integer  $k \geq 1$ ,*

$$d_{\text{TV}} \left( \mathcal{L}(V), \bigotimes_c \text{Pois}(\lambda_c) \right) \geq \mathbb{P}(\text{Pois}(\lambda) \geq k) - N_{\text{orig}} e^{-\lambda} \lambda_{\max}^k. \quad (10.14)$$

For the microscopic run field and an origin with at most  $X$  values, this specializes to

$$d_{\text{TV}} \left( \mathcal{L}((R_L(U+j))_{j=1}^H), \bigotimes_{j=1}^H \text{Pois}(a) \right) \geq \mathbb{P}(\text{Pois}(\tau_H) \geq k) - X e^{-\tau_H} a^k. \quad (10.15)$$

*Proof.* The actual support has at most  $N_{\text{orig}}$  points. A target configuration of total count at least  $k$  has probability  $e^{-\lambda} \prod_c \lambda_c^{v_c}/v_c! \leq e^{-\lambda} \lambda_{\max}^k$ . Use the event of configurations with total count at least  $k$  outside the actual support. Its actual probability is zero, and its target probability is at least the right-hand side.  $\square$

The same elementary support-counting argument appears in [26, Remark 7.8b] for local run vectors after a completely multiplicative realization has been fixed. Its empirical Poisson conclusion concerns the scalar count, not that full vector. The proposition here allows an arbitrary finite source and a chosen cell partition; it supplies no correlation cancellation.

If  $\tau_H \rightarrow \tau > 0$  and  $X \leq H^C$ , any fixed  $k > C$  prevents microscopic total-variation convergence. For example,  $H \asymp \sqrt{X}$  is already obstructed by  $k = 3$ . More generally, for  $m \rightarrow \infty$  balanced cells with  $\lambda_{\max} = O(1/m)$  and total intensity converging to a positive constant, a necessary condition is  $\log N_{\text{orig}} / \log m \rightarrow \infty$ . It is not sufficient. The scalar count and weak limits after changing spatial resolution are not excluded.

The obstruction can be substantial at finite size. For  $X = 2^{20}$  possible uniform origins,  $H = 2^{14}$ ,  $L = 14$  and  $\tau_H = 1$ , choosing  $k = 2$  gives the universal lower bound

$$1 - \left(2 + \frac{1}{256}\right) e^{-1} = 0.262804088590 \dots$$

This holds for every deterministic sign sequence, not only Liouville; a good count histogram does not certify a microscopic field comparison. It is a different obstruction from Proposition 3.6: here the original source has too few possible configurations, whereas there the independent prime source is not finite in that sense and the proposed Poisson field creates artificial same-site collisions. In both cases a scalar conclusion does not determine an admissible labelled resolution.

## 10.6 What remains arithmetic for Liouville

For  $\varepsilon(n) = \lambda(n) = (-1)^{\Omega(n)}$ , every identity above is exact, but the required packet estimates are not proved here. The nontrivial supports are even. Under random Haar prime signs,

$$\mathbb{E} \prod_i F(n_i) = \mathbf{1}_{\{\prod_i n_i \text{ is a square}\}},$$

whereas for Liouville  $\prod_i \lambda(n_i) = \lambda(\prod_i n_i)$ . A nonsquare product does not vanish by averaging nonexistent random prime coordinates. Cancellation must come from the chosen law of the origin. Fixing all prime signs to  $-1$  is singular relative to their infinite Haar product and is not a permitted global density change.

Tao's logarithmically averaged two-point Chowla theorem [32] is directly relevant to the first even sector. It treats fixed affine forms and a logarithmic averaging law; it does not by substitution give the uniform growing-shift estimates needed to sum an entire degree-two packet. Nor does it control the remaining even degrees. Results on sign patterns [22] and odd logarithmically averaged Chowla correlations [33] provide relevant context, but neither odd correlations nor a fixed-degree estimate supplies all the growing even degrees in (D.2). Even after a two-point contribution has been controlled, the sum of degrees at least four remains an additional arithmetic input, followed by higher occurrence packets. Logarithmic averages are admitted by specifying another law of  $U$ ; their conclusion must not be relabelled as a uniform average.

## 11 Scope, consequences and further questions

### 11.1 What the method isolates

The common interface is useful because the same three errors can be verified in different ways. Haar private coordinates make conditional marginal errors vanish; products of several nonuniform prime values make them relatively small. Coordinate erasure bounds joint events by removing the constraints that rank defects prevent from being independent. Character filtering ignores

relations invisible to the prescribed motif. Rational probes can reduce a high-dimensional block problem to primitivity of a single valuation vector.

The target is part of the problem, not a fixed conclusion imposed on every model. Under  $L^2$ , prime sites can retain non-negligible microscopic mass even though their arithmetic density vanishes. A finite irregular environment can create exact points and clusters along its multiplier orbits. Those sources are retained rather than deleted at a fictitious Haar cost. Conversely, a weaker  $a^{3/2}$  pair bound can be sufficient on  $O(N \log \log N)$  exceptional pairs. Uniform quadratic decorrelation of every pair is unnecessary.

## 11.2 Limits which the examples distinguish

Global empirical equidistribution does not determine shrinking-target statistics. The logarithmic cusp densities in Section 3 can have all finite density moments and still contribute a nonvanishing or divergent prime source. The corrected process target captures that source. Proposition 3.5 shows that the universal Haar-composite rule fails below  $L^2$  on prime-square masks; it is not an impossibility statement about individual densities or targets. Proposition 3.6 separates a second issue below  $L^2$ : a Poisson field with the correct site intensities can still create too many artificial multiple points at a site. Its count may remain close to Poisson. The restricted aggregation bound (3.20) illustrates one remedy for independent prime-only sources, not an extension to arbitrary mixed masks.

One-block independence is also different from an occurrence process. Turk's estimate, as recalled by Najnudel, already supplies independence within logarithmic Steinhaus blocks. It neither constructs the sparse conditional graph across all moving blocks nor determines which cluster relations or prime sources survive the chosen targets. The preceding proofs address those separate tasks.

Complete recorded values require complete marked marginal laws and a graph for those values. Their aggregation in Stein sums does not give permission to record the whole prime environment for free. Random targets depending on a retained finite environment are handled by its kernel, including any induced duplication. Targets adapted to unretained prime values are not covered by deterministic uniformity.

The lattice classification is for a corrected prefix count with Haar outside a fixed exceptional set and a common shrinking arc. Its initial-ideal interpretation uses standard binomial and monomial ideal theory. It gives finite rational certificates once exact relations are known, not a numerical algorithm to recognize arbitrary real phases. Distinct relations can produce the same count law while retaining different spatial clusters.

## 11.3 Arithmetic interfaces beyond the proved branches

Several extensions are naturally posed by the existing estimates, but are not inputs to the present results. Below  $L^2$ , the immediate task is to specify a source resolution and a target: coarsen observations, keep selected Bernoulli germs exactly, or impose a small-collision hypothesis. A hybrid target must retain dependencies between populations sharing prime coordinates; the prime-only calculation does not make prime and composite contributions independent. Nonuniform prime laws in finite groups require aperiodicity and surjectivity conditions adapted to torsion; a bounded density alone may be concentrated on a proper subgroup. Extending the  $L^2$  theorem from  $F(n)$  to general rational observations requires replacing both the disjoint-basis argument and the prime-power mass decomposition. The current growing-support bounds do not cover logarithmic word lengths under arbitrary persistent biases.

For patterns with substantial short overlaps, the local cluster target must be determined before the long-range arithmetic transfer; rare-word cluster approximation in Markov models provides relevant antecedents [29,30]. For deterministic signs, Proposition 10.3 identifies interme-

diates observations between one count and a microscopic field. Its packet norm contracts under coarsening, while Proposition 10.4 rules out certain resolutions before any arithmetic estimate is attempted. Cancellation of the remaining even packets is an additional mathematical question, not a consequence of the random model.

## 11.4 Binary patterns and finite-colour process targets

The estimates in *Long runs and rare patterns* [26, Section 3 and technical companion, Appendix A] supply the population-level arithmetic input to Proposition 8.2. Its complete binary run comparison [26, Theorem 7.7, Corollary 7.8 and technical companion, Appendix F] serves a different purpose: it retains base-contained interior starts, full exact excesses and actual signs on the same spatial grid. Its hypotheses include diverging intensity and conditioning in the prescribed hard prime field; the run from 1 is excluded. Neither that comparison nor its contraction under a common observation replaces the variable-base population estimates used here. The source-corrected and finite-environment theorems of Sections 3–5 remain independent of that binary input.

*Records, scale flows, and Gaussian limits* [27] studies records, censoring, occupation, scale flows and spatial fields from a complete binary comparison. Its Input 2.1 and Corollary 2.2 retain the preceding hypotheses, while its direct shell route (Theorem 2.3 and technical companion, Proposition A.2) also has a domain with intensity at least one, without requiring divergence. The finite-colour target of its Section 8 agrees with (9.5). This agreement identifies a target, not a transfer of every process law to every group or prime bias. The additional macroscopic, tail and ordering estimates remain necessary. A common observation preserves distance to an exact moving target; it does not itself establish a weak limit or transfer unbounded moments.

The boundary analysis of *Long runs and rare patterns* is distinct from the fixed-environment classification here: it involves every prime up to a growing run length. Our classification fixes the exceptional environment before the size limit. Similarly, a compound-Poisson count determined by  $L_\xi$  does not assert independent spatial increments. The marked germ process already retains each arithmetic cluster's spatial configuration; descendants of one germ remain linked even in disjoint observation windows. Further record or occupation laws for such clusters require analysis of that target and a suitable arithmetic transfer, rather than replacement of its offspring by independent points.

The resulting principle is constructive: the comparison preserves the sources that carry microscopic mass and the relations that create repeated hits, while discarding information only at a proved cost. Choosing the intensity, the cluster mark and the observation resolution together is what makes the separate arithmetic estimates usable.

## A One-window defects and deletion at a moving cutoff

This appendix supplies the one-window estimates used by the pattern branch. The deep two-window profile remains the population analysis of [26], used in Proposition 8.2. The arguments here concern weighted defects and the actual mass of the selected windows.

### A.1 An integer Runge bound for a fixed power

**Lemma A.1** (Split product and fixed power). *Fix a prime  $\ell$ . Let  $u_1, \dots, u_s \in \mathbb{Z}$  be distinct,  $s \geq 1$ , with  $|u_i| \leq R$ ,  $R \geq 1$ , and let  $1 \leq e_i \leq \ell - 1$ . Suppose  $D = \sum_i e_i$  is divisible by  $\ell$ . If  $X, Y \in \mathbb{Z}$ ,  $X \geq 2R$ , and*

$$\prod_{i=1}^s (X + u_i)^{e_i} = Y^\ell,$$

then  $X \leq (C_\ell DR)^{C_\ell D}$  for a constant depending only on  $\ell$ .

*Proof.* Write  $D = \ell k$ ,  $G(T) = \prod_i (T + u_i)^{e_i}$ , and expand the positive root at infinity:

$$G(T)^{1/\ell} = T^k \prod_i (1 + u_i/T)^{e_i/\ell}.$$

Let  $P(T)$  be its polynomial part. For a total series degree  $j \leq k$ , denominators divide  $\ell^j \prod_i j_i!$ , so  $D_0 = \ell^k k!$  is a common denominator.

On  $|z| = (2R)^{-1}$ , the analytic product  $\prod_i (1 + u_i z)^{e_i/\ell}$ , with branches equal to one at zero, has modulus at most  $(3/2)^k$ . Cauchy's bound gives, for  $X \geq 4R$ ,

$$|G(X)^{1/\ell} - P(X)| \leq \frac{2(3/2)^k (2R)^{k+1}}{X}. \quad (\text{A.1})$$

The positive root is an integer:  $Y$  for odd  $\ell$  and  $|Y|$  for even  $\ell$ . If it differs from  $P(X)$ , integrality of  $X$  gives separation at least  $1/D_0$ , and (A.1) proves the bound.

Otherwise  $X$  is a root of the integer polynomial  $D_0^\ell (G - P^\ell)$ . This polynomial is nonzero because the root multiplicities  $e_i$  of  $G$  are not divisible by  $\ell$ . Its coefficients are bounded by  $(C_\ell DR)^{C_\ell D}$ , using  $D_0 \leq (\ell k)^k$  and the same coefficient estimates. The elementary root bound for a nonzero integer polynomial concludes. Increase the constant to absorb  $2R \leq X < 4R$ .  $\square$

The integer hypotheses are necessary. For  $G(T) = T(T+1)$ , the real numbers  $X_m = (\sqrt{1+4m^2} - 1)/2$  are unbounded and satisfy  $G(X_m) = m^2$ . This is the same integrality issue made explicit in [26, technical companion, Appendix A.1]; the present statement includes  $X, Y \in \mathbb{Z}$ ,  $R \geq 1$  and a nonempty factor family.

## A.2 Coding defects and exponential weights

An integer  $n$  is  $\ell$ -defective at threshold  $B$  when  $v_p(n) \equiv 0 \pmod{\ell}$  for every prime  $p > B$ . In a window  $V_x$  of length  $B \asymp \log N$ , let  $m_{\ell,B}(x)$  count such integers.

**Lemma A.2** (Weighted defects). *For fixed  $\ell$  and  $B$  in a fixed logarithmic band,*

$$\max_{x \in I_N} m_{\ell,B}(x) \ll_\ell \frac{\log N}{\log \log N}, \quad \#\{x : m_{\ell,B}(x) > 0\} \leq N^{1/\ell+o(1)}.$$

For every fixed  $z > 1$ ,

$$\sum_{x \in I_N} (z^{m_{\ell,B}(x)} - 1) \leq N^{1/\ell+o(1)}. \quad (\text{A.2})$$

*Proof.* Every defective integer has a unique representation  $a^\ell u$ , with  $u$   $\ell$ -free and  $B$ -smooth. Up to  $3N$ , their number is at most

$$(3N)^{1/\ell} \prod_{p \leq B} (1 + p^{-1/\ell} + \dots + p^{-(\ell-1)/\ell}) = N^{1/\ell+o(1)}.$$

Each integer belongs to at most  $B = N^{o(1)}$  windows, giving the window count.

For the maximum, associate to the  $m$  defects in one window the columns

$$((v_p(n) \bmod \ell)_{p \leq B}, 1) \in \mathbb{F}_\ell^{\pi(B)+1}.$$

A nonzero word of their kernel selects exponents in  $\{1, \dots, \ell-1\}$  with sum divisible by  $\ell$ , and a product which is an  $\ell$ th power. Applying Lemma A.1 to the distinct offsets, with total degree at most  $(\ell-1)$  times the word weight, forces minimum weight at least  $c_\ell \log N / \log \log N$ .

The kernel has codimension at most  $r = \pi(B) + 1 = O(\log N / \log \log N)$ . Choose  $t \asymp_\ell \log N / \log \log N$  with  $2t$  strictly below its minimum distance. If  $m < 2t$ , the result follows. Otherwise disjoint Hamming balls of radius  $t$  around codewords give

$$\binom{m}{t}(\ell - 1)^t \leq \ell^r.$$

Since  $r/t$  is bounded and  $\binom{m}{t} \geq (m/t)^t$ , this implies  $m = O_\ell(t)$ . The maximum bound makes  $z^m = N^{o(1)}$ ; multiply by the number of defective windows to prove (A.2).  $\square$

For  $K = H \times \mathbb{T}^d$ , let  $u_B(x)$  count vertices defective for at least one prime in the fixed set

$$\{2\} \cup \{\ell : \ell \text{ prime}, \ell \mid |H|\}.$$

The lemma gives

$$\mathcal{U}_{z,H,B} := \sum_x (z^{u_B(x)} - 1) \leq N^{1/2+o(1)}. \quad (\text{A.3})$$

For a nondefective vertex, the gcd of the valuations of its large prime factors is coprime to the exponent of  $H$ . Bézout's identity makes the corresponding map onto  $H$  surjective. At least one nonzero valuation suffices on the torus. Haar components in the product are independent, and distinct vertices of a window use disjoint large-prime families. All retained values are therefore jointly Haar.

Under (8.4), projection yields  $\mathbb{P}(\mathbf{F}_x \in A) \leq C_0 a z^{u_B(x)}$ , with equality to  $a$  when  $u_B(x) = 0$ . For  $u \geq 1$ ,  $z^u \leq z(z^u - 1)/(z - 1)$ . Thus every deterministic deleted set  $D$  satisfies

$$\sum_{x \in D} \mathbb{P}(\mathbf{F}_x \in A) \leq a|D| + \frac{C_0 z}{z - 1} a \mathcal{U}_{z,H,B}. \quad (\text{A.4})$$

This pays the true probability mass of deletion. It does not use the false bound  $z^u - 1 \geq 1$  for every  $u \geq 1$  when  $1 < z < 2$ .

### A.3 Windows containing many large prime squares

Put

$$t_B(x) = \#\{n \in V_x : \exists p > B, p^2 \mid n\}, \quad D_{B,r} = \{x : t_B(x) \geq r\}.$$

For  $1 \leq r \leq B$  and an auxiliary cutoff  $Z_0 > B$ , the finite bound is

$$\frac{|D_{B,r}|}{N} \leq \binom{B}{r} \left( \sum_{B < p \leq Z_0} p^{-2} \right)^r + \frac{1}{N} \binom{B}{r} \pi(Z_0)^r + 3B \sum_{p > Z_0} p^{-2}. \quad (\text{A.5})$$

The last term covers windows containing a multiple of  $p^2$  for  $p > Z_0$ . For the others, choose  $r$  positions and one prime  $B < p \leq Z_0$  whose square divides each value. These primes are distinct:  $p > B$  cannot divide two values in the same window. The Chinese remainder theorem gives at most  $N / \prod p^2 + 1$  starts. Sum over choices, and then drop distinctness to majorize the two sums by the first two displayed terms. The  $+1$  is retained.

Write  $U = \log N$ ,  $V = \log \log N$  and  $r_N = \lceil \sqrt{2U/V} \rceil$ . Fix  $c < c_0 < 1/\sqrt{2}$  and choose  $Z_0 = e^{c_0 \sqrt{UV}}$ . Using  $\sum_{p > B} p^{-2} \ll (B \log B)^{-1}$  and  $\binom{B}{r} \leq (eB/r)^r$ , the three terms in (A.5) are at most

$$e^{-(1/\sqrt{2}+o(1))\sqrt{UV}}, \quad N^{-1+c_0\sqrt{2}+o(1)}, \quad e^{-c_0\sqrt{UV}+O(V)}.$$

Hence  $|D_{B,r_N}|/N \ll_c e^{-c\sqrt{\log N \log \log N}}$ . The prime estimates are standard consequences of [35,31].

#### A.4 Rankin's bound at the conditioning cutoff

Let  $Y = e^{\gamma\sqrt{\log N \log \log N}}$  for fixed  $\gamma > 0$ , and let  $\Psi(X, Y)$  count integers up to  $X$  with prime factors at most  $Y$ . Put

$$u = \frac{\log N}{\log Y}, \quad \delta = \frac{\log u}{\log Y}.$$

For large  $N$ ,  $\delta \rightarrow 0$ , and Rankin's inequality gives

$$\Psi(3N, Y) \leq (3N)^{1-\delta} \prod_{p \leq Y} (1 - p^{-1+\delta})^{-1}.$$

Terms of order at least two in the logarithm of the Euler product are summable. Partial summation using  $\pi(t) \ll t/\log t$  bounds the first-order terms by a constant times

$$\int_{\log 2}^{\log Y} \frac{e^{\delta v}}{v} dv = O(u + \log \log Y).$$

Up to  $1/\delta$ , use  $e^{\delta v} \leq e$ ; above it, put  $w = \delta v$  and bound  $\int_1^{\log u} e^w/w dw$  by  $u$ . Consequently,

$$\frac{\Psi(3N, Y)}{N} \leq e^{-u \log u + O(u + \log \log Y)} = \exp \left\{ - \left( \frac{1}{2\gamma} + o(1) \right) \sqrt{\log N \log \log N} \right\}. \quad (\text{A.6})$$

The calculation is uniform for  $\gamma$  in a fixed compact subset of  $(0, \infty)$ , and for replacing  $3N$  by a fixed multiple of  $N$ .

An integer with no prime  $p > Y$  of valuation exactly one is either  $Y$ -smooth or divisible by  $p^2$  for some  $p > Y$ . Contaminated windows are therefore at most

$$B\Psi(3N, Y) + B \sum_{p > Y} \left\lfloor \frac{3N}{p^2} \right\rfloor.$$

At  $\gamma = 1/\sqrt{2}$ , their proportion is  $O_c(e^{-c\sqrt{\log N \log \log N}})$  for every  $c < 1/\sqrt{2}$ ; the factor  $B \asymp \log N$  is absorbed by the exponent margin.

Combine this selection with (A.5) to obtain the good set in Theorem 8.4. Simple pivots above  $Y > 2B$  make the blocks and local unions exactly uniform conditionally on small primes. Their true deleted mass is still controlled by (A.4), not only by the number of deleted windows.

## B Growing-order Bonferroni and abundant private primes

This appendix proves the counting estimate used for nonuniform prime densities. The auxiliary average is over integers in an interval; it introduces no new random source into the multiplicative model. Congruence-counting errors are retained as the truncation order grows.

### B.1 A finite estimate

Let  $\mathcal{P}$  be a finite set of  $P$  primes and put

$$\omega_{\mathcal{P}}(n) = \sum_{p \in \mathcal{P}} \mathbf{1}_{\{p|n\}}, \quad A = \sum_{p \in \mathcal{P}} p^{-1} > 0.$$

Let  $I$  be any interval of  $L_0$  consecutive integers. Fix  $0 < \theta < 1$ , write  $s = 1 - \theta$ , and let  $\ell \geq 2$  be even.

**Lemma B.1** (Truncated transform). *For the uniform average  $\mathbb{E}_I$  on  $I$ ,*

$$\mathbb{E}_I \theta^{\omega_{\mathcal{P}}(n)} \leq \prod_{p \in \mathcal{P}} (1 - s/p) + \frac{(sA)^{\ell+1}}{(\ell+1)!} + \frac{1}{L_0} \sum_{j=1}^{\ell} s^j \binom{P}{j}, \quad (\text{B.1})$$

with  $\binom{P}{j} = 0$  for  $j > P$ .

*Proof.* For  $w \in \mathbb{N}_0$ , set  $B_\ell(w) = \sum_{j=0}^\ell (-s)^j \binom{w}{j}$ . The binomial expansion, or Taylor's formula with remainder when  $w > \ell$ , gives

$$0 \leq B_\ell(w) - (1-s)^w \leq s^{\ell+1} \binom{w}{\ell+1}. \quad (\text{B.2})$$

The parity of  $\ell$  is essential: the derivative of order  $\ell+1$  has negative sign and modulus bounded by its value at zero. For  $w \leq \ell$ , the expansion is complete.

The binomial moments satisfy

$$\mathbb{E}_I \binom{\omega_{\mathcal{P}}}{j} = \sum_{\substack{S \subseteq \mathcal{P} \\ |S|=j}} \mathbb{E}_I \mathbf{1}_{\{\prod_{p \in S} p | n\}}.$$

Each divisibility count differs from  $L_0 / \prod_{p \in S} p$  by at most one, even when the modulus exceeds the interval length. Therefore

$$\left| \mathbb{E}_I \binom{\omega_{\mathcal{P}}}{j} - e_j((1/p)_{p \in \mathcal{P}}) \right| \leq L_0^{-1} \binom{P}{j}, \quad (\text{B.3})$$

where  $e_j$  is the elementary symmetric polynomial.

Only to evaluate the reference polynomial, introduce independent Bernoulli variables of parameters  $1/p$ , with sum  $S$ . They satisfy

$$\mathbb{E} \binom{S}{j} = e_j((1/p)), \quad \mathbb{E} \theta^S = \prod_p (1 - s/p), \quad \mathbb{E} \binom{S}{\ell+1} \leq \frac{A^{\ell+1}}{(\ell+1)!}.$$

Average (B.2) for this reference sum and compare each coefficient through (B.3). This proves (B.1). Independence has not been attributed to actual divisibilities.  $\square$

On  $\{\omega_{\mathcal{P}} < \theta A\}$ ,  $\theta^{\omega_{\mathcal{P}}} \geq \theta^{\theta A}$ . Markov's inequality and  $\prod_p (1 - s/p) \leq e^{-sA}$  imply

$$\begin{aligned} \frac{\#\{n \in I : \omega_{\mathcal{P}}(n) < \theta A\}}{L_0} &\leq e^{-I(\theta)A} \\ &+ \theta^{-\theta A} \left[ \frac{((1-\theta)A)^{\ell+1}}{(\ell+1)!} + \frac{1}{L_0} \sum_{j=1}^{\ell} (1-\theta)^j \binom{P}{j} \right], \end{aligned} \quad (\text{B.4})$$

where  $I(\theta) = 1 - \theta + \theta \log \theta > 0$ .

## B.2 Choose the prime band and truncation together

With  $U = \log N$ ,  $V = \log \log N$ , choose

$$Y = U^4, \quad \ell_N = 2[4V], \quad Z = N^{1/(4\ell_N)}, \quad \mathcal{P} = \{p : Y < p \leq Z\}.$$

Thus  $\ell_N = 8V + O(1)$  and  $Z^{\ell_N} = N^{1/4}$ . Eventually  $Z > Y$ . Mertens' formula [35,31] gives

$$A_N = \sum_{Y < p \leq Z} p^{-1} = \log \log Z - \log \log Y + o(1) = V - 2 \log V - \log 128 + o(1). \quad (\text{B.5})$$

In particular  $A_N \leq V$  eventually and  $A_N \sim V$ .

For  $L_0 \asymp N$ , the congruence remainder in (B.4), before multiplication by  $\theta^{-\theta A_N}$ , is at most

$$L_0^{-1} \sum_{j=1}^{\ell_N} (1-\theta)^j \binom{P}{j} \ll \frac{(1+P)^{\ell_N}}{N} \leq 2^{\ell_N} N^{-3/4}.$$

The extra factor is at most  $e^{V/e}$  because  $-\theta \log \theta \leq 1/e$ . The final remainder is  $N^{-3/4+o(1)}$ .

For the truncation remainder,  $n! \geq (n/e)^n$  gives

$$\frac{((1-\theta)A_N)^{\ell_N+1}}{(\ell_N+1)!} \leq \left(\frac{eA_N}{\ell_N+1}\right)^{\ell_N+1} \leq (e/8)^{\ell_N+1}.$$

Even after multiplication by  $e^{V/e}$ , this is  $O(e^{-8V})$ . Both remainders are smaller than  $e^{-I(\theta)A_N}$ , since  $0 < I(\theta) < 1$ .

**Lemma B.2** (Abundance of prime factors in the band). *For fixed  $0 < \theta < 1$ , the preceding cutoffs satisfy, uniformly over intervals of length  $L_0 \asymp N$ ,*

$$\frac{\#\{n \in I : \omega_{Y,Z}(n) < \theta A_N\}}{L_0} \ll_{\theta} e^{-I(\theta)A_N} = (\log N)^{-I(\theta)+o(1)}. \quad (\text{B.6})$$

Translations play no role in (B.3). The lemma therefore applies to  $x + s_j$  in intervals of length  $N$  and to the prefix of length  $M \asymp N$ . For an arbitrary mask, no concentration relative to the mask's own cardinality is claimed: its bad set is bounded by the bad set of the whole prefix.

### B.3 From available factors to private prime coordinates

For  $k$  sampled positions, deleting sites where one position has fewer than  $t_N = \lfloor \theta A_N \rfloor$  factors in the band removes at most  $O(kNe^{-I(\theta)A_N})$  sites. Squares of primes above  $Y$  contribute at most

$$k \sum_{p>Y} \left\lfloor \frac{M}{p^2} \right\rfloor \ll \frac{kN}{Y \log Y}.$$

A large prime dividing two positions must divide their difference. The corresponding site count is at most

$$\binom{k}{2} \frac{\log M}{\log Y} \left( \frac{N}{Y} + 1 \right).$$

After these deletions, the simple prime factors of each position are private in the full sampled support. The selected coordinate sets are deterministic and disjoint between positions, justifying products of densities before integrating the other prime values.

Under a spectral envelope  $\rho < 1$ , mixing costs  $e^{-\theta(-\log \rho)A_N}$ , while insufficient factorization costs  $e^{-I(\theta)A_N}$ . The common exponent is their minimum. The decreasing function  $I(\theta)$  and increasing function  $-\theta \log \rho$  have one intersection, improving the convenient choice  $\theta = 1/2$  without establishing the optimal rate of the true law.

The cutoff change is part of the proof. Retaining  $Z = N^{1/4}$  from a second-moment argument and then increasing the moment order would not pay the number of congruence errors. The identity  $Z^{\ell_N} = N^{1/4}$  is what permits a growing truncation.

## C Arithmetic dependencies and the variable-base ledger

### C.1 Arithmetic inputs and contextual comparisons

The deep arithmetic input is the published Version 3 of *Long runs and rare patterns* [26], whose article and technical companion form one proof package. The population ledger below uses Section 3 of the article before reweighting and capping; its Runge–Pell arguments are given in the technical companion, Appendix A. These estimates are not replaced by a final binary numerical bound or by the complete run-field comparison. Each row below specifies the population, its dimension bound, its counting estimate and the operation on its weight.

The population estimates are a consumed proof input to Proposition 8.2. Marked Chen–Stein comparison and the other classical tools retain their own local attributions. The binary typical-dictionary and microscopic results, as well as the process laws in Version 2 of *Records, scale flows, and Gaussian limits* [27], are contextual comparisons discussed in Section 11.4; they are not additional hypotheses of the source, rational-observation or fixed-environment theorems.

The factor denoted  $E_N = N^{o(1)}$  below depends on fixed logarithmic-band parameters and on fixed  $z$ , but is chosen independently of the cap  $T$ . An arbitrarily small  $N^\varepsilon$  may absorb it after the arithmetic exponent is fixed with a margin.

## C.2 Population-by-population transfer

Use the source article's rational/residual decomposition  $\rho = \sigma + \tau$  and its canonical sector partition. Write  $Q_z = z^B$ ; all relations and dimensions in this table are still over  $\mathbb{F}_2$ . “Hosts” are ordered pairs supporting the indicated population, not individual codewords. The variable  $q$  in the rational row is the height of a reduced rational channel, not the alphabet size.

Table 1: Population estimates of [26] used in the variable-base profile.

| Part     | Source result                                                                            | Host and dimension information                                                                                                                             | Weight and cap operation                                                                                                                        |
|----------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| Rational | Lemma 3.2 (dimension/height); Proposition 3.8 and (3.12) (masses); (3.6) (decomposition) | $\sigma = m - 1 \leq B/q$ ; units are spaced by channel height. Sum the rational families before combining with residual sectors.                          | $E_N N(z^{B/3} + z^{B/2})$ . Leave uncapped; both powers are absorbed by $NQ_z^{2/3}$ .                                                         |
| 1        | Proposition 3.10, first shallow population; host bound Proposition 3.7                   | Small certificate product implies $\tau = o(B)$ . Full relation hosts total $N^{3/2+o(1)}$ ; rational mass handled separately.                             | $E_N(N^{3/2} + Nz^{B/2})$ . Use $z^\sigma \leq z(z^\sigma - 1)/(z - 1)$ when $\sigma > 0$ ; no cap needed.                                      |
| 2        | Proposition 3.10, small-height channel population; Proposition 3.7                       | Again $\tau = o(B)$ and the same unrestricted host bound; retain the rational contributions from channel enumeration.                                      | $E_N(N^{3/2} + Nz^{B/2} + Nz^{B/3})$ . Leave uncapped.                                                                                          |
| 3        | Proposition 3.10, remaining shallow population; Proposition 3.7                          | $\sigma = o(B)$ and $\tau \leq B/6 + o(B)$ on $N^{3/2+o(1)}$ hosts.                                                                                        | $E_N N^{3/2} Q_z^{1/6}$ . Leave uncapped.                                                                                                       |
| 4        | Proposition 3.17, aligned deep-core exclusion                                            | A positive-density aligned residual core gives a short codeword; the integer Runge bound contradicts its height.                                           | This sector is empty for sufficiently large size with fixed parameters. It contributes zero independently of $z, T$ .                           |
| 5        | Lemma 3.16 and the paragraph after Proposition 3.17                                      | A bounded nontrivial component gives $N^{1+o(1)}$ hosts; $\tau \leq D^\# + c^\# \leq 2B/3 + o(B)$ .                                                        | $E_N NQ_z^{2/3}$ . Leave uncapped. The bounded-component/Pell estimates are imported, not reproved here.                                        |
| 6        | Lemma 3.18, Proposition 3.19; Lemma 3.14 for partners                                    | Two defects restrict one start to $N^{o(1)}$ choices; the partner has at most $N^{1/2+o(1)}$ choices. Thus $N^{1/2+o(1)}$ hosts and $\tau \leq B + o(B)$ . | Pointwise weight $\leq E_N Q_z$ . Cap to $E_N \min(T, Q_z)$ using $\min(T, E_N Q_z) \leq E_N \min(T, Q_z)$ . Total $E_N N^{1/2} \min(T, Q_z)$ . |

| Part   | Source result                                    | Host and dimension information                                                                                                                                                                          | Weight and cap operation                                                                                                                                 |
|--------|--------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 7      | Equation (3.14) and the sector-7 paragraph       | With $j = s + e_k \geq T_B + 1$ and $D^\# \leq 2$ , the exact deficit yields $\tau \leq 2B/3 + O(1)$ . The same deep-host argument gives $N^{1+o(1)}$ hosts.                                            | $E_N N Q_z^{2/3}$ . Leave uncapped.                                                                                                                      |
| 8a     | Lemmas 3.20–3.22; first part of Proposition 3.23 | On a slice with larger start $X$ , $m_j = B - 3j - 1 \geq 2$ gives two small kernels. Shifted energy counts $X^{2/3+\varepsilon} B^2$ starts and each has $X^{o(1)}$ partners. Also $\tau \leq B + 2$ . | Weight $\leq z^2 Q_z$ . The counted population and partner bounds precede $T$ . Total $X^{2/3+\varepsilon} E_X \min(T, Q_z)$ after a fixed $z^2$ factor. |
| 8b     | Lemma 3.21 and second part of Proposition 3.23   | $m_j \leq 1$ implies $3\tau \leq 2B + 8$ . One small-kernel container gives $X^{3/4+o(1)}$ hosts including the partner estimate.                                                                        | Weight $\leq z^{8/3} Q_z^{2/3}$ . Leave uncapped; the slice sum is absorbed by $E_N N Q_z^{2/3}$ .                                                       |
| Values | Proposition 3.24; cap Proposition 3.25           | $\rho_{\text{rel}} \leq \rho_{\text{val}} \leq \rho_{\text{rel}} + 2$ . New full-value relation hosts are counted by Proposition 3.7 without block-parity restrictions.                                 | Use the inequality below; the extra indicator costs $N^{3/2+o(1)}$ independently of $T$ .                                                                |

The full-value step is explicitly

$$\min\{T, z^{\rho_{\text{val}}} - 1\} \leq z^2 \min\{T, z^{\rho_{\text{rel}}} - 1\} + (z^2 - 1) \mathbf{1}_{\{\rho_{\text{val}} > 0\}}.$$

If the full kernel is zero, both weights vanish. Otherwise write  $z^{\rho_{\text{val}}} - 1 \leq z^2(z^{\rho_{\text{rel}}} - 1) + (z^2 - 1)$  and use subadditivity of the cap. The new indicator is not assumed to be contained in the relative-sign host set.

For terminal slices, the source article keeps the original canonical data when exchanging the larger and smaller coordinates for counting. Its partner bound needs only the upper height bound, not comparability of the two starts. The constants are uniform in the fixed logarithmic band. Summing slices geometrically gives the terminal term  $N^{2/3+\varepsilon} \min(T, Q_z)$ . These facts are the reason cap uniformity survives; no cap-dependent counting population is introduced.

This ledger separates two operations that must not be conflated. Replacing  $2^\rho$  by  $z^\rho$  changes only a weight on a binary dimension. Passing to another value group instead uses the integer-host and unimodular-minor arguments of Section 8. In particular, it does not rely on a false comparison between nullities in different finite characteristics.

### C.3 Scientific dependency map

| Result      | Inputs actually used                                                             | Retained conclusion and main restriction                                                  |
|-------------|----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| Theorem 2.2 | Classical process Chen–Stein [3,4], disintegration and common intensity parts.   | Marked field on one prelimit space; an exact graph for marked collections.                |
| Theorem 8.4 | Proposition 8.2 and Lemmas A.2 and 8.3; integer minors and Appendix A selection. | Fixed $H \times \mathbb{T}^d$ , logarithmic length, controlled erasure and small overlap. |

| Result                     | Inputs actually used                                                                                        | Retained conclusion and main restriction                                                                                      |
|----------------------------|-------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Corollary 9.1              | Projection cost of a dictionary and Theorem 8.4.                                                            | Growing dictionaries; structured lifts measured in their quotient.                                                            |
| Theorem 6.1                | Characters, gcd certificates, support collisions, Rankin and prime graphs.                                  | Shrinking rational targets, equality/inversion clusters; no deep binary profile.                                              |
| Theorem 7.2                | $L^\infty$ mixing, rank-two domination, Lemma B.2, first-active marks.                                      | Permanent independent prime biases; a small power of $\log N$ sampled values.                                                 |
| Theorem 3.1                | Convolution, disjoint bases, exceptional-pair classification, mixed edges and Bonferroni.                   | $F(n)$ on masks up to $CN$ with targets $O(N^{-1})$ ; true prime intensities.                                                 |
| Proposition 3.5            | A power-singular density, independent prime-square hits and prime counting.                                 | Sharpness below $L^2$ for this target and arbitrary masks, not for every possible formulation.                                |
| Proposition 3.6            | Independent prime sites, exact cusp mass, a multiple-point event and the classical scalar Stein factor.     | Below $L^2$ : correct-intensity field obstruction with a Poisson count; (3.20) is restricted to independent prime-only cells. |
| Theorem 4.1                | Core decomposition, fixed semigroup sums, two-ceiling ledger (4.14).                                        | Fixed finite exceptional environment; exact points and a conditional cluster kernel.                                          |
| Theorem 5.1                | Prelimit transfer specialized to Haar, fixed-depth phase separation, coprime-core counts, reciprocal tails. | Corrected prefix count and common shrinking arc; no uniform irrational-environment limit rate.                                |
| Corollary 5.3 and (5.12)   | Integer multiplier order; divisibility and classical monomial-ideal finiteness.                             | Explicit parameters from complete exact certificates, not numerical guesses of phase relations.                               |
| Propositions 10.3 and 10.4 | Exact sign expansions, finite multivariate inversion and finite-source support.                             | Resolution-adapted conditional transfer; further correlation estimates are required for Liouville.                            |

## C.4 Quantifiers that remain separate

For Haar patterns,  $B \asymp \log N$ ,  $\mathfrak{m} \leq N^{1/2-\eta}$  and  $\Omega \rightarrow 0$  give (8.10). For rational observations under Haar,  $\log k = o(\sqrt{\log N \log \log N})$  gives the full-mark subexponential budget; phase-only subpolynomial growth follows a separate certificate-size budget. Under bounded prime densities, fixed  $K$  and  $k = o((\log N)^\kappa)$  use  $\kappa < \min\{I(\theta), -\theta \log \rho_K\}$ . Under  $L^2$ ,  $K_2, C, A$  are fixed, each target has mass at most  $A/N$ , and the prime-source intensity may diverge.

The finite exceptional set remains fixed. Its composite-core target unions and prime-core unions use different ceilings; only the former are layer-truncated. The lattice limit fixes the environment before the size limit and then removes a finite multiplier cutoff. None of these fixed constants becomes an admissible growing parameter by substitution into an implicit constant. The logarithmic rates are sufficient budgets, not claims of optimal convergence speed.

## D Finite identities for visible packets and aggregation

This appendix proves the finite algebra behind the resolution statements in Section 10. It applies to any deterministic sign sequence and any law of the origin. The identities do not supply asymptotic cancellation of their remaining correlations.

## D.1 Even spectrum, degree compression and a diameter band

For these identities use the marker convention  $I_L(n) = R_L(n+1)$ ; then  $Z = \sum_{i=0}^{H-1} I_L(U+i)$ . On  $L+1$  signs,

$$I_L(x_0, \dots, x_L) = a \sum_{\substack{S \subseteq \{0, \dots, L\} \\ |S| \text{ even}}} (-1)^{\mathbf{1}_{\{0 \in S\}}} \prod_{j \in S} x_j. \quad (\text{D.1})$$

The formula follows by expanding the two admissible words  $(-, +, \dots, +)$  and its negative. Every nonempty even character has coefficient of modulus  $a$ ; all odd coefficients vanish. The unsigned count and every function of it are invariant under simultaneous sign reversal. Estimates for odd correlations alone therefore do not calibrate this observable.

Let  $y_0(n) = -\varepsilon(n)$ ,  $y_j(n) = \varepsilon(n+j)$  for  $1 \leq j \leq L$ , and  $M_L(n) = \#\{j : y_j(n) = -1\}$ . Define the coefficient-normalized binary Krawtchouk polynomials

$$\mathcal{K}_d(m; B) = \sum_{r=0}^d (-1)^r \binom{m}{r} \binom{B-m}{d-r}, \quad B = L+1.$$

They equal  $\binom{B}{d} K_d(m; 1/2, B)$  in the convention of [14, Eq. 18.23.3]. The finite generating product

$$\prod_{j=0}^L (1 + ty_j) = (1-t)^{M_L} (1+t)^{B-M_L} = \sum_{d=0}^B \mathcal{K}_d(M_L; B) t^d$$

gives, by its even part at  $t=1$ ,

$$\frac{\mathbb{E}Z}{\tau_H} - 1 = \sum_{\substack{2 \leq d \leq L+1 \\ d \text{ even}}} \frac{1}{H} \sum_{i=0}^{H-1} \mathbb{E} \mathcal{K}_d(M_L(U+i); L+1). \quad (\text{D.2})$$

This compresses the first moment, not the joint law of windows. The sum remains signed. With  $S_L = -\varepsilon(n) + \sum_{j=1}^L \varepsilon(n+j)$ ,

$$\mathcal{K}_2 = \frac{S_L^2 - B}{2}, \quad \mathcal{K}_4 = \frac{S_L^4 - (6B-8)S_L^2 + 3B(B-2)}{24}.$$

A uniform bound  $\epsilon$  for the distinct two-point correlations used here gives a degree-two contribution at most  $\binom{L+1}{2} \epsilon$ . It says nothing by itself about the signed sum of the remaining even degrees.

Summation over starts produces another filter. If an even support shape has minimum 0 and diameter  $d \leq L$ , its volume coefficient is  $L-d-1$ . In particular, an entire diameter band telescopes.

**Proposition D.1** (A diameter band reduces to the endpoints). *For  $L \geq 3$ , put*

$$G_L(n) = \varepsilon(n) \varepsilon(n+L-1) \mathbf{1}_{\{\varepsilon(n+1)=\dots=\varepsilon(n+L-2)\}}.$$

*The Fourier contribution to  $Z$  of all even support shapes of diameter  $L-1$  is exactly*

$$Z^{(\text{diam}=L-1)}(U) = \frac{1}{8} (G_L(U+H) - G_L(U)). \quad (\text{D.3})$$

*It is a signed Fourier contribution, not a nonnegative count.*

*Proof.* For such a shape only translations 0, 1 fit inside a word. Their signs in (D.1) are respectively minus and plus. Summing even subsets containing both endpoints over the  $L-2$  interior signs gives  $2^{L-3} G_L(n)$ . Multiply by  $a = 2^{-L}$  and telescope the sum over  $H$  starts.  $\square$

For a law of  $U$  satisfying  $|\mathbb{E}f(U+h) - \mathbb{E}f(U)| \leq \eta_U(h)$  for all  $|f| \leq 1$ , its expected contribution is at most  $\eta_U(H)/8$ . Uniform  $U$  on  $X$  consecutive integers has  $\eta_U(H) \leq 2H/X$ , hence the bound  $H/(4X)$ . The cancellation is only for this first scalar contraction; it cannot be inserted without proof into a product of constrained windows or a general Stein test. The following subsections give the complete shape formula, the remaining coefficient mass and a high-degree counterexample.

## D.2 A high-degree counterexample

For even  $B$ , consider the probability on  $y \in \{-1, 1\}^B$

$$\mu_c(y) = 2^{-B} \left( 1 + c \prod_{j=0}^{B-1} y_j \right), \quad -1 \leq c \leq 1.$$

Every proper marginal is independent and fair, and every odd correlation vanishes. Nevertheless,

$$\mathbb{P}_{\mu_c}(y_0 = \cdots = y_{B-1}) = 2^{1-B}(1+c).$$

This event is forbidden for  $c = -1$  and has twice its fair mass for  $c = 1$ . The example is a finite sign law, not a multiplicative model. It shows why good proper marginals cannot remove a highest-degree character which the observable still sees; it does not contradict the erasure theorem, which has additional source and rank hypotheses.

Writing  $S = \sum_j y_j = B - 2M$ , Newton's identities applied to  $y_j^2 = 1$  give

$$\mathcal{K}_2(M; B) = \frac{S^2 - B}{2}, \quad \mathcal{K}_4(M; B) = \frac{S^4 - (6B - 8)S^2 + 3B(B - 2)}{24}.$$

The polynomial coefficient normalization differs from the normalized Krawtchouk convention by  $\binom{B}{d}$ , as stated in Section 10; the finite product verifies it independently of a normalization convention.

## D.3 The full diameter coefficient

For an even support shape  $S$  with  $\min S = 0$  and  $\max S = d \leq L$ , write  $\chi_S(n) = \prod_{s \in S} \varepsilon(n + s)$ . It can occur in a word at translations  $j = 0, \dots, L - d$ . Translation zero includes the marker and has sign minus; all the others have sign plus. Its contribution to the count in the marker convention is exactly

$$a \left[ - \sum_{i=0}^{H-1} \chi_S(U + i) + \sum_{j=1}^{L-d} \sum_{i=0}^{H-1} \chi_S(U + i + j) \right].$$

A shift of  $j$  in a sum of  $H$  terms bounded by one changes it by at most  $2j$ . Thus this equals

$$a(L - d - 1) \sum_{i=0}^{H-1} \chi_S(U + i) + \mathcal{R}_S(U), \quad |\mathcal{R}_S(U)| \leq a(L - d)(L - d + 1). \quad (\text{D.4})$$

For  $d = L - 1$ , the volume coefficient is zero. Summing before taking absolute values over all even subsets containing both endpoints gives

$$\sum_{\substack{S \subseteq \{0, \dots, L-1\} \\ 0, L-1 \in S, |S| \text{ even}}} \chi_S(n) = 2^{L-3} G_L(n),$$

including  $L = 3$ , when the interior-constancy condition has one coordinate and is automatic. Formula (D.3) follows exactly.

There is one normalized even shape of diameter one and  $2^{d-2}$  of diameter  $d \geq 2$ . Direct summation, for  $L \geq 3$ , gives

$$\sum_{d=1}^L |L - d - 1| \# \{S : \min S = 0, \max S = d, |S| \text{ even}\} = 2^{L-1} - 1. \quad (\text{D.5})$$

The filter reduces a band to endpoints but does not make the total absolute coefficient mass polynomial in  $L$ .

For the particular quartic shape  $S = \{0, 1, 2, 3\}$ ,  $L \geq 3$ , define

$$\overline{C}_4 = H^{-1} \sum_{i=0}^{H-1} \mathbb{E}[\varepsilon(U+i)\varepsilon(U+i+1)\varepsilon(U+i+2)\varepsilon(U+i+3)].$$

Its contribution is

$$\tau_H(L-4)\overline{C}_4 + O(\tau_H L^2/H).$$

This weight belongs to the first-moment packet. A subtraction or a connected-kernel construction elsewhere may assign the same geometric support another weight; such a calculation cannot be substituted without identifying the packet. Nor is separate cancellation of this quartic term claimed necessary: different shapes and degrees can cancel in the prescribed signed sum.

#### D.4 Proof of the scalar factorial-moment bound

This is a finite binomial-inversion implementation of the factorial-moment method; [6] provides its classical Poisson-convergence context. The constants below are checked directly.

*Proof of Lemma 10.1.* For  $0 \leq r < K$ , truncate binomial inversion at order  $K-1$ :

$$T_r(Z) = \sum_{k=r}^{K-1} \frac{(-1)^{k-r}}{r!(k-r)!} \mathbb{E}(Z)_k, \quad |\mathbb{P}(Z=r) - T_r(Z)| \leq \frac{\mathbb{E}(Z)_K}{r!(K-r)!}.$$

For an integer  $z < K$ , the expansion is exact. For  $z \geq K$ , the absolute truncated sum is  $\binom{z}{r} \binom{z-r-1}{K-r-1}$ , at most  $\binom{z}{K} \binom{K}{r}$ , since their ratio is  $(K-r)/(z-r) \leq 1$ . Apply the same estimate to  $P \sim \text{Pois}(\tau)$  and bound the two tails by  $\mathbb{E}(Z)_K/K!$  and  $\tau^K/K!$ . After summing over  $r < K$  and dividing by two, the coefficients of discrepancies of order  $k < K$  are  $2^{k-1}/k!$ . The remainders and tails total at most

$$\frac{2^{K-1}}{K!} (\mathbb{E}(Z)_K + \tau^K).$$

Use  $\mathbb{E}(Z)_K \leq \tau^K + |\mathbb{E}(Z)_K - \tau^K|$ . □

#### D.5 Multivariate binomial inversion

For  $z \in \mathbb{N}_0^m$ , let  $|z| = \sum_c z_c$ ; write  $\binom{z}{\alpha} = \prod_c \binom{z_c}{\alpha_c}$  and  $\alpha! = \prod_c \alpha_c!$ . Let  $Z = (Z_c)_{c=1}^m$  be any nonnegative-integer vector, and  $\lambda_c \geq 0$  with  $\lambda = \sum_c \lambda_c > 0$ . Define

$$m_\alpha = \mathbb{E} \binom{Z}{\alpha}, \quad m_\alpha^P = \prod_c \frac{\lambda_c^{\alpha_c}}{\alpha_c!}.$$

**Lemma D.2** (Finite multivariate Bonferroni bound). *For every  $K \geq 1$  with  $\mathbb{E} \binom{|Z|}{K} < \infty$ ,*

$$d_{\text{TV}} \left( \mathcal{L}(Z), \bigotimes_c \text{Pois}(\lambda_c) \right) \leq \frac{1}{2} \sum_{1 \leq |\alpha| \leq K} 2^{|\alpha|} |m_\alpha - m_\alpha^P| + \frac{(2\lambda)^K}{K!}. \quad (\text{D.6})$$

*There is no dimension-dependent constant.*

*Proof.* For a configuration  $z$  with  $|z| < K$ , truncate binomial inversion at total degree  $K - 1$ :

$$T_z = \sum_{\substack{\alpha \geq z \\ |\alpha| < K}} (-1)^{|\alpha| - |z|} \binom{\alpha}{z} m_\alpha.$$

For a deterministic value  $n \geq z$ , use  $\binom{n}{\alpha} \binom{\alpha}{z} = \binom{n}{z} \binom{n-z}{\alpha-z}$ . Vandermonde's identity makes the sum of terms of residual degree  $j$  equal to  $\binom{n}{z} \binom{|n| - |z|}{j}$ . The one-variable truncated identity then bounds the pointwise remainder by

$$\sum_{|\alpha|=K} \binom{n}{\alpha} \binom{\alpha}{z}.$$

For  $n$  not coordinatewise at least  $z$ , both sides of the relevant inversion vanish. Averaging proves the remainder bound with  $m_\alpha$ .

Sum it over  $|z| < K$ . For  $|\alpha| = K$ , the coefficient is  $\sum_{z \leq \alpha, |z| < K} \binom{\alpha}{z} = 2^K - 1$ . Adding the tail bound  $\mathbb{P}(|Z| \geq K) \leq \mathbb{E} \binom{|Z|}{K} = \sum_{|\alpha|=K} m_\alpha$  changes this to  $2^K$ . Apply the same argument to the product of Poisson variables. Each discrepancy of degree  $k < K$  appears with coefficient  $2^k$  before the factor  $1/2$  in total variation. For degree  $K$ , use

$$\sum_{|\alpha|=K} m_\alpha \leq \frac{\lambda^K}{K!} + \sum_{|\alpha|=K} |m_\alpha - m_\alpha^P|, \quad \sum_{|\alpha|=K} m_\alpha^P = \frac{\lambda^K}{K!}.$$

This gives (D.6), including its displayed remainder.  $\square$

For completeness, let a cell tuple  $\mathbf{c}$  have multiplicities  $\alpha_c$ , total  $k$ . Its factorial moment is  $\mathbb{E} \prod_c (Z_c)_{\alpha_c}$ . There are  $k!/\alpha_c!$  ordered tuples with those multiplicities. Therefore

$$\sum_{|\alpha|=k} |m_\alpha - m_\alpha^P| = \frac{1}{k!} \sum_{\mathbf{c} \in \{1, \dots, m\}^k} \left| \mathbb{E} \prod_c (Z_c)_{\alpha_c} - \prod_i \lambda_{c_i} \right|.$$

For run counts, expand only compatible starts within each labelled tuple, using (10.2). Their correlation discrepancy is bounded by  $\tau_H^k \mathcal{A}_k^{\mathcal{P}}$ . All missing independent placements across all labels are counted once; their total is at most  $\tau_H^k k(k-1)L/H$ . Substitution in (D.6) proves Proposition 10.3 with no loss in the number of cells.

The norm contraction under merging cells follows directly from combining their signed placement sums. It compares packet norms, not lower bounds on total variation. Finite computations of these norms or of Krawtchouk sectors can diagnose cancellation and resolution loss, but do not establish their asymptotic smallness.

## Availability and formalization status

The source package contains the complete LaTeX manuscript, its frozen bibliography, compilation instructions and finite diagnostic programs. The external arithmetic input to the long-pattern branch is identified in Section C, with exact references to the source article and its technical companion.

The results of this article are not claimed to be formalized in Lean. The formal development associated with the binary source has the scope stated in that work; it does not formalize the arguments of this article. Finite diagnostics detect algebraic and implementation errors; they do not replace the asymptotic proofs.

## Declaration on artificial-intelligence assistance

The research was organized and directed by the author with substantial assistance from GPT-6 Astra (OpenAI), the only language-model system used for this article. It was used to explore candidate lemmas and proof routes, draft, translate and reorganize the manuscript, write diagnostic programs, check references and build artifacts, and conduct adversarial reviews through separate instances.

The author determines the claims and scope and accepts responsibility for the arguments, references and any remaining errors. AI-assisted checks are not independent human peer review. The mathematical claims rest on the printed arguments and the stated external results, not on model output or the success of finite diagnostic tests.

## Competing interests

The author declares no competing interests.

## References

- [1] N. Álvarez, V. Becher and M. Mereb. *Poisson generic sequences*. International Mathematics Research Notices 2023(24): 20970–20987. 2023. doi: [10.1093/imrn/rnac234](https://doi.org/10.1093/imrn/rnac234). arXiv: [2202.01632](https://arxiv.org/abs/2202.01632).
- [2] Nicolás Álvarez, Verónica Becher, Eda Cesaratto, Martín Mereb, Yuval Peres and Benjamin Weiss. *Poisson genericity in numeration systems with exponentially mixing probabilities*. Transactions of the American Mathematical Society 379(5): 3403–3423. 2026. doi: [10.1090/tran/9592](https://doi.org/10.1090/tran/9592). arXiv: [2411.04116](https://arxiv.org/abs/2411.04116).
- [3] Richard Arratia, Larry Goldstein and Louis Gordon. *Two moments suffice for Poisson approximations: The Chen–Stein method*. Annals of Probability 17(1): 9–25. 1989. doi: [10.1214/aop/1176991491](https://doi.org/10.1214/aop/1176991491).
- [4] 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).
- [5] Matthias Aschenbrenner and Wai-Yan Pong. *Orderings of monomial ideals*. Fundamenta Mathematicae 181(1): 27–74. 2004. doi: [10.4064/fm181-1-2](https://doi.org/10.4064/fm181-1-2). arXiv: [math/0305384](https://arxiv.org/abs/math/0305384).
- [6] A. D. Barbour. *Poisson convergence and random graphs*. Mathematical Proceedings of the Cambridge Philosophical Society 92(2): 349–359. 1982. doi: [10.1017/S0305004100059995](https://doi.org/10.1017/S0305004100059995).
- [7] A. D. Barbour and Peter Hall. *On the rate of Poisson convergence*. Mathematical Proceedings of the Cambridge Philosophical Society 95(3): 473–480. 1984. doi: [10.1017/S0305004100061806](https://doi.org/10.1017/S0305004100061806).
- [8] Debmalya Basak, Kunjakanan Nath and Alexandru Zaharescu. *Poisson behavior for chains of small quadratic non-residues*. International Mathematics Research Notices 2024(4): 3356–3390. 2024. doi: [10.1093/imrn/rnad168](https://doi.org/10.1093/imrn/rnad168).
- [9] Yeow Meng Chee, Han Mao Kiah, Punarbasu Purkayastha and Chengmin Wang. *Cross-bifix-free codes within a constant factor of optimality*. IEEE Transactions on Information Theory 59(7): 4668–4674. 2013. doi: [10.1109/TIT.2013.2252952](https://doi.org/10.1109/TIT.2013.2252952).
- [10] Stefanka Chukova and Leda D. Minkova. *Characterization of the Pólya–Aeppli process*. Stochastic Analysis and Applications 31(4): 590–599. 2013. doi: [10.1080/07362994.2013.798994](https://doi.org/10.1080/07362994.2013.798994).
- [11] Régis de la Bretèche, Victor Y. Wang and Max Wenqiang Xu. *Random multiplicative functions and making squares from polynomial values*. 2026. arXiv: [2607.06398v1](https://arxiv.org/abs/2607.06398v1). <https://arxiv.org/abs/2607.06398v1>.
- [12] Geoffrey K. Eagleson. *Some simple conditions for limit theorems to be mixing*. Theory of Probability and Its Applications 21(3): 637–643. 1977. doi: [10.1137/1121078](https://doi.org/10.1137/1121078).
- [13] David Eisenbud and Bernd Sturmfels. *Binomial ideals*. Duke Mathematical Journal 84(1): 1–45. 1996. doi: [10.1215/S0012-7094-96-08401-X](https://doi.org/10.1215/S0012-7094-96-08401-X). arXiv: [alg-geom/9401001](https://arxiv.org/abs/alg-geom/9401001).
- [14] NIST Digital Library of Mathematical Functions. *NIST Digital Library of Mathematical Functions: Chapter 18, Orthogonal Polynomials*. 2026-09-15. <https://dlmf.nist.gov/18>.
- [15] Sébastien Gouëzel. *Variations around Eagleson’s theorem on mixing limit theorems for dynamical systems*. Ergodic Theory and Dynamical Systems 40(12): 3368–3374. 2020. doi: [10.1017/etds.2019.42](https://doi.org/10.1017/etds.2019.42). arXiv: [1803.11450](https://arxiv.org/abs/1803.11450).
- [16] Adam J. Harper, Kannan Soundararajan and Max Wenqiang Xu. *Distribution of random multiplicative functions in short intervals, with proper normalization*. 2026. arXiv: [2606.29040v1](https://arxiv.org/abs/2606.29040v1). <https://arxiv.org/abs/2606.29040v1>.
- [17] Nicolai Haydn and Sandro Vaienti. *The compound Poisson distribution and return times in dynamical systems*. Probability Theory and Related Fields 144: 517–542. 2009. doi: [10.1007/s00440-008-0153-y](https://doi.org/10.1007/s00440-008-0153-y). arXiv: [0804.1032](https://arxiv.org/abs/0804.1032).
- [18] Michael Hochman and Nicolò Paviato. *A threshold for Poisson behavior of non-stationary product measures*. Ergodic Theory and Dynamical Systems 46(4): 1009–1019. 2026. doi: [10.1017/etds.2025.10258](https://doi.org/10.1017/etds.2025.10258). arXiv: [2501.11423](https://arxiv.org/abs/2501.11423).
- [19] Shizuo Kakutani. *On equivalence of infinite product measures*. Annals of Mathematics 49(1): 214–224. 1948. doi: [10.2307/1969123](https://doi.org/10.2307/1969123).

- [20] Yuri Kifer and Ariel Rapaport. *Poisson and compound Poisson approximations in conventional and nonconventional setups*. Probability Theory and Related Fields 160: 797–831. 2014. doi: [10.1007/s00440-013-0541-9](https://doi.org/10.1007/s00440-013-0541-9). arXiv: [1211.5238](https://arxiv.org/abs/1211.5238).
- [21] 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).
- [22] Kaisa Matomäki, Maksym Radziwiłł and Terence Tao. *Sign patterns of the Liouville and Möbius functions*. Forum of Mathematics, Sigma 4, e14. 2016. doi: [10.1017/fms.2016.6](https://doi.org/10.1017/fms.2016.6). arXiv: [1509.01545](https://arxiv.org/abs/1509.01545).
- [23] Ezra Miller and Bernd Sturmfels. *Combinatorial Commutative Algebra*. Graduate Texts in Mathematics. Springer. New York. 2005. doi: [10.1007/b138602](https://doi.org/10.1007/b138602).
- [24] Joseph Najnudel. *On consecutive values of random completely multiplicative functions*. Electronic Journal of Probability 25, 59. 2020. doi: [10.1214/20-EJP456](https://doi.org/10.1214/20-EJP456). arXiv: [1702.01470](https://arxiv.org/abs/1702.01470).
- [25] Françoise Pène and Benoît Saussol. *Spatio-temporal Poisson processes for visits to small sets*. Israel Journal of Mathematics 240(2): 625–665. 2020. doi: [10.1007/s11856-020-2074-0](https://doi.org/10.1007/s11856-020-2074-0). arXiv: [1803.06865](https://arxiv.org/abs/1803.06865).
- [26] 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).
- [27] 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).
- [28] Ariel Rapaport. *Compound Poisson statistics for multiple returns in shrinking cylinders for mixing processes*. Ergodic Theory and Dynamical Systems 36(5): 1616–1643. 2016. doi: [10.1017/etds.2014.139](https://doi.org/10.1017/etds.2014.139). arXiv: [1404.3516](https://arxiv.org/abs/1404.3516).
- [29] Gesine Reinert and Sophie Schbath. *Compound Poisson and Poisson process approximations for occurrences of multiple words in Markov chains*. Journal of Computational Biology 5(2): 223–253. 1998. doi: [10.1089/cmb.1998.5.223](https://doi.org/10.1089/cmb.1998.5.223).
- [30] Etienne Roquain and Sophie Schbath. *Improved compound Poisson approximation for the number of occurrences of any rare word family in a stationary Markov chain*. Advances in Applied Probability 39(1): 128–140. 2007. doi: [10.1239/aap/1175266472](https://doi.org/10.1239/aap/1175266472).
- [31] Terence Tao. *Mertens’ theorems*. 2013-12-11. <https://terrytao.wordpress.com/2013/12/11/mertens-theorems/>.
- [32] Terence Tao. *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*. Forum of Mathematics, Pi 4, e8. 2016. doi: [10.1017/fmp.2016.6](https://doi.org/10.1017/fmp.2016.6). arXiv: [1509.05422](https://arxiv.org/abs/1509.05422).
- [33] Terence Tao and Joni Teräväinen. *Odd order cases of the logarithmically averaged Chowla conjecture*. Journal de théorie des nombres de Bordeaux 30(3): 997–1015. 2018. doi: [10.5802/jtnb.1062](https://doi.org/10.5802/jtnb.1062). [https://www.numdam.org/item/JTNB\\_2018\\_\\_30\\_3\\_997\\_0/](https://www.numdam.org/item/JTNB_2018__30_3_997_0/).
- [34] Terence Tao. *Products of consecutive integers with unusual anatomy*. 2026. arXiv: [2603.27990v2](https://arxiv.org/abs/2603.27990v2). <https://arxiv.org/abs/2603.27990v2>.
- [35] Gérald Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. 3 ed. Graduate Studies in Mathematics. American Mathematical Society. Providence, RI. 2015. doi: [10.1090/gsm/163](https://doi.org/10.1090/gsm/163).
- [36] J. Turk. *Multiplicative properties of integers in short intervals*. Indagationes Mathematicae (Proceedings) 83(4): 429–436. 1980. doi: [10.1016/1385-7258\(80\)90044-X](https://doi.org/10.1016/1385-7258(80)90044-X).
- [37] Guangqu Zheng. *Four-moment criteria for Poisson convergence on Poisson and Rademacher chaoses*. 2026. arXiv: [2608.12451v1](https://arxiv.org/abs/2608.12451v1). <https://arxiv.org/abs/2608.12451v1>.