

Long runs and rare patterns

of a random completely multiplicative function

Brice Pouly

Version 3

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

Abstract

Let f be a random completely multiplicative function with independent symmetric signs at the primes. We study rare constant runs and prescribed words at logarithmic lengths, despite long-range multiplicative identities. A uniform weighted two-window square-relation estimate, with critical bound $O_\epsilon(N^{5/3+\epsilon})$, separates exact rational relations from residual components.

At critical intensity, this estimate gives Poisson comparisons retaining positions, exact excess lengths and signs. For dictionaries, uniform bounds under size and overlap conditions are complemented by typical bounds obtained by averaging the distance after fixing the dictionary. In a quantitatively controlled regime of diverging intensity, an exact one-word coupling by largest-odd-prime pivots compares the full signed run field under prescribed small-prime conditioning. For fixed admissible window parameters and almost every fixed realization of f , run-start counts in prescribed near-macroscopic windows satisfy an empirical Poisson law along dyadic scales.

In the exceptionally rare regime, a relative marked comparison separates the boundary event from bulk occurrences. Conditional on existence, the first location, sign and overshoot have a two-source lattice law with scale-dependent weights. The two overshoots use different clocks: prime rank at the boundary and integer distance in the bulk.

Keywords. Random multiplicative functions; long runs; rare patterns; Poisson approximation; square-product relations; boundary effects.

MSC 2020. 11N64; 60F05; 60C05; 11D45; 94B05.

Contents

1	Introduction	2
2	Affine events and private prime coordinates	8
3	A uniform two-window relation estimate	12
4	From arithmetic relations to Poisson laws	29
5	Rare word fields, exact marks and the Poisson staircase	32
6	Arithmetic information and resolved conditional laws	43
7	Finite prefixes and the arithmetic boundary	48
8	What the method isolates	62
	Formalization and Palomar registrations	64
	References	65

1 Introduction

In a sequence of independent fair signs, a long constant run is rare for a simple reason: each additional sign must agree with the preceding one. A left-maximal run of length at least L starts with probability 2^{-L} , so N potential starts have natural intensity $N2^{-L}$. The transition between many runs and no runs occurs at $L = \log_2 N + O(1)$; recording starts rather than all constant windows removes the local clustering within a run [5,3,20].

Logarithmic blocks also underlie the original Erdős–Rényi law of large numbers [17]; the distributional and marked run laws used here are the later results cited above.

Now impose complete multiplicativity. Choose the signs $f(p)$ independently at the primes and set

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

In particular $f(n^2) = 1$. This is the completely multiplicative, sometimes called *extended Rademacher*, model, not the squarefree-supported model. Random multiplicative models go back to Wintner [42]; the completely multiplicative convention here is fixed by (1.1). The identity $f(2n) = f(2)f(n)$ persists at every scale. Distant blocks therefore cannot simply be treated as independent: the effect of these identities must be counted. Nevertheless, the critical run statistics have the independent-sign Poisson form.

There is a second, quite different mechanism at the left boundary. Requiring $f(1), \dots, f(L)$ to be equal is exactly the requirement that every prime sign up to L be positive. Its probability is $2^{-\pi(L)}$, not 2^{-L} . Thus a sufficiently exceptional run can be produced either by many potential starts in the bulk or by a single arithmetically cheaper event at the boundary. Our results quantify both mechanisms and their competition.

The same approach also distinguishes many prescribed words at once. Two sources of dependence must be separated: words can overlap because a suffix matches a prefix, and distant occurrences can interact because a product of their integer arguments is a square. The first is a property of the dictionary; the second is a property of the arithmetic sequence. Keeping them separate is the organizing principle of the paper.

1.1 Main results

For $x \geq 2$ and $L \geq 1$, let

$$J_{x,L} = \mathbf{1}_{\{f(x-1)=-f(x), f(x)=\dots=f(x+L-1)\}}. \quad (1.2)$$

Write $I_N = [N, 2N) \cap \mathbb{Z}$, $Z_{N,L} = \sum_{x \in I_N} J_{x,L}$ and $\lambda_N = N2^{-L}$. On a start, let ℓ_x be the maximal rightward run length. Every such length is almost surely finite. Define

$$K_{x,e} = \mathbf{1}_{\{J_{x,L}=1, \ell_x=L+e\}}, \quad K_{x,e}^s = K_{x,e} \mathbf{1}_{\{f(x)=s\}}, \quad e \in \mathbb{N}_0, s \in \{\pm 1\}.$$

Let $\nu_{\text{geo}}\{e\} = 2^{-e-1}$ and $\nu_{\pm} = (\delta_{-1} + \delta_{+1})/2$. A *word* is a prescribed finite sign sequence, a *dictionary* is a set of such words, and *pattern* is the broader term for the corresponding occurrence observable. All logarithms without an indicated base are natural. All total-variation distances below use the convention $d_{\text{TV}}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)|$.

Theorem 1.1 (Critical sign-resolved run field). *Fix $C > 0$. Uniformly for integer lengths $|L - \log_2 N| \leq C$,*

$$d_{\text{TV}}\left(\mathcal{L}((K_{x,e}^s)_{x,e,s}), \bigotimes_{x,e,s} \text{Pois}(2^{-L-e-2})\right) \leq \exp\left[-\left(\frac{1}{\sqrt{2}} + o_C(1)\right) \sqrt{\log N \log \log N}\right]. \quad (1.3)$$

Here (x, e, s) ranges over $I_N \times \mathbb{N}_0 \times \{\pm 1\}$, and the comparison is on the common countable lattice. In particular the same upper bound applies to $d_{\text{TV}}(\mathcal{L}(Z_{N,L}), \text{Pois}(\lambda_N))$.

Since L is integral, changing it by one divides λ_N by two; λ_N need not converge. This is why the approximation retains its moving parameter. For continuous positions, the consequence is a weak limit. Along a subsequence with $\lambda_N \rightarrow \lambda \in (0, \infty)$,

$$\sum_{x \in I_N} \sum_{e \geq 0} \sum_{s \in \{\pm 1\}} K_{x,e}^s \delta_{(x/N, e, s)} \implies \text{PPP}(\lambda dt \otimes \nu_{\text{geo}} \otimes \nu_{\pm})$$

weakly on finite point configurations on $[1, 2] \times \mathbb{N}_0 \times \{\pm 1\}$. The nonempty configurations on a deterministic spatial grid and those of a diffuse Poisson process are singular. Consequently the last convergence is weak, not in total variation. The lattice comparison is stronger in the appropriate prelimit space: a single coupling controls positions, excesses, signs, threshold counts and order statistics simultaneously. In the bulk, positive and negative long runs have equal limiting intensities, despite the deterministic positivity of every square.

A complete prefix field at growing intensity. There is also a comparison beyond bounded critical intensity. Let $L = \lfloor \log_2 M \rfloor - a_M$, with $a_M \geq 0$ integral and $a_M = o(\log M)$, and put $\Lambda = (M - L)2^{-L} \rightarrow \infty$. For an event of cost $I_M = -\log \mathbb{P}(\mathcal{A}_M)$ measurable in the hard prime field $\mathcal{F}_{Y_M^\sharp}$, the budget $I_M + \log \Lambda \leq V_M - c\nu_M$ gives

$$d_{\text{TV}}\left(\mathcal{L}((K_{x,e}^s)_{2 \leq x \leq M-L+1, e \geq 0, s} \mid \mathcal{A}_M), \bigotimes_{x,e,s} \text{Pois}(2^{-L-e-2})\right) \ll e^{-c'\nu_M} + M^{-1/3+\varepsilon}$$

for $0 < c' < c$ and $0 < \varepsilon < 1/3$. The exact hard scales are defined in (4.4). Theorem 7.7 retains every site, exact excess and actual sign; it requires only the base window to be contained and excludes the run from 1. Its selected-pivot coupling does not replace the finer critical rate above or the relative low-intensity estimates below. The full proof of Theorem 7.7 is given in the companion, Appendix F. Restricting the prefix at $4N$ gives the same one-factor hard-

field budget on $[N, 2N)$ (Corollary 7.7a); no change of rescaled grids or enlargement of the conditioning field is involved. As a consequence, for each fixed pair $0 < \alpha < 1$ and $\tau > 0$, Corollary 7.8a fixes almost every realization of f and obtains a spatial empirical $\text{Pois}(\tau)$ law for run-start counts along dyadic scales, in windows of $M \exp\{-\alpha V_M + O(1)\} = M^{1-o(1)}$ start sites.

Let R_M be the longest constant run in $(f(1), \dots, f(M))$.

Theorem 1.2 (Longest run and competing rare-event sources). *At critical lengths $|L - \log_2 M| \leq C$,*

$$\mathbb{P}(R_M < L) = \exp(-M2^{-L}) + o_C(1). \quad (1.4)$$

Moreover,

$$R_M = \log_2 M + O(\log \log M) \quad \text{almost surely.} \quad (1.5)$$

Suppose instead that $L \rightarrow \infty$, $L = O(\log M)$, and $\lambda_M = M2^{-L} \rightarrow 0$. Then

$$\mathbb{P}(R_M \geq L) = 2^{-\pi(L)} + M2^{-L} + o(2^{-\pi(L)} + M2^{-L}). \quad (1.6)$$

Let $X_{M,L}$ be the leftmost start of a length- L constant window contained in $[1, M]$, on the event $R_M \geq L$. If $L - \log_2 M - \pi(L) \rightarrow s \in \mathbb{R}$, then

$$\mathcal{L}(X_{M,L}/M \mid R_M \geq L) \implies \frac{1}{1+2^{-s}}\delta_0 + \frac{2^{-s}}{1+2^{-s}}\text{Leb}_{[0,1]}. \quad (1.7)$$

The boundary branch is more precise than its atom at macroscopic position zero: with conditional probability tending to one its start is exactly 1.

The centring in (1.7) compares two exact costs, $L - \log_2 M$ and $\pi(L)$, rather than replacing the prime count by a first-order asymptotic at a precision where that replacement is invalid. Theorem 7.11 strengthens the location law to a lattice comparison with moving source weights, sign and overshoot. Its geometric clock counts integers in the bulk and prime ranks at the boundary; no phase-convergent subsequence is needed for that prelimit comparison.

The third result treats many words at once, with their absolute signs prescribed. For a dictionary $\mathcal{W}_N \subset \{\pm 1\}^{B_N}$ of $m_N \geq 1$ distinct words, define its occurrence field by

$$I_{x,w} = \mathbf{1}_{\{f(x-1+j)=w_j \ (0 \leq j < B_N)\}}, \quad x \in I_N, \ w \in \mathcal{W}_N.$$

Write $w \sim_d v$ if the suffix $(w_d, \dots, w_{B_N-1})$ equals the prefix $(v_0, \dots, v_{B_N-d-1})$, and set

$$\Omega(\mathcal{W}_N) = \frac{1}{m_N} \sum_{w,v \in \mathcal{W}_N} \sum_{d=1}^{B_N-1} 2^{-d} \mathbf{1}_{\{w \sim_d v\}}. \quad (1.8)$$

Theorem 1.3 (Uniform and typical dictionaries of rare words). (i) Uniform dictionaries. Fix $C > 0$ and $0 < \eta < 1/2$. For deterministic dictionaries, suppose

$$m_N \leq N^{1/2-\eta}, \quad |B_N - \log_2(Nm_N)| \leq C, \quad \Omega(\mathcal{W}_N) \rightarrow 0.$$

Then

$$d_{\text{TV}} \left(\mathcal{L}((I_{x,w})_{x \in I_N, w \in \mathcal{W}_N}), \bigotimes_{x \in I_N, w \in \mathcal{W}_N} \text{Pois}(2^{-B_N}) \right) \rightarrow 0. \quad (1.9)$$

The same conclusion holds after any deterministic restriction of the site-word index set. The rare-word field is also asymptotically equivalent in total variation to the field formed from independent fair signs.

(ii) Typical dictionaries. Fix a positive logarithmic band for B_N and a compact subinterval of $(0, \infty)$ for $Nm_N 2^{-B_N}$. Choose $\mathcal{W}_N$ uniformly among the size- m_N dictionaries, independently of f . The two distances in part (i) tend to zero in probability over this choice, with the dictionary held fixed when forming the laws. No restriction $m_N \leq N^{1/2-\eta}$ is required in this part.

The total target intensity is $Nm_N 2^{-B_N}$, while each word has mean $N 2^{-B_N}$. The uniform assertion is proved quantitatively in Theorem 5.1; it includes explicit marker dictionaries with no overlaps (Corollary 5.6). The typical assertion, proved in Theorem 5.3, averages the full error rather than only the overlap weight. It permits $m_N \asymp N$ at $B_N = 2 \log_2 N + O(1)$, or any fixed polynomial dictionary size in a suitable logarithmic band. In each case the expected total number of occurrences remains bounded even though the number of possible types grows. Both position and word identity are retained.

The same typical conclusion holds for uniformly sampled affine cosets in $\mathbb{F}_2^{B_N}$ (Corollary 5.4). A matrix of $O((\log N)^2)$ bits can therefore describe an admissible random candidate containing order N words. This is a high-probability result over the description, not a deterministic certificate for an arbitrary matrix.

For a single word counted up to sign, take $\mathcal{W} = \{a, -a\}$. The run-start word $a = (-1, +1, \dots, +1)$ has overlap weight 2^{-L} . For a specified dictionary, nonvanishing weighted overlap is outside the small-error regime of the uniform bound; short-period words illustrate the possible local clusters. The compound-Poisson law in Corollary 5.9 describes the clusters produced by counting every constant window rather than only run starts.

The results share arithmetic profiles, but use different comparisons. The raw profile controls run fields; a marginal cap controls dictionaries uniformly, while averaging the word selection gives the typical theorem. For a growing prefix run field, the extra reciprocal-pivot estimate controls the changes actually made by one conditional forcing. The relative comparison resolves the competition between the exact boundary event and the bulk. The distinction between uniform and typical concerns the choice of observable, not a change of multiplicative model.

1.2 The mechanism and the arithmetic estimate

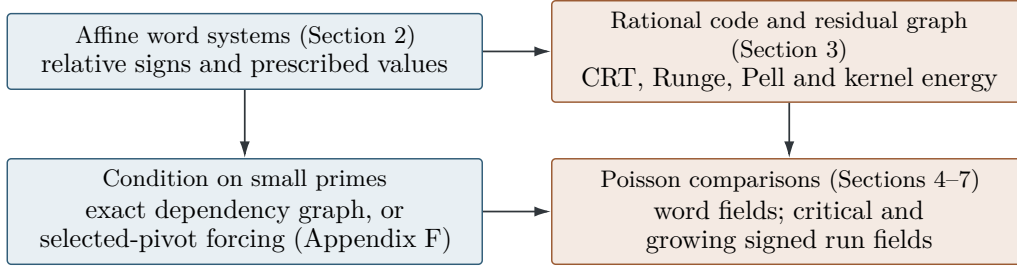


Figure 1: The arithmetic and probabilistic interfaces. The weighted two-window estimate bounds the separated-pair term. Private prime coordinates supply correct marginals and an exact conditional graph. For the growing prefix run field, selected pivots instead construct a full one-word conditional law; the directed change sets need not form a dependency graph. Local overlap is checked separately in both routes.

Write prime signs as $f(p) = (-1)^{X_p}$ with independent uniform $X_p \in \mathbb{F}_2$. A pattern up to sign prescribes L affine linear conditions on these prime coordinates. For two separated windows, let $\rho_L(x, y)$ be the dimension of the homogeneous relation space of their stacked systems. It depends on the windows, not on the prescribed signs. The affine Fourier identity gives

$$\left| \mathbb{P}(I_x = I_y = 1) - 2^{-2L} \right| \leq 2^{-2L} (2^{\rho_L(x, y)} - 1).$$

Thus the useful arithmetic object counts *all nonzero relations*, with their full multiplicity, rather than merely the pairs that carry one. For prescribed absolute words the relevant kernel omits the two block-parity conditions. Its dimension increases by at most two. A host count valid without parity restrictions shows that the same weighted bound survives; this is the first bridge from relative signs to the dictionary theorem.

The second bridge is probabilistic: at a fixed site the dictionary words are mutually exclusive, so the chance of a joint occurrence cannot exceed a single-site probability. Capping the relation weight at this ceiling, before summing over pairs, is what permits dictionaries of size $N^{1/2-\eta}$.

The uniform estimate proved in section 3 is, for fixed $0 < \delta < 1$ and $B = L + 1 \asymp \log M$,

$$\sum_{\substack{M^\delta \leq x, y < M \\ |x-y| > L}} (2^{\rho_L(x, y)} - 1) \ll_\varepsilon M^\varepsilon \left(M^{3/2} 2^{B/6} + M 2^{2B/3} + M^{2/3} 2^B \right). \quad (1.10)$$

The fixed logarithmic band and δ are understood in the implicit constant. This single geometry includes the dyadic and bounded-ratio estimates by restriction. At criticality the bound is $M^{5/3+\varepsilon}$.

Two devices make the proof work. Exact rational coincidences are first removed as a systematic binary code. The remaining relations are resolved by the graph of prime coordinates larger than B . For the hardest terminal population, determinants force many small odd kernels into the same short window. Counting the kernels separately is wasteful; a shifted pair-energy estimate detects the necessary concentration. Finally one conditions on prime signs below a larger cutoff Y . Most starts then have exact conditional marginal 2^{-L} , and disjoint prime supports give an exact conditional dependency graph. Chen–Stein is applied *after* this arithmetic separation, not as a substitute for it. For the prefix coupling, one forces a single prescribed word and charges only the prime coordinates actually modified. A prime shared by two supports need not be selected by that forcing; the finite example in section F.4 makes this distinction

explicit. The reciprocal-pivot estimate reduces this cost without discarding the two-window profile. Its analytic bound is also a consequence of the classical largest-prime-factor reciprocal asymptotics of Erdős–Ivić–Pomerance [16]; Appendix F gives both the direct uniform Rankin proof and the square-part reduction to that antecedent. The contribution of the coupling is its exact realization of the conditional law and the arithmetic control of the resulting changes, not a new general Palm–Stein principle or a new reciprocal asymptotic.

1.3 Relation to earlier work

The independent-sign theory supplies the run-start normalization, the geometric excess law and the lattice extreme-value comparison [19,22,3]. Dependency-graph Poisson approximation is classical [4,5,13]. Poisson approximation for occurrences of prescribed words and the role of self-overlap are also classical [20,21]. For several words, Reinert–Schbath [33] give Poisson-process and compound-Poisson approximations for clumps in Markov chains; Roquain–Schbath [35] treat overlapping rare word families. Thus a joint approximation with word types is not itself a new probabilistic principle. The contribution here is to verify joint, site-labelled comparisons in a nonstationary arithmetic sequence with exact multiplicative identities, uniformly for logarithmic word length and a growing dictionary.

Najnudel proves eventual independence of fixed-length blocks and almost-sure equidistribution of fixed words in roots-of-unity models [30]. For Steinhaus values, the stronger threshold of Turk reported in [30, arXiv version 2, Remark 2.3] already permits blocks of length at most any fixed multiple of the logarithm of their starting point. Logarithmic length alone is therefore not the distinction here. In the binary model, values at squares are deterministic, so the Steinhaus criterion cannot be read as independence of fair signs in every block. Our question instead concerns simultaneous, site-labelled fields of many rare occurrences and the weighted interactions between their moving windows. The uniform theorem makes explicit which part of that control is insensitive to the affine right-hand side; the typical theorem uses the additional averaging supplied by a random choice of those sides. Fixed-pattern and sign-complexity questions also arise for deterministic multiplicative functions [10,28,27].

An arithmetic predecessor for consecutive rare patterns is the work of Basak–Nath–Zaharescu [8] on chains of small quadratic non-residues. Their published abstract states Poisson behaviour in initial ranges of size $(\log p)^A$, $A > 2$, for almost all prime moduli in short intervals. This varies a quadratic character’s prime modulus, rather than drawing independent prime signs as here. This comparison is limited to the stated model and observable; no inclusion or exhaustive separation of the chain-length regimes is asserted.

The random-origin perspective itself has precedents. Pandey–Wang–Xu [32, arXiv version 4, Theorem 1.2] obtain complex Gaussian limits for normalized short-interval sums of typical Steinhaus functions. Their good sets have measure $1 - o(1)$ at each size; the common full-probability set along the prescribed dyadic scales in Corollary 7.8a uses summability in addition to a field comparison. Its observable is a rare binary run-start count and its mode of convergence is total variation to a finite-mean Poisson law. Poisson-generic sequences [1] instead concern the distribution over long words of their occurrence counts in a fixed sequence. The extensions to invariant exponentially mixing measures [2] do not directly apply to the nonstationary multiplicative sign sequence. These distinctions concern the observable and the quantifiers, not the invention of empirical observation.

Square-product paucity and Diophantine rigidity are important tools in random multiplicative functions. Related applications include short-interval central limit theorems and random Chowla problems [11,37,24,14]. For fixed polynomial values, de la Bretèche, Wang and Xu use square-product counts and Evertse–Silverman or Pell estimates [9]. Munsch–Shparlinski–

Sun–Xiao [29, Theorem 1.3] prove an unconditional square-product bound, uniform in the translation of one interval, for each fixed number of factors. The constants depend on that fixed number; the result does not directly control all relation weights in two moving windows of logarithmic length. Tao’s study of products of consecutive integers with unusual anatomy provides a particularly close comparison for the smooth-number and Diophantine mechanisms [38]. These works do not provide the growing-window profile (1.10) as an immediate specialization. The properly normalized short-interval sums of Harper, Soundararajan and Xu [23] concern a different observable and, in their stated model, Steinhaus prime values rather than binary run indicators.

We use Runge’s method in an explicitly proved split-product form [36,40,41]; for fixed-degree component counts, bounded-height square-class localization and a Pell equation suffice (Lemma 3.12). The uniform polynomial-height Pell count is already in [15, arXiv text, Lemma 4] and [38, Lemma 2.10 and Remark 2.12]; its proof is included to make the required uniformity explicit. Uniform integral-point theorems such as Evertse–Silverman [18] provide stronger, height-free information, but are not inputs to our proof. The microscopic analysis uses the uniform prime-divisor lower bound of Laishram–Shorey [26] and the pointwise result of Balasubramanian–Shorey [6]. The geometric marked target, its thinning chain and its Gaussian limits are probabilistic consequences once the arithmetic comparison is established; they are not presented as new properties of Poisson processes.

1.4 Reading guide and scope

Section 2 develops the affine encoding and the one-window bounds. Section 3 proves the final uniform two-window estimate and its extension to prescribed values. Sections 4–6 extract scalar, pattern, marked and conditional consequences. Section 7 handles the finite prefix, the selected-pivot signed comparison, the boundary and the relative crossover. A reader primarily interested in probability may take the uniform, prescribed-value and capped two-window profiles as arithmetic inputs (Theorem 3.1 and propositions 3.24 and 3.25), then continue with Sections 4–7. The two-window proof includes its Diophantine counting arguments, using standard ideal and unit theory for the Pell step. Section 7 adds the explicitly stated prime-divisor and smooth-product inputs needed near the boundary. The companion expands the technical steps called from the article, rather than giving a second proof of the relation estimate. The exact prime-cutoff definitions in Section 4 are not needed to read the three numbered introductory statements; they are needed for the quantitative growing-intensity comparison.

We do not assert optimality of the exponent $5/3$ or of the prime-cutoff ranges. The one-factor hard-cutoff comparison is proved for the base-contained signed run field when $\Lambda \rightarrow \infty$, not for an arbitrary word dictionary, a larger conditioning field or every intensity regime. The size exponent in the uniform dictionary theorem is likewise a sufficient bound, not an optimality statement; the typical theorem has different quantifiers.

2 Affine events and private prime coordinates

A run event is a linear system over $\mathbb{F}_2$. Its probability is determined by rank and compatibility; its dependence on a second event is determined by relations between the two systems. This elementary observation lets us keep arithmetic and probability separate. For $n \geq 1$, let

$$v(n) = (v_p(n) \bmod 2)_p \in \mathbb{F}_2^{(\mathcal{P})}.$$

Write the prime signs additively: $X = (X_p)_{p \in \mathcal{P}}$ denotes the sequence of independent Bernoulli variables, uniform on $\mathbb{F}_2$, defined by $f(p) = (-1)^{X_p}$. Then

$$f(n) = (-1)^{\langle X, v(n) \rangle}.$$

For a start x , introduce the L rows

$$\begin{aligned} a_{x,0} &= v(x-1) + v(x), & b_{x,0} &= 1, \\ a_{x,j} &= v(x) + v(x+j), & b_{x,j} &= 0, \quad 1 \leq j < L. \end{aligned}$$

Let A_x be the matrix with rows $a_{x,j}$ and $b_x = (1, 0, \dots, 0)$. Then

$$J_{x,L} = \mathbf{1}_{\{A_x X = b_x\}}.$$

For a tuple $\mathbf{x} = (x_1, \dots, x_k)$, stack the matrices and the right-hand sides. Set

$$\mathcal{R}(\mathbf{x}) = \ker(A_{\mathbf{x}}^\top), \quad \rho(\mathbf{x}) = \dim \mathcal{R}(\mathbf{x}),$$

and

$$q_{\mathbf{x}}(c) = \langle c, b_{\mathbf{x}} \rangle.$$

Lemma 2.1 (Affine Fourier identity). *For every tuple $\mathbf{x} = (x_1, \dots, x_k)$,*

$$2^{kL} \mathbb{P}(J_{x_1,L} = \dots = J_{x_k,L} = 1) = \sum_{c \in \mathcal{R}(\mathbf{x})} (-1)^{q_{\mathbf{x}}(c)}.$$

Consequently,

$$\mathbb{P}(J_{x_1,L} = \dots = J_{x_k,L} = 1) = 2^{-kL} \eta(\mathbf{x}) 2^{\rho(\mathbf{x})},$$

where $\eta(\mathbf{x}) = 1$ if $q_{\mathbf{x}}$ vanishes on $\mathcal{R}(\mathbf{x})$, and 0 otherwise. Finally,

$$\left| \eta(\mathbf{x}) 2^{\rho(\mathbf{x})} - 1 \right| \leq 2^{\rho(\mathbf{x})} - 1. \quad (2.1)$$

Proof. For a matrix with $m = kL$ rows,

$$\mathbf{1}_{\{AX=b\}} = 2^{-m} \sum_{c \in \mathbb{F}_2^m} (-1)^{\langle c, AX-b \rangle}.$$

Taking expectations, the terms with $A^\top c \neq 0$ vanish. A character sum over a vector space equals the cardinality of that space if the character is trivial, and 0 otherwise. If $\eta = 0$, then q is nontrivial, hence $\rho \geq 1$, and $1 \leq 2^\rho - 1$. $\square$

2.1 The boundary map of a tree

Put $B = L + 1$, $I_L = \{-1, 0, \dots, L-1\}$ and $V_x = \{x + i : i \in I_L\}$. The rows of A_x are the edge vectors of the star with edges $\{x-1, x\}$ and $\{x, x+j\}$, $1 \leq j < L$. For a finite tree, the map taking an edge set to the vertices of odd degree is a bijection onto the even vertex subsets. Thus a relation among the rows of A_x selects an even subset of V_x whose product is a square.

Lemma 2.2 (Square-product interpretation). *For disjoint windows $V_{x_1}, \dots, V_{x_k}$, the relation space is in bijection with tuples of even subsets $S_i \subset V_{x_i}$ such that $\prod_i \prod_{n \in S_i} n$ is a square. The affine character is $\sum_i \mathbf{1}_{\{x_i-1 \in S_i\}}$ modulo two. More generally, the edge relations of any tree are identified with its square-product even vertex subsets.*

Proof. The sum of the selected edge vectors is the sum of the valuation vectors at the odd-degree vertices. It vanishes exactly when their product is a square. The boundary map is injective because a tree has no nonempty even degree edge subgraph, and its image has dimension $|V| - 1$, the dimension of the even-subset space. The affine character records the root leaf edges. $\square$

2.2 One-window arithmetic

For $H \geq 2$ define the odd large-prime kernel

$$K_H(n) = \prod_{\substack{p > H \\ v_p(n) \text{ odd}}} p.$$

Call n H -defective when $K_H(n) = 1$. If an interval has diameter less than H , every nondefective vertex has a prime coordinate appearing there and nowhere else in the interval. Such a coordinate is a private pivot. Notice that the definition uses odd valuations, not merely the presence of a large prime divisor.

Lemma 2.3 (Split-product Runge bound). *There is an absolute C_0 such that the following holds. Let $d = 2k \geq 2$, let $\gamma_1, \dots, \gamma_d$ be distinct integers with $|\gamma_i| \leq R$, where $R \geq 1$, and set $G(T) = \prod_i (T + \gamma_i)$. If $U \in \mathbb{Z}$, $U \geq 2R$, and $G(U)$ is an integer square, then*

$$U \leq (C_0 d R)^{2d}. \quad (2.2)$$

Proof. Expand $\sqrt{G(T)}$ at infinity and let $P(T)$ be its polynomial part. Its coefficients lie in $2^{-d}\mathbb{Z}$ and, for $U > 4R$, $|\sqrt{G(U)} - P(U)| \leq (CdR)^d/U$. If the difference is nonzero, integrality of U and $\sqrt{G(U)}$ makes its modulus at least 2^{-d} . Otherwise the nonzero integer polynomial $2^{2d}(G - P^2)$ vanishes at U ; its coefficient height is at most $(CdR)^{2d}$, and the integral-root bound applies. The range $2R \leq U \leq 4R$ is immediate. All constants and the growing-degree coefficient estimates are given in Lemma A.1. This is a direct split-product implementation of Runge's method, not an unqualified invocation of a fixed-degree theorem [36,40,41]. $\square$

Proposition 2.4 (Few defects, with their exponential weight). *Fix $0 < c_1 < c_2$. Uniformly for integers $c_1 \log X \leq H \leq c_2 \log X$, an interval of $H + 1$ integers in $[X/2, 3X]$ contains $O_{c_1, c_2}(\log X / \log \log X)$ H -defective integers. If m_u counts these defects in $[u, u + H]$, then*

$$\sum_{X \leq u < 2X} (2^{m_u} - 1) \leq X^{1/2 + o_{c_1, c_2}(1)}. \quad (2.3)$$

Proof. Write each defective integer as sa^2 , where s is squarefree and H -smooth. Associate to the m defects in a window the columns $((v_p(s))_{p \leq H}, 1) \in \mathbb{F}_2^{\pi(H)+1}$. A nonzero kernel word selects an even number d of distinct integers with square product. The Runge bound gives $X/2 \leq (CdH)^{2d}$; consequently its Hamming weight is at least $c \log X / \log H$ for a fixed $c > 0$. Let t be an integer of order $\log X / \log H$ with $2t + 1 \leq c \log X / \log H$. If the kernel is zero, $m \leq \pi(H) + 1$. If $m < 2t$, the desired bound is again immediate. Otherwise disjoint Hamming balls about its words give

$$\sum_{j \leq t} \binom{m}{j} \leq 2^{\pi(H)+1}, \quad (m/t)^t \leq 2^{\pi(H)+1}.$$

Since $(\pi(H) + 1)/t = O_{c_1, c_2}(1)$, it follows that $m = O_{c_1, c_2}(\log X / \log \log X)$.

On the other hand the number of defective integers up to $3X$ is at most

$$\sqrt{3X} \sum_{\substack{s \text{ squarefree} \\ P^+(s) \leq H}} s^{-1/2} = \sqrt{3X} \prod_{p \leq H} (1 + p^{-1/2}) = X^{1/2+o(1)}.$$

Each belongs to at most $H + 1 = X^{o(1)}$ windows. Only $X^{1/2+o(1)}$ windows therefore have a defect, and the pointwise bound makes each weight 2^{m_u} equal to $X^{o(1)}$. This proves the sum. $\square$

For $0 < \delta < 1$, write $U_{M,\delta} = [\lceil M^\delta \rceil, M) \cap \mathbb{Z}$. If $B \asymp \log M$, a dyadic slice $X \leq x < 2X$ in this interval has $B \asymp_\delta \log X$. Applying proposition 2.4 on these slices gives

$$\sum_{x \in U_{M,\delta}} (2^{m_B(x)} - 1) \leq M^{1/2+o(1)}, \quad m_B(x) = \#\{n \in V_x : K_B(n) = 1\}. \quad (2.4)$$

The same assertion holds for windows whose width and defect threshold are any two comparable quantities in a fixed logarithmic band: use an enclosing window of width equal to a fixed multiple of the threshold and subdivide it into a bounded number of windows before the coding argument. In the applications below the threshold can always be chosen at least the width.

Corollary 2.5 (Uniform first moment and an affine bound). *Put $p = 2^{-L}$ and $m_B^\circ(x) = \#\{n \in V_x \setminus \{x-1\} : K_B(n) = 1\}$. For any right-hand side $b \in \mathbb{F}_2^L$,*

$$|\mathbb{P}(A_x X = b) - p| \leq p(2^{m_B^\circ(x)} - 1), \quad \mathbb{P}(A_x X = b) \leq 2^{-L + \max(m_B(x) - 1, 0)}. \quad (2.5)$$

For every deterministic $A \subset I_N$ and B in a fixed logarithmic band,

$$\mathbb{E} \sum_{x \in A} J_{x,L} = |A|p + O(pN^{1/2+o(1)}).$$

The analogous error over any subset of $U_{M,\delta}$ is $O(pM^{1/2+o(1)})$.

Proof. A private pivot excludes its vertex from any relation. An even subset of m defective vertices has at most $\max(m-1, 0)$ degrees of freedom. Thus $\rho(x) \leq \max(m_B(x)-1, 0) \leq m_B^\circ(x)$. Apply Lemma 2.1, which is unchanged by the affine right-hand side, and sum proposition 2.4 or (2.4). Restricting a nonnegative error sum to a mask can only decrease it. $\square$

For $Y \geq 2$ write $\mathcal{F}_Y = \sigma(f(p) : p \text{ prime}, p \leq Y)$ for the information carried by the small prime signs.

Corollary 2.6 (Prescribed values in one window). *Let M_x have rows $v(n)$, $n \in V_x$, and let $w \in \{\pm 1\}^B$. Write $I_{x,w} = \mathbf{1}_{\{(f(x-1+j))_{0 \leq j < B} = w\}}$ and $p_B = 2^{-B}$. Then*

$$|\mathbb{P}(I_{x,w} = 1) - p_B| \leq p_B(2^{m_B(x)} - 1). \quad (2.6)$$

If, for every vertex $n \in V_x$, there is a prime $p > Y$, where $Y > B$, such that $v_p(n)$ is odd, then $\mathbb{P}(I_{x,w} = 1 \mid \mathcal{F}_Y) = p_B$ for every realization of $\mathcal{F}_Y$. Fix $0 < c_1 < c_2$. Uniformly for $c_1 \log N \leq B \leq c_2 \log N$, every deterministic start mask $A \subset I_N$ and dictionary $\mathcal{W} \subset \{\pm 1\}^B$ satisfy

$$\mathbb{E} \sum_{x \in A} \sum_{w \in \mathcal{W}} I_{x,w} = |A| |\mathcal{W}| 2^{-B} + O_{c_1, c_2}(|\mathcal{W}| 2^{-B} N^{1/2+o_{c_1, c_2}(1)}).$$

The expectation is unconditional. The threshold and error are uniform in $B, A, \mathcal{W}$, which may vary with N ; empty masks and dictionaries are allowed. The mask restricts starts, not vertices:

every occurrence uses all B values $x - 1, \dots, x + B - 2$, without truncation at the mask or at the boundary of I_N .

Proof. For a fixed window of positive integers, take a prime cutoff at least as large as its largest vertex, and at least Y when conditioning. The word event in the infinite product is the inverse image of the corresponding finite-cylinder event, so their probabilities agree exactly. Only the prime signs are independent; the integer values retain all their multiplicative relations. In a relation among the rows of M_x , every nondefective vertex is excluded by its private coordinate. Hence the nullity is at most $m_B(x)$, without the even-subset restriction used for relative signs. Finite Fourier inversion gives (2.6). On a good support, the projection of M_x onto primes above Y has an identity minor of order B , so every prescribed value vector has probability 2^{-B} on each positive-mass assignment atom of $\mathcal{F}_Y$. Sum the nonnegative bound using proposition 2.4, then use linearity of expectation. The same threshold works for every mask and every dictionary; no independence between occurrences is needed. $\square$

Lemma 2.7 (Private pivots on a graph). *Let G be a finite graph. In each connected component suppose all vertices except possibly one have distinct private prime coordinates in the whole graph. The kernel of its edge vectors, projected onto these coordinates, is contained in its cycle space. In particular its rank is at least $|E(G)| - \beta_1(G)$, where $\beta_1(G) = |E(G)| - |V(G)| + c(G)$. For a tree this rank is $|E(G)|$.*

Proof. A vanishing edge combination has even degree at every pivoted vertex. The sum of degrees in a connected component is even, so its possible unpivoted vertex has even degree as well. The selected edge set is therefore in the cycle space, of dimension $\beta_1(G)$. $\square$

Lemma 2.8 (Local run pairs). *If $0 < |x - y| < L$, then $J_{x,L}J_{y,L} = 0$. Over ordered pairs with $|x - y| = L$ in a dyadic block or in $U_{M,\delta}$, the homogeneous weights $2^{\rho(x,y)}$ have sum at most $N^{1+o(1)}$ or $M^{1+o(1)}$, respectively.*

Proof. An overlapping start requires two consecutive signs within the first constant run to be opposite. At distance L , the two start trees form a tree on an interval of diameter $2L$. With threshold $2L + 1$, private pivots and proposition 2.4 bound its relation dimension by $O(\log N / \log \log N)$ on a slice. There are $O(N)$ such pairs; summing slices gives the macroscopic assertion. $\square$

3 A uniform two-window relation estimate

The transfer in Theorem 4.1 separates four costs: weighted one-window defects, supports removed at the prime cutoff, shared-prime graph edges, and the weighted relations estimated here. Only the last requires the two-window arithmetic. We sum $2^{\rho(x,y)} - 1$, not merely count pairs with a relation, by separating exact rational identities from residual prime relations in one macroscopic geometry. Word overlaps are treated separately in Theorem 5.1.

Fix $0 < \delta < 1$ and $0 < \beta_- < \beta_+ < \infty$. Throughout this section,

$$U = U_{M,\delta}, \quad \beta_- \log M \leq B = L + 1 \leq \beta_+ \log M, \quad Q_B = 2^B. \quad (3.1)$$

A factor $M^{o(1)}$ is uniform in this band, with dependence only on the fixed parameters (and on an explicitly fixed component size when one is used). All these parameters are fixed before M . Every occurrence lies in

$$M^\delta - 1 \leq n \leq M + B \leq 2M. \quad (3.2)$$

Theorem 3.1 (Uniform weighted relation profile). *For every $\varepsilon > 0$,*

$$R_{2,\delta}(M, L) := \sum_{\substack{x, y \in U \\ |x-y| > L}} (2^{\rho_L(x, y)} - 1) \\ \ll_{\delta, \beta_-, \beta_+, \varepsilon} M^\varepsilon \left(M^{3/2} Q_B^{1/6} + M Q_B^{2/3} + M^{2/3} Q_B \right) \quad (3.3)$$

$$\ll_{\delta, \beta_-, \beta_+, \varepsilon} M^\varepsilon \left(M^{5/3} + M^{2/3} Q_B \right). \quad (3.4)$$

The defining sum is over ordered pairs. In particular $R_2(N, L) \ll_{C, \varepsilon} N^{5/3+\varepsilon}$ when $|L - \log_2 N| \leq C$ in a dyadic block. The same conclusion, with a constant depending on the fixed upper ratio, holds on $[N, \kappa N)$.

The last assertions follow by restriction from (3.4) at $M = 2N$ or $M = \kappa N$ and, for example, $\delta = 1/2$; the lower endpoint is then at least M^δ for large N . No sum of independent block laws is being taken. The proof of the estimates occupies the rest of this section.

There are three obstacles to counting the relations. Exact rational identities may create many of them for one identifiable reason; we remove that subspace first. Residual relations are then controlled by combining rarity of their hosts with a bound on their dimension. In the terminal case neither resource is sufficient separately: determinants force several small kernels into one window, and it is this simultaneous concentration, rather than the size of the candidate set, that must be counted. The partition below distinguishes which of these resources is available.

3.1 Rational identities and the residual graph

For a separated pair, use the even-subset description in Lemma 2.2. An exact rational relation can be present before any primes are exposed. For example, with $x = 100$, $y = 67$ and $L = 3$, the windows contain the two units $(99, 66)$ and $(102, 68)$ of ratio $3 : 2$. Their product is 6732^2 . Each selected subset in a window has two vertices, so this gives a nonzero homogeneous relation; its two root occurrences also make the affine character even. In fact the individual ranks are three and the stacked rank is five, giving joint probability $1/32$, rather than $1/64$ for independent starts.

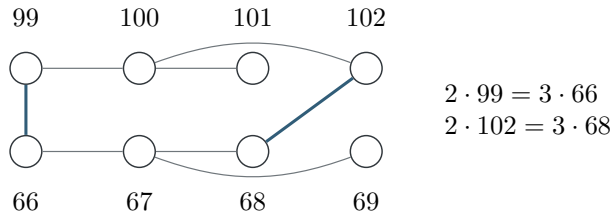


Figure 2: An exact two-unit relation. The thick edges are rational units; the thin edges encode the affine start trees. Even subsets of the units produce relations. This macroscopic example also illustrates why counting all prime coincidences without quotienting the rational subspace loses information.

For coprime $a, b \geq 1$, set $h = by - ax$ and

$$\mathcal{M}_{a,b}(h) = \{(i, j) \in I_L^2 : ai - bj = h\}, \quad m(a, b, h) = |\mathcal{M}_{a,b}(h)|.$$

A member of $\mathcal{M}_{a,b}(h)$ is an exact unit, because $a(x + i) = b(y + j)$. Even subsets of the exact units form a subspace $\mathcal{S}_{a,b,h} \subset \mathcal{R}(x, y)$.

Lemma 3.2 (Rational code and height). *If $m = m(a, b, h) \geq 1$, then $\dim \mathcal{S}_{a,b,h} = m - 1$. If $m \geq 2$, then*

$$\max(a, b) \leq L < B.$$

This conclusion is independent of the ratio between the endpoints x and y .

Proof. For an exact unit, coprimality gives integers t with $x + i = bt$ and $y + j = at$. An even collection E of units satisfies

$$\prod_{(i,j) \in E} (x + i)(y + j) = (ab)^{|E|} \left(\prod_{(i,j) \in E} t \right)^2,$$

which is a square. The even subsets of an m -element set have dimension $m - 1$. Distinct solutions of $ai - bj = h$ differ by (b, a) ; since both offsets remain in an interval of diameter L , one has $a, b \leq L$. $\square$

Fix once and for all $A = 3$. Call (x, y) aligned if there are coprime $a, b \leq B^A$ such that

$$|by - ax| < (a + b)B. \quad (3.5)$$

Lemma 3.3 (Uniqueness of the macroscopic channel). *For M sufficiently large in terms of $(\delta, \beta_-, \beta_+)$, an aligned pair has at most one reduced fraction (a, b) satisfying (3.5). Every channel with at least two exact units is aligned and therefore has this unique fraction.*

Proof. If (a, b) and (a', b') both satisfy (3.5), then

$$|(a'b - ab')x| < 4B^{2A+1}.$$

But $x \geq M^\delta$ and $B = O(\log M)$, so the right-hand side is $o(M^\delta)$. The integer $a'b - ab'$ is therefore zero for large M , and reducedness gives $(a, b) = (a', b')$. If a channel has two units, Lemma 3.2 gives $a, b \leq L$, and for any unit $h = ai - bj$ satisfies $|h| < (a + b)B$. $\square$

When the canonical aligned channel has $m \geq 2$, define

$$\mathcal{S}_{\text{rat}} = \mathcal{S}_{a,b,h}, \quad \sigma = m - 1.$$

Otherwise define $\mathcal{S}_{\text{rat}} = 0$ and $\sigma = 0$. In all cases put

$$\tau = \rho - \sigma.$$

Thus no contraction is made when $m_{\text{ex}} \leq 1$.

Remark 3.4 (Disparate scales). Suppose, for example, that $x \asymp M^\delta$ and $y \asymp M$. If there were two exact units, then $a, b \leq L$ and $x + i = bt$, $y + j = at$, so

$$y \leq L(x + B) + B = o(M),$$

a contradiction. Hence every genuinely disparate-scale pair has $m_{\text{ex}} \leq 1$. It remains in the ordinary large-prime graph and is treated by the small-product, shallow-core or deep-core bounds below; no illegitimate valuation identity or code contraction is used.

Now form the graph $G_{>B}(x, y)$ on the occurrence set $\mathcal{O} = V_x \sqcup V_y$. For a prime $p > B$, let $\mathcal{O}_p$ be the occurrences at which v_p is odd. Each block has diameter $B - 1 < p$, so $|\mathcal{O}_p| \leq 2$, with at most one occurrence in each block. If $|\mathcal{O}_p| = 1$, pin that vertex to zero; if $|\mathcal{O}_p| = 2$,

join the two vertices; parallel equations are deduplicated. A vertex with no large-prime label is defective. Let D be the number of defective vertices and c the number of nontrivial unpinned components.

Lemma 3.5 (Resolution and square class of components). *The solution space of the $p > B$ equations is canonically isomorphic, after one representative is chosen in each component, to*

$$\mathbb{F}_2^D \oplus \mathbb{F}_2^c.$$

In particular, $D + 2c \leq 2B$. Every nontrivial unpinned component C meets both blocks and has a unique representation

$$\prod_{n \in C} n = d_C z_C^2,$$

where d_C is positive, squarefree, and B -smooth.

Proof. A pinned component is zero. All variables in an unpinned component containing an edge are equal and contribute one free coordinate. An isolated defective vertex contributes one free coordinate. Every nontrivial component uses at least two occurrences, giving $D + 2c \leq 2B$.

Every edge is bipartite. For each $p > B$, an unpinned component contains either both odd occurrences of p or none; a single odd occurrence would pin it. Hence the total p -valuation of the product of the component is even for every $p > B$. Its squarefree part is therefore supported on $p \leq B$, and uniqueness is immediate. $\square$

Assume $m \geq 2$. An exact unit has the form $x + i = bt$, $y + j = at$ with $a, b < B$. For every $p > B$, $p \nmid ab$ and therefore

$$v_p(x + i) \equiv v_p(t) \equiv v_p(y + j) \pmod{2}.$$

If $K_B(t) = 1$, the two occurrences are defective; otherwise they form an isolated unpinned component of size two. Distinct exact units use disjoint occurrences. Let d be the number of exact units of the first type and $n = m - d$ the number of the second type, and define

$$D^\# = D - d, \quad c^\# = c - n,$$

removing the n exact components from the residual component set $\mathcal{C}_{\text{res}}$. If $m < 2$, set $D^\# = D$, $c^\# = c$, and let $\mathcal{C}_{\text{res}}$ be all nontrivial unpinned components. Notice that $D^\#$ is a dimensional correction, not literally a number of remaining defective vertices.

Lemma 3.6 (Quotient core). *Uniformly for all separated pairs,*

$$\tau \leq D^\# + c^\#, \quad D^\# + 2c^\# \leq 2B, \quad D^\# = O(B/\log B).$$

Proof. For $m \geq 2$, the m exact-unit vectors are independent in the large-prime solution space and the even code has dimension $m - 1$. The parity functional of the first block descends to the quotient and is nonzero there; hence the even-first-block subspace has dimension $D + c - m$. The full relation quotient embeds in that subspace, and the small-prime equations can only decrease dimension. Therefore

$$\tau \leq D + c - m = (D - d) + (c - n) = D^\# + c^\#.$$

For $m < 2$, the estimate is the direct inclusion $\mathcal{R} \subset W_{>B}$. The inequality $D^\# + 2c^\# \leq 2B$ follows from $D + 2c \leq 2B$. The defect estimate follows from proposition 2.4, applied to each window on its own dyadic slice; (3.1) makes the constants uniform. $\square$

From each component of $\mathcal{C}_{\text{res}}$, choose canonically its smallest large-prime label and then the lexicographically first edge carrying that label. This gives

$$(i_s, j_s, p_s), \quad 1 \leq s \leq c^\#,$$

with distinct offsets in each block, distinct primes $p_s > B$, and $p_s \mid x + i_s$, $p_s \mid y + j_s$. When $m \geq 2$, these cells are not exact; when $m = 1$, only the at most two components meeting the two occurrences of the unique exact unit can be exceptions. Put

$$P^\# = \prod_{s=1}^{c^\#} p_s,$$

with the empty product equal to one.

The dimension decomposition gives the exact identity

$$2^\rho - 1 = (2^\sigma - 1) + 2^\sigma(2^\tau - 1). \quad (3.6)$$

We first sum the rational term. The residual term will be divided into exhaustive geometric sectors.

3.2 Hosts and rational mass

The elementary prime and divisor estimates used below are

$$\begin{aligned} \pi(B) &\ll B/\log B, & \sum_{p \leq B} p^{-1/2} + B \sum_{p > B} p^{-3/2} &\ll \sqrt{B}/\log B, \\ \sum_{p > B} p^{-2} &\ll (B \log B)^{-1}, & \omega(n) + \log \tau(n) &\ll_K \log M / \log \log M \quad (1 \leq n \leq M^K). \end{aligned} \quad (3.7)$$

$$(3.8)$$

Here $\tau(n)$ in (3.8) is the divisor function, not the residual dimension. The last bound is used only for fixed K . Partial summation of the prime number theorem proves the prime estimates; the divisor estimate is standard [39,31]. In particular,

$$\prod_{p \leq B} (1 + cp^{-1/2}) = M^{o(1)}, \quad 2^{\pi(B)} = M^{o(1)} \quad (c \text{ fixed}), \quad (3.9)$$

while a squarefree B -smooth integer has size at most

$$\prod_{p \leq B} p = \exp(O(B)) = M^{O_{\beta+}(1)}. \quad (3.10)$$

Let $H_{2,\delta}(M, L)$ count separated ordered pairs in U^2 with $\rho_L(x, y) > 0$. We also need $H_{2,\delta}^{\text{val}}(M, B)$, the number of such pairs for which some nonempty subset of $V_x \sqcup V_y$ has square product, with no parity requirement in either window.

Proposition 3.7 (Host bound). *Uniformly in the fixed band,*

$$H_{2,\delta}(M, L) \leq H_{2,\delta}^{\text{val}}(M, B) \leq M^{3/2+o(1)}.$$

Proof. For each pair counted by $H_{2,\delta}^{\text{val}}$, choose a nonempty square-product subset and its first selected occurrence, ordering first the x -block, then the y -block, and then the offsets increasingly. Suppose first that this occurrence is $n = x + i$; the other orientation costs a factor two. Put

$K = K_B(n)$. For every $p \mid K$, the parity of the selected square product forces an occurrence of odd p -valuation in the other block. It is unique there because $p > B$ and the block diameter is $B - 1$. Forgetting compatibility between primes, there are at most $B^{\omega(K)}$ assignments of the primes of K to offsets of the other block.

A fixed assignment determines one residue class for y modulo every $p \mid K$, hence one class modulo K . The interval $U_{M,\delta}$ has length less than M , so it contains at most $M/K + 1$ representatives. By (3.2), $K \leq n \leq 2M$, and therefore

$$M/K + 1 \leq 3M/K,$$

including $K = 1$. The orientation and selected offset cost at most $2B$, and hence

$$H_{2,\delta}^{\text{val}}(M, B) \ll BM \sum_{n \leq 2M} \frac{B^{\omega(K_B(n))}}{K_B(n)}. \quad (3.11)$$

Write uniquely $n = a^2 ur$, where u is squarefree and B -smooth and $r = K_B(n)$ is squarefree with all prime factors $> B$. Summing over a first gives

$$\begin{aligned} \sum_{n \leq 2M} \frac{B^{\omega(K_B(n))}}{K_B(n)} &\leq \sqrt{2M} \sum_u u^{-1/2} \sum_r B^{\omega(r)} r^{-3/2} \\ &\leq \sqrt{2M} \prod_{p \leq B} (1 + p^{-1/2}) \prod_{p > B} (1 + Bp^{-3/2}). \end{aligned}$$

For this host estimate, elementary integer sums suffice:

$$\sum_{p \leq B} p^{-1/2} \leq \sum_{n=1}^B n^{-1/2} \leq 2\sqrt{B}, \quad B \sum_{p > B} p^{-3/2} \leq B \int_B^\infty t^{-3/2} dt = 2\sqrt{B}.$$

Since $\log(1+t) \leq t$, the weighted kernel sum is at most $\sqrt{2M} \exp(4\sqrt{B})$. Under $B \leq \beta_+ \log M$, $B \exp(4\sqrt{B}) = M^{o(1)}$, uniformly in the length. Substitution in (3.11) proves the claim. The sharper prime-sum bounds are not needed at this step. $\square$

Proposition 3.8 (Rational mass). *The contributions with canonical height $q = \max(a, b)$ satisfy*

$$\sum_{q \geq 3} (2^\sigma - 1) \leq MQ_B^{1/3} M^{o(1)}, \quad \sum_{q=2} (2^\sigma - 1) \leq MQ_B^{1/2} M^{o(1)}. \quad (3.12)$$

The geometry-only sum, without translation parameters, satisfies

$$\sum_{\substack{(a,b)=1, \max(a,b) \geq 2, h \in \mathbb{Z} \\ m(a,b,h) \geq 2}} 2^{m(a,b,h)-1} \leq Q_B^{1/2} M^{o(1)}. \quad (3.13)$$

Proof. At height q there are $O(q)$ reduced pairs (a, b) and $O(qB)$ possible nonempty values of h . A unit (i_0, j_0) gives $x = bt - i_0$, $y = at - j_0$, with $O(M/q)$ choices of t . Units are spaced by (b, a) , so $m - 1 \leq B/q$, and $m \geq 2$ implies $q \leq L$. Summing $O(MqB2^{B/q})$ for $q \geq 3$ gives the first bound; $q = 2$ gives the second. At $q = 1$, separation excludes a unit. Omitting the $O(M/q)$ translation factor and summing the same geometries proves (3.13). $\square$

Remark 3.9 (Why the dyadic interval helps, and what it does not remove). The starts $(x, 2x)$ cannot both belong to $[N, 2N)$. This removes the volumetric $2 : 1$ family, not all ratio-two occurrences in two windows: $(N - 1, 2N - 2)$ can still occur when the starts are N and $2N - 1$.

Our master estimate keeps the volumetric branch and therefore also applies to the finite prefix. In a dyadic restriction, a more precise channel count improves this particular term, but that improvement is not needed for (3.4).

3.3 The partition and the shallow sectors

The partition balances two resources: the number of host pairs and the weight carried by the relation space at each host. The rational mass has already been separated. For the residual mass, a small dimension makes many hosts affordable; when the dimension is large, a bounded component or a defect must instead make the hosts rare. The first tests exploit these alternatives, with Runge excluding the aligned deep case. In the terminal sector neither a one-kernel host count nor the full relation weight suffices separately: several small kernels in one short window must be counted together. The assembly in §3.6 records the cost obtained from each branch.

For a nonaligned pair let $\tilde{k}$ be the rank of the small-prime and two block-parity equations after solving all large-prime equations. Then

$$\tau = D^\# + c^\# - \tilde{k}. \quad (3.14)$$

Set $T_B = \lfloor (B-2)/3 \rfloor$. After removing $2^\sigma - 1$ in (3.6), apply the following tests successively:

Sector	Defining test, after the preceding tests have failed
1	$P^\# \leq M$.
2	$\sigma > 0$ and the canonical channel has $q \leq \sqrt{\log B}$.
3	$c^\# \leq B/6$.
4	The pair is aligned.
5	$c^\# \leq 2B/3$.
6	$D^\# \geq 3$.
7	Write $c^\# = B - s$; then $s + \tilde{k} > T_B$.
8	$s + \tilde{k} \leq T_B$.

The tests are complementary at each step. In sectors 5–8 the pair is nonaligned, $\sigma = 0$, and $D^\# = D$, $c^\# = c$. In sectors 7–8 one also has $c^\# > 2B/3$ and $D^\# \leq 2$. Thus this is a disjoint partition of all residual mass, including pairs with very different heights.

Proposition 3.10 (Direct bounds in sectors 1–3). *Up to a common factor $M^{o(1)}$, the residual masses of sectors 1, 2 and 3 are bounded, respectively, by*

$$M^{3/2} + MQ_B^{1/2}, \quad M^{3/2} + MQ_B^{1/2} + MQ_B^{1/3}, \quad M^{3/2}Q_B^{1/6}.$$

Proof. In sector 1 the canonical certificate contains exactly $c^\#$ distinct primes, all greater than B . Hence

$$(B+1)^{c^\#} \leq P^\# \leq M, \quad c^\# \leq \frac{\log M}{\log(B+1)}, \quad 2^{c^\#} = M^{o(1)}. \quad (3.15)$$

The last estimate is uniform in the fixed positive logarithmic band. Together with $D^\# = O(B/\log B)$ and $\tau \leq D^\# + c^\#$, it gives $2^\tau = M^{o(1)}$ pointwise. When $\sigma = 0$, only relational hosts contribute, and their number is $M^{3/2+o(1)}$ by proposition 3.7. When $\sigma > 0$, use

$$2^\sigma(2^\tau - 1) \leq 2(2^\sigma - 1)M^{o(1)}$$

and sum the rational mass in proposition 3.8. The two branches give $M^{o(1)}(M^{3/2} + MQ_B^{1/2})$, without a critical-scale substitution or any restriction on the ratio between the starts.

In sector 2, write the canonical channel as $h = by - ax$. If a prime $p > B$ joins a nonexact cell (i, j) , it divides $a(x + i) - b(y + j) = ai - bj - h \neq 0$. The existence of an exact unit bounds this integer by $O(qB)$. Thus every residual certificate prime is at most a fixed multiple of qB . Consequently $c^\# \leq \pi(CqB) = o(B)$ uniformly for $q \leq \sqrt{\log B}$. Also $D^\# = o(B)$, so $\tau = o(B)$. Summing the systematic weights by (3.12) bounds this sector. The displayed $M^{3/2}$ term is harmless and also covers a partition convention that sends a zero-dimensional rational code here.

In sector 3, $\sigma = 0$ or $q > \sqrt{\log B}$, hence $\sigma = o(B)$. The quotient estimate gives $2^\sigma(2^\tau - 1) \leq Q_B^{1/6} M^{o(1)}$. Only relational hosts contribute. Multiply by their $M^{3/2+o(1)}$ count. $\square$

The untreated pairs now have a large certificate product and more than $B/6$ residual components. The next step either rules out their alignment or extracts a bounded component whose square class supplies a rigid Diophantine equation.

3.4 Diophantine counting and deep cores

A bounded component forces a product of a fixed number of shifted values to lie in one square class. We need to count solutions only in a polynomial height range. Two of the factors already suffice: after localizing their square classes, they satisfy a Pell equation. This bounded-height argument keeps the uniformity we need without a height-free integral-point theorem. The uniform polynomial-height Pell estimate is classical: see [15, arXiv text, Lemma 4] and [38, Lemma 2.10 and Remark 2.12]. We retain a self-contained proof with the explicit height dependence required here; the lemma is not a separate priority claim.

Lemma 3.11 (Uniform Pell count). *Fix $K_0 > 0$. Let A, C be positive squarefree integers with $A/C \notin \mathbb{Q}^{\times 2}$, and let $e \neq 0$. If $A, C, |e| \leq M^{K_0}$, then the number of integer solutions of*

$$Az^2 - Cw^2 = e, \quad |z|, |w| \leq M^{K_0},$$

is

$$\exp\left(O_{K_0}\left(\frac{\log M}{\log \log M}\right)\right) = M^{o(1)}.$$

Proof. The detailed reduction and uniformity checks are given in Lemma A.2. $\square$

Lemma 3.12 (Split products in a polynomial range). *Fix $d \geq 2$ and $K > 0$. Let $h_1, \dots, h_d$ be distinct integers and e a positive squarefree integer with $|h_i|, e \leq M^K$. The number of integer pairs (X, Y) satisfying*

$$|X| \leq M^K, \quad X + h_i > 0 \ (1 \leq i \leq d), \quad \prod_{i=1}^d (X + h_i) = eY^2$$

is at most $\exp(O_{d,K}(\log M / \log \log M))$.

Proof. Put $\Delta = \prod_{i < j} (h_i - h_j) \neq 0$ and $S = \{p : p \mid e\Delta\}$. If $p \notin S$, it divides at most one factor $X + h_i$ and has even valuation in the product. Hence the squarefree part of every factor is supported on S . In particular

$$X + h_1 = su^2, \quad X + h_2 = tv^2, \quad su^2 - tv^2 = h_1 - h_2 \neq 0,$$

where s, t are positive squarefree integers supported on S . There are at most $4^{|S|}$ choices of (s, t) . We need retain only $s, t \leq 2M^K$, and $u, v \leq (2M^K)^{1/2}$. For $s \neq t$, their quotient is not a rational square, so Lemma 3.11 bounds the solutions uniformly. For $s = t$, the equation

$s(u-v)(u+v) = h_1 - h_2$ is bounded by the divisor estimate. Each pair (u, v) determines X , and then at most two values of Y . Finally $|e\Delta| \leq M^{O_{d,K}(1)}$, so $|S| = O_{d,K}(\log M / \log \log M)$; the divisor and Pell bounds have the same exponential scale. Multiplying the bounds proves the claim. $\square$

This argument counts bounded-height solutions, not all integral points. The distinction is precisely what makes a single Pell equation sufficient. The square-class and height bookkeeping is recorded in section A.3.

Lemma 3.13 (Normalization of a component). *Let I, J be two nonempty sets of offsets, with $|I| + |J| \leq K_0$, and let d be a positive squarefree B -smooth integer. If*

$$P_I(x)Q_J(y) = dz^2, \quad P_I(X) = \prod_{i \in I} (X + i), \quad Q_J(Y) = \prod_{j \in J} (Y + j), \quad (3.16)$$

then, after fixing x, I, J, d , there is a positive squarefree integer $e \leq M^{O_{K_0}(1)}$ such that

$$Q_J(y) = ev^2. \quad (3.17)$$

The integer e is the squarefree part of $dP_I(x)$.

Proof. In the group $\mathbb{Q}^\times / \mathbb{Q}^{\times 2}$, the equality (3.16) gives

$$[Q_J(y)] = [dP_I(x)].$$

The positive squarefree representative of the right-hand class is e ; the squarefree part of $Q_J(y)$ is therefore e . Since $|I| \leq K_0$, we have $dP_I(x) \leq M^{O_{K_0}(1)}$. $\square$

Lemma 3.14 (One-sided count of a component). *Fix K_0 . For x, I, J, d fixed as in Lemma 3.13, the number of $y \in U_{M,\delta}$ satisfying (3.16) is*

$$\begin{cases} O(M^{1/2}), & |J| = 1, \\ M^{O_{K_0}(1)}, & |J| \geq 2. \end{cases}$$

The same assertion holds after exchanging the two blocks.

Proof. Lemma 3.13 gives $Q_J(y) = ev^2$. For $|J| = 1$, the equation $y + j = ev^2$ leaves at most $1 + \sqrt{3M/e} = O(M^{1/2})$ values. For $|J| \geq 2$, use Lemma 3.12. Its degree is at most K_0 , all shifted factors are positive, $y \leq M$, and $e \leq M^{O_{K_0}(1)}$. Thus a single fixed polynomial-height exponent covers every coefficient and variable. The symmetric statement follows by interchanging the blocks. $\square$

Lemma 3.15 (Two singleton components). *Fix two offsets i, j . The number of triples (x, y, d) , with $x, y \in U_{M,\delta}$, d positive, squarefree and B -smooth, satisfying*

$$(x + i)(y + j) = dz^2$$

is

$$M \exp\left(O\left(\frac{\sqrt{B}}{\log B}\right)\right). \quad (3.18)$$

Proof. Set $X = x + i$, $Y = y + j$. A prime-by-prime analysis of $XY = dz^2$ gives a unique parametrization

$$X = ecu^2, \quad Y = (d/e)cv^2,$$

where $e \mid d$, where c is squarefree and $(c, d) = 1$. For fixed e , the conditions $1 \leq X, Y \leq 3M$ force

$$c \leq C_e := \frac{3M}{\max(e, d/e)}.$$

Dropping the “squarefree” and “coprime to d ” conditions, which only increases the count, the number of triples (u, v, c) is at most

$$\begin{aligned} \sum_{c \leq C_e} \left(1 + \sqrt{\frac{3M}{ec}}\right) \left(1 + \sqrt{\frac{3Me}{dc}}\right) &\ll C_e + \sqrt{\frac{M}{e}} \sqrt{C_e} + \sqrt{\frac{Me}{d}} \sqrt{C_e} + \frac{M}{\sqrt{d}} \sum_{c \leq C_e} \frac{1}{c} \\ &\ll \frac{M \log M}{\sqrt{d}}. \end{aligned}$$

Here we used $\max(e, d/e) \geq \sqrt{d}$. After summing over $e \mid d$, the number of solutions for fixed d is

$$\ll \frac{M \tau(d) \log M}{\sqrt{d}}.$$

Finally

$$\sum_{\substack{d \text{ squarefree} \\ p \mid d \Rightarrow p \leq B}} \frac{\tau(d)}{\sqrt{d}} = \prod_{p \leq B} \left(1 + \frac{2}{\sqrt{p}}\right) = \exp\left(O\left(\frac{\sqrt{B}}{\log B}\right)\right).$$

The factor $\log M$ is absorbed into the exponential $M^{o(1)}$. □

Lemma 3.16 (Hosts of a bounded component). *Fix K_0 . The number of separated pairs (x, y) whose non-aligned residual graph contains a nontrivial unpinned component meeting both blocks and of size at most K_0 is*

$$M^{1+o_{K_0}(1)}. \tag{3.19}$$

If this component has exactly two vertices, one in each block, we have the sharper bound

$$M \exp\left(O\left(\frac{\sqrt{B}}{\log B}\right)\right).$$

Proof. Let I, J be the offset sets of the component. Lemma 3.5 provides, with a unique coefficient d , equation (3.16). There are only $B^{O_{K_0}(1)} = M^{o(1)}$ choices of the two offset sets. We separate three branches.

Two degrees at least two. If $|I|, |J| \geq 2$, fix x , the offsets and d . Lemma 3.14 leaves $M^{o_{K_0}(1)}$ values of y . Summing over the M values of x and the $2^{\pi(B)} = M^{o(1)}$ coefficients d gives $M^{1+o_{K_0}(1)}$ hosts.

A single degree equal to one. If $|I| = 1 < |J|$, the same argument fixes x and uses the mobile side of degree $|J| \geq 2$. If $|J| = 1 < |I|$, we exchange the two blocks. We again obtain $M^{1+o_{K_0}(1)}$ after summing over d .

Two singletons. If $|I| = |J| = 1$, Lemma 3.15 applies directly. It already counts the triples (x, y, d) and, in particular, performs the simultaneous summation

$$\sum_{d \text{ } B\text{-smooth, squarefree}} \frac{\tau(d)}{\sqrt{d}}.$$

No factor $2^{\pi(B)}$ must be reintroduced in this branch. After the union over the B^2 offset pairs, absorbed into the exponential, the sharp bound remains

$$M \exp\left(O\left(\frac{\sqrt{B}}{\log B}\right)\right).$$

The three branches give both assertions. $\square$

Proposition 3.17 (No aligned residual core of linear density). *For every fixed $\alpha > 0$, no aligned pair in $U_{M,\delta}^2$ satisfies $c^\# \geq \alpha B$ once M is sufficiently large in terms of the fixed parameters and α .*

Proof. Let m_{ex} be the number of exact units of the canonical channel. If $m_{\text{ex}} \geq 2$, the nondefective exact components have already been removed in $\mathcal{C}_{\text{res}}$. If $m_{\text{ex}} = 1$, make no height or valuation assumption and remove directly the at most two graph components containing its two occurrences. If $m_{\text{ex}} = 0$, remove nothing. In every case, from $c^\# \geq \alpha B$ there remain at least $\alpha B/2$ nontrivial components containing no exact unit. Their total size is at most $2B$, so at least $m \geq \alpha B/4$ have size at most $K_\alpha = \lceil 8/\alpha \rceil$.

For each such component C_t form the column

$$\Phi(C_t) = \left((v_p(\prod_{n \in C_t} n) \bmod 2)_{p \leq B}, |C_t \cap V_x| \bmod 2, |C_t \cap V_y| \bmod 2 \right) \in \mathbb{F}_2^{\pi(B)+2}.$$

The kernel of the matrix of these m columns has codimension at most $r = \pi(B) + 2$. Choose

$$t = \left\lceil C_\alpha \frac{B}{\log B \log \log B} \right\rceil$$

with C_α large. If no nonzero kernel word had weight at most $2t$, the Hamming bound would give $\sum_{j \leq t} \binom{m}{j} \leq 2^r$. Since $m \gg_\alpha B$ and $t = o(m)$,

$$\log \binom{m}{t} \geq t \log(m/t) - O(t) = (C_\alpha + o_\alpha(1)) \frac{B}{\log B},$$

which exceeds $r \log 2$ when C_α is fixed sufficiently large. Thus a nonzero word uses

$$O_\alpha\left(\frac{B}{\log B \log \log B}\right)$$

components.

Let I, J be the unions of their offsets. The last two coordinates of Φ make $|I|$ and $|J|$ even; the remaining coordinates, together with the large-prime equations internal to each component, show that

$$\prod_{i \in I} (x + i) \prod_{j \in J} (y + j)$$

is a square. Put $d_0 = |I| + |J|$; then

$$d_0 \ll_\alpha \frac{B}{\log B \log \log B}.$$

For the canonical channel (a, b, h) define

$$G(T) = \prod_{i \in I} (T + ai) \prod_{j \in J} (T + h + bj).$$

There are no root collisions: one would be an exact unit, and all components containing such a unit were removed. At $T = ax$,

$$G(ax) = a^{|I|} b^{|J|} \prod_{i \in I} (x + i) \prod_{j \in J} (y + j)$$

is an integer square. All roots have modulus $O(B^{A+1})$, while $ax \geq x \geq M^\delta$; in particular $ax \geq 2R$ for the root radius $R = O(B^{A+1})$ once M is large. Lemma 2.3 gives

$$ax \leq (Cd_0 B^{A+1})^{2d_0}.$$

The logarithm of the right-hand side is $O_\alpha(B/\log \log B) = o(B)$, whereas

$$\log(ax) \geq \delta \log M \geq (\delta/\beta_+)B.$$

This contradiction closes the aligned deep sector directly at scale M . $\square$

We now work in the nonaligned complement. A component of size at most $2/\alpha$ exists whenever $c^\# \geq \alpha B$, since the total number of occurrences is $2B$. The preceding component count therefore gives $M^{1+o(1)}$ hosts in every fixed positive-density deep sector. For sector 5, $\tau \leq D^\# + c^\# \leq 2B/3 + o(B)$, so its mass is $MQ_B^{2/3}M^{o(1)}$.

Lemma 3.18 (Two defects in one block). *The number of starts $x \in U_{M,\delta}$ whose window contains two distinct B -defective occurrences is $M^{o(1)}$.*

Proof. Write $x + i_1 = d_1 u^2$ and $x + i_2 = d_2 v^2$, with d_1, d_2 squarefree and B -smooth. If $d_1 \neq d_2$, then d_1/d_2 is not a square and

$$d_1 u^2 - d_2 v^2 = i_1 - i_2 \neq 0$$

is covered by Lemma 3.11; all coefficients and variables are in fixed polynomial ranges by (3.10) and (3.2). If $d_1 = d_2 = d$, then $d \mid i_1 - i_2$ and

$$(u - v)(u + v) = (i_1 - i_2)/d,$$

which has $M^{o(1)}$ solutions by the divisor bound. There are $4^{\pi(B)} B^2 = M^{o(1)}$ choices of coefficients and offsets. $\square$

Proposition 3.19 (Sector 6). *For nonaligned deep-core pairs with $c^\# > 2B/3$ and $D^\# \geq 3$,*

$$R_{\text{res}} \leq M^{1/2} Q_B M^{o(1)}.$$

Proof. One block contains at least two defects, so its start has $M^{o(1)}$ choices by Lemma 3.18. Since $c^\# > 2B/3$, a residual component has bounded size. Fix its offsets and square class at a cost $M^{o(1)}$. The start in the other block is then counted by Lemma 3.14: at most $M^{1/2}$ choices when its component degree is one, and $M^{o(1)}$ otherwise. Hence the number of hosts is $M^{1/2+o(1)}$.

Finally $D^\# + 2c^\# \leq 2B$ and $D^\# = O(B/\log B)$ imply $D^\# + c^\# \leq B + o(B)$. Therefore $2^\tau \leq Q_B M^{o(1)}$, and multiplication proves the result. $\square$

For sector 7, integrality gives $s + \tilde{k} \geq T_B + 1$. Thus (3.14) and $D^\# \leq 2$ imply $\tau \leq B + 1 - T_B = 2B/3 + O(1)$. The same deep-host count bounds its mass by $MQ_B^{2/3}M^{o(1)}$.

3.5 Terminal kernels: count clustering, not just candidates

The remaining issue is concentration within a window. On a slice with larger start of size X , one kernel below a constant multiple of $\sqrt{XB}$ allows $X^{3/4+o(1)}$ candidate starts (Lemma 3.21); this count is too large when paired with the full relation weight Q_B . If the component constraints force two such kernels in the same window, the shifted energy (Lemma 3.22) reduces the count to the $X^{2/3+o(1)}$ scale. If only one small kernel is forced, the residual dimension is smaller, and the one-kernel container is sufficient. Lemma 3.20 supplies the small kernels; proposition 3.23 combines the two populations. The gain comes from counting their concentration, not from shrinking the container.

It remains to consider sector 8. Write $j = s + \tilde{k} \leq T_B$. If a component has size $2 + e_t$, then $\sum_t e_t \leq 2B - 2c^\# = 2s$. At most $2s$ components have size greater than two, so at least $B - 3s \geq B - 3j \geq 2$ are isolated size-two components. For such a component $\{x + i_t, y + j_t\}$ put $R_t = K_B(x + i_t)$.

Lemma 3.20 (Terminal kernels and determinants). *For every isolated size-two component,*

$$R_t = K_B(x + i_t) = K_B(y + j_t) > 1.$$

Kernels belonging to distinct components are coprime. For $s \neq t$,

$$R_s R_t \mid \Delta_{s,t} := (x + i_s)(y + j_t) - (x + i_t)(y + j_s),$$

and

$$0 < |\Delta_{s,t}| \leq 4MB$$

for large M .

Proof. An isolated unpinned component contains exactly the two occurrences carrying every one of its $p > B$ labels, so the full kernels agree and are nontrivial. A prime cannot label two components, because each block contains at most one odd occurrence of that prime; hence the kernels are coprime. Divisibility of the determinant follows term by term.

If the determinant vanished, the two occurrence ratios would be equal. Writing their common reduced ratio as b/a gives $x + i_s = bu_s$, $y + j_s = au_s$ and similarly with u_t . Subtraction yields

$$i_s - i_t = b(u_s - u_t), \quad j_s - j_t = a(u_s - u_t).$$

Distinct components have distinct offsets, so $u_s \neq u_t$, and $a, b \leq L$. The two cells would therefore be two units of an aligned exact channel, contradicting nonalignment. Finally

$$|\Delta_{s,t}| \leq |x + i_s||j_t - j_s| + |y + j_s||i_s - i_t| \leq 4MB$$

by (3.2). □

Slice terminal pairs by $X \leq \max(x, y) < 2X$, where $M^\delta/2 \leq X \leq M$. For counting, call the larger coordinate z and the smaller one its partner; keep each pair's original canonical data. The two possible orientations cost only a factor two, and no symmetry of the canonical selection is assumed. On this slice $R_s R_t \ll XB$. Set $T_X = (CXB)^{1/2}$ with an absolute constant large enough for this inequality. At most one isolated kernel exceeds T_X .

Lemma 3.21 (Partners and a one-dimensional container). *On a terminal slice, a fixed value of either start has $X^{o(1)}$ possible partners. The number of possible values of either coordinate, and hence of the larger start, is at most $X^{3/4+o(1)}$. All constants are uniform in the slice.*

Proof. Fix either start and write it as x , with partner y , for this count only; the original canonical data of the pair are unchanged. Choose two isolated components at a cost $O(B^4) = X^{o(1)}$. Their distinct coprime kernels $R, S > 1$ are fixed by x . The other occurrences have the unique forms $y + j = Ruz^2$, $y + v = Su'w^2$ with independent squarefree B -smooth coefficients u, u' . Thus

$$Ruz^2 - Su'w^2 = j - v \neq 0.$$

The coefficients Ru, Su' are squarefree, of size $O(X)$, and their quotient is nonsquare because the large-prime supports of R, S are disjoint. The Pell bound gives $X^{o(1)}$ partners per pair of smooth coefficients, of which there are $4^{\pi(B)} = X^{o(1)}$. This argument needs only $y \ll X$, not $y \asymp X$.

At least one isolated kernel is at most T_X . Integers with such a kernel have the unique representation $n = ua^2r$, where u is squarefree and B -smooth and $r = K_B(n) \leq T_X$. Therefore

$$\#\{n \leq 3X : K_B(n) \leq T_X\} \leq \sqrt{3X} \prod_{p \leq B} (1 + p^{-1/2}) \sum_{r \leq T_X} r^{-1/2} \ll X^{3/4+o(1)}. \quad (3.20)$$

Each such integer lies in at most $B = X^{o(1)}$ windows, whether the coordinate is the first or the second start. Their union also bounds the larger starts. Since $B/\log X$ stays in a fixed compact interval, all constants are uniform. $\square$

Lemma 3.22 (Shifted small-kernel energy). *For $T \ll (XB)^{1/2}$ put $A_T(x) = \#\{i \in I_L : K_B(x+i) \leq T\}$. Uniformly on every slice,*

$$\sum_{X \leq x < 2X} \binom{A_T(x)}{2} \ll_{\varepsilon} X^{2/3+\varepsilon} B^2. \quad (3.21)$$

Proof. Let $W_B = \prod_{p \leq B} (1 + p^{-1/2}) = X^{o(1)}$. Kernels $r \leq R_0 = X^{1/3}$ occur at at most

$$\sqrt{3X} W_B \sum_{r \leq R_0} r^{-1/2} = X^{2/3+o(1)}$$

integers in $[X/2, 3X]$. Each can anchor at most $O(B^2)$ pairs of occurrences and windows, so these integers give the required bound.

For the remaining pairs fix $1 \leq h \leq B$ and dyadic ranges $r = K_B(n) \asymp R$, $s = K_B(n+h) \asymp S$, with $R, S \geq X^{1/3}$ and $R, S \ll (XB)^{1/2}$. Set $q = n/r$, $q' = (n+h)/s$. Both are a square times a squarefree B -smooth integer, with $q \asymp X/R$ and $q' \asymp X/S$. To count their choices explicitly, put

$$\mathcal{Q}_B(T) = \{a^2u \leq T : u \text{ squarefree}, P^+(u) \leq B\}.$$

The squarefree decomposition is unique, and therefore

$$|\mathcal{Q}_B(T)| \leq \sum_{\substack{u \leq T \\ \text{squarefree} \\ P^+(u) \leq B}} \left\lfloor \sqrt{T/u} \right\rfloor \leq \sqrt{T} W_B.$$

For suitable absolute constants in the height bounds, the number of cofactor pairs (q, q') is thus at most

$$C\sqrt{X/R}W_B\sqrt{X/S}W_B = \frac{CXW_B^2}{\sqrt{RS}}.$$

Fix such a pair and write $g = (q, q')$. The equation

$$qr + h = q's$$

requires $g \mid h$. When it has a solution (r_0, s_0) , all solutions are

$$r = r_0 + \frac{q'}{g}t, \quad s = s_0 + \frac{q}{g}t, \quad t \in \mathbb{Z}.$$

The box $r \asymp R, s \asymp S$ contains at most

$$1 + C \min\{Rg/q', Sg/q\} \leq 1 + CBRs/X$$

solutions, since $g \leq h \leq B$. Dropping any remaining squarefreeness or coprimality restrictions only enlarges this upper bound. Multiplying the number of cofactor choices by the number of solutions for each gives

$$X^{o(1)} \left(\frac{X}{\sqrt{RS}} + B\sqrt{RS} \right) \leq X^{2/3+o(1)}.$$

Here $X/\sqrt{RS} \leq X^{2/3}$, while $B\sqrt{RS} \ll X^{1/2}B^{3/2} = X^{1/2+o(1)}$, as $B \asymp \log X$. There are $X^{o(1)}$ pairs of dyadic ranges. Sum over h and note that a shifted pair lies in at most B windows. This proves (3.21). $\square$

Proposition 3.23 (Terminal mass). *Sector 8 contributes at most*

$$M^\varepsilon \left(M^{2/3}Q_B + M^{3/4}Q_B^{2/3} \right).$$

Proof. On a fixed larger-start slice put $j = s + \tilde{k}$ and $m_j = B - 3j - 1$. The sector test gives $m_j \geq 1$. By equation (3.14), $\tau + j = B + D^\#$ with $D^\# \leq 2$. There are at least $B - 3j$ isolated kernels, all but at most one below T_X , so each window has at least m_j small kernels.

First collect all pairs with $m_j \geq 2$, without separating the values of j . Their larger start z satisfies $(A_{T_X/2}(z)) \geq 1$. The energy estimate counts at most $X^{2/3+o(1)}B^2$ such starts, and the uniform partner bound contributes only $X^{o(1)}$. Since $2^\tau - 1 \leq 4Q_B$, the total mass is at most $X^{2/3+o(1)}Q_B$; the factor B^2 is subpolynomial.

In the complementary population, $m_j \leq 1$ implies $3j \geq B - 2$. Consequently $3\tau \leq 2B + 8$, and $2^\tau - 1 \leq 8Q_B^{2/3}$. At least one small kernel remains, so the one-dimensional container and partner bound give at most $X^{3/4+o(1)}$ pairs. Their mass is at most $X^{3/4+o(1)}Q_B^{2/3}$.

The dyadic slices of the larger coordinate exhaust the terminal pairs. Their logarithmic bands are uniform, and geometric summation proves the claim. The partner estimate uses only its upper height bound, never comparability with X . $\square$

3.6 Assembly and lower bounds

Proof of Theorem 3.1. The completed estimates are as follows; a factor $M^{o(1)}$ is suppressed.

Population	Mechanism	Weighted mass
Rational	Exact-unit enumeration	$MQ_B^{1/3} + MQ_B^{1/2}$
1	Small product and rational mass	$M^{3/2} + MQ_B^{1/2}$
2	Small channel height	$M^{3/2} + MQ_B^{1/2} + MQ_B^{1/3}$
3	Host count and rank	$M^{3/2}Q_B^{1/6}$
4	Runge and a short codeword	0
5	Bounded components	$MQ_B^{2/3}$
6	Two defects and a component	$M^{1/2}Q_B$
7	Exact rank deficit	$MQ_B^{2/3}$
8	Shifted kernel energy and Pell	$M^{2/3}Q_B + M^{3/4}Q_B^{2/3}$

Every entry is bounded by one of the three monomials in (3.3). The partition is exhaustive and (3.6) is exact, so summation proves that profile. Finally

$$M^{3/2}Q_B^{1/6} = (M^{5/3})^{5/6}(M^{2/3}Q_B)^{1/6}, \quad MQ_B^{2/3} = (M^{5/3})^{1/3}(M^{2/3}Q_B)^{2/3};$$

weighted arithmetic–geometric mean proves (3.4). $\square$

The exponent $5/3$ is an upper bound, not an optimality claim. In a dyadic block, the ordered pairs $x = 2t, y = 3t$, $N/2 \leq t < 2N/3$, and their reversals give

$$R_2(N, L) \geq \left(\frac{N}{3} + O(1)\right) \left(2^{\lfloor (B+1)/3 \rfloor - 1} - 1\right). \quad (3.22)$$

On $U_{M,\delta}$, the pairs $x = t, y = 2t$ and their reversals give

$$R_{2,\delta}(M, L) \geq (M - 2M^\delta + O(1)) \left(2^{\lfloor B/2 \rfloor - 1} - 1\right). \quad (3.23)$$

These follow simply by counting the even code of exact units, regardless of any extra residual relations. At criticality the lower exponents are $4/3$ and $3/2$, respectively. The one-dimensional container (3.20) is itself of order $X^{3/4+o(1)}$ when $T \asymp \sqrt{XB}$: choose r prime near T and $n = ra^2 \asymp X$. What improves the terminal estimate is not a smaller container, but the fact that its elements cannot cluster sufficiently often in one short window, as quantified by (3.21).

The weighted relation estimate is now available. A Poisson comparison still requires the conditional marginal probabilities, an exact dependency graph and the local overlap counts. Before constructing that graph, we record two forms of the estimate adapted to prescribed words.

3.7 From relative signs to prescribed values

There is little extra arithmetic in prescribing the sign of a word rather than identifying it with its negative. What changes is the parity restriction on square-product subsets.

For separated windows let $\rho_B^{\text{val}}(x, y)$ be the nullity of the transpose of the matrix with all $2B$ valuation rows $v(n)$, $n \in V_x \sqcup V_y$. Thus its kernel consists of all square-product subsets, without requiring even cardinality in each block. Write $R_{2,\delta}^{\text{val}}(M, B)$ for the sum of $2\rho_B^{\text{val}}(x, y) - 1$ over the same separated pairs as $R_{2,\delta}$, and $R_2^{\text{val}}(N, B)$ for its dyadic analogue.

Proposition 3.24 (Prescribing the global signs). *Uniformly under the hypotheses of Theorem 3.1,*

$$R_{2,\delta}^{\text{val}}(M, B) \leq 4R_{2,\delta}(M, B-1) + 3H_{2,\delta}^{\text{val}}(M, B), \quad (3.24)$$

$$R_{2,\delta}^{\text{val}}(M, B) \ll_{\varepsilon} M^{\varepsilon} \left(M^{3/2} Q_B^{1/6} + M Q_B^{2/3} + M^{2/3} Q_B \right) \ll_{\varepsilon} M^{\varepsilon} (M^{5/3} + M^{2/3} Q_B). \quad (3.25)$$

The same conclusion holds on a dyadic block and on bounded-ratio subintervals, with the corresponding ambient scale.

Proof. Imposing the two block-parity equations on the full square-relation kernel gives exactly the relative-sign kernel, by the tree boundary bijection. Therefore

$$\rho_{B-1}(x, y) \leq \rho_B^{\text{val}}(x, y) \leq \rho_{B-1}(x, y) + 2.$$

If the full kernel is zero both weights vanish. Otherwise

$$2^{\rho_B^{\text{val}}} - 1 \leq 4(2^{\rho_{B-1}} - 1) + 3.$$

The host term includes every separated ordered pair admitting a nonempty square-product subset of the full vertex occurrences, with no even-cardinality condition in either block. In particular it may count a pair whose relative-sign kernel is zero. Sum the inequality using exactly this unrestricted host estimate from proposition 3.7. Its $M^{3/2+o(1)}$ term is absorbed by the first monomial in (3.3). Restriction proves the other geometries. The extra two affine constraints therefore cost only a fixed factor and a sparse set of new relational pairs, not an exponential factor in the window length. $\square$

3.8 Capping the relation weight

The full relation weight is useful when estimating moments. For the probability that one of many mutually exclusive words occurs, however, its largest values can be wasteful: a joint probability cannot exceed either marginal. We record a version of the arithmetic estimate that preserves this elementary ceiling.

For $T \geq 1$ define

$$R_{2,\delta}^{[T]}(M, L) = \sum_{\substack{x, y \in U_{M,\delta} \\ |x-y| > L}} \min\{T, 2^{\rho_L(x,y)} - 1\}.$$

The notation $R_{2,\delta}^{\text{val},[T]}(M, B)$ has the same meaning with the full-value nullity ρ_B^{val} ; omission of δ denotes the corresponding dyadic sum.

Proposition 3.25 (Capped two-window profile). *Under the fixed-band hypotheses of Theorem 3.1, uniformly also for $T \geq 1$,*

$$R_{2,\delta}^{[T]}(M, L) \ll_{\varepsilon} M^{\varepsilon} \left(M^{3/2} Q_B^{1/6} + M Q_B^{2/3} + M^{2/3} \min\{T, Q_B\} \right), \quad (3.26)$$

$$R_{2,\delta}^{\text{val},[T]}(M, B) \ll_{\varepsilon} M^{\varepsilon} \left(M^{3/2} Q_B^{1/6} + M Q_B^{2/3} + M^{2/3} \min\{T, Q_B\} \right). \quad (3.27)$$

The same bounds hold on dyadic and bounded-ratio subintervals. The implicit constants are independent of T .

Proof. Write $2^{\rho} - 1 = (2^{\sigma} - 1) + 2^{\sigma}(2^{\tau} - 1)$ and use $\min(T, a+b) \leq a + \min(T, b)$ for $a, b \geq 0$. The rational contribution and sectors 1–5 and 7 are left uncapped; the assembly table bounds their sum by $M^{o(1)}(M^{3/2} Q_B^{1/6} + M Q_B^{2/3})$. In sector 6, the proof of its host bound gives $M^{1/2+o(1)}$

pairs, and the pointwise weight is at most $Q_B M^{o(1)}$. Here the factor is chosen independently of T . The inequality $\min(T, EQ_B) \leq E \min(T, Q_B)$ for $E \geq 1$ therefore bounds this contribution by $M^{1/2+o(1)} \min(T, Q_B)$.

For sector 8, use the two populations in the proof of proposition 3.23. On a larger-start slice, the first population has at most $X^{2/3+o(1)}$ pairs and pointwise weight at most $4Q_B$. All geometric and partner bounds were fixed before T , so

$$\min\{T, 2^\tau - 1\} \leq 4 \min\{T, Q_B\}$$

gives capped mass $X^{2/3+o(1)} \min\{T, Q_B\}$ with a common subpolynomial factor independent of T . Leave the complementary population uncapped: its mass $X^{3/4+o(1)} Q_B^{2/3}$ is absorbed by $X^{1+o(1)} Q_B^{2/3}$. Geometric summation proves (3.26).

Finally, the two parity constraints give the pointwise inequality

$$\min\{T, 2^{\rho_B^{\text{val}}} - 1\} \leq 4 \min\{T, 2^{\rho_B - 1} - 1\} + 3 \mathbf{1}_{\{\rho_B^{\text{val}} > 0\}}.$$

The last indicator has sum $M^{3/2+o(1)}$ by the full-value host bound. This proves (3.27); restriction proves the other geometries. $\square$

The cap changes only the dense terminal contribution. Its benefit will come from choosing T according to a probability, rather than from improving the uncapped exponent $5/3$.

4 From arithmetic relations to Poisson laws

The arithmetic estimate is now available independently of the probability argument. The remaining difficulty is not a lack of pair information, but the absence of a sparse unconditional dependency graph. Conditioning on the small prime signs supplies exactly such a graph.

4.1 An explicit transfer inequality

Let $\mathcal{F}_Y = \sigma(X_p : p \leq Y)$, with $Y > 2B$. Let D_Y be the starts in I_N for which some vertex of V_x is Y -defective, and put $G_Y = I_N \setminus D_Y$. We deliberately use the whole support in this removal rule, so that the same convention works for marks and overlaps. For $x \in G_Y$, the private-pivot lemma gives, for every environment,

$$\mathbb{P}(A_x X = b \mid \mathcal{F}_Y) = 2^{-L} \quad \text{for every } b \in \mathbb{F}_2^L. \quad (4.1)$$

After conditioning, an event at x depends only on the prime coordinates

$$\Pi_x(Y) = \{p > Y : \exists n \in V_x, v_p(n) \text{ odd}\}.$$

Join starts whose coordinate sets intersect. Disjoint collections with no edge between them use disjoint independent coordinates. This is an exact conditional dependency graph, not an approximate decorrelation assertion.

Write $E_Y(A)$ for its ordered edge count restricted to a mask A , and set

$$\mathcal{M}_B(A) = \sum_{x \in A} (2^{m_B(x)} - 1), \quad D_Y(A) = A \cap D_Y, \quad \mathcal{M}_B = \mathcal{M}_B(I_N).$$

All defect counts use the complete support, including $x - 1$. By Corollary 2.5, the actual probability mass removed from A is at most $p(\mathcal{M}_B(A) + |D_Y(A)|)$, where $p = 2^{-L}$. Deleting

the corresponding Poisson target coordinates costs at most $p|D_Y(A)|$. The edge count obeys

$$E_Y(I_N) \ll B^2 \left(\frac{N^2}{Y \log Y} + N \log \log N + \frac{N}{\log N} \right). \quad (4.2)$$

Indeed a prime $p > Y$ can occur in at most $O(B(N/p + 1))$ windows; sum the square of this quantity over $Y < p \leq 3N$ and use partial summation.

We use the dependency-graph form of the Arratia–Goldstein–Gordon theorem [4,5]. For indicators (I_α) with neighbourhoods B_α containing α , let

$$b_1 = \sum_{\alpha} \sum_{\beta \in B_\alpha} p_\alpha p_\beta, \quad b_2 = \sum_{\alpha} \sum_{\substack{\beta \in B_\alpha \\ \beta \neq \alpha}} \mathbb{E}(I_\alpha I_\beta), \quad \mu = \sum_{\alpha} p_\alpha.$$

An exact dependency graph has no third error term. The scalar Poisson distance is at most $C(1 \wedge \mu^{-1})(b_1 + b_2)$; the joint indicator-to-independent-Poisson distance is at most $C(b_1 + b_2)$. These assertions also apply on every fibre of $\mathcal{F}_Y$.

Theorem 4.1 (Masked arithmetic-to-Poisson transfer). *For every deterministic $A \subset I_N$, put $\mu_A = |A \cap G_Y|p$ and $R_2(A) = \sum_{x,y \in A, |x-y|>L} (2^{\rho(x,y)} - 1)$. Then*

$$\begin{aligned} & \mathbb{E} d_{\text{TV}} \left(\mathcal{L} \left(\sum_{x \in A} J_{x,L} \mid \mathcal{F}_Y \right), \text{Pois}(|A|p) \right) \\ & \ll p(\mathcal{M}_B(A) + 2|D_Y(A)|) + (1 \wedge \mu_A^{-1})p^2(|A| + E_Y(A) + R_2(A)). \end{aligned} \quad (4.3)$$

The multiplier is interpreted as one if $\mu_A = 0$. Replacing it by one gives a valid joint-field bound on the same finite site lattice.

Proof. Delete bad actual sites and their target coordinates. The natural coupling costs at most $p(\mathcal{M}_B(A) + 2|D_Y(A)|)$. On each remaining fibre, the marginals are exactly p , and the graph above is a dependency graph. Thus $b_1 \leq p^2(|A| + E_Y(A))$. Overlapping starts are incompatible. At distance L , the union is a tree with all its vertices possessing private coordinates above Y , since its diameter is $2L < Y$. Its joint probability is therefore p^2 . For separated pairs the affine Fourier identity gives, after averaging,

$$\mathbb{E} b_2 \ll p^2(|A| + E_Y(A) + R_2(A)).$$

Apply the scalar or process theorem on each fibre, average, and add the deletions. The target is deterministic, so mixing also gives the same unconditional bound. $\square$

The denominator uses the mask's own mean. For a sparse mask it cannot be replaced by the ambient intensity Np . At bounded ambient intensity, all nonnegative sums are bounded by their full-block values, so the critical estimate is uniform over all deterministic masks.

4.2 The two prime-cutoff scales

The leading scales are already visible without any special functions. Set $S_N = \sqrt{\log N \log \log N}$ and $Y = e^{aS_N}$. The normalized number of bad starts has leading upper exponent $1/(2a)$, while the normalized edge count has exponent a . Their balance is $a = 1/\sqrt{2}$. For rare conditioning events and growing intensities we retain the smaller second-order terms. The following definitions keep these terms explicit.

For $\nu \geq e$, let $u(\nu) \geq 1$ solve $e^u/u = \nu$, and define

$$D(\nu) = \nu u(\nu) - \text{Ei}(u(\nu)), \quad H_N = \log N.$$

For sufficiently large N , define V_N and $\widehat{V}_N$ by

$$V_N = D(H_N/V_N), \quad \widehat{V}_N = 2D(H_N/\widehat{V}_N), \quad \nu_N = H_N/V_N, \quad \widehat{\nu}_N = H_N/\widehat{V}_N. \quad (4.4)$$

The relevant solutions have $H_N/V_N, H_N/\widehat{V}_N \geq e$. Since $D'(\nu) = u(\nu) > 0$, each equation has exactly one such solution for large N . Put $Y_N^\sharp = \lfloor e^{V_N} \rfloor$ and $\widehat{Y}_N = \lfloor e^{\widehat{V}_N} \rfloor$.

Proposition 4.2 (Uniform cutoff estimates). *Let $w \asymp \sqrt{H_N \log H_N}$, $\nu_w = H_N/w$ and $Y = \lfloor e^w \rfloor$. Uniformly in a fixed logarithmic band,*

$$\frac{|D_Y|}{N} \leq e^{-D(\nu_w)+o(\nu_w)}, \quad \frac{E_Y(I_N)}{N^2} \leq e^{-w+o(\nu_w)} + N^{-1+o(1)}. \quad (4.5)$$

The maximum degree Δ_Y of the graph on all starts, including bad ones, satisfies the same second bound after normalization by N . Moreover,

$$V_N^2 = \frac{H_N}{2}(\log H_N + \log \log H_N - \log 2 - 2 + o(1)), \quad (4.6)$$

$$\widehat{V}_N^2 = H_N(\log H_N + \log \log H_N - \log 4 - 2 + o(1)). \quad (4.7)$$

In particular $\widehat{V}_N \sim \sqrt{2}V_N$ and $\log \log N = o(\nu_N)$.

Proof. If $A(X, Y)$ counts Y -defective integers up to X , their unique square-smooth factorization gives, for $0 < \zeta < 1/2$,

$$A(X, Y)/X \leq X^{-\zeta} \prod_{p \leq Y} (1 + p^{-1+\zeta}).$$

Choose $\zeta = u(\nu_w)/w$. The prime number theorem and partial summation give $\sum_{p \leq e^w} \log(1 + p^{-1+\zeta}) = \text{Ei}(u(\nu_w)) + o(\nu_w)$. This proves the first bound, because $|D_Y| \leq BA(3N, Y)$ and $\log B = o(\nu_w)$. The edge bound follows from (4.2). Each window uses $O(BH_N/w)$ primes above Y , each meeting $O(B(N/Y + 1))$ other windows, which proves the maximum-degree bound. The expansions follow by substituting the asymptotic expansion of $\text{Ei}(u)$ into (4.4). Proposition B.1 gives the uniform remainder calculation and both expansions in detail. $\square$

Theorem 4.3 (Hard deletion and soft retention). *Fix $0 < \beta_- < \beta_+ < \infty$. Uniformly when $\beta_- \log N \leq L + 1 \leq \beta_+ \log N$, for every $\varepsilon > 0$,*

$$d_{\text{TV}}(\mathcal{L}(Z_{N,L}), \text{Pois}(\lambda_N)) \ll \min \left\{ 1, \lambda_N \left(e^{-V_N+o(\nu_N)} + N^{-1/3+\varepsilon} \right) \right\}, \quad (4.8)$$

$$d_{\text{TV}}(\mathcal{L}(Z_{N,L}), \text{Pois}(\lambda_N)) \ll \min \left\{ 1, e^{-(\widehat{V}_N - \log^+ \lambda_N)/2+o(\widehat{\nu}_N)} + N^{-1/3+\varepsilon} \right\}. \quad (4.9)$$

The first estimate also bounds the mean conditional distance given $\mathcal{F}_{Y_N^\sharp}$, and the second given $\mathcal{F}_{\widehat{Y}_N}$. Neither saddle is asserted to be optimal for the distribution itself.

Proof. At the hard cutoff, $|D_Y| = o(N)$ and $\mu_{I_N} \asymp \lambda_N$. The master estimate and $Q_B = 2N/\lambda_N$ give

$$R_2(N, L)/N^2 \ll N^{-1/3+\varepsilon}(1 + \lambda_N^{-1}).$$

Its contribution to (4.3) is $O(\lambda_N N^{-1/3+\varepsilon})$. The defect mass contributes $\lambda_N N^{-1/2+o(1)}$ and the cutoff terms contribute $\lambda_N e^{-V_N+o(\nu_N)}$. This proves (4.8).

For the second estimate retain the exceptional indicators in the scalar sum instead of coupling all of them to zero. The scalar Stein equation has factors $\|f\|_\infty \ll 1 \wedge \lambda^{-1/2}$ and $\|\Delta f\|_\infty \leq 1 \wedge \lambda^{-1}$ [25,34]. Telescope only the good-start terms, using their exact conditional marginal. Bound the exceptional part by its total first moment times $\|f\|_\infty$. Exceptional neighbours of a good start stay in the neighbour sums; their mean is controlled after averaging, not pointwise. Set $\ell = \log^+ \lambda$ and $K = \max(1, \lambda) = e^\ell$. The same two Stein factors, at the same cutoff, give for every $\lambda > 0$ the upper bound

$$e^{\ell/2 - D(H_N/w) + o(H_N/w)} + e^{\ell - w + o(H_N/w)} + KN^{-1/3 + \varepsilon/2}. \quad (4.10)$$

Here a smaller arithmetic exponent absorbs the subpolynomial terms. The detailed calculation is in section B.3. At $w = \widehat{V}_N$ choose a common nonnegative remainder $\eta_N \widehat{v}_N$, $\eta_N \rightarrow 0$, in the two exponentials and put $q = \exp(-(w - \ell)/2 + \eta_N \widehat{v}_N)$. If $q + N^{-1/3 + \varepsilon} \geq 1$, the bound $d_{TV} \leq 1$ suffices. Otherwise $q < 1$ implies $\ell \leq w$, the second exponential is at most q , and $K \leq e^w \leq N^{\varepsilon/2}$ for all sufficiently large N . This proves (4.9). In particular intensities below one are covered directly, without switching to a different conditioning argument. All thresholds precede the length in the fixed band; the fibrewise proof gives the claimed mean distance for the full $\mathcal{F}_{\widehat{Y}_N}$. $\square$

The hard estimate is more accurate at bounded intensity. The soft estimate permits larger intensities: $\log^+ \lambda_N \leq \widehat{V}_N - c\widehat{v}_N$ for fixed $c > 0$ suffices. In particular the hard-cutoff endpoint $\log \lambda_N = V_N + O(1)$ is covered, but no convergence at the exact new endpoint $\log \lambda_N = \widehat{V}_N$ is inferred.

Write $(z)_2 = z(z - 1)$ for the second falling factorial.

Corollary 4.4 (First two moments). *At critical lengths,*

$$\begin{aligned} \mathbb{E}Z_{N,L} &= \lambda_N + O(N^{-1/2 + o(1)}), \\ \mathbb{E}(Z_{N,L})_2 &= \lambda_N^2 + O_\varepsilon(N^{-1/3 + \varepsilon}), \\ \text{Var } Z_{N,L} &= \lambda_N + O_\varepsilon(N^{-1/3 + \varepsilon}). \end{aligned}$$

Proof. The first assertion is Corollary 2.5. Expand the second factorial moment over ordered distinct starts. Overlaps cost $O(NLp^2)$; touching pairs cost $p^2N^{1+o(1)}$ by Lemma 2.8. Separated pairs differ from p^2 by at most $p^2(2^{\rho(x,y)} - 1)$, whose sum is $O_\varepsilon(N^{-1/3 + \varepsilon})$. The variance identity concludes. This argument proves the moments directly: total variation alone would not imply convergence of unbounded moments. $\square$

5 Rare word fields, exact marks and the Poisson staircase

The two-window estimate controls dependence between distant words. Their local dependence has a different origin: a suffix of one word may coincide with a prefix of another. Keeping these two issues separate gives a joint law for a growing dictionary of prescribed words, and then a sign-resolved law for long runs.

5.1 A dictionary of prescribed words

Joint word and clump approximations are classical in Markov sequences [33]. Here the local overlap condition is kept separate from the arithmetic estimate for separated windows. Let $\mathcal{W} \subset \{\pm 1\}^B$ be a nonempty deterministic set of distinct words, and write $m = |\mathcal{W}|$, $p_B = 2^{-B}$

and $\Lambda = Nmp_B$. All signs, including the first one, are prescribed. Define

$$I_{x,w} = \mathbf{1}_{\{f(x-1+j)=w_j \ (0 \leq j < B)\}}, \quad x \in I_N, \quad w \in \mathcal{W}.$$

The first letter is at $x - 1$; this indexing matches the left-border convention for runs. For $1 \leq d < B$ write $w \sim_d v$ when $w_{d+j} = v_j$ for $0 \leq j < B - d$. The directed overlap weight is

$$\Omega(\mathcal{W}) = \frac{1}{m} \sum_{w,v \in \mathcal{W}} \sum_{d=1}^{B-1} 2^{-d} \mathbf{1}_{\{w \sim_d v\}}. \quad (5.1)$$

The normalization makes $\Lambda\Omega$ the expected scale of ordered local pairs, up to an absolute factor. It includes self-overlaps and cross-overlaps; controlling each word separately would miss the latter. For example, $(+, +, -)$ followed one site later by $(+, -, +)$ is compatible: the two occurrences fit into $(+, +, -, +)$. The overlap is a constraint on the words themselves, before the values of the multiplicative function are considered.

Theorem 5.1 (Joint Poisson approximation for a word dictionary). *Fix $0 < \beta_- < \beta_+ < \infty$ and $0 < \varepsilon < 1/3$. Uniformly for $\beta_- \log N \leq B \leq \beta_+ \log N$ and all dictionaries $\mathcal{W}$,*

$$\begin{aligned} \mathbb{E} d_{\text{TV}} \left(\mathcal{L}((I_{x,w})_{x \in I_N, w \in \mathcal{W}} \mid \mathcal{F}_{Y_N^\#}), \bigotimes_{x \in I_N, w \in \mathcal{W}} \text{Pois}(p_B) \right) \\ \ll_{\beta_-, \beta_+, \varepsilon} \Lambda\Omega(\mathcal{W}) + \Lambda(1 + \Lambda)e^{-V_N + o(\nu_N)} \\ + N^{-1/3+\varepsilon} \left(\Lambda + \Lambda^{11/6} m^{1/6} + \Lambda^{4/3} m^{2/3} \right). \end{aligned} \quad (5.2)$$

The implied constant and the little-oh term are independent of the words and of m . In particular, fix $0 < \eta < 1/2$ and $C > 0$. If

$$m_N \leq N^{1/2-\eta}, \quad |B_N - \log_2(Nm_N)| \leq C, \quad \Omega(\mathcal{W}_N) \longrightarrow 0, \quad (5.3)$$

then the entire site-and-word-labelled field converges in total variation to the independent Poisson field. Its error is

$$O_{C,\eta} \left(\Omega(\mathcal{W}_N) + e^{-V_N + o(\nu_N)} + N^{-\eta/2} \right). \quad (5.4)$$

Every deterministic restriction of the site-word index set, and every measurable function of the field, satisfies the same bound.

Proof. Use $Y = Y_N^\#$ and remove starts whose B -vertex support contains a Y -defective integer. Let D_Y^{val} denote this set and put $\mathcal{M}_B^{\text{val}} = \sum_{x \in I_N} (2^{m_B(x)} - 1)$ and $a = mp_B$. Distinct words at a site are mutually exclusive, so $0 < a \leq 1$ and $\Lambda = Na$. By Corollary 2.6, deleting the actual and target coordinates costs at most

$$a\mathcal{M}_B^{\text{val}} + 2a|D_Y^{\text{val}}|.$$

All retained word events have conditional marginal p_B . Join all words at the same site and all words at sites with intersecting prime supports above Y , adding edges for overlapping supports if necessary. After conditioning this is an exact dependency graph. Its first Poisson term is $O(a^2(NB + E_Y))$, where E_Y counts ordered pairs of distinct sites sharing a prime above Y .

For $y = x + d$ with $1 \leq d < B$, two specified words are incompatible unless $w \sim_d v$. In the compatible case they prescribe all $B + d$ values in the union. Every vertex of a good union has

a private coordinate above Y , because its diameter is less than $2B < Y$. Hence

$$\mathbb{P}(I_{x,w}I_{x+d,v} = 1 \mid \mathcal{F}_Y) = 2^{-(B+d)} \mathbf{1}_{\{w \sim_d v\}}. \quad (5.5)$$

Summing ordered local pairs costs at most $2\Lambda\Omega(\mathcal{W})$. Different words at the same site contribute zero to the second term.

For separated good sites put $I_x^{\mathcal{W}} = \sum_{w \in \mathcal{W}} I_{x,w}$. This is an indicator, not a count with possible multiplicities, and $\mathbb{P}(I_x^{\mathcal{W}} = 1) = a$. Fourier inversion on the valuation rows also gives $\mathbb{P}(I_x^{\mathcal{W}} I_y^{\mathcal{W}} = 1) \leq a^2 2^{\rho_B^{\text{val}}(x,y)}$. Combining the two bounds *before* summing over sites yields

$$\begin{aligned} \sum_{w,v \in \mathcal{W}} \mathbb{P}(I_{x,w} I_{y,v} = 1) &= \mathbb{P}(I_x^{\mathcal{W}} I_y^{\mathcal{W}} = 1) \\ &\leq \min\{a, a^2 2^{\rho_B^{\text{val}}(x,y)}\} \\ &\leq a^2 \left(1 + \min\{a^{-1}, 2^{\rho_B^{\text{val}}(x,y)} - 1\}\right). \end{aligned} \quad (5.6)$$

Only the error summands have been grouped: the Poisson theorem is still applied to the full family indexed by (x, w) . No word or position label is discarded. Its averaged second term over separated graph edges is therefore at most $a^2(E_Y + R_2^{\text{val}, [1/a]}(N, B))$. We have proved the finite transfer inequality

$$\mathbb{E} d_{\text{TV}} \ll a(\mathcal{M}_B^{\text{val}} + |D_Y^{\text{val}}|) + a^2 \left(NB + E_Y + R_2^{\text{val}, [1/a]}(N, B) \right) + \Lambda\Omega(\mathcal{W}). \quad (5.7)$$

Here $\mathcal{M}_B^{\text{val}} \leq N^{1/2+o(1)}$, $|D_Y^{\text{val}}|/N \leq e^{-V_N+o(\nu_N)}$, and $E_Y/N^2 \leq e^{-V_N+o(\nu_N)} + N^{-1+o(1)}$. Use proposition 3.25 with $T = N/\Lambda$ and $Q_B = 2^B = Nm/\Lambda \geq T$. Its three contributions, after multiplication by a^2 , are bounded by

$$N^{-1/3+\varepsilon} \left(\Lambda^{11/6} m^{1/6} + \Lambda^{4/3} m^{2/3} + \Lambda \right).$$

The smaller polynomial errors are absorbed with a smaller intermediate ε : for example, $m \geq \Lambda/N$ implies $\Lambda^2/N \leq N^{-1/3} \Lambda^{4/3} m^{2/3}$. This proves (5.2). Under (5.3), Λ is bounded above and below and $B/\log N$ stays in one fixed compact interval. The polynomial term is $O(N^{-\eta/2})$ on taking $\varepsilon = \eta/6$. Restriction and measurable contraction cannot increase total variation. $\square$

The exponent $1/2$ is a sufficient uniform bound, not an optimality claim. The remaining restriction comes from the nonterminal term $NQ_B^{2/3}$: after normalization it gives $N^{-1/3+\varepsilon} m^{2/3}$ at bounded intensity. The terminal term has instead been capped by the marginal probability. This distinction matters: a large raw relation count need not produce a proportionally large probability of seeing a dictionary word.

Corollary 5.2 (Replacement by independent signs). *Let (ξ_n) be independent fair signs and let $I_{x,w}^{\text{iid}}$ be the same word indicators formed from them. Under (5.3),*

$$d_{\text{TV}} \left(\mathcal{L}((I_{x,w})_{x,w}), \mathcal{L}((I_{x,w}^{\text{iid}})_{x,w}) \right) \longrightarrow 0.$$

In particular the vector of counts of the individual words is jointly close to independent Poisson variables of mean $N2^{-B}$, even when the number of words tends to infinity.

Proof. For the independent sequence the graph needs only overlapping supports. Equation (5.5) is exact without conditioning, and its process Poisson error is $O(\Lambda\Omega(\mathcal{W}) + \Lambda^2 B/N)$. Compare both fields with the same independent Poisson field and apply Theorem 5.1. Summing sites is a measurable contraction. $\square$

5.2 Typical dictionaries: average before taking the worst case

The uniform theorem protects against an adversarial choice of words. A randomly chosen dictionary asks a different question: can its particular selection avoid most of the arithmetic dependence? The useful average is not the law obtained by forgetting the dictionary. We first fix the selected words, compare the resulting labelled laws, and only then average the distances.

For a dictionary $\mathcal{W}$ and a deterministic start mask $A \subset I_N$, write

$$\Delta_{N,A,Y}(\mathcal{W}) = \mathbb{E} d_{\text{TV}} \left(\mathcal{L}((I_{x,w})_{x \in A, w \in \mathcal{W}} \mid \mathcal{F}_Y), \bigotimes_{x \in A, w \in \mathcal{W}} \text{Pois}(2^{-B}) \right). \quad (5.8)$$

The expectation in this definition is over the prime environment only. At $A = I_N$ and $Y = Y_N^\sharp$, abbreviate it by $\Delta_N(\mathcal{W})$.

Theorem 5.3 (Typical word dictionaries). *Fix $0 < \beta_- < \beta_+ < \infty$, $0 < \lambda_- < \lambda_+ < \infty$ and $0 < \varepsilon < 1/3$. For all sufficiently large N , uniformly for integers B, m with*

$$\beta_- \log N \leq B \leq \beta_+ \log N, \quad 1 \leq m \leq 2^B, \quad \lambda_- \leq \Lambda = Nm2^{-B} \leq \lambda_+,$$

choose $\mathcal{W}$ uniformly among the m -element subsets of $\{\pm 1\}^B$, independently of f . There is a deterministic bound

$$\mathbb{E}_{\mathcal{W}} \Delta_N(\mathcal{W}) \leq r_N, \quad r_N \ll_{\beta_-, \beta_+, \lambda_-, \lambda_+, \varepsilon} e^{-V_N + o(\nu_N)} + N^{-1/3+\varepsilon}. \quad (5.9)$$

The little-oh is uniform in B and m . For every $t > 0$,

$$\mathbb{P}_{\mathcal{W}}\{\Delta_N(\mathcal{W}) > t\} \leq \min\{1, r_N/t\}. \quad (5.10)$$

Thus $t_N \rightarrow 0$ and $r_N/t_N \rightarrow 0$ give a class of dictionaries of proportion tending to one, each with mean conditional distance at most t_N . Every deterministic restriction of its site-word coordinates obeys the same bound. Moreover the mean, over $\mathcal{W}$, of the unconditional distance between its multiplicative and independent-sign word fields is at most $r_N + O(\Lambda^2 B/N)$.

Proof. Put $p_B = Q^{-1}$, $Q = 2^B$, $a = m/Q$, and $b = m(m-1)/(Q(Q-1))$. For fixed words u, v ,

$$\mathbb{P}_{\mathcal{W}}(u, v \in \mathcal{W}) = b + (a-b)\mathbf{1}_{\{u=v\}}, \quad 0 \leq b \leq a^2 \leq a. \quad (5.11)$$

In particular, if $\mathbf{f}_x = (f(x-1+j))_{0 \leq j < B}$, independence of the selection from f gives

$$\mathbb{E}_{\mathcal{W}} \mathbb{P}_f(\mathbf{f}_x, \mathbf{f}_y \in \mathcal{W}) = b + (a-b)\mathbb{P}_f(\mathbf{f}_x = \mathbf{f}_y) \leq a^2 + a\mathbb{P}_f(\mathbf{f}_x = \mathbf{f}_y). \quad (5.12)$$

For disjoint supports, let $\kappa(x, y)$ be the nullity of the B valuation vectors for the ratios of corresponding values. Sending a coefficient vector t to (t, t) injects its relation space into the full $2B$ -value relation space. Therefore

$$\mathbb{P}_f(\mathbf{f}_x = \mathbf{f}_y) = 2^{-B} 2^{\kappa(x, y)} \leq p_B 2^{\rho_B^{\text{val}}(x, y)}. \quad (5.13)$$

For good supports at distance $1 \leq d < B$ and $Y > 2B$, all $B+d$ values in their union are conditionally uniform. Their equality graph has d chains and B independent edges. Thus $\mathbb{P}_f(\mathbf{f}_x = \mathbf{f}_{x+d} \mid \mathcal{F}_Y) = p_B$.

Apply the process Chen–Stein bound to each fixed dictionary on each prime fibre, using the same site-word graph as in Theorem 5.1. Average this *nonnegative error bound* over $\mathcal{W}$. The mean real deletion cost at each bad site is exactly a , regardless of its arithmetic word law; the

target cost is at most a . Products of good marginals cost $O(a^2(NB + E_Y))$. Distinct words at the same site cannot occur together. Equations (5.12)–(5.13) bound all remaining pairs: their baseline is absorbed since $ap_B \leq a^2$, and their arithmetic excess costs $ap_B R_2^{\text{val}}(N, B)$. More generally, for every deterministic mask A this proves

$$\mathbb{E}_{\mathcal{W}} \Delta_{N,A,Y}(\mathcal{W}) \ll a |D_Y^{\text{val}} \cap A| + a^2(|A|B + E_Y(A)) + ap_B R_2^{\text{val}}(A, B). \quad (5.14)$$

Here $E_Y(A)$ counts ordered distinct-site pairs sharing a nonzero valuation-parity coordinate above Y , and $R_2^{\text{val}}(A, B)$ sums over ordered pairs in A at distance at least B . All sums may include deleted sites. The companion, section C.1, gives the full finite bookkeeping. In particular no weighted defect term is removed from the *uniform* theorem: its removal here uses the dictionary average.

At $Y = Y_N^\sharp$, the deletion and edge bounds are those of proposition 4.2. The coarse full-value profile gives

$$ap_B R_2^{\text{val}}(N, B) \ll_\varepsilon \Lambda N^{-1/3+\varepsilon} \left(1 + \frac{\Lambda}{m}\right).$$

Together with $a^2 NB = \Lambda^2 B/N$, this proves (5.9). Markov's inequality proves (5.10). Restrictions contract the distance for each fixed dictionary, before either expectation. For independent signs there are only local edges; the same selection calculation gives mean Poisson error $O(\Lambda^2 B/N)$. Comparing the two fields with their common Poisson target proves the last assertion. $\square$

For example, $m_N \asymp N$ and $B_N = 2 \log_2 N + O(1)$ are allowed. More generally $m_N \asymp N^\alpha$ for any fixed $\alpha \geq 0$ is allowed with $B_N = (1+\alpha) \log_2 N + O(1)$ and a suitable fixed band. There is no uniform assertion over all dictionaries of those sizes. The dictionary is selected independently of the prime signs and then held fixed; neither adaptation to the observed sequence nor an almost-sure conclusion over all scales is part of the theorem.

Corollary 5.4 (Random affine dictionaries). *Identify sign words with $\mathbb{F}_2^B$. Choose an $r \times B$ matrix A_0 uniformly among the matrices of row rank r , where $0 \leq r \leq B$, and choose $b_0 \in \mathbb{F}_2^r$ independently and uniformly. The random dictionary*

$$\mathcal{W} = \{u \in \mathbb{F}_2^B : A_0 u = b_0\}, \quad m = 2^{B-r},$$

is independent of f . Under the band and intensity hypotheses of Theorem 5.3, all its conclusions hold for this ensemble as well. In particular $r = \log_2 N + O(1)$ and $B = 2 \log_2 N + O(1)$ give dictionaries of order N specified by $r(B+1) = O((\log N)^2)$ bits.

Proof. Each codimension- r kernel has the same number of ordered row bases, and uniform b_0 chooses uniformly among its cosets. A fixed word is included with probability $a = 2^{-r}$. For distinct u, v , invariance under invertible linear maps gives

$$\mathbb{P}(A_0(u - v) = 0) = \frac{2^{B-r} - 1}{2^B - 1}, \quad \mathbb{P}(u, v \in \mathcal{W}) = 2^{-r} \frac{2^{B-r} - 1}{2^B - 1} = \frac{m(m-1)}{Q(Q-1)}.$$

These are precisely the one- and two-point probabilities used in (5.14); no higher inclusion probability was used. They transfer the nonnegative deletion, pair and overlap costs, not an identity between the two ensembles' average total-variation distances. The dictionary remains fixed while bounding each distance; only its bound is then averaged over the actual matrix and offset. The cases $r = 0$ and $r = B$ obey the same formulas. $\square$

The corollary supplies a short random description and a high-probability statement, not a deterministic construction certified at every scale. The proof uses the one- and two-point

sampling identities of the affine ensemble. A short matrix description by itself does not certify the Poisson approximation for a chosen dictionary or control its overlap weight.

The following exact identity controls local overlaps alone. It is useful for the uniform theorem, but does not replace averaging the full distance as in Theorem 5.3.

Corollary 5.5 (Most dictionaries have small overlap). *Let $B \geq 1$ and $1 \leq m \leq 2^B$. Choose an m -element subset $\mathcal{W}$ uniformly among the subsets of $\{\pm 1\}^B$. This auxiliary choice is independent of f . Then*

$$\mathbb{E}_{\mathcal{W}} \Omega(\mathcal{W}) = \frac{m(B-1)}{2^B}. \quad (5.15)$$

Under the first two conditions in (5.3), the fraction of dictionaries with $\Omega(\mathcal{W}) > N^{-1/2}$ is $O_{C,\eta}(\log N/\sqrt{N})$. Every dictionary outside this exceptional set has the field approximation (5.4).

Proof. Put $Q = 2^B$. At each proper displacement $1 \leq d < B$, there are 2^d self-compatible words and 2^{B+d} ordered compatible pairs. The number of distinct compatible pairs is therefore $2^d(Q-1)$. A word belongs to the dictionary with probability m/Q ; a distinct ordered pair belongs with probability $m(m-1)/(Q(Q-1))$. After multiplication by $2^{-d}/m$, the expected contribution of this displacement is exactly

$$\frac{2^{-d}}{m} \left(2^d \frac{m}{Q} + 2^d(Q-1) \frac{m(m-1)}{Q(Q-1)} \right) = \frac{m}{Q}.$$

Sum over the $B-1$ displacements to get (5.15). This also covers $m=1$; for $B=1$ the sum is empty. At criticality $m2^{-B} = \Lambda/N = O_C(1/N)$ and $B = O_\eta(\log N)$. Markov's inequality proves the exceptional-fraction estimate. The result is a statement about the size of a class of deterministic dictionaries; no additional randomness is required in Theorem 5.1. $\square$

Corollary 5.6 (Explicit dictionaries with no overlaps). *For all sufficiently large B , there is an explicitly described family $\mathcal{C}_B \subset \{0,1\}^B$ of size at least $2^B/(32B)$ with no proper suffix of any word equal to a prefix of any word, including itself. Every subset of $\mathcal{C}_B$, after identifying 0,1 with $-1,+1$, has $\Omega = 0$. In particular deterministic examples exist throughout the range (5.3).*

Proof. This is the binary marker construction of Chee, Kiah, Purkayastha and Wang [12, Section III]; we include the short argument and the coarse size bound needed here. Set $k = \lceil \log_2(2B) \rceil$ and take all words

$$0^k 1 u 1, \quad |u| = B - k - 2, \quad u \text{ contains no } 0^k.$$

The only occurrence of 0^k in such a word starts at its first position: the first following letter and the last letter are both 1. An overlap of length at most k is impossible because the prefix is all zeros and the suffix ends in 1. A longer proper overlap would put 0^k at a later position, also impossible. There are 2^{B-k-2} choices before the restriction on u . A union bound over at most B positions excludes at most the fraction $B2^{-k} \leq 1/2$. Thus $|\mathcal{C}_B| \geq 2^{B-k-3} \geq 2^B/(32B)$. At the critical dictionary length, $m/2^B = O(1/N)$ and $B = O(\log N)$, so any required m words can be selected from this family for large N . $\square$

For example, one may select $\lfloor N^{2/5} \rfloor$ such words at length $B = (7/5) \log_2 N + O(1)$. Their full labelled field has a Poisson approximation. The coding construction is classical; the conclusion for the multiplicative sequence is the application of the theorem above.

5.3 Words up to a common sign

For $a = (a_0, \dots, a_L) \in \{\pm 1\}^{L+1}$ put $I_x(a) = I_{x,a} + I_{x,-a}$. Define

$$\mathcal{O}(a) = \{1 \leq d \leq L : \exists \eta \in \{\pm 1\}, a_{d+j} = \eta a_j \ (0 \leq j \leq L-d)\}, \quad \Theta(a) = \sum_{d \in \mathcal{O}(a)} 2^{-d}. \quad (5.16)$$

For the two-word dictionary $\{a, -a\}$ one has $\Omega(\{a, -a\}) = \Theta(a)$: every compatible shift has two directed sign choices.

Corollary 5.7 (Poisson law for low-overlap patterns). *Fix $C > 0$ and $\varepsilon > 0$. Uniformly in $a = a_N$ and $|L - \log_2 N| \leq C$,*

$$d_{\text{TV}} \left(\mathcal{L} \left(\sum_{x \in I_N} I_x(a) \right), \text{Pois}(N2^{-L}) \right) \ll_{C,\varepsilon} \Theta(a) + e^{-V_N + o_C(\nu_N)} + N^{-1/3+\varepsilon}. \quad (5.17)$$

The same bound holds under any deterministic spatial mask, with target mean $|A_N|2^{-L}$.

Proof. Apply Theorem 5.1 with $B = L+1$, $m = 2$, and merge the two word coordinates. Their events are disjoint and their target means add. $\square$

If the least compatible shift $d_*(a_N)$ tends to infinity, then $\Theta(a_N) \leq 2^{1-d_*(a_N)} \rightarrow 0$. For the run-start word $(-1, +1, \dots, +1)$ only $d = L$ is compatible. Periodic words can instead retain a short overlap and need declustering. Notice that the preceding dictionary theorem also separates a and $-a$: their asymptotic independence is not assumed in this corollary.

The uniform and typical dictionary comparisons control the rare occurrences of each prescribed word. We return to constant runs and keep their attributes together: a start, its full length, its sign and all thresholds it exceeds.

5.4 Exact marks without a growing-mark loss

Almost surely every rightward constant run is finite. Indeed constancy on $[x, \infty)$ would make all prime signs beyond x equal, an event of probability zero; take a countable union over x . Thus

$$J_{x,L} = \sum_{e \geq 0} K_{x,e}, \quad K_{x,e} = \mathbf{1}_{\{f(x-1)=-f(x), f(x)=\dots=f(x+L+e-1), f(x+L+e)=-f(x)\}}. \quad (5.18)$$

This system has $L + e + 1$ independent homogeneous rows on a good support and conditional marginal $p_e = 2^{-L-e-1}$. Its homogeneous rows are nested in those for a larger mark, although the terminal affine right-hand side changes. Marks at the same start are mutually exclusive. We may retain the sign as well:

$$K_{x,e}^s = K_{x,e} \mathbf{1}_{\{f(x)=s\}}, \quad p_{e,s} = 2^{-L-e-2}, \quad s \in \{\pm 1\}. \quad (5.19)$$

On a good support this event prescribes every value, so its conditional probability is $p_{e,s}$. The exact identities are $\sum_s K_{x,e}^s = K_{x,e}$ and $\sum_{e,s} K_{x,e}^s = J_{x,L}$.

Theorem 5.8 (Finite exact-mark comparison). *Let $E \geq 0$, $Q = L + E + 1$, and assume $\beta_- \log N \leq L+1 \leq Q+1 \leq \beta_+ \log N$. For $Y > 2(Q+1)$ let $D_Y(Q)$ be the starts whose support $\{x-1, \dots, x+Q-1\}$ contains a Y -defective integer, and let $E_Y(Q)$ be the ordered graph-edge*

count for these maximal supports. Then

$$\mathbb{E} d_{\text{TV}} \left(\mathcal{L}((K_{x,e})_{x \in I_N, 0 \leq e \leq E} \mid \mathcal{F}_Y), \bigotimes_{x \in I_N, 0 \leq e \leq E} \text{Pois}(p_e) \right) \ll p \mathcal{M}_B + p |D_Y(Q)| + p^2 (NQ + E_Y(Q) + R_2(N, L)), \quad p = 2^{-L}. \quad (5.20)$$

Exactly the same upper bound holds for the sign-resolved family $(K_{x,e}^s)_{x \in I_N, 0 \leq e \leq E, s \in \{\pm 1\}}$, compared with independent Poisson variables of means $p_{e,s}$. There is no polynomial factor in the number of marks. At $Y = Y_N^\sharp$, the right side is at most

$$C_\varepsilon \lambda_N (1 + \lambda_N) \left(e^{-V_N + o(\nu_N)} + N^{-1/3+\varepsilon} \right). \quad (5.21)$$

Proof. On good maximal supports all marginals are p_e . Apply the process Poisson theorem to the marked family, allowing all marks on neighbouring supports to be neighbours, and all marks at the same site to be neighbours. The two identities

$$\sum_{e=0}^E p_e \leq p, \quad \sum_{e,f=0}^E \mathbb{P}(K_{x,e} K_{y,f} = 1) \leq \mathbb{P}(J_{x,L} J_{y,L} = 1) \quad (5.22)$$

remove any double sum over marks before estimating the pair mass. Local pairs cost $O(NQp^2)$ and separated pairs cost $p^2(E_Y(Q) + R_2(N, L))$. Bad actual marks at one site have total mass at most that of $J_{x,L}$, so their deletion is bounded by $p \mathcal{M}_B + p |D_Y(Q)|$, regardless of which late vertex caused removal. This proves (5.20). The bounds $|D_Y(Q)| \leq (Q+1)A(3N, Y)$ and (4.2) with B replaced by $Q+1$, together with Theorem 3.1 and proposition 4.2, give (5.21).

For the signed version, the same-site alternatives are still disjoint and both sums in (5.22) remain valid after summing signs. It remains to check local pairs before that summation. Put $a = L + e$, $b = L + f$, and $d = y - x > 0$. If $d < a$ the pair is impossible. If $d = a$, the two prescribed value strings share two vertices, require the second run sign to be the opposite of the first, and have probability $4p_{e,s}p_{f,t}$ when compatible. If $d = a + 1$, they share one vertex, require equal run signs, and have probability $2p_{e,s}p_{f,t}$. For $d > a + 1$ their actual supports are disjoint, and a good local union has probability $p_{e,s}p_{f,t}$. These assertions follow by counting the prescribed values in a union whose private coordinates make them uniform. Thus the local sum is still $O(NQp^2)$ and all other terms are unchanged. $\square$

The use of (5.22) is essential. An exact mark $e < E$ does not imply a threshold event at length $L + E + 1$; a bad vertex can occur beyond its support. We bound deletion by the base event, not the longest event.

The exact tail identity is

$$\{\exists x \in I_N, e > E : K_{x,e} = 1\} = \{Z_{N,L+E+1} > 0\}. \quad (5.23)$$

The first-moment estimate bounds this probability by $\lambda_N 2^{-(E+1)}(1 + N^{-1/2+o(1)})$ whenever the enlarged length is in a fixed logarithmic band. The independent target tail has mass $\lambda_N 2^{-(E+1)}$. Couple both tails to zero. At criticality choose $E = \lceil V_N / \log 2 \rceil$. This is $o(\log N)$; (5.21) and (5.23) prove the full-field estimate, including signs, in Theorem 1.1. Riemann sums and geometric tail tightness give its weak diffuse limit.

Corollary 5.9 (An explicit compound-Poisson cluster law). *Let $C_{N,L}$ count all constant length- L windows whose left endpoint lies in I_N , without requiring left maximality. At criticality its*

law is at distance $o(1)$ from that of

$$\sum_{i=1}^{P_N} H_i, \quad P_N \sim \text{Pois}(\lambda_N), \quad \mathbb{P}(H_i = h) = 2^{-h}, \quad h \geq 1,$$

with independent H_i and P_N . The target probability generating function is $\exp[\lambda_N(z/(2-z)-1)]$ for $0 \leq z \leq 1$.

Proof. A maximal run of length $L+e$ contributes $e+1$ windows. Thus, apart from runs crossing an endpoint of I_N , $C_{N,L} = \sum_{x \in I_N, e \geq 0} (e+1)K_{x,e}$. A discrepancy implies a constant length- L window starting at N or at $2N-1$. Each such event has $L-1$ relative constraints; private pivots and the pointwise defect bound give probability at most $2^{-(L-1)+O(\log N/\log \log N)} = N^{-1+o(1)}$. The full marked-field comparison contracts under the measurable weighted sum. Its target is the stated compound Poisson law, since $H = e+1$. No moment convergence is inferred from this contraction. $\square$

5.5 Growing depth: distinguish positions from counts

Put $b_N = \lfloor \log_2 N \rfloor$, $\vartheta_N = \log_2 N - b_N$. For $d_N \in \mathbb{N}_0$ with $d_N = o(\log N)$, set

$$L_N = b_N - d_N, \quad \Lambda_N = N2^{-L_N} = 2^{\vartheta_N + d_N}.$$

Let $K_{x,q}^{\text{ex}}$ denote the exact-length event for a left-maximal run of length q (zero for $q < 1$). The moving field and its target are

$$\mathcal{E}_{N,d_N} = \sum_{x \in I_N} \sum_{r \geq -d_N} K_{x,b_N+r}^{\text{ex}} \delta_{(x/N,r)}, \quad \mathbb{E} \Pi_{N,d_N} \{(x/N,r)\} = 2^{-b_N-r-1}, \quad (5.24)$$

where the target atoms are independent Poisson variables. Their total mean is Λ_N . Define exact-level and threshold counts by

$$X_{N,r} = \sum_{x \in I_N} K_{x,b_N+r}^{\text{ex}}, \quad T_{N,m} = \sum_{r \geq m} X_{N,r} = Z_{N,b_N+m}.$$

For the target write $P_{N,r} = \Pi_{N,d_N}([1,2] \times \{r\})$ and $S_{N,m} = \sum_{r \geq m} P_{N,r}$. Thus $P_{N,r} \sim \text{Pois}(2^{\vartheta_N-r-1})$ independently.

The preceding estimates bound the mean, over the small-prime environment, of a conditional total-variation distance. Restricting this average to an event $C_N \in \mathcal{F}_Y$ multiplies the error by at most $1/\mathbb{P}(C_N)$; this is the stable-kernel principle of Lemma 6.1. It is the first information cost below.

Theorem 5.10 (Two levels of joint approximation). *Fix $c > 0$ and $0 < \varepsilon < 1/3$. Let C_N be an arithmetic conditioning event of positive probability and put $I_N^{\text{info}} = -\log \mathbb{P}(C_N)$.*

(i) *Labelled field. If $C_N \in \mathcal{F}_{Y_N^\#}$, a sufficient condition for full-field total variation to tend to zero is*

$$I_N^{\text{info}} + \log \Lambda_N + \log(1 + \Lambda_N) \leq V_N - c\nu_N. \quad (5.25)$$

At a movable cutoff, if $C_N \in \mathcal{F}_{\widehat{Y}_N}$, the sufficient condition improves to

$$I_N^{\text{info}} + \log \Lambda_N \leq \widehat{V}_N/2 - c\widehat{\nu}_N. \quad (5.26)$$

In the nontrivial range, the latter comparison has upper bound

$$e^{I_N^{\text{info}} + \log \Lambda_N - \widehat{V}_N/2 + o(\widehat{\nu}_N)} + e^{I_N^{\text{info}} + 2 \log \Lambda_N - \widehat{V}_N + o(\widehat{\nu}_N)} + e^{I_N^{\text{info}} \Lambda_N (1 + \Lambda_N) N^{-1/3 + \varepsilon + o(1)}}. \quad (5.27)$$

(ii) Aggregated exact levels. If $C_N \in \mathcal{F}_{Y_N^\#}$ and

$$I_N^{\text{info}} + \log \Lambda_N \leq V_N - c\nu_N, \quad (5.28)$$

then, for every $0 < c' < c$ and $0 < \varepsilon < 1/3$,

$$d_{\text{TV}} \left(\mathcal{L}((X_{N,r})_{r \geq -d_N} \mid C_N), \bigotimes_{r \geq -d_N} \text{Pois}(2^{\vartheta_N - r - 1}) \right) \ll e^{-c'\nu_N} + N^{-1/3 + \varepsilon}. \quad (5.29)$$

Exactly the same distance holds for the full threshold paths $(T_{N,m})_{m \geq -d_N}$ and $(S_{N,m})_{m \geq -d_N}$.

Proof. Conditioning multiplies any averaged kernel error by at most $e^{I_N^{\text{info}}}$. For (i), apply Theorem 5.8 with base L_N and then remove the mark tail using (5.23). For the hard bound one may take $E = \lceil (I_N^{\text{info}} + V_N + \log(2 + \Lambda_N)) / \log 2 \rceil$. Under (5.25) this is $o(\log N)$ and both tail costs are absorbed by the displayed error. At a free cutoff $Y = e^w$, the leading process errors are

$$e^{I_N^{\text{info}} \Lambda_N} e^{-D(H_N/w) + o(H_N/w)}, \quad e^{I_N^{\text{info}} \Lambda_N^2} e^{-w + o(H_N/w)}.$$

Choose $w = \widehat{V}_N$, so $D(H_N/w) = w/2$, and take an equally large $o(\log N)$ tail cutoff. This proves (5.27) and (5.26). The hard choice remains sharper at bounded intensity; the movable choice improves the admissible growing intensity.

For (ii), aggregate sites *before* solving the multivariate Poisson Stein equation. If $q_e = 2^{-(e+1)}$ are the normalized category masses, Barbour's Lemmas 2–3 [7, p. 179], as applied in the companion, bound the directional second-difference factor by

$$C \frac{1 + \log^+(2\Lambda_N)}{\Lambda_N \sqrt{q_e q_f}}.$$

The mark sum remains uniform because $\sum_{e \geq 0} \sqrt{q_e} = 1 + \sqrt{2}$. The full argument and the precise finite-cutoff bound are given in Theorem C.1. It yields a leading error $e^{I_N^{\text{info}} \Lambda_N (1 + \log^+(2\Lambda_N))} e^{-V_N + o(\nu_N)}$ and a polynomial remainder after the enlarged-length arithmetic estimate. All subpolynomial factors in that remainder are absorbed by choosing a smaller initial ε . Choose a constant strictly between c' and c for the tail cutoff. The scalar relative estimate at the shifted length bounds both tails by $O(e^{-c'\nu_N})$; see Theorem C.1. Finally, on finite-mass sequences, the map $X_r \mapsto \sum_{s \geq r} X_s$ is a measurable bijection with inverse $X_r = T_r - T_{r+1}$. Total variation is preserved, not merely contracted. $\square$

Part (ii) retains every statistic of the multiset of lengths, but not the positions. A maximal coupling identifies all those retained statistics at once, with no union bound over thresholds. A field coupling from (i) also retains positions. The ranges are sufficient bounds for these methods, not phase transitions. In particular the full one-factor labelled range is not proved by these comparisons.

The fixed enlarged-cutoff bound remains useful for events exposed up to $\widehat{Y}_N$. When their information can be exposed at a smaller admissible cutoff, proposition 6.2 optimizes the two labelled costs as a function of that information.

Corollary 5.11 (Sign-resolved growing fields). *Both conclusions of Theorem 5.10 remain valid after retaining the sign of each run. The labelled target has atom means 2^{-b_N-r-2} for each $s \in \{\pm 1\}$, and the aggregated target has independent means 2^{ϑ_N-r-2} at each pair (r, s) .*

Proof. For the labelled result use the signed version of Theorem 5.8 and the unchanged tail identity. For aggregation, the normalized category weights are $q_{e,s} = 2^{-e-2}$ and $\sum_{e,s} \sqrt{q_{e,s}} = \sqrt{2}(1 + \sqrt{2})$. The directional Stein bound therefore has the same summability. At the largest retained support, the full value-relation space dominates each mixed signed kernel by extension of coefficient vectors by zero; proposition 3.24 supplies the same two-window envelope. The retained support has $L + E + 2$ vertices, differing by only one from the relative-mark convention. Since $E = O(V_N) = o(\log N)$, its polynomial remainder is absorbed exactly as in part (ii). The signed local-pair calculation and the same tail bound finish the proof. $\square$

5.6 Target kernels, extremes and the Gaussian lower edge

The independent target levels give, exactly,

$$S_{N,m} \sim \text{Pois}(2^{\vartheta_N-m}), \quad (5.30)$$

$$S_{N,m+j} \mid S_{N,m} = n \sim \text{Bin}(n, 2^{-j}), \quad (5.31)$$

$$S_{N,m-j} \mid S_{N,m} = n \stackrel{d}{=} n + \text{Pois}(2^{\vartheta_N-m}(2^j - 1)), \quad (5.32)$$

when the lower levels are retained. The whole forward path is the number of survivors among n independent geometric lifetimes, each with $\mathbb{P}(E = e) = 2^{-(e+1)}$. These are identities for the Poisson target, not exact finite- N Markov assertions for the multiplicative sequence.

For any integer $m \geq 0$, the dyadic excess maximum $M_{N,L} = \max\{\ell_x - L : x \in I_N, J_{x,L} = 1\}$, with empty value $-\infty$, satisfies the exact identity

$$\{M_{N,L} \leq m\} = \{Z_{N,L+m+1} = 0\}. \quad (5.33)$$

Thus the scalar estimate at $L' = L + m + 1$, with $\mu = N2^{-L'}$, gives

$$\mathbb{P}(M_{N,L} \leq m) = e^{-\mu} + O\left(\min\{1, \mu(e^{-V_N+o(\nu_N)} + N^{-1/3+\varepsilon})\}\right)$$

whenever $L' + 1$ is in a fixed logarithmic band. In particular one-hit upper tails are relative when $N2^{-L'} \rightarrow 0$. For the k th largest run length among starts in I_N , along a subsequence $\vartheta_N \rightarrow \vartheta$,

$$\mathbb{P}(R_N^{[k]} - b_N < m) \longrightarrow e^{-2^{\vartheta-m}} \sum_{j=0}^{k-1} \frac{(2^{\vartheta-m})^j}{j!}. \quad (5.34)$$

This is just the event that fewer than k starts exceed the threshold.

Theorem 5.12 (Joint Poisson–Gaussian limit). *Assume $d_N \rightarrow \infty$ and the hypotheses of part (ii) of Theorem 5.10. Put*

$$G_N(j) = \frac{Z_{N,b_N-d_N+j} - \Lambda_N 2^{-j}}{\sqrt{\Lambda_N}}.$$

Along $\vartheta_N \rightarrow \vartheta$, for every fixed J and every finite $R \subset \mathbb{Z}$, conditionally on C_N ,

$$((G_N(j))_{0 \leq j \leq J}, (X_{N,r})_{r \in R}) \implies ((G(j))_{0 \leq j \leq J}, (P_r^{(\vartheta)})_{r \in R}),$$

where $\mathbb{E}G(j)G(k) = 2^{-\max(j,k)}$, the variables $P_r^{(\vartheta)} \sim \text{Pois}(2^{\vartheta-r-1})$ are independent, and the Gaussian and Poisson vectors are independent. The normalized Gaussian sequence $H(j) =$

$2^{j/2}G(j)$ is the stationary $AR(1)$ chain $H(j+1) = 2^{-1/2}H(j) + 2^{-1/2}\zeta_{j+1}$, where $(\zeta_j)_{j \geq 1}$ are independent standard normal random variables, independent of $H(0) \sim N(0, 1)$.

Proof. Transfer to the independent target by (5.29). Choose a fixed $r_0 < \min R$ (any fixed r_0 if R is empty). The lower-edge sums restricted to $r < r_0$ are independent of the displayed critical levels and satisfy the multivariate Poisson central limit theorem. Their covariance after normalization tends to $2^{-\max(j,k)}$. The centered sum over $r \geq r_0$ has bounded variance and hence vanishes after division by $\sqrt{\Lambda_N}$. This proves the joint limit and independence. The covariance of H is $2^{-|j-k|/2}$, giving the stated Gaussian autoregression. $\square$

For fixed compact sets of levels, the untruncated centered field has the phase-indexed target

$$\text{PPP} \left(dt \otimes \sum_{r \in \mathbb{Z}} 2^{\vartheta_N - r - 1} \delta_r \right).$$

Riemann sums give the corresponding Laplace approximation without assuming that ϑ_N converges. A phase-convergent subsequence turns it into a fixed weak limit; its intensity is locally finite although it diverges towards $r = -\infty$. Full formulas and conditional configuration statements are in section D.4. No continuous run lengths or canonical coupling simultaneously over all N is asserted.

6 Arithmetic information and resolved conditional laws

There are two different conditioning operations. An arithmetic event fixes information about prime signs before the Poisson comparison. Conditioning on a count or another observable then divides by the probability of that observable. Keeping the two costs separate avoids treating a rare state as though an absolute error resolved it automatically.

The first table records the sufficient budgets obtained directly by the scalar and dependency-graph methods; it is not a list of the strongest available hard-field conclusions. Here $I = -\log \mathbb{P}(C)$, $c > 0$ is fixed, $Y^\# = \lfloor e^V \rfloor$ and $\hat{Y} = \lfloor e^{\hat{V}} \rfloor$; ν and $\hat{\nu}$ belong to those cutoffs. The following rows concern the dyadic domain $I_N = [N, 2N) \cap \mathbb{Z}$ and the scalar or dependency-graph comparisons of Sections 4–6. Their moving-field intensities Λ are at least one.

Observable	Sufficient budget	Prime field
Scalar, hard	$I + \log^+ \lambda \leq V - c\nu$	$\mathcal{F}_{Y^\#}$
Scalar, soft	$2I + \log^+ \lambda \leq \hat{V} - c\hat{\nu}$	$\mathcal{F}_{\hat{Y}}$
All exact-level counts	$I + \log \Lambda \leq V - c\nu$	$\mathcal{F}_{Y^\#}$
All sites, graph / hard	$I + \log \Lambda + \log(1 + \Lambda) \leq V - c\nu$	$\mathcal{F}_{Y^\#}$
All sites, graph / enlarged	$I + \log \Lambda \leq \hat{V}/2 - c\hat{\nu}$	$\mathcal{F}_{\hat{Y}}$
All sites, graph / adapted	$\log \Lambda \leq \ell_{\text{lab},N}(I) - c\nu_N(I)$	$\mathcal{F}_{Y_N(I)}$

The last row uses (6.4)–(6.5), with $0 \leq I \leq V_N$. Exact marks and signs can be retained in all site-labelled rows. The inequalities have fixed negative margins; equality at the boundary and optimality for the model are not asserted. A further conditioning on an observed count has the separate probability cost analysed below.

The one-factor hard-field conclusion. The selected-pivot theorem improves the available site-labelled range both on prefixes and, by restriction, on dyadic blocks:

Observable and domain	Sufficient budget	Prime field
All prefix sites, exact excesses and signs	$I_M + \log \Lambda \leq V_M - c\nu_M$	$\mathcal{F}_{Y_M^\#}$
All dyadic sites, exact excesses and signs	$I_N^{\text{info}} + \log \Lambda_N \leq V_N - c\nu_N$	$\mathcal{F}_{Y_N^\#}$

The first row is Theorem 7.7, with $\Lambda = (M - L)2^{-L} \rightarrow \infty$ and $L = \lfloor \log_2 M \rfloor - a_M$, $0 \leq a_M = o(\log M)$. The second is Corollary 7.7a, with $\Lambda_N = N2^{-L_N} \rightarrow \infty$ and $L_N = \lfloor \log_2 N \rfloor - d_N$, $0 \leq d_N = o(\log N)$. Both use the exact means 2^{-L-e-2} on their respective integer index sets. The dyadic conclusion is obtained by restricting a prefix of size $4N$ without changing L_N or the integer start labels. Thus, in this growing-intensity corridor, keeping the sites does not require the extra intensity factor of the direct hard graph bound. This does not enlarge the prime field. The direct routes remain useful for their critical or relative rates and for events measurable only at a larger or adapted cutoff; no dictionary extension is implicit.

6.1 Stable kernels and information cost

Lemma 6.1 (Stable product lift). *Let W take values in a standard Borel space, let $\mathcal{F}$ be a sigma-field, and let ν be a deterministic target law. If*

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(W \mid \mathcal{F}), \nu) \leq \varepsilon,$$

then for every $\mathcal{F}$ -measurable random element V ,

$$d_{\text{TV}}(\mathcal{L}(V, W), \mathcal{L}(V) \otimes \nu) \leq \varepsilon.$$

For $C \in \mathcal{F}$ with $\mathbb{P}(C) > 0$, the conditional bound is at most $\varepsilon/\mathbb{P}(C)$, both for W and for the joint product lift with V .

Proof. Disintegrate the actual and product laws over $\mathcal{F}$. For a test set, the difference of its conditional sections is bounded by the conditional total-variation distance. Integrate, or integrate over C and divide by $\mathbb{P}(C)$. Countable configuration spaces used here have measurable kernels; the standard-Borel formulation permits the additional recorded variables. $\square$

For $C_N \in \mathcal{F}_{Y_N^\#}$ write $I_N^{\text{info}} = -\log \mathbb{P}(C_N)$. The hard scalar estimate becomes

$$d_{\text{TV}}(\mathcal{L}(Z_{N,L} \mid C_N), \text{Pois}(\lambda_N)) \ll \min\{1, e^{I_N^{\text{info}}} \lambda_N (e^{-V_N + o(\nu_N)} + N^{-1/3+\varepsilon})\}. \quad (6.1)$$

For $C_N \in \mathcal{F}_{\widehat{Y}_N}$ the soft proof gives instead

$$d_{\text{TV}}(\mathcal{L}(Z_{N,L} \mid C_N), \text{Pois}(\lambda_N)) \ll \min\{1, e^{I_N^{\text{info}} - (\widehat{V}_N - \log^+ \lambda_N)/2 + o(\widehat{\nu}_N)} + e^{I_N^{\text{info}}} N^{-1/3+\varepsilon+o(1)}\}. \quad (6.2)$$

Thus the sufficient scalar budgets are

$$I_N^{\text{info}} + \log^+ \lambda_N \leq V_N - c\nu_N, \quad 2I_N^{\text{info}} + \log^+ \lambda_N \leq \widehat{V}_N - c\widehat{\nu}_N.$$

The hard route can be preferable at a larger information cost, but it requires measurability with respect to its smaller prime field. These budgets measure probabilities of events, not the number of primes exposed. For instance, specifying k independent prime signs costs $k \log 2$; a rank- k compatible affine cylinder has the same cost.

At small intensity the hard estimate is relative: under $I_N^{\text{info}} \leq V_N - c\nu_N$ and $\lambda_N \rightarrow 0$,

$$\mathbb{P}(Z_{N,L} > 0 \mid C_N) = \lambda_N(1 + o(1)), \quad \mathbb{P}(Z_{N,L} = 1 \mid C_N) = \lambda_N(1 + o(1)). \quad (6.3)$$

The polynomial terms are negligible since $I_N^{\text{info}} = o(\log N)$. At large intensity, either valid scalar budget gives a conditional Gaussian limit after centering by λ_N and dividing by $\sqrt{\lambda_N}$. The target Poisson normal-approximation error must still be added to the arithmetic error; the quantitative and relative forms are in section D.

6.2 Choosing the labelled cutoff after the information cost

In the dependency-graph proof, retaining all sites makes two errors compete. Increasing the prime cutoff removes shared-prime edges, but deletes more supports. After conditioning on an event of cost I , the leading admissible logarithmic intensity is

$$g_{N,I}(w) = \min\{D(H_N/w) - I, (w - I)/2\}.$$

The information cost enters the two constraints with different slopes. For $0 \leq I \leq V_N$, define

$$2D(H_N/w_N(I)) = w_N(I) + I, \quad V_N \leq w_N(I) \leq \hat{V}_N, \quad (6.4)$$

and put

$$\nu_N(I) = \frac{H_N}{w_N(I)}, \quad Y_N(I) = \lfloor e^{w_N(I)} \rfloor, \quad \ell_{\text{lab},N}(I) = \frac{w_N(I) - I}{2}. \quad (6.5)$$

The root is unique: $2D(H_N/w) - w - I$ is strictly decreasing, with endpoint values $V_N - I$ and $-I$. It is the crossing of a decreasing and an increasing constraint, so $\ell_{\text{lab},N}(I)$ is the maximum of $g_{N,I}$ on this cutoff interval.

Proposition 6.2 (Information-adapted labelled comparison). *Let C_N be events of positive probability, put $I_N^{\text{info}} = -\log \mathbb{P}(C_N)$, and assume $0 \leq I_N^{\text{info}} \leq V_N$. For integer $L_N \geq 1$ put $\Lambda_N = N2^{-L_N} \geq 1$. Suppose, for fixed $c > 0$,*

$$C_N \in \mathcal{F}_{Y_N(I_N^{\text{info}})}, \quad \log \Lambda_N \leq \ell_{\text{lab},N}(I_N^{\text{info}}) - c\nu_N(I_N^{\text{info}}). \quad (6.6)$$

For every $0 < c' < c$ and $0 < \varepsilon < 1/3$,

$$d_{\text{TV}} \left(\mathcal{L}((K_{x,e}^s)_{x \in I_N, e \geq 0, s \in \{\pm 1\}} \mid C_N), \bigotimes_{x,e,s} \text{Pois}(2^{-L_N - e - 2}) \right) \ll_{c,c',\varepsilon} e^{-c'\nu_N(I_N^{\text{info}})} + N^{-1/3+\varepsilon}. \quad (6.7)$$

All site labels, signs and exact excesses are retained. The estimate also holds after any deterministic restriction or measurable contraction of this field.

Proof. Write $w = w_N(I_N^{\text{info}})$, $\nu = H_N/w$ and $\ell = \log \Lambda_N$. The finite signed comparison at this cutoff has leading errors $\exp(I_N^{\text{info}} + \ell - D(H_N/w) + o(\nu))$ and $\exp(I_N^{\text{info}} + 2\ell - w + o(\nu))$. By (6.4) and (6.6), their exponents are at most $-c\nu + o(\nu)$ and $-2c\nu + o(\nu)$. Take $E = 3\lceil \hat{V}_N / \log 2 \rceil$. Since $I_N^{\text{info}} + \ell \leq w \leq \hat{V}_N$, the enlarged support is $\log_2 N + O(V_N)$, uniformly in the parameters, and the real conditional and target tails are $O(e^{-2\hat{V}_N})$. This larger auxiliary truncation disappears from the final field. The polynomial remainder is at most $e^{I_N^{\text{info}}} \Lambda_N (1 + \Lambda_N) N^{-1/3+\varepsilon/2+o(1)}$. Since $I_N^{\text{info}} + 2\ell \leq w = O(V_N) = o(\log N)$, it is absorbed by the stated remainder. The free-cutoff estimates are uniform on $[V_N, \hat{V}_N]$; the detailed calculation, including the support and tail checks, is in section C.5. $\square$

At zero information cost, $w_N(0) = \widehat{V}_N$ and $\ell_{\text{lab},N}(0) = \widehat{V}_N/2$, recovering the enlarged-cutoff labelled range. For $0 < I < V_N$, the allowed leading value is strictly larger than $\widehat{V}_N/2 - I$. This is a comparison of proof budgets for events measurable at the required cutoff, not a transfer of events from a larger sigma-field into a smaller one. If $I_N^{\text{info}}/V_N \rightarrow \theta \in [0, 1]$, then

$$\frac{\ell_{\text{lab},N}(I_N^{\text{info}})}{V_N} \rightarrow \frac{\sqrt{\theta^2 + 8} - 3\theta}{4}. \quad (6.8)$$

The exact saddle, rather than this first-order limit, is used in (6.6) to retain its $\nu_N(I_N^{\text{info}})$ margin.

Admissible information matters. For a fixed event known only to belong to $\mathcal{F}_{\lfloor e^{w_0} \rfloor}$, with $w_0 \in [V_N, \widehat{V}_N]$, optimize over $w \geq w_0$, not over all cutoffs. The best leading budget on this interval is attained at $w = \max\{w_0, w_N(I)\}$. Since $w \geq w_0$, the original event belongs to $\mathcal{F}_{\lfloor e^w \rfloor}$ by upward inclusion of the prime sigma-fields; no new measurability assumption is required. The sufficient condition there is $\log \Lambda_N \leq g_{N,I}(w) - cH_N/w$ with that event measurable in $\mathcal{F}_{\lfloor e^w \rfloor}$. The proof is the same finite comparison. These optimized upper bounds do not identify an intrinsic phase boundary of the multiplicative model, nor do they use the scalar soft-exception estimate as a field theorem.

The exact prefix comparison of Theorem 7.7, and its dyadic restriction Corollary 7.7a, use the hard field. Deep starts and signed tails are checked independently in Lemma F.1. Its partition readouts require no mesh hypothesis. Process limits still require their own target, ordering and censoring estimates. Neither that contraction nor this information-adapted graph calculation silently changes the admissible prime field of the other route.

6.3 The exact resolution cost

Lemma 6.3 (Sharp conditioning inequality). *Let μ, ν be probability laws on $E \times F$ with $d_{\text{TV}}(\mu, \nu) \leq \varepsilon$. Let $B \subset E$ have $p = \nu_E(B) > \varepsilon$ and write $q = \mu_E(B)$. Then $q > 0$ and*

$$d_{\text{TV}}(\mu_F(\cdot | B), \nu_F(\cdot | B)) \leq \frac{\varepsilon}{\max(p, q)} \leq \frac{\varepsilon}{p}. \quad (6.9)$$

Proof. The bound $|p - q| \leq \varepsilon$ gives positivity. First suppose $q \leq p$. For any measurable $D \subset F$ put $A = B \times D$. Since $\mu(A) \leq q$,

$$\frac{\mu(A)}{q} - \frac{\nu(A)}{p} \leq \frac{\mu(A) - \nu(A) + p - q}{p} = \frac{\nu(B \times D^c) - \mu(B \times D^c)}{p} \leq \frac{\varepsilon}{p}.$$

The supremum of the positive differences equals total variation for probability laws. If $q \geq p$, interchange the two laws. $\square$

The denominator expresses the correct scale: convergence after conditioning requires $\varepsilon/p \rightarrow 0$, not merely $p > \varepsilon$. This lemma applies to an entire future path at once.

Let $\lambda_{N,m} = 2^{\theta_{N-m}}$ and $p_{N,m}(n) = e^{-\lambda_{N,m}} \lambda_{N,m}^n / n!$. Write $\varepsilon_{N,m,C}^{\text{path}}$ for the total-variation error between the actual complete future threshold path, under C_N , and the target staircase based at m . The aggregated proof works equally for an arbitrary base threshold in a fixed logarithmic band (including $\lambda_{N,m} < 1$). In its nontrivial information range it gives

$$\varepsilon_{N,m,C}^{\text{path}} \ll e^{I_N^{\text{info}}} \lambda_{N,m} (1 + \log^+(2\lambda_{N,m})) e^{-V_N + o(\nu_N)} + N^{-1/3+\varepsilon}, \quad (6.10)$$

with an arbitrarily small fixed exponent slack absorbing the enlarged length and the subpolynomial information factor.

Theorem 6.4 (Resolved thinning kernels). *If $p_{N,m}(n) > \varepsilon_{N,m,C}^{\text{path}}$, then*

$$d_{\text{TV}}(\mathcal{L}((Z_{N,b_N+m+j})_{j \geq 0} \mid C_N, Z_{N,b_N+m} = n), \text{Thin}_n) \leq \frac{\varepsilon_{N,m,C}^{\text{path}}}{p_{N,m}(n)}, \quad (6.11)$$

where Thin_n is the path of n independent geometric lifetimes from (5.31). If a fixed lower segment is also retained, the same inequality uses the comparison error based at its lowest level and gives the reverse immigration kernel (5.32) jointly with the future.

Proof. Apply Lemma 6.3 to the present count and the entire retained path. The conditional target is exactly the geometric-lifetime path by independent Poisson splitting. For lower levels, the increments below the present level are independent Poisson variables. A finite intensity half-line comparison cannot control the whole infinite reverse path, whose total target mass diverges. $\square$

For every integer $n \geq 1$ and every $\lambda > 0$, the local formula is

$$p_\lambda(n) = \frac{e^{-\delta_n}}{\sqrt{2\pi n}} e^{-\lambda h(n/\lambda)}, \quad 0 \leq \delta_n \leq \frac{1}{12n}, \quad h(v) = v \log v - v + 1.$$

The correction is independent of λ ; its relative error is at most $1/(12n)$. For $n = 0$ use the exact mass $e^{-\lambda}$ instead. The telescoping proof is in section D. For $n_N \rightarrow \infty$, a sufficient condition for the right side of (6.11) to vanish is

$$I_N^{\text{info}} + \log^+ \lambda_{N,m} + \log(1 + \log^+(2\lambda_{N,m})) + \frac{1}{2} \log n_N + \lambda_{N,m} h(n_N/\lambda_{N,m}) \leq V_N - c\nu_N. \quad (6.12)$$

In the central regime $n_N = \lambda_{N,m} + t_N \sqrt{\lambda_{N,m}} + O(1)$, $|t_N| = o(\lambda_{N,m}^{1/6})$, this reduces, with a fixed margin, to

$$I_N^{\text{info}} + \frac{3}{2} \log \lambda_{N,m} + \frac{1}{2} t_N^2 \leq V_N - c\nu_N.$$

The coefficient $3/2$ is a sufficient error budget, not a threshold of the model. Its derivation, including the common-band check, is in section D.

6.4 Quenched comparison on geometric scales

Corollary 6.5 (Almost-sure environmental control). *Let $N_k \asymp q^k$, $q > 1$, and suppose $\log \Lambda_{N_k} \leq V_{N_k} - c\nu_{N_k}$. For each $0 < \beta < c$, the conditional total-variation distance, given $\mathcal{F}_{Y_{N_k}^\dagger}$, between the complete aggregated staircase and its deterministic target is at most $e^{-\beta\nu_{N_k}}$ for all sufficiently large k , almost surely.*

Proof. The fibrewise proof of Theorem 5.10(ii) gives mean distance $e^{-c\nu_N + o(\nu_N)} + N^{-1/3+o(1)}$. Markov's inequality at $e^{-\beta\nu_N}$ gives a summable series along $N_k \asymp q^k$, since $\nu_{N_k} \asymp \sqrt{k/\log k}$. Borel–Cantelli concludes; no independence between scales is used. $\square$

The same argument applies to the dyadic graph-based labelled field in the sufficient range of that estimate, and to scalar kernels uniformly over the $O(\log N)$ lengths of any fixed logarithmic band with a fixed error margin. A maximal coupling exists separately in each qualified environment and scale. This does not furnish a canonical coupling over all scales.

7 Finite prefixes and the arithmetic boundary

A finite prefix is not a union of independent dyadic experiments. Its retained bulk must be treated in one Poisson comparison, using the macroscopic relation estimate. The starts near the origin need a different argument: there the initial run costs one sign per prime, rather than one sign per integer. The two candidate masses are $\alpha_L = 2^{-\pi(L)}$ at the boundary and $\lambda_M = M2^{-L}$ in the bulk. To compare them, we must bound the other microscopic starts, discard the intervening region, and approximate the bulk with error $o(\lambda_M)$. A critical-scale error of size $o(1)$ is not sufficient for this last task.

The limiting source weights arise by normalizing these two masses. The stronger conclusion of Theorem 7.11 also distinguishes their marks: both excess laws are geometric, but one counts new primes and the other new integer positions. This distinction is arithmetic, not a property of the common geometric distribution.

7.1 The exact boundary event

Set

$$\begin{aligned} C_L &= \mathbf{1}_{\{f(1)=\dots=f(L)\}}, & \mathbf{B}_L &= (C_L, (J_{x,L})_{2 \leq x \leq 2L^2}), \\ B_L^{\text{mic}} &= C_L + \sum_{2 \leq x \leq 2L^2} J_{x,L}, & q_L &= \mathbb{P}(B_L^{\text{mic}} > 0). \end{aligned}$$

Because $f(1) = 1$, the event $C_L = 1$ prescribes precisely the prime signs up to L . Hence

$$\mathbb{P}(C_L = 1) = 2^{-\pi(L)} \quad (7.1)$$

exactly. The quantity q_L is different: it includes all other microscopic starts. The next result shows that their combined mass is smaller by a fixed exponential factor in $\pi(L)$.

Theorem 7.1 (Microscopic non-vacancy and localization). *Let $\delta_* = 28271/332640$. As $L \rightarrow \infty$,*

$$\sum_{x=2}^{2L^2} \mathbb{P}(J_{x,L} = 1) \leq 2^{-(1+\delta_*-o(1))\pi(L)}. \quad (7.2)$$

Consequently, for every fixed $0 < \delta_0 < \delta_*$,

$$q_L = 2^{-\pi(L)}(1 + O_{\delta_0}(2^{-\delta_0\pi(L)})). \quad (7.3)$$

In particular $\delta_0 = 1/12$ is admissible. Conditional on microscopic non-vacancy, the exact border is the unique recorded start with probability $1 - O_{\delta_0}(2^{-\delta_0\pi(L)})$.

Proof. Let M_x be the matrix whose rows are $v(n)$, $n \in V_x$. The start matrix is $A_x = D_x M_x$, where

$$D_x z = (z_{x-1} + z_x, z_x + z_{x+1}, \dots, z_x + z_{x+L-1}).$$

The kernel of D_x consists of constant vectors. Therefore

$$\text{rank } A_x \geq \text{rank } M_x - 1, \quad \mathbb{P}(J_{x,L} = 1) \leq 2^{-\text{rank } A_x}. \quad (7.4)$$

If $2 \leq x \leq L-1$, both $2x$ and $2x-2$ lie in the proposed constant run. The first forces $f(2) = 1$, and the second forces $f(2) = -1$. Thus these starts are impossible. For the three transition starts $x \in \{L, L+1, L+2\}$, primes between $2\sqrt{B}$ and B , together with prime vertices above B , give an identity minor of size $(2 - o(1))\pi(B)$. The explicit choice of rows is in Lemma E.1.

For $B + 2 \leq x \leq B^2 + 2$, the uniform Laishram–Shorey bound gives

$$\omega \left(\prod_{n \in V_x} n \right) \geq (2 - o(1))\pi(B).$$

After discarding the prime divisors at most B , at least $t_x \geq (1 - o(1))\pi(B)$ primes $q > B$ remain. All vertices are at most $B^2 + B < q^2$, so such a prime occurs to valuation one at one vertex only; two such primes cannot occupy the same vertex. Their columns span a t_x -dimensional coordinate subspace.

Use the medium primes $B/11 < p \leq B$. Their total odd-valuation incidence count satisfies

$$E_x \geq (H_{11} - 1 - o(1))\pi(B), \quad H_{11} = \sum_{j=1}^{11} j^{-1}. \quad (7.5)$$

Here the loss from even valuations is only $O(1)$ over the entire prime range: if $p^2 \mid n \leq B^2 + B$ then $n = ap^2$ with $a \leq 122$, and for each a there are only $O(1)$ possibilities in an interval of length B . This global count avoids paying one lost incidence per prime.

Delete the private rows. Each deleted row removes at most one medium incidence, each remaining row has weight at most two, each column has weight at most eleven, and a pair of medium primes cannot occur together in two rows. The associated incidence graph is therefore simple. The elementary simple-incidence estimate Lemma E.2 gives rank at least $E/12$. Passing to the quotient by the private coordinate subspace yields

$$\text{rank } M_x \geq t_x + \frac{(E_x - t_x)_+}{12} \geq \left(1 + \frac{H_{11} - 2}{12} - o(1)\right) \pi(B).$$

Since $(H_{11} - 2)/12 = 28271/332640$, (7.4) proves the desired uniform surplus in the quadratic range.

For $x > B^2 + 2$, the pointwise Balasubramanian–Shorey consequence Lemma E.3 gives

$$m_B(x) \leq B - g_L, \quad g_L = \frac{B}{\log B}(\log \log B - \log \log \log B - \theta_0 + o(1)), \quad \frac{g_L}{\pi(B)} \rightarrow \infty. \quad (7.6)$$

The affine bound then gives $\mathbb{P}(J_{x,L} = 1) \leq 2^{-g_L}$. Sum the four ranges; their polynomial number of starts is absorbed in the $o(\pi(B))$ exponent. Since $\pi(B) = \pi(L) + O(1)$, this proves (7.2). Equation (7.1) and a union bound prove the remaining conclusions. The complete incidence count, the uniform literature input and the choice $K = 11$ are detailed in section E, with local product thresholds distinguished from the thresholds of the asymptotic inputs. $\square$

7.2 Discarding deep starts without assuming independence

Lemma 7.2 (Two-defect windows at intermediate heights). *Assume $B \asymp \log M$. The number of starts $x \in [2L^2, 2M] \cap \mathbb{Z}$ with at least two B -defective vertices is at most $\exp(O(\log M / \log \log M))$. The same bound holds for every subpopulation; in particular it is uniform on each $[X, 2X]$ with $2L^2 \leq X \leq M$.*

Proof. Two such vertices have forms $d_1 a^2, d_2 b^2$ differing by a nonzero integer of modulus at most L , with squarefree B -smooth d_i . There are at most $4^{\pi(B)} = \exp(O(\log M / \log \log M))$ coefficient pairs and $O(B^2)$ offset pairs. For $d_1 \neq d_2$, use the uniform Pell bound at ambient scale M ; the coefficients and variables are polynomially bounded in M . For equal coefficients use $(a - b)(a + b) = (i - j)/d_1$ and the divisor bound. A solution determines the start once the

offsets are fixed. This counts the entire population up to height $2M$ in a single Pell box, not once per dyadic interval. $\square$

Proposition 7.3 (Deep-start and mesoscopic estimates). *For each fixed $0 < \delta < 1$ and fixed logarithmic band,*

$$\sum_{2 \leq x < M^\delta} \mathbb{P}(J_{x,L} = 1) \ll \lambda_M M^{\delta-1} + \exp(-c \log M / \log \log M), \quad \lambda_M = M 2^{-L}, \quad (7.7)$$

where $c > 0$ depends only on that band and δ . More precisely,

$$\mathbb{P}(\exists x \in (2L^2, M^\delta) : J_{x,L} = 1) \leq M^\delta 2^{-L} + C_1 e^{C_2 L / \log L} 2^{-g_L}. \quad (7.8)$$

If $L \rightarrow \infty$, $L = O(\log M)$ and $\lambda_M \rightarrow 0$, the latter bound is $o(q_L + \lambda_M)$.

Proof. For any deterministic $A \subset [2L^2, 2M] \cap \mathbb{Z}$, starts with at most one defect have probability at most 2^{-L} . The remaining starts are counted together by Lemma 7.2, and (7.6) bounds each probability by 2^{-g_L} . Uniformly in A ,

$$\sum_{x \in A} \mathbb{P}(J_{x,L} = 1) \leq |A| 2^{-L} + C_1 e^{C_2 L / \log L} 2^{-g_L}.$$

The constants depend only on the fixed logarithmic band. Adding the microscopic range, controlled by Theorem 7.1, proves (7.7); restricting this mask proves (7.8). No exterior factor counting height slices is needed. For the last assertion, $M^\delta 2^{-L} = o(\lambda_M)$, while $g_L - \pi(L) \gg L \log \log L / \log L$ dominates the sieve prefactor. Thus the second term is $o(2^{-\pi(L)}) = o(q_L)$. $\square$

The same argument gives cutoff stability. For $2L^2 \leq K \leq 2^L$,

$$\mathbb{P}(\exists x \in (2L^2, K] : J_{x,L} = 1) \leq K 2^{-L} + O(e^{CL / \log L} 2^{-g_L}). \quad (7.9)$$

One may use the artificial ambient scale 2^L in Lemma 7.2. If $K = o(2^{L-\pi(L)})$, the entire bound is $o(q_L)$. Hence the boundary asymptotic and conditional uniqueness remain valid with cutoff K , and the two conditional microscopic configurations, padded by zeros, are at total-variation distance $o(1)$ by Lemma 6.3.

7.3 Global laws and the longest run

Theorem 7.4 (Finite-prefix Poisson approximation). *In a fixed logarithmic band, for every $\varepsilon > 0$,*

$$\begin{aligned} d_{\text{TV}} \left(\mathcal{L} \left(\sum_{2 \leq x < M} J_{x,L} \right), \text{Pois}(M 2^{-L}) \right) \\ \ll \min\{1, \lambda_M (e^{-V_M + o(\nu_M)} + M^{-1/3+\varepsilon})\} + e^{-c \log M / \log \log M}. \end{aligned} \quad (7.10)$$

For $W_{M,L} = C_L + \sum_{2 \leq x \leq M-L+1} J_{x,L}$, enlarge the constant multiplying the existing deep-start remainder and add

$$2^{-\pi(L)} + O(L 2^{-L}). \quad (7.11)$$

In particular $\mathbb{P}(R_M < L) = e^{-M 2^{-L}} + o_C(1)$ at critical lengths.

Proof. Retain the single interval $U_{M,1/2}$. The proof of Theorem 4.1 uses only support heights $O(M)$, the one-window mass, the edge count and the separated-pair profile. These are supplied

by (2.4), (4.2) at upper scale M , and Theorem 3.1. Thus (4.8) applies on the retained interval with mean $|U_{M,1/2}|2^{-L}$. Restoring deep starts costs (7.7); the deterministic target-mean discrepancy is $O(\lambda_M M^{-1/2})$. This proves (7.10).

The event $W_{M,L} = 0$ is exactly $R_M < L$: every long run contained in the prefix either begins at one or has a left-maximal start at most $M - L + 1$. The full interior count differs only if the border event occurs or a start lies in the excluded mask $M - L + 2 \leq x < M$. This mask has at most L sites and lies above $2L^2$ for large M . The mask bound in the proof of proposition 7.3 gives total mass

$$L2^{-L} + O(e^{CL/\log L} 2^{-gL}) \leq L2^{-L} + O(e^{-c \log M / \log \log M}).$$

Absorb the second term into the remainder already present in (7.10). The exact border probability then proves (7.11) and the void law. $\square$

Corollary 7.5 (Almost-sure asymmetric envelopes). *For every fixed $\varepsilon > 0$, with probability one, eventually*

$$R_M - \log_2 M \geq -\log_2(\log \log M) - O(1), \quad (7.12)$$

$$R_M - \log_2 M \leq \log_2(\log M) + (1 + \varepsilon) \log_2(\log \log M) + O_\varepsilon(1). \quad (7.13)$$

In particular $R_M = \log_2 M + O(\log \log M)$ almost surely.

Proof. Set $M_k = 2^k$ and, for sufficiently large k , choose

$$L_k^- = \lfloor k - \log_2(\log k) - 1 \rfloor, \quad L_k^+ = \lceil k + \log_2 k + (1 + \varepsilon) \log_2(\log k) \rceil.$$

The lower intensity satisfies $2 \log k \leq M_k 2^{-L_k^-} < 4 \log k$, so the target void probability is at most k^{-2} . The upper intensity is at most $1/[k(\log k)^{1+\varepsilon}]$, a summable sequence. Every error in (7.10)–(7.11) is also summable: the cutoff error is a polynomial in k times $e^{-c\sqrt{k} \log k}$, and the deep-start and border errors are at most $e^{-c'k/\log k}$. Borel–Cantelli gives $L_k^- \leq R_{2^k} < L_k^+$ for all sufficiently large k , almost surely. For $2^k \leq M < 2^{k+1}$ use $R_{2^k} \leq R_M \leq R_{2^{k+1}}$. The change from k to $\log M / \log 2$ and from k to $k + 1$ costs only a bounded additive term. This proves both envelopes. No independence between different scales has been used, and no matching lower bound for the upper fluctuations is asserted. $\square$

Theorem 7.6 (Macroscopic exact-mark transport). *Let $a_M = o(\log M)$ be a nonnegative integer, $b_M = \lfloor \log_2 M \rfloor$, $L_M = b_M - a_M$, and $\lambda = M 2^{-L_M}$. Define*

$$\mathcal{E}_{M,a_M}^{\text{base}} = \sum_{x=2}^{M-L_M+1} \sum_{r \geq -a_M} K_{x,b_M+r}^{\text{ex}} \delta_{(x/M,r)}.$$

Its target has independent Poisson masses 2^{-b_M-r-1} on the same lattice, and total intensity $(M - L_M)2^{-L_M}$. For $A_M \in \mathcal{F}_{Y_M^\sharp}$ with $\mathbb{P}(A_M) > 0$, put $I_M^{\text{info}} = -\log \mathbb{P}(A_M)$; the full-field distance is bounded by

$$C_\varepsilon e^{I_M^{\text{info}}} \left[\lambda(1 + \lambda)(e^{-V_M + o(\nu_M)} + M^{-1/3+\varepsilon}) + \lambda M^{-1/2} + e^{-c \log M / \log \log M} \right]. \quad (7.14)$$

It tends to zero under the hard labelled budget $I_M^{\text{info}} + \log \lambda + \log(1 + \lambda) \leq V_M - c' \nu_M$. The movable-cutoff labelled budget and the aggregated one-factor budget of Theorem 5.10 also hold, with N replaced by M and exact-level target means $(M - L_M)2^{-b_M-r-1}$.

Proof. On $[\lceil M^{1/2} \rceil, M - L_M + 1]$ use the finite marked comparison with the macroscopic master profile; the identities (5.22) still collapse the separated-pair term to the base threshold event. Restore lower starts with (7.7) and condition using Lemma 6.1. Remove the exact-mark tail using (5.23) and the global scalar bound. A cutoff $E = \lceil (I_M^{\text{info}} + V_M + \log(2 + \lambda)) / \log 2 \rceil$ is $o(\log M)$ whenever the bound is nontrivial; the deep-start tail error is absorbed by the last term of (7.14). This proves the displayed bound. For the movable cutoff repeat with $w = \hat{V}_M$. For aggregation, sum over positions before applying the directional Stein factor; the deep-start error is negligible under its information budget. Section C.6 details the finite-cutoff calculation. $\square$

The superscript “base” is important: only the base L_M -window is required to be contained. The exact right endpoint of a larger marked run may lie beyond M . This is not the list of right-censored maximal runs of the finite sequence. All measurable functionals of this base-contained field transfer under the displayed labelled comparison. Theorem 7.7 below improves the hard-cutoff budget for its growing-intensity signed field. The proof just given is kept independent, as are the scalar and deep-start estimates on which that theorem relies. Aggregated couplings by themselves do not retain locations.

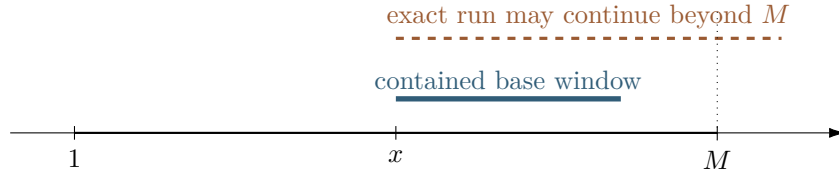


Figure 3: Containment of a base window does not censor an exact mark. The longest-run void event instead uses the threshold-specific upper start $M - L + 1$. The two conventions agree where stated, but are not interchangeable.

Two coordinate conventions. The site-indexed convention in the next subsection is related to the prefix convention by the exact relabelling $j = x - 1$:

Quantity	Prefix convention	Site-indexed convention
Interior starts	$2 \leq x \leq M - L + 1$	$1 \leq j \leq n, n = M - L$
Total target intensity	$(M - L)2^{-L}$	$n2^{-L}$
Spatial coordinate	x/M	j/n
Integer centre	$b_M = \lfloor \log_2 M \rfloor$	$\tilde{b}_M = \lfloor \log_2 n \rfloor$

The rescaled positions differ uniformly by $O((L + 1)/M)$, but this does not identify distinct grids in total variation. An exact-length level r centred at b_M becomes $r + b_M - \tilde{b}_M$ in the other convention. The centres need not be equal: at $M = 2^k$ and positive $L = o(M)$, they are k and $k - 1$ for large k . The limiting phase 1 therefore retains a level shift relative to phase 0. Full marks are unchanged by this relabelling; observing their censored growth is a separate process question.

7.4 The full same-grid signed prefix comparison

The graph-based transport above remains useful for its stated ranges. We now keep the entire signed prefix field at one factor of intensity by constructing the exact conditional law after forcing one occurrence. Only the primes chosen for that forcing enter its change set.

Use the hard scales of (4.4) with $N = M$ and set

$$L = \lfloor \log_2 M \rfloor - a_M, \quad n = M - L, \quad p = 2^{-L}, \quad \Lambda = np, \quad a_M \in \mathbb{N}_0, \quad a_M = o(\log M). \quad (7.15)$$

Here $\mathbb{N}_0 = \{0, 1, \dots\}$, as elsewhere in this article. For $1 \leq j \leq n$, $e \geq 0$ and $\sigma \in \{-1, +1\}$, put

$$I_{j,e,\sigma} = K_{j+1,e}^\sigma = \mathbf{1}_{\{j+1 \text{ is a run start, } \ell_{j+1}=L+e, f(j+1)=\sigma\}}, \quad \gamma_{e,\sigma} = 2^{-e-2}. \quad (7.16)$$

Thus the raw exact word occupies the integers $j, \dots, j + L + e + 1$. Its base is contained in $[1, M]$, but its full rightward length may continue beyond M . The run starting at 1 is not indexed. Write $\mathbf{I}_M = (I_{j,e,\sigma})_{j,e,\sigma}$, and let $\mathbf{P}_M$ have independent coordinates $P_{j,e,\sigma} \sim \text{Pois}(p\gamma_{e,\sigma})$ on the *same* index set. Both are finite-mass configurations on a countable lattice; the target has total mean Λ , whereas the source has at most one atom at a site.

Let $\mathcal{A}_M \in \mathcal{F}_{Y_M^\#}$ have positive probability, put $I_M = -\log \mathbb{P}(\mathcal{A}_M)$, and assume, for fixed $c > 0$,

$$\Lambda \rightarrow \infty, \quad I_M + \log \Lambda \leq V_M - c\nu_M. \quad (7.17)$$

In particular $\Lambda \geq 1$ for all sufficiently large M .

Theorem 7.7 (The full same-grid signed comparison). *Under (7.15)–(7.17), for every $0 < c' < c$ and $0 < \varepsilon < 1/3$,*

$$\Delta_M^{\text{mic}} := d_{\text{TV}}(\mathcal{L}(\mathbf{I}_M \mid \mathcal{A}_M), \mathcal{L}(\mathbf{P}_M)) \ll_{c,c',\varepsilon} e^{-c'\nu_M} + M^{-1/3+\varepsilon}. \quad (7.18)$$

All spatial labels, full exact excesses and actual signs are retained simultaneously. No partition or mesh condition is needed for this comparison with the moving same-grid target.

The pivot notation in the proof agrees with $P^*(m) = P^+(\text{core}(m))$ in [9, eq. (5)], with value 1 on squares.

Proof mechanism. For an integer $m \geq 1$ define

$$\kappa_{\text{odd}}(m) = \max\{q \text{ prime} : v_q(m) \text{ is odd}\}, \quad \kappa_{\text{odd}}(m) = 1 \quad \text{if } m \text{ is a square.}$$

The pivot is not the ordinary largest prime factor. Use the finite cutoff

$$E_M^* = \left\lceil \frac{I_M + V_M + \log(2 + M2^{-L})}{\log 2} \right\rceil, \quad Q = L + E_M^* + 1. \quad (7.19)$$

On retained maximal supports each vertex has a private pivot above $Y_M^\#$. To force a type (e, σ) , modify only the pivots of its $L + e + 2$ raw vertices, including both boundaries; the unused pivots of the maximal support remain unchanged. This produces the *entire* law conditioned on that prescribed signed word. It leaves the nonpivot law, including the small-prime environment, unchanged. This is proposition F.7.

The deterministic set $\mathcal{D}_j$ of sites possibly changed by this forcing need not be symmetric. The dimension-free Palm–Stein inequality of Lemma F.2 charges its size, not every edge of the all-large-prime graph. The uniform reciprocal estimate

$$\sum_{\substack{m \leq M + O(\log M) \\ \kappa_{\text{odd}}(m) > Y_M^\#}} \frac{1}{\kappa_{\text{odd}}(m)} \leq M \exp\{-2V_M + o(\nu_M)\} \quad (7.20)$$

gives $\sum_j |\mathcal{D}_j| \ll n^2 e^{-2V_M + o(\nu_M)} + n(Q+1)^2$. Proposition F.5 proves this estimate by a uniform Rankin bound and a convex saddle calculation. Precision $o(\nu_M)$, not merely $o(V_M)$, preserves the fixed information margin.

With the deletion and shifted-tail cost $d_{0,M}$ of (F.17), the finite ledger is

$$\Delta_M^{\text{mic}} \ll d_{0,M} + np^2 + e^{I_M} p^2 \left\{ nQ + \sum_j |\mathcal{D}_j| + R_{2,U}^{\text{val}}(M, Q+1) \right\}. \quad (7.21)$$

The directed-footprint term costs $e^{I_M + 2 \log \Lambda - 2V_M + o(\nu_M)} \leq e^{-I_M - 2c\nu_M + o(\nu_M)}$. Deletion costs $e^{I_M + \log \Lambda - V_M + o(\nu_M)}$; the full-value relation profile contributes $M^{-1/3 + \varepsilon + o(1)}$. Section F gives the complete coupling and ledger, including the signed local bound and all deletions. Its tail input is the independent scalar theorem Theorem 7.4, proved independently above. No process limit or Palm-void conclusion is used. $\square$

Restriction to a dyadic block. The one-factor hard-field budget is not confined to prefix geometry. Restricting a larger prefix retains the original integer start labels.

Corollary 7.7a (Dyadic signed field at the one-factor budget). *Let $L_N = \lfloor \log_2 N \rfloor - d_N$, where $d_N \in \mathbb{N}_0$ and $d_N = o(\log N)$, and put $\Lambda_N = N2^{-L_N} \rightarrow \infty$. Suppose $C_N \in \mathcal{F}_{Y_N^\#}$ has positive probability and, for a fixed $c > 0$,*

$$I_N^{\text{info}} := -\log \mathbb{P}(C_N), \quad I_N^{\text{info}} + \log \Lambda_N \leq V_N - c\nu_N.$$

For $0 < c' < c$ and $0 < \varepsilon < 1/3$, with $I_N = [N, 2N) \cap \mathbb{Z}$,

$$d_{\text{TV}} \left(\mathcal{L}((K_{x,e}^s)_{x \in I_N, e \geq 0, s = \pm 1} \mid C_N), \bigotimes_{x \in I_N, e \geq 0, s = \pm 1} \text{Pois}(2^{-L_N - e - 2}) \right) \ll_{c, c', \varepsilon} e^{-c' \nu_N} + N^{-1/3 + \varepsilon}.$$

The coordinates retain the full exact excess and actual sign at every start; the target uses the same integer labels.

Proof. Take $M = 4N$ and keep the same length L_N . Then

$$\lfloor \log_2 M \rfloor - L_N = d_N + 2 = o(\log M), \quad (M - L_N)2^{-L_N} = (4 - L_N/N)\Lambda_N.$$

Every base starting in I_N is contained in $[1, M]$ for large N ; no restriction is imposed on its full rightward length.

The hard scales are stable by monotonicity alone. Write $v(H) = V_{e_H}$ and $\nu(H) = H/v(H)$. The saddle equation is $H = \nu(H)D(\nu(H))$. On the relevant positive branch both D and $zD(z)$ are strictly increasing. For $H \leq K$, first compare the solutions $\nu(H), \nu(K)$ and then their D -values to obtain

$$v(H) \leq v(K), \quad \nu(H) \leq \nu(K) \leq \frac{K}{H} \nu(H).$$

At $K = H + \log 4$, this gives $\nu_{4N}/\nu_N \rightarrow 1$ and $Y_{4N}^\# \geq Y_N^\#$. In particular C_N belongs to the hard field at $4N$, although the hypothesis still requires it to be measurable in the original field at N .

Choose $c' < a < c$. Eventually $L_N \leq N$, so the prefix intensity is between Λ_N and $4\Lambda_N$. The preceding inequalities give

$$V_{4N} - a\nu_{4N} - (V_N - c\nu_N) \geq \left(c - a - a \frac{\log 4}{\log N} \right) \nu_N \rightarrow \infty.$$

They therefore absorb the bounded increase $\log 4$ in log intensity: $I_N^{\text{info}} + \log((4N - L_N)2^{-L_N}) \leq V_{4N} - a\nu_{4N}$. Apply Theorem 7.7 at $4N$ with margin a and rate c' , then project onto $x \in I_N$. Total variation contracts under this projection, $e^{-c'\nu_{4N}} \leq e^{-c'\nu_N}$, and the polynomial error has the stated order. The coordinate means are unchanged. $\square$

The passage is a restriction of integer coordinates, not an approximate identification of two rescaled grids. It improves the available dyadic hard-field range at growing intensity. It neither enlarges the permitted prime sigma-field nor proves a comparison for arbitrary dictionaries; these require their own hypotheses and estimates.

Corollary 7.8 (Every measurable readout of the retained field). *Under Theorem 7.7, every measurable map Φ_M into a standard Borel space satisfies*

$$d_{\text{TV}}(\mathcal{L}(\Phi_M(\mathbf{I}_M) \mid \mathcal{A}_M), \mathcal{L}(\Phi_M(\mathbf{P}_M))) \leq \Delta_M^{\text{mic}}. \quad (7.22)$$

The same holds for any common Markov kernel, including auxiliary marking followed by a measurable readout. The map or kernel may depend on M .

Proof. Pull an output event back to the input space. For a kernel K_M , apply the variational formula to $K_M(\cdot, B) \in [0, 1]$. The formula follows by integrating level sets; indicator tests give equality. $\square$

There is no continuity or number-of-queries restriction on this comparison: an adaptive interrogation determined by the same retained data is one measurable readout. Every deterministic partition, including the singleton partition, inherits (7.18) without a smallest-cell penalty. Multiple target atoms receive separate auxiliary marks; a rule initially defined only for simple configurations is extended measurably to the other configurations.

This universality concerns an observation of the specified field, not a universal limit law. It supplies no new starts, the missing boundary run, an arbitrary word field, unbounded moments or an independence statement. The recorded full lengths can be read beyond M , but the rest of the future sequence is not recorded. Grid-to-continuum limits, censoring and rare post-selection retain their own estimates. In particular neither bounded nor vanishing Λ is added to (7.17) by this corollary; Theorems 1.1 and 7.9 retain their separate critical and relative roles.

Empirical Poisson laws for almost every realization

The comparison above averages over the prime signs. A summable error also permits a different order of quantifiers: first sample the entire function f , then fix it and randomize only the spatial origin. The following consequence concerns a count, not a comparison of the whole local field after f has been fixed.

Corollary 7.8a (Spatial empirical law for a fixed realization). *Fix $0 < \alpha < 1$ and $\tau > 0$. With V_M as in (4.4), for all sufficiently large k set*

$$M_k = 2^k, \quad L_k = k - \left\lfloor \frac{\alpha V_{M_k}}{\log 2} \right\rfloor, \quad p_k = 2^{-L_k},$$

$$n_k = M_k - L_k, \quad h_k = \lfloor \tau 2^{L_k} \rfloor, \quad N_k = n_k - h_k + 1.$$

Let U_k be uniform on $\{0, \dots, n_k - h_k\}$, independently of f , and put

$$Z_k(f, u) = \sum_{j=u+1}^{u+h_k} J_{j+1, L_k}(f), \quad \mu_k^f = \frac{1}{N_k} \sum_{u=0}^{n_k-h_k} \delta_{Z_k(f, u)}.$$

Then, for almost every f in the model (1.1),

$$d_{\text{TV}}(\mu_k^f, \text{Pois}(\tau)) = d_{\text{TV}}(\mathcal{L}_{U_k}(Z_k(f, U_k)), \text{Pois}(\tau)) \longrightarrow 0.$$

Here h_k counts potential start sites, the two signs are aggregated, and a run is counted only at its left-maximal start. In particular,

$$h_k p_k \longrightarrow \tau, \quad h_k = M_k \exp\{-\alpha V_{M_k} + O(1)\} = M_k^{1-o(1)}.$$

The full-probability set is asserted for the fixed pair (α, τ) ; the conclusion is along the specified dyadic scales.

The starts read by Z_k range from $u + 2$ to $u + h_k + 1$. Thus every base of length L_k is contained in the global prefix $[1, M_k]$, and the initial run from 1 is excluded. A full run is not censored at M_k , although its excess is not used by this count. In $\mathcal{L}_{U_k}$ only the origin is random; no independence between the origins at different scales is required.

Proof. We first check the arithmetic budget and then do the frequency calculation solely in the independent target. The hard-scale expansions in proposition 4.2 give

$$V_{M_k} \asymp \sqrt{k \log k}, \quad \nu_{M_k} \asymp \sqrt{k / \log k}, \quad V_{M_k} / \nu_{M_k} \longrightarrow \infty.$$

Writing $\Lambda_k = n_k p_k$, the definitions imply

$$\log \Lambda_k = \alpha V_{M_k} + O(1), \quad h_k / n_k = \exp\{-\alpha V_{M_k} + O(1)\}.$$

Hence $\Lambda_k \rightarrow \infty$, $1 \leq h_k \leq n_k$ eventually, and (7.17) holds with $\mathcal{A}_{M_k}$ equal to the whole probability space and any fixed $c > 0$, for all sufficiently large k . Theorem 7.7 consequently gives a sequence Δ_k with

$$d_{\text{TV}}(\mathcal{L}(\mathbf{I}_{M_k}), \mathcal{L}(\mathbf{P}_{M_k})) \leq \Delta_k, \quad \sum_k \Delta_k < \infty, \quad \sum_k h_k / n_k < \infty.$$

Indeed, for fixed admissible c' and ε , the bound is $O(e^{-c'\nu_{M_k}} + 2^{-k(1/3-\varepsilon)})$, and both terms are summable.

Sum the target over the signs and excesses at a site. Since $\sum_{e,\sigma} \gamma_{e,\sigma} = 1$, the resulting variables

$$X_{k,j} = \sum_{e \geq 0, \sigma = \pm 1} P_{k,j,e,\sigma} \sim \text{Pois}(p_k), \quad 1 \leq j \leq n_k,$$

are independent. For $0 \leq u < N_k$ and $r \in \mathbb{N}_0$, define

$$S_{k,u} = \sum_{j=u+1}^{u+h_k} X_{k,j}, \quad F_{k,r} = \frac{1}{N_k} \sum_{u=0}^{N_k-1} \mathbf{1}_{\{S_{k,u}=r\}}, \quad q_{k,r} = \text{Pois}(h_k p_k)\{r\}.$$

Every $S_{k,u}$ has exactly the displayed Poisson law, so $\mathbb{E}F_{k,r} = q_{k,r}$. The indicators at origins u, v are independent if $|u - v| \geq h_k$. For the other pairs the upper bound on their covariance is $1/4$: $\text{Var}(A - B) \geq 0$ and $\text{Var}(A), \text{Var}(B) \leq 1/4$ give this directly for indicators A, B . Counting at most $N_k(2h_k - 1)$ ordered pairs gives

$$\text{Var}(F_{k,r}) \leq \frac{2h_k - 1}{4N_k} \ll \frac{h_k}{n_k},$$

uniformly in r . Indeed $h_k/n_k \rightarrow 0$ gives $N_k = n_k - h_k + 1 \geq n_k/2$ eventually, so $\sum_k h_k/N_k < \infty$. In particular, overlapping windows have not been treated as independent.

Almost surely all runs are finite, so summing $I_{j,e,\sigma}$ over (e, σ) gives J_{j+1, L_k} . The empirical frequency is therefore the same measurable function of the source field as $F_{k,r}$ is of the target field. For each fixed $r \in \mathbb{N}_0$ and $\eta > 0$, Corollary 7.8 and Chebyshev's inequality in the target yield

$$\mathbb{P}_f\left(\left|\mu_k^f\{r\} - q_{k,r}\right| > \eta\right) \leq \Delta_k + \frac{2h_k - 1}{4N_k\eta^2}.$$

This is one application to the event defined by the entire empirical frequency: the error Δ_k is not multiplied by the number of origins. The bound is summable. Borel–Cantelli and a countable intersection over $r \in \mathbb{N}_0$ and $\eta = 1/m$, $m \geq 1$, show that, almost surely, $\mu_k^f\{r\} - q_{k,r} \rightarrow 0$ for every r .

Finally, $0 \leq \tau - h_k p_k < p_k \rightarrow 0$, so $q_{k,r} \rightarrow q_r := \text{Pois}(\tau)\{r\}$. For any probability μ on $\mathbb{N}_0$ and any $R \in \mathbb{N}_0$, normalization of the masses gives

$$d_{\text{TV}}(\mu, q) \leq \sum_{r=0}^R |\mu\{r\} - q_r| + \sum_{r>R} q_r.$$

First let $k \rightarrow \infty$ with R fixed, and then let $R \rightarrow \infty$. This proves the asserted total-variation convergence. Neither an unbounded source moment nor a joint coupling of the targets at all scales is used, and Borel–Cantelli requires no independence between the source events at different scales. $\square$

Scope and comparison. The normalization $h_k 2^{-L_k} \rightarrow \tau$ is essential: the window-size estimate with an unspecified $O(1)$ alone does not fix the limiting mean. The inequality $\alpha < 1$ leaves room in the arithmetic budget, whereas $\alpha > 0$ makes the target variance summable. This statement fixes the entire f , unlike Corollary 6.5. Najnudel's empirical block laws [30] use fixed block length. Pandey–Wang–Xu [32, arXiv version 4, Theorem 1.2 and footnote 1] already study a typical multiplicative function from a uniform spatial origin: their normalized Steinhaus sums have a complex Gaussian limit, on sets of functions of measure $1 - o(1)$ at each size. The present statement instead counts rare left-maximal binary runs, converges in total

variation to a finite-mean Poisson law, and has a single full-probability set for each fixed (α, τ) along the specified dyadic scales. Poisson genericity samples long words in a fixed sequence [1,2]; it is a different empirical observable. No all-scale or simultaneous all-parameter assertion, and no assertion about Liouville, follows here. No result about records, occupation or scale flows is an input. The next observation shows why the restriction to a count is substantive.

Remark 7.8b (A support obstruction for the fixed local field). Keep the scales of Corollary 7.8a. For any fixed f , let

$$\mathbf{Z}_k(f, u) = (J_{u+i+1, L_k}(f))_{1 \leq i \leq h_k}, \quad \hat{\mu}_k^f = \frac{1}{N_k} \sum_{u=0}^{N_k-1} \delta_{\mathbf{Z}_k(f, u)}, \quad Q_k = \bigotimes_{i=1}^{h_k} \text{Pois}(p_k).$$

Then

$$\liminf_{k \rightarrow \infty} d_{\text{TV}}(\hat{\mu}_k^f, Q_k) \geq 1 - e^{-\tau}(1 + \tau) > 0.$$

Indeed, the support $\mathcal{S}_k(f)$ of $\hat{\mu}_k^f$ contains at most N_k configurations. For large k , $p_k \leq 1$ and every $z \in \mathbb{N}_0^{h_k}$ with $|z| = \sum_i z_i \geq 2$ satisfies

$$Q_k\{z\} = e^{-h_k p_k} \frac{p_k^{|z|}}{\prod_i z_i!} \leq p_k^2.$$

Testing the complement of $\mathcal{S}_k(f)$ within $\{|z| \geq 2\}$ gives

$$d_{\text{TV}}(\hat{\mu}_k^f, Q_k) \geq 1 - e^{-h_k p_k}(1 + h_k p_k) - N_k p_k^2.$$

Now $h_k p_k \rightarrow \tau$ and $N_k p_k^2 = M_k^{-1+o(1)} \rightarrow 0$. This holds for every fixed realization, not merely almost every one. Thus the count law can converge even though the full local vector, expressed relative to its sampled origin, cannot have the analogous product-Poisson approximation in total variation.

7.5 Relative bulk approximation and the crossover

Assume from now on $L \rightarrow \infty$, $L = O(\log M)$ and $\lambda_M = M 2^{-L} \rightarrow 0$. Fix $0 < \delta < 1$ and put $H_M = \lceil M^\delta \rceil$, $I_{M,L}^{\text{bulk}} = [H_M, M - L + 1] \cap \mathbb{Z}$ and $\Xi_{M,L}^{\text{bulk}} = \sum_{x \in I_{M,L}^{\text{bulk}}} J_{x,L} \delta_{x/M}$. Let $\Pi_{M,L}^{\text{bulk, lat}}$ have independent Poisson masses $p = 2^{-L}$ at those same sites; its intensity is $\lambda_M^{\text{bulk}} = |I_{M,L}^{\text{bulk}}| p \sim \lambda_M$. Fix an upper-band constant A_0 with $L \leq A_0 \log M$ for all large M . The rare-intensity assumption supplies the positive lower band. Constants and little-oh terms in this subsection may depend on δ , A_0 and the displayed error exponent, but not on M or L .

Theorem 7.9 (Relative signed marked bulk kernel). *Let $\mathcal{K}_{M,L}^{\text{bulk}} = (K_{x,e}^s)_{x \in I_{M,L}^{\text{bulk}}, e \geq 0, s \in \{\pm 1\}}$ and let $\mathcal{P}_{M,L}^{\text{bulk}}$ have independent Poisson coordinates of means 2^{-L-e-2} on this same index set. At $Y = Y_M^\sharp$, for every fixed $0 < \varepsilon < 1/3$,*

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\mathcal{K}_{M,L}^{\text{bulk}} | \mathcal{F}_Y), \mathcal{L}(\mathcal{P}_{M,L}^{\text{bulk}})) \ll_{\delta, A_0, \varepsilon} \lambda_M \left(e^{-V_M + o(\nu_M)} + M^{-1/3+\varepsilon} \right). \quad (7.23)$$

In particular, summing marks and signs gives

$$\mathbb{E} d_{\text{TV}}(\mathcal{L}(\Xi_{M,L}^{\text{bulk}} | \mathcal{F}_Y), \mathcal{L}(\Pi_{M,L}^{\text{bulk, lat}})) = o(\lambda_M). \quad (7.24)$$

Proof. For a finite mark cutoff E , repeat the signed comparison of Theorem 5.8 on this single macroscopic interval. With $Q = L + E + 1$ and $p = 2^{-L}$ its error is at most

$$Cp(\mathcal{M}_B + |D_Y(Q)|) + Cp^2(MQ + E_Y(Q) + R_{2,\delta}(M, L)).$$

The signed marks are summed before the separated-pair bound, so the relation sum is at the base length L . Divide by $Mp = \lambda_M$. The defect and support terms become $M^{-1/2+o(1)}$ and $e^{-V_M+o(\nu_M)}$, respectively. The edge term becomes $\lambda_M(e^{-V_M+o(\nu_M)} + M^{-1+o(1)})$. By (3.4) and $2^B = 2/p$, the relation term is $O_\varepsilon(M^{-1/3+\varepsilon}(1 + \lambda_M))$. The local term pQ is smaller because $\lambda_M \rightarrow 0$.

Take $E = \lceil V_M / \log 2 \rceil$. Then $E = o(\log M)$ and the base and maximal support lengths stay in a fixed logarithmic band. The target tail has mass at most $\lambda_M 2^{-E-1}$. The actual tail is the event that some retained site starts a run at length $L + E + 1$; the macroscopic masked first-moment bound gives the same upper bound times $1 + M^{-1/2+o(1)}$. Couple both tails to zero and average. Since $\lambda_M \leq 1$ for large M , the preceding bounds give (7.23). Measurable contraction proves (7.24). $\square$

Every coordinate of $\mathbf{B}_L$ depends only on prime signs up to $2L^2 + L < Y_M^\#$. The stable product lift therefore gives

$$d_{\text{TV}} \left(\mathcal{L}(\mathbf{B}_L, \Xi_{M,L}^{\text{bulk}}), \mathcal{L}(\mathbf{B}_L) \otimes \mathcal{L}(\Pi_{M,L}^{\text{bulk,lat}}) \right) = o(\lambda_M) = o(q_L + \lambda_M). \quad (7.25)$$

This is asymptotic factorization at the indicated error scale, not exact finite- M independence.

Theorem 7.10 (Two-source rare tail and conditional location). *Under the preceding rare-intensity assumptions,*

$$\mathbb{P}(R_M \geq L) = q_L + \lambda_M + o(q_L + \lambda_M). \quad (7.26)$$

Let $X_{M,L}^*$ be the least start of a contained run of length at least L , on the event that one exists. If $L - \log_2 M - \pi(L) \rightarrow s \in \mathbb{R}$, then

$$\mathcal{L}(X_{M,L}^*/M \mid R_M \geq L) \implies \frac{1}{1 + 2^{-s}} \delta_0 + \frac{2^{-s}}{1 + 2^{-s}} \text{Leb}_{[0,1]}. \quad (7.27)$$

The source-resolved version assigns a microscopic start its location divided by L^2 , and a bulk start its location divided by M ; its limit is the same mixture with the two source labels retained. The weights also have their evident degenerate limits when $s \rightarrow \pm\infty$.

Proof. The mesoscopic region has mass $o(q_L + \lambda_M)$ by proposition 7.3. Under the product target in (7.25), a simultaneous microscopic and bulk hit has probability $O(q_L \lambda_M)$, and at least two bulk hits have probability $O(\lambda_M^2)$. Both are negligible on the combined scale. This proves (7.26). Given a microscopic hit, Theorem 7.1 puts the start at one with probability $1 - o(1)$. Given exactly one bulk target point, its location is uniform on the common lattice. Finally

$$\frac{\lambda_M}{q_L} = 2^{-(L - \log_2 M - \pi(L))} (1 + o(1)) \longrightarrow 2^{-s}.$$

Divide the two one-hit masses by (7.26) and pass from the uniform grid to Lebesgue measure weakly. All total-variation comparisons precede this final diffuse limit. $\square$

7.6 The prime clock

Let $p_1 = 2 < p_2 < \dots$. Almost surely $J_\partial = \min\{j : f(p_j) = -1\}$ is finite and the maximal initial positive run has length $\ell_\partial = p_{J_\partial} - 1$: every smaller integer has only positive prime

factors, whereas that negative prime changes sign. On this full-measure event and $C_L = 1$, put $G_L = J_\partial - \pi(L) - 1$; assign $G_L = 0$ elsewhere whenever a globally defined random variable is needed. Then

$$\mathbb{P}(G_L = k \mid C_L = 1) = 2^{-(k+1)}, \quad \mathbb{P}(\ell_\partial - L \geq h \mid C_L = 1) = 2^{-[\pi(L+h) - \pi(L)]}. \quad (7.28)$$

Thus prime rank, not integer distance, is its geometric clock. The boundary theorem makes the law of G_L given $B_L^{\text{mic}} > 0$ $O(2^{-\delta_0 \pi(L)})$ -close to ν_{geo} . The prime number theorem gives $\ell_\partial/L \rightarrow 1$ conditionally, but the unscaled integer overshoot is not tight: for $L = m! + 1$ the following $m - 1$ integers contain no prime.

7.7 A lattice crossover with two geometric clocks

The two branches have the same geometric law but different clocks: additional integers in a bulk run, and additional positive prime signs at the initial boundary. A moving discrete target states their joint law without requiring a limiting ratio of the two rare masses.

Write $\alpha_L = 2^{-\pi(L)}$ and $\lambda^{\text{bulk}} = \lambda_M^{\text{bulk}}$. On $\{R_M \geq L\}$, let $X^* = X_{M,L}^*$ and define

$$\Gamma_{M,L} = \begin{cases} (\partial, 1, +1, G_L), & X^* = 1, \\ (\text{bulk}, X^*, f(X^*), \ell_{X^*} - L), & X^* \in I_{M,L}^{\text{bulk}}, \\ \dagger, & \text{otherwise.} \end{cases}$$

Here $\dagger$ records an exceptional start in the intervening region. Let $\nu_\pm = (\delta_{-1} + \delta_{+1})/2$, and let $\text{Unif}(I_{M,L}^{\text{bulk}})$ be the uniform law on that integer grid.

Theorem 7.11 (Sign-resolved, two-clock crossover). *Under the rare-intensity assumptions of Theorem 7.10,*

$$d_{\text{TV}}(\mathcal{L}(\Gamma_{M,L} \mid R_M \geq L), Q_{M,L}) \rightarrow 0, \quad (7.29)$$

where the deterministic probability law on the same discrete state space is

$$\begin{aligned} Q_{M,L} &= \frac{\alpha_L}{\alpha_L + \lambda^{\text{bulk}}} \delta_{(\partial, 1, +1)} \otimes \nu_{\text{geo}} \\ &\quad + \frac{\lambda^{\text{bulk}}}{\alpha_L + \lambda^{\text{bulk}}} \delta_{\text{bulk}} \otimes \text{Unif}(I_{M,L}^{\text{bulk}}) \otimes \nu_\pm \otimes \nu_{\text{geo}}. \end{aligned} \quad (7.30)$$

No convergence of $\lambda^{\text{bulk}}/\alpha_L$ is needed. In particular,

$$\mathbb{P}(f(X^*) = +1 \mid R_M \geq L) = \frac{\alpha_L + \lambda^{\text{bulk}}/2}{\alpha_L + \lambda^{\text{bulk}}} + o(1). \quad (7.31)$$

The bulk excess in (7.29) may instead be measured using the right-censored run length in the finite word $(f(1), \dots, f(M))$.

Proof. Fix $K \geq 1$. Count the consecutive positive signs among the first K primes after L , capped at K , and set the result to zero off $C_L = 1$. This is a genuine finite-cylinder random variable, agreeing almost surely with $G_L \wedge K$. For fixed K , Bertrand's postulate places these primes below $2^K L < Y_M^\sharp$ for large M . The stable lift applied to Theorem 7.9 therefore retains the microscopic vector and this clock jointly with *all* signed bulk marks, at error $o(\lambda_M)$. No bulk-mark truncation is needed here. The interior microscopic and mesoscopic errors are $o(\alpha_L + \lambda_M)$ by Theorem 7.1 and proposition 7.3.

Write $\alpha = \alpha_L$ and $b = \lambda^{\text{bulk}}$. In the product target, keep either a boundary hit with no bulk point, or no boundary hit with exactly one bulk point. This event has exact mass

$$e^{-b}[\alpha + (1 - \alpha)b].$$

Its conditional source weights are $\alpha/[\alpha + (1 - \alpha)b]$ and $(1 - \alpha)b/[\alpha + (1 - \alpha)b]$. Their mixture differs from the weights in (7.30) by at most

$$\frac{\alpha^2 b}{(\alpha + b)[\alpha + (1 - \alpha)b]} \leq b.$$

The discarded simultaneous and multiple hits cost $O(\alpha b + b^2) = o(\alpha + b)$. Given the one bulk point, its site, sign and excess have the independent laws in (7.30). Normalize all comparison errors by $\alpha + b$, not by the smaller source mass. This proves the comparison with the boundary clock capped at K , uniformly in the moving weights.

Under the actual law, a recorded boundary tail $G_L > K$ has conditional probability at most 2^{-K-1} , since $\mathbb{P}(C_L = 1, G_L > K) = \alpha 2^{-K-1}$ and $\mathbb{P}(R_M \geq L) \geq \alpha$. The target boundary tail is also at most 2^{-K-1} . Let $M \rightarrow \infty$ first and then $K \rightarrow \infty$. This proves (7.29); projecting the sign gives (7.31).

For the last assertion, under the one-hit bulk target a mark is censored at site x only if $e > M - x + 1 - L$. The average of this geometric tail is

$$\frac{1}{|I_{M,L}^{\text{bulk}}|} \sum_{x \in I_{M,L}^{\text{bulk}}} 2^{-(M-x+2-L)} \leq \frac{1}{|I_{M,L}^{\text{bulk}}|} = o(1).$$

The total-variation comparison transfers this estimate, so replacing that mark changes the conditional law by $o(1)$. $\square$

After scaling sites by M , convergence of the ratio recovers the mixture in Theorem 7.10, now with signs and both overshoot clocks retained. The spatial limit is weak; the stronger statement (7.29) is entirely on the prelimit lattice. Its $o(1)$ error is on the combined source scale. It does not supply a relative estimate for the smaller source when its mixture weight tends to zero.

7.8 Affine information at the boundary

Theorem 7.12 (Affine-conditioned crossover). *Let $A_M = \{G_M X(Y_M^\sharp) = b_M^{\text{aff}}\}$ be a compatible affine cylinder of rank r_M , and assume compatibility also with $C_L = 1$. Set*

$$U_L = \text{span}\{e_p : p \leq L\}, \quad \kappa_M = \dim(\text{row } G_M \cap U_L), \quad d_M = \pi(L) - \kappa_M.$$

Whenever $L \leq Y_M^\sharp$, the finite identity $\mathbb{P}(C_L = 1 \mid A_M) = 2^{-d_M}$ holds exactly. For the asymptotic conclusions, let $M, L \rightarrow \infty$ with $L = O(\log M)$ and $\lambda_M = M 2^{-L} \rightarrow 0$. Fix $c > 0$ and assume $r_M \log 2 \leq V_M - c \nu_M$ eventually. The affine equations, their number and their right-hand sides may all vary along the sequence. Then the rare-tail formula is

$$\mathbb{P}(R_M \geq L \mid A_M) = 2^{-d_M} + \lambda_M + o(2^{-d_M} + \lambda_M).$$

If $L - \log_2 M - d_M \rightarrow s$, the conditional location law given A_M and $R_M \geq L$ is (7.27) with the same weights. Microscopic non-vacancy under A_M is concentrated on the border.

Proof. The stacked cylinder and border system has rank $r_M + \pi(L) - \kappa_M$, proving the exact mass ratio. Write $I = r_M \log 2$. Apply the factor e^I to the quantitative bound (7.23) *before* taking a limit. Its normalized cutoff term is at most $e^{-c\nu_M + o(\nu_M)}$, and its polynomial term vanishes because $I = O(V_M) = o(\log M)$. Hence the conditional bulk error is $o(\lambda_M)$. The microscopic interior mass, after conditioning and division by 2^{-d_M} , is at most $O(e^I 2^{-\delta_0 \pi(L) - \kappa_M}) = o(1)$, since $I = O(V_M) = o(\pi(L))$. Similarly the two terms of (7.8) become $o(\lambda_M)$ and $o(2^{-d_M})$. The conditional stable lift and the exact one-hit normalization in the proof of Theorem 7.11 now apply, with total normalization $2^{-d_M} + \lambda_M^{\text{bulk}}$. $\square$

The parameter κ_M counts linear consequences supported wholly on the border coordinates, not input equations merely mentioning them. No condition on future primes is needed for the rare tail, microscopic concentration or location laws. The geometric future mark requires a separate condition. Let $q_1 < q_2 < \dots$ be the primes after L . If, for each fixed K , eventually in M with a threshold that may depend on K ,

$$(\text{row } G_M + U_L) \cap \text{span}\{e_{q_1}, \dots, e_{q_K}\} = \{0\},$$

then the geometric boundary mark and the full mixture of Theorem 7.11 survive under A_M , with α_L replaced by 2^{-d_M} . First fix K and pass to the size limit, then let $K \rightarrow \infty$; no common threshold for all K or neutrality for a growing $K(M)$ is assumed. In general, if the extended affine stack is compatible, $\mathbb{P}(G_L \geq K \mid A_M, C_L = 1)$ equals $2^{-[K - \dim((\text{row } G_M + U_L) \cap \text{span}\{e_{q_1}, \dots, e_{q_K}\})]}$; if incompatible, the tail is zero. This separates the information that changes the border mass from information that changes its future mark. Section E.7 records the rank calculation.

8 What the method isolates

Prime signs are independent, but their products satisfy exact identities at arbitrarily large distances. The proof uses both facts. It counts the homogeneous relation spaces created by multiplication, then uses the prime environment either for an exact dependency graph or for an exact one-word conditional coupling. In the latter route, potential dependence is separated from the changes actually imposed by the forcing. The three-term profile (3.3), and its full-value extension (3.25), are the arithmetic outputs. The independent Poisson field is obtained only after the separate marginal, local-overlap and prime-support estimates have been verified.

8.1 A class of observables, rather than one count

The dictionary theorem separates local word combinatorics from nonlocal arithmetic. Changing the prescribed signs does not change a valuation matrix. Prescribing their absolute orientation removes at most two parity conditions in a pair of windows; the host count controls the additional relations. Consequently the same profile handles a growing dictionary and retains the identity and position of every rare word. Explicit marker families make its overlap condition zero. For a typical dictionary, one can do more: averaging before the worst-case estimate reduces pair dependence to equality of the two complete words. The finite collision bound removes the uniform cardinality restriction while retaining the same labels and bounded total intensity. Short suffix-prefix matches can instead create local clusters independently of any multiplicative relation; such dictionaries require a separate analysis.

Long runs provide two complementary examples. Counting left-maximal starts gives a Poisson field with geometric excesses and fair signs. Counting all constant windows retains the clusters, and the field contracts to the explicit compound-Poisson law of Corollary 5.9. Both

statements come from the same comparison, not from two unrelated moment arguments. The asymmetric almost-sure bounds in Corollary 7.5 use its summable quantitative errors; they do not require independence across scales.

The boundary is a different arithmetic regime. Its initial positive cylinder imposes $\pi(L)$ independent prime constraints, whereas other microscopic starts have an additional rank cost. The competition between $2^{-\pi(L)}$ and $M2^{-L}$ therefore produces a change in the location and sign of a rare long run. The two-clock theorem retains more than the limiting source weights: it gives a prelimit lattice law, with prime-rank overshoot at the boundary and integer overshoot in the bulk. The logistic formula is a consequence of normalizing these two masses. The substantive inputs are the microscopic rank surplus, the relative marked bulk kernel and the negligible region between them.

8.2 Arithmetic and probabilistic limits of the estimates

Three terms reach $N^{5/3}$ at criticality in the present upper bound. Improving only the terminal kernel energy would therefore not improve the total exponent. The rational families in (3.22)–(3.23) provide lower exponents $4/3$ and $3/2$. Sharpening this gap is an arithmetic question. It is distinct from improving the leading subexponential error, which is governed by exceptional supports and shared prime coordinates. For growing dictionaries the relevant quantity is not the full relation sum, but its capped version. Mutual exclusivity at one site gives a probability ceiling, and proposition 3.25 preserves it through the terminal count. This removes the terminal restriction on dictionary size without improving the uncapped $5/3$ exponent. The remaining term $NQ_B^{2/3}$ gives the sufficient exponent $1/2$ in (5.3). Extending this uniform result requires control of the nonterminal contribution or additional information about the dictionary. The typical theorem provides the latter through its sampling law, not through a smaller universal relation profile. A short affine description alone is not a guarantee: the high-probability result concerns a uniformly sampled full-rank matrix and an independent uniform right-hand side.

The Diophantine step uses only polynomial-height boxes. Localizing the squarefree parts to the coefficient and root differences reduces the count to two factors and a classical uniform Pell estimate; a height-free bound on all integral points is not needed. The self-contained proof fixes the required uniformity, rather than claiming priority for that counting principle.

Different proof routes give sufficient budgets, not intrinsic prices for retaining a label. In the hard corridor with $\Lambda \rightarrow \infty$, Theorem 7.7 and Corollary 7.7a retain all sites at the one-factor budget, on a prefix and by restriction on a dyadic block. Selected-prime reciprocals replace the all-large-prime footprint. The tables in section 6 separate these available ranges from the direct scalar and graph-based estimates. Local resolution still requires an error negligible relative to the event being resolved. At information-dependent cutoffs, proposition 6.2 balances deletion and graph costs; the event's measurability must be retained along with its information cost.

What the Palm complement does and does not prove. Section G separates regular target configurations, exact signed finite identities and consequences of the microscopic comparison. Prescribed presence does not by itself control the absence of further runs. The configuration-mass identity and Theorem 7.7 imply the averaged one-sided Palm deficit Corollary G.5; a stronger squared Walsh norm is not asserted to vanish. Conversely, Theorem G.9 exhibits diverging all-order absolute activity while microscopic distance tends to zero. Rare configurations may dominate an exponential moment without preventing ordinary total-variation approximation. The proof by selected pivots never requires that activity to be small.

8.3 Further questions

Dictionaries with substantial short overlaps need a cluster analysis compatible with the non-local arithmetic dependence. The compound-Poisson word laws for Markov sources in [33,35] provide context, not an automatic transfer. Our constant-window law is one model case, not a classification of periodic or interacting patterns. Extending the selected-coordinate coupling to nested word families, other prime alphabets or larger conditioning fields would likewise require its own exact forcing law, arithmetic footprint and tail estimates.

Records, censoring and occupation times require additional conventions about the order and recognition of occurrences. Their process-level analysis is outside the scope of this article. A common measurable readout of the same-grid comparison does not by itself supply a new Gaussian limit, a topology of convergence or asymptotic independence. Similarly, the binary run-field theorem is not a comparison for every prime-value distribution or for arbitrary persistent cluster targets. These are boundaries of the results proved here, not assertions that the corresponding questions are unresolved in all other models.

Formalization and Palomar registrations

A companion formalization in Lean 4 covers the principal arithmetic constructions and probabilistic conclusions of this article and its technical companion. The development uses Lean 4.34.0 and Mathlib v4.34.0 and is available in the [paper-c-lean repository](#). It includes explicit definitions of the random model, conditioning events, observables and comparison laws, together with proofs and reproducible verification instructions. The repository’s correspondence between the manuscripts and Lean statements records the precise coverage, hypotheses and proof substitutions.

The formalization is relative to seven explicit literature inputs: scalar Poisson Stein factors; finite Arratia–Goldstein–Gordon process comparison; the prime number theorem remainder; multivariate Poisson Stein solution bounds; the uniform prime-divisor estimate of Laishram–Shorey; Shorey’s square-product exclusion; and the Nicolas–Robin divisor bound. These propositions are hypotheses of the formal theorems that use them; their proofs are not supplied by this development. In particular, the analytic Stein solution used in Lemma F.2 remains an input to its Lean proof. The finite coupling argument is formally checked under this premise, while the semigroup construction given in the companion is outside that formal proof. This is a limitation of formal coverage, not an additional assumption in the written argument.

The formal proofs contain no proof placeholders and use only Lean’s standard foundational axioms `propext`, `Classical.choice` and `Quot.sound`. This foundational check does not discharge the literature hypotheses above. The written proofs and the formal development provide complementary accounts, whose precise correspondence is documented in the repository.

Seven Palomar registrations cover a selection of 89 declarations from the development. All entries below are version 1; the counts refer to the registered selection, not to the total number of results in the repository.

Scope	Count	Palomar identifier
Critical Poisson fields and conditional transfers	16	PALOMAR-2026-09-20-000003
Word dictionaries, exact marks and compound clusters	11	PALOMAR-2026-09-20-000004
Threshold staircases and the Poisson–Gaussian bridge	29	PALOMAR-2026-09-20-000007
Microscopic boundary, prefixes and longest runs	7	PALOMAR-2026-09-20-000011
Microscopic–bulk crossover and conditional affine laws	11	PALOMAR-2026-09-20-000012
Microscopic signed fields and empirical Poisson laws	6	PALOMAR-2026-09-21-000002
Regular configurations, Palm deficits and cumulant obstructions	9	PALOMAR-2026-09-21-000003

Each registration identifies a precise set of formal statements and a fixed repository revision. Comparator checks agreement between the selected statements and their proved counterparts; the proofs are checked by Lean and independently replayed through NanoDa. Palomar’s separate editorial review assesses the correspondence with the informal account and the mathematical interest of the submission. That review is automated and does not constitute independent human peer review. The registrations document the verification of the selected formal results under their stated hypotheses; they do not certify every argument or claim in the article and companion. The verification and review protocol is specified in [PalomarPolicy](#).

Declaration on artificial-intelligence assistance

The research was organized and directed by the author with substantial assistance from language-model systems, including Claude Fable 5 and Claude Opus 4.8 (Anthropic), and GPT-5.6 Sol and GPT-6 Astra (OpenAI). These systems were used to explore candidate lemmas and proof routes, draft and reorganize manuscripts, assist with the separate Lean development, write diagnostic programs, check references and build artifacts, and conduct adversarial reviews through separate instances.

GPT-6 Astra assisted in particular with the selected-pivot and Palm arguments, the empirical corollary, the dyadic restriction and local-field obstruction, and the exposition, bibliography and preparation of this edition. This attribution preserves the other models’ historical contributions. 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 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”. In: *International Mathematics Research Notices* 2023.24 (2023), pp. 20970–20987. doi: [10.1093/imrn/rnac234](#). arXiv: [2202.01632](#).

- [2] N. Álvarez, V. Becher, E. Cesaratto, M. Mereb, Y. Peres, and B. Weiss. “Poisson genericity in numeration systems with exponentially mixing probabilities”. In: *Transactions of the American Mathematical Society* 379.5 (2026), pp. 3403–3423. doi: [10.1090/tran/9592](https://doi.org/10.1090/tran/9592). arXiv: [2411.04116](https://arxiv.org/abs/2411.04116).
- [3] R. Arratia, L. Gordon, and M. S. Waterman. “The Erdős–Rényi law in distribution, for coin tossing and sequence matching”. In: *Annals of Statistics* 18.2 (1990), pp. 539–570. doi: [10.1214/aos/1176347615](https://doi.org/10.1214/aos/1176347615).
- [4] R. Arratia, L. Goldstein, and L. Gordon. “Two moments suffice for Poisson approximations: The Chen–Stein method”. In: *Annals of Probability* 17.1 (1989), pp. 9–25. doi: [10.1214/aop/1176991491](https://doi.org/10.1214/aop/1176991491).
- [5] R. Arratia, L. Goldstein, and L. Gordon. “Poisson approximation and the Chen–Stein method”. In: *Statistical Science* 5.4 (1990), pp. 403–424. doi: [10.1214/ss/1177012015](https://doi.org/10.1214/ss/1177012015).
- [6] R. Balasubramanian and T. N. Shorey. “Squares in products from a block of consecutive integers”. In: *Acta Arithmetica* 65.3 (1993), pp. 213–220. doi: [10.4064/aa-65-3-213-220](https://doi.org/10.4064/aa-65-3-213-220).
- [7] A. D. Barbour. “Stein’s method and Poisson process convergence”. In: *Journal of Applied Probability* 25A (1988), pp. 175–184. doi: [10.2307/3214155](https://doi.org/10.2307/3214155).
- [8] D. Basak, K. Nath, and A. Zaharescu. “Poisson behavior for chains of small quadratic non-residues”. In: *International Mathematics Research Notices* 2024.4 (2024), pp. 3356–3390. doi: [10.1093/imrn/rnad168](https://doi.org/10.1093/imrn/rnad168).
- [9] R. de la Bretèche, V. Y. Wang, and M. W. Xu. *Random multiplicative functions and making squares from polynomial values*. 2026. arXiv: [2607.06398v1](https://arxiv.org/abs/2607.06398v1). url: <https://arxiv.org/abs/2607.06398v1>.
- [10] Y. Buttkewitz and C. Elsholtz. “Patterns and complexity of multiplicative functions”. In: *Journal of the London Mathematical Society* 84.3 (2011), pp. 578–594. doi: [10.1112/jlms/jdr026](https://doi.org/10.1112/jlms/jdr026).
- [11] S. Chatterjee and K. Soundararajan. “Random multiplicative functions in short intervals”. In: *International Mathematics Research Notices* 2012.3 (2012), pp. 479–492. doi: [10.1093/imrn/rnr023](https://doi.org/10.1093/imrn/rnr023).
- [12] Y. M. Chee, H. M. Kiah, P. Purkayastha, and C. Wang. “Cross-bifix-free codes within a constant factor of optimality”. In: *IEEE Transactions on Information Theory* 59.7 (2013), pp. 4668–4674. doi: [10.1109/TIT.2013.2252952](https://doi.org/10.1109/TIT.2013.2252952).
- [13] L. H. Y. Chen and A. Röllin. “Approximating dependent rare events”. In: *Bernoulli* 19.4 (2013), pp. 1243–1267. doi: [10.3150/12-BEJSP18](https://doi.org/10.3150/12-BEJSP18).
- [14] J. Chinis and B. Shala. “Random Chowla’s conjecture for Rademacher multiplicative functions”. In: *Transactions of the American Mathematical Society* 378.11 (2025), pp. 8025–8053. doi: [10.1090/tran/9457](https://doi.org/10.1090/tran/9457).
- [15] J. Cilleruelo and M. Z. Garaev. “Concentration of points on two and three dimensional modular hyperbolas and applications”. In: *Geometric and Functional Analysis* 21 (2011), pp. 892–904. doi: [10.1007/s00039-011-0127-6](https://doi.org/10.1007/s00039-011-0127-6). arXiv: [1007.1526](https://arxiv.org/abs/1007.1526).
- [16] P. Erdős, A. Ivić, and C. Pomerance. “On sums involving reciprocals of the largest prime factor of an integer”. In: *Glasnik Matematički, Series III* 21(41) (1986), pp. 283–300. url: <https://math.dartmouth.edu/~carlp/PDF/54.pdf>.
- [17] P. Erdős and A. Rényi. “On a new law of large numbers”. In: *Journal d’Analyse Mathématique* 23 (1970), pp. 103–111. doi: [10.1007/BF02795493](https://doi.org/10.1007/BF02795493).
- [18] J.-H. Evertse and J. H. Silverman. “Uniform bounds for the number of solutions to $Y^n = f(X)$ ”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 100.2 (1986), pp. 237–248. doi: [10.1017/S0305004100066068](https://doi.org/10.1017/S0305004100066068).

- [19] A. Földes. “The limit distribution of the length of the longest head-run”. In: *Periodica Mathematica Hungarica* 10.4 (1979), pp. 301–310. doi: [10.1007/BF02020027](https://doi.org/10.1007/BF02020027).
- [20] A. P. Godbole. “Poisson approximations for runs and patterns of rare events”. In: *Advances in Applied Probability* 23.4 (1991), pp. 851–865. doi: [10.2307/1427680](https://doi.org/10.2307/1427680).
- [21] A. P. Godbole and A. A. Schaffner. “Improved Poisson approximations for word patterns”. In: *Advances in Applied Probability* 25.2 (1993), pp. 334–347. doi: [10.2307/1427656](https://doi.org/10.2307/1427656).
- [22] L. Gordon, M. F. Schilling, and M. S. Waterman. “An extreme value theory for long head runs”. In: *Probability Theory and Related Fields* 72.2 (1986), pp. 279–287. doi: [10.1007/BF00699107](https://doi.org/10.1007/BF00699107).
- [23] A. J. Harper, K. Soundararajan, and M. W. Xu. *Distribution of random multiplicative functions in short intervals, with proper normalization*. 2026. arXiv: [2606.29040v1](https://arxiv.org/abs/2606.29040v1). url: <https://arxiv.org/abs/2606.29040v1>.
- [24] O. Klurman, I. D. Shkredov, and M. W. Xu. “On the random Chowla conjecture”. In: *Geometric and Functional Analysis* 33.3 (2023), pp. 749–777. doi: [10.1007/s00039-023-00641-y](https://doi.org/10.1007/s00039-023-00641-y).
- [25] K. Krokowski. “Poisson approximation of Rademacher functionals by the Chen–Stein method and Malliavin calculus”. In: *Communications on Stochastic Analysis* 11.2 (2017), pp. 195–222. doi: [10.31390/cosa.11.2.05](https://doi.org/10.31390/cosa.11.2.05). arXiv: [1505.01417](https://arxiv.org/abs/1505.01417).
- [26] S. Laishram and T. N. Shorey. “Number of prime divisors in a product of consecutive integers”. In: *Acta Arithmetica* 113.4 (2004), pp. 327–341. doi: [10.4064/aa113-4-3](https://doi.org/10.4064/aa113-4-3).
- [27] A. P. Mangerel. “On equal consecutive values of multiplicative functions”. In: *Discrete Analysis*, 12 (2024). doi: [10.19086/da.125450](https://doi.org/10.19086/da.125450).
- [28] K. Matomäki, M. Radziwiłł, and T. Tao. “Sign patterns of the Liouville and Möbius functions”. In: *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).
- [29] M. Munsch, I. E. Shparlinski, Y.-C. Sun, and Y. Xiao. *Large sieve inequality for sums of Legendre symbols over short intervals*. 2026. arXiv: [2604.23661v3](https://arxiv.org/abs/2604.23661v3). url: <https://arxiv.org/abs/2604.23661v3>.
- [30] J. Najnudel. “On consecutive values of random completely multiplicative functions”. In: *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).
- [31] J.-L. Nicolas and G. Robin. “Majorations explicites pour le nombre de diviseurs de N ”. In: *Canadian Mathematical Bulletin* 26.4 (1983), pp. 485–492. doi: [10.4153/CMB-1983-078-5](https://doi.org/10.4153/CMB-1983-078-5).
- [32] M. Pandey, V. Y. Wang, and M. W. Xu. “Partial sums of typical multiplicative functions over short moving intervals”. In: *Algebra & Number Theory* 18.2 (2024), pp. 389–408. doi: [10.2140/ant.2024.18.389](https://doi.org/10.2140/ant.2024.18.389). arXiv: [2207.11758](https://arxiv.org/abs/2207.11758).
- [33] G. Reinert and S. Schbath. “Compound Poisson and Poisson process approximations for occurrences of multiple words in Markov chains”. In: *Journal of Computational Biology* 5.2 (1998), pp. 223–253. doi: [10.1089/cmb.1998.5.223](https://doi.org/10.1089/cmb.1998.5.223).
- [34] A. Röllin. “On the optimality of Stein factors”. In: *Probability Approximations and Beyond*. Ed. by A. Barbour, H. P. Chan, and D. Siegmund. Vol. 205. Lecture Notes in Statistics. Springer, 2012, pp. 61–72. doi: [10.1007/978-1-4614-1966-2_5](https://doi.org/10.1007/978-1-4614-1966-2_5). arXiv: [0706.0879](https://arxiv.org/abs/0706.0879).
- [35] E. Roquain and S. Schbath. “Improved compound Poisson approximation for the number of occurrences of any rare word family in a stationary Markov chain”. In: *Advances in Applied Probability* 39.1 (2007), pp. 128–140. doi: [10.1239/aap/1175266472](https://doi.org/10.1239/aap/1175266472).
- [36] C. Runge. “Ueber ganzzahlige Lösungen von Gleichungen zwischen zwei Veränderlichen”. In: *Journal für die reine und angewandte Mathematik* 100 (1887), pp. 425–435. doi: [10.1515/crll.1887.100.425](https://doi.org/10.1515/crll.1887.100.425).
- [37] K. Soundararajan and M. W. Xu. “Central limit theorems for random multiplicative functions”. In: *Journal d’Analyse Mathématique* 151.1 (2023), pp. 343–374. doi: [10.1007/s11854-023-0331-y](https://doi.org/10.1007/s11854-023-0331-y).

- [38] T. Tao. *Products of consecutive integers with unusual anatomy*. 2026. arXiv: [2603.27990v2](https://arxiv.org/abs/2603.27990v2). url: <https://arxiv.org/abs/2603.27990v2>.
- [39] G. Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. 3rd ed. Graduate Studies in Mathematics 163. Providence, RI: American Mathematical Society, 2015. doi: [10.1090/gsm/163](https://doi.org/10.1090/gsm/163).
- [40] P. G. Walsh. “A quantitative version of Runge’s theorem on Diophantine equations”. In: *Acta Arithmetica* 62.2 (1992), pp. 157–172. doi: [10.4064/aa-62-2-157-172](https://doi.org/10.4064/aa-62-2-157-172).
- [41] P. G. Walsh. “Corrections to ‘A quantitative version of Runge’s theorem on Diophantine equations’”. In: *Acta Arithmetica* 73.4 (1995), pp. 397–398. doi: [10.4064/aa-73-4-397-398](https://doi.org/10.4064/aa-73-4-397-398).
- [42] A. Wintner. “Random factorizations and Riemann’s hypothesis”. In: *Duke Mathematical Journal* 11.2 (1944), pp. 267–275. doi: [10.1215/S0012-7094-44-01122-1](https://doi.org/10.1215/S0012-7094-44-01122-1).

Technical companion to

*Long runs and rare patterns of a random
completely multiplicative function*

Brice Pouly

Version 3

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

Abstract

This companion completes the proof package for the article, Version 3. Appendices A–E give the growing-degree Runge and bounded-height Pell calculations, prime cutoffs, scalar smoothing, marked and dictionary comparisons, conditional kernels, and the boundary and relative-bulk estimates. The common macroscopic square-product argument remains in the article.

Appendix F proves the full same-grid signed prefix comparison. It assembles the existing arithmetic inputs and an independently proved scalar tail, then gives the finite Palm–Stein inequality, the uniform reciprocal-pivot estimate and the exact one-word conditional coupling. Appendix G is a separate methodological complement: regular target configurations, signed Palm-void identities and an arithmetic obstruction to global absolute-cumulant smallness. Its vanishing Palm deficit is a consequence of Appendix F, never an input to its proof.

Reading convention. The two documents form one proof package. Appendix letters and bibliographic numbers are local; unqualified numbered statements refer to the article. All unlabelled logarithms are natural. The scope, explicit literature inputs and limitations of the Lean development, together with its seven Palomar registrations, are described in the article’s section “Formalization and Palomar registrations”.

Contents

External inputs and internal deductions	2
How to use this companion	3
A Runge coefficients and uniform Diophantine inputs	4
B Prime cutoffs and scalar smoothing	8
C Marked fields and spatial aggregation	11
D Conditional laws and the Poisson target	19
E The microscopic boundary and relative factorization	22
F Microscopic comparison by largest-odd-prime pivots	28
G Private kernels, signed Palm voids, and the absolute-cumulant obstruction	35
References	43

External inputs and internal deductions

This register identifies the non-elementary inputs, their precise role, and the calculations proved in this package. References are local to this companion; statement links lead to the relevant document. Classical finite-dimensional linear algebra, conditioning and elementary inequalities are not separate imported hypotheses. The equation numbers (2.14)–(2.15) cited for Krokowski refer to arXiv version 3; the cited Lemma 4 of Cilleruelo–Garaev refers to their arXiv text.

Input and source	Form consumed	Use and internal deduction
Prime number theorem; standard prime sums [18]	$\pi(t) \sim t/\log t$; elementary weighted consequences	Proposition B.1 proves both cutoff balances; Lemma F.3 and proposition F.5 prove the uniform reciprocal-pivot estimate.
Finite-dependency Poisson approximation [1,2]	Scalar and process theorems; $b_3 = 0$ for an exact graph	Theorems 4.1, 5.1, and 5.8; the supremum convention for variation is half the sources' norm.
Scalar Stein factors [12, (2.14)–(2.15)]	Zeroth and first differences, of orders $\lambda^{-1/2}$ and λ^{-1}	Lemma B.2 proves the exceptional-set decomposition; no Malliavin input is used.
Directional Poisson factor [4, p. 179, Lemmas 2–3]	Weighted quadratic Hessian bound and $ \Delta_{ij}g \leq 1$, for the same solution	Equation (C.3) and Theorem C.1; Hessian conjugation, independent Poisson filling and geometric mark summation are explicit.
Palm–Stein method [4,8]	Immigration–death generator and conditional-law identity	Lemma F.2 proves the finite total-variation bound with no graph assumption; the exact arithmetic forcing is proved in proposition F.7.
Quadratic ideals and units [11]; divisors [14]	Ideal factorization, unit orbits and $\tau(n) = \exp(O(\log n/\log \log n))$	Lemma A.2 gives a self-contained proof of the classical uniform count; split-product localization is internal.
Uniform polynomial-height Pell count [9, arXiv text, Lemma 4]; [17, Lemma 2.10, Remark 2.12]	$\exp(O_K(\log M/\log \log M))$ solutions in polynomial-height boxes	The estimate itself has these antecedents; Appendix A.2 fixes the normalization and height dependence needed here.
Largest-prime reciprocal asymptotics [10, eq. (1.3)]	$S_+(x) = x \exp\{-2V_x + o(\nu_x)\}$ at the precision used here	Appendix F.3 deduces the odd-pivot bound by $n = a^2v$. The direct Rankin proof is retained; no separate analytic priority is claimed.
Laishram–Shorey [13, eq. (14), p. 331]	$\omega(\prod_{j < B}(n+j)) \geq (2-\varepsilon)\pi(B)$, uniformly for $n > B$	Appendix E.2 records the quantifiers; the private-coordinate and medium-prime rank arguments are internal.
Balasubramanian–Shorey [3], as recorded in [16, pp. 55–56]	Square-product exclusion with $P^+(b) \leq k$ and the stated height/density conditions	Lemma E.3 gives the substitutions and contrapositive. No size bound on the smooth coefficient is added.
Berry's normal approximation bound [5]	Third-absolute-moment normal approximation	Appendix D uses the explicit Poisson specialization; the marked comparison and joint limiting independence are proved here.
Binary marker construction [7]	A classical cross-bifix-free family	Corollary 5.6 proves the particular binary construction and its size bound. This illustrative application is not an input to the arithmetic theorem.

Runge’s split-product estimate, the Hamming-ball argument, the raw and capped two-window profiles, the simple-incidence rank lemma and the relative bulk normalization are proved in the article–companion pair. Evertse–Silverman is contextual only. The selected-pivot comparison and Palm complements are proved in Appendices F–G, with the latter using, not establishing, the microscopic comparison. The table concerns the mathematical inputs to the written arguments. In particular the directional estimate is obtained directly from Barbour’s two lemmas, while the arithmetic comparison and Poisson filling are proved here.

Written inputs and formal coverage. This table describes the written proof. The scope and limitations of the Lean development, including its seven explicit literature inputs, are recorded in the article’s section “Formalization and Palomar registrations”. In particular, the directional input supplies the entrywise bound and the weighted quadratic bound for the same Stein solution. An entrywise bound by one is not an unweighted Euclidean operator-norm bound. The derivation in (C.3) uses the weighted and entrywise estimates.

How to use this companion

This companion and the article form a single proof package. The article gives the objects, main arguments and conclusions; the entries below identify the calculation behind each excursion and the point at which the main argument resumes. Unqualified numbers refer to the article; lettered numbers refer to this companion.

Entry from the article	Calculation here	Return to the article
Lemma 2.3	Coefficients, integrality and the growing-degree remainder in Lemma A.1.	Weighted defects, proposition 2.4.
Lemmas 3.11 and 3.12	Uniform ideal and unit counts, then localization to two square classes in A.2–A.3.	Bounded components and the deep-core estimates.
Proposition 4.2 and Theorem 4.3	Weighted prime sums, degree bounds and scalar retention in B.1–B.3.	Scalar laws or the fields of Section 5.
Theorems 5.1 and 5.3	Marginal cap, selection probabilities and collision bounds in C.1.	Uniform, typical and affine dictionaries.
Theorem 5.10(ii)	The two Stein bounds, Poisson filling and the single tail event in C.3–C.4.	The staircase; information-adapted labels in C.5 and proposition 6.2.
Theorem 6.4	Poisson atom estimates and the resolution cost in D.1–D.2.	Conditional laws in Section 6.
Theorem 7.1	Prime incidences, quotient rank and the post-quadratic bound in E.1–E.5.	The microscopic mass and the prefix law.
Theorems 7.9 and 7.11	The normalized error table, finite-cylinder clock and affine conditioning in E.6–E.7.	The two-source comparison and its affine variant.

The prefix coupling and its optional complement. Theorem 7.7 enters Appendix F. The signed input register Lemma F.1 points to the one-window and two-window estimates, local signed ranks, deep-start bound and independently proved scalar tail. The argument then proceeds through the finite Palm–Stein inequality, uniform Rankin estimate, reciprocal pivots

and exact forcing. It returns to the full same-grid field and Corollary 7.8. The restriction to the dyadic field in Corollary 7.7a needs only the stability of the hard scales under $M = 4N$ and projection of integer labels. No result of Appendix G is used along this route. Appendix G separately proves regular-cloud statements, finite signed Palm identities and the absolute-cumulant obstruction. Its vanishing Palm deficit is downstream of Appendix F. The empirical consequence Corollary 7.8a is proved in the article, immediately after Corollary 7.8: it uses concentration of overlapping-window frequencies in the independent target and the summable microscopic error, with no additional arithmetic input. Remark 7.8b then gives a finite-support obstruction to the analogous local-vector comparison after fixing f .

Support and target conventions. For interior starts, the following probabilities hold on good supports conditionally on the exposed prime signs.

Event	Values	Independent equations	Target probability
Run of length at least L	$L + 1$	L	2^{-L}
Absolute word of length B	B	B	2^{-B}
Exact run, excess e	$L + e + 2$	$L + e + 1$	2^{-L-e-1}
Exact run, excess e , sign s	$L + e + 2$	$L + e + 2$	2^{-L-e-2}

Summing the last row over signs gives 2^{-L-e-1} , and then summing over $e \geq 0$ gives 2^{-L} . With $e \leq E$, the maximal support has $Q + 1$ values, $Q = L + E + 1$. This is distinct from the base support $B = L + 1$ and from the exponential profile parameter $Q_B = 2^B$.

Two different summations. For the graph-based spatially labelled comparison of Section 5, marks are summed without weights before the pair estimate, leaving the relation profile at the base length. For an aggregated vector, the directional Stein weights are inserted first, and the maximal-length profile is required. Section C.3 keeps this distinction explicit before any asymptotic absorption. Its signed local factor can be four, not two. Full marks of base-contained starts are not right-censored run lengths.

In the selected-pivot comparison of Appendix F the signed low types are instead controlled on maximal supports and summed with their exact weights. Its dimension-free bound has no inverse minimum type probability. This is a third ledger, not a use of an aggregated Stein factor on an unrecorded spatial field.

A Runge coefficients and uniform Diophantine inputs

This appendix supplies the growing-degree Runge calculation and the bounded-height Pell argument used in the article. The two types of uniformity are different: the Runge degree is allowed to grow, whereas a component's degree and its polynomial-height exponent are fixed before the ambient scale. No height-free integral-point estimate is needed.

A.1 The split-product calculation

Lemma A.1 (Coefficient-level Runge estimate). *Under the hypotheses of Lemma 2.3, including $U \in \mathbb{Z}$, the bound $U \leq (C_0 d R)^{2d}$ holds with an absolute constant C_0 .*

Proof. The γ_ν being distinct and $d \geq 2$, we have $R \geq 1$ and G is not the square of a polynomial. Set

$$F(z) = \prod_{\nu=1}^d (1 + \gamma_\nu z)^{1/2} = \sum_{m \geq 0} c_m z^m,$$

where each square root is the analytic branch equal to 1 at 0, and

$$P(T) = \sum_{m=0}^k c_m T^{k-m}.$$

For $m \leq k$, the identity

$$\binom{1/2}{j} = (-1)^{j-1} \frac{\text{Cat}_{j-1}}{2^{2j-1}} \quad (j \geq 1)$$

implies $2^{2j} \binom{1/2}{j} \in \mathbb{Z}$ and $|\binom{1/2}{j}| \leq 1$. More explicitly,

$$c_m = \sum_{j_1 + \dots + j_d = m} \prod_{\nu=1}^d \binom{1/2}{j_\nu} \gamma_\nu^{j_\nu}.$$

The number of weak compositions of m into d parts is $\binom{m+d-1}{d-1} \leq 2^{m+d}$. Since $m \leq k = d/2$ and $|\gamma_\nu| \leq R$, we obtain simultaneously

$$2^d c_m \in \mathbb{Z}, \quad |c_m| \leq 2^{m+d} R^m \leq (8R)^d, \quad 0 \leq m \leq k. \quad (\text{A.1})$$

In particular $2^d P \in \mathbb{Z}[T]$, uniformly as d grows with the ambient scale.

If $2R \leq U \leq 4R$, the conclusion is immediate after enlarging C_0 . Assume therefore $U > 4R$. On the circle $|z| = (2R)^{-1}$,

$$|F(z)| \leq (3/2)^{d/2} =: M_d.$$

Cauchy's formula gives $|c_m| \leq M_d (2R)^m$. Since $2R/U < 1/2$,

$$\begin{aligned} |\sqrt{G(U)} - P(U)| &= U^k \left| \sum_{m \geq k+1} c_m U^{-m} \right| \\ &\leq U^k M_d \sum_{m \geq k+1} (2R/U)^m \\ &\leq \frac{2M_d (2R)^{k+1}}{U} \leq \frac{(C_1 d R)^d}{U}. \end{aligned} \quad (\text{A.2})$$

If the difference is nonzero, $\sqrt{G(U)}$ is an integer and, by (A.1), $P(U) \in 2^{-d}\mathbb{Z}$. It is therefore at least 2^{-d} in absolute value. Inequality (A.2) then implies

$$U \leq 2^d (C_1 d R)^d \leq (C_0 d R)^{2d}.$$

Assume finally $\sqrt{G(U)} = P(U)$. The polynomial

$$Q(T) = 2^{2d} G(T) - (2^d P(T))^2$$

has integer coefficients. It is nonzero, since G is not a square. The expansion at infinity gives $\sqrt{G(T)} - P(T) = O(T^{-1})$, whence

$$\deg Q \leq k - 1.$$

The coefficient of T^r in G is a symmetric polynomial containing at most $\binom{d}{r} \leq 2^d$ monomials, each of modulus at most R^d after the uniform upper bound $R \geq 1$. By (A.1), every coefficient of $2^d P$ is at most $(16R)^d$ in modulus. A coefficient of $(2^d P)^2$ is the sum of at most $d + 1$

products of two such coefficients. Consequently

$$H(2^{2d}G) \leq (8R)^{2d}, \quad H((2^dP)^2) \leq (d+1)(16R)^{2d},$$

and therefore, for an absolute constant C_2 ,

$$H(Q) \leq (C_2 d R)^{2d}.$$

This estimate is uniform in d ; it hides no step where d would be fixed. An integer zero of a nonzero integer polynomial has modulus at most $1 + H(Q)$, by the elementary Cauchy bound. After enlarging C_0 , we again obtain (2.2). $\square$

The integrality of U is used precisely in the separation by 2^{-d} . Without it the statement is false: for $G(T) = T(T+1)$ the real numbers $U_m = (\sqrt{1+4m^2} - 1)/2$ tend to infinity and satisfy $G(U_m) = m^2$.

A.2 Pell equations with varying coefficients

The uniform polynomial-height count below is a classical Pell estimate; see [9, arXiv text, Lemma 4] and [17, Lemma 2.10 and Remark 2.12]. For comparison with $x^2 - Dy^2 = E$, multiply $Az^2 - Cw^2 = e$ by A and set $x = Az$, $D = AC$, $E = Ae$. The nonsquare and polynomial-height conditions are preserved. The proof below is retained to expose its uniformity, not as a claim that this counting result is new.

Lemma A.2 (Uniform Pell count in a polynomial range). *Fix $K_0 > 0$. Let A, C be positive squarefree integers such that A/C is not a rational square, and let $e \neq 0$. Assume*

$$A, C, |e| \leq M^{K_0}.$$

The number of integer solutions (z, w) of

$$Az^2 - Cw^2 = e, \quad |z|, |w| \leq M^{K_0},$$

is

$$\exp\left(O_{K_0}\left(\frac{\log M}{\log \log M}\right)\right). \quad (\text{A.3})$$

Proof. Set $g = (A, C)$, $A = gA_0$, $C = gC_0$. Since A, C are squarefree, A_0, C_0 are squarefree, coprime, and their prime supports are disjoint. If a solution exists, $g \mid e$; write $e = ge_0$. After multiplying by A_0 and substituting $X = A_0z$, the equation becomes

$$X^2 - Dw^2 = m_0, \quad D = A_0C_0, \quad m_0 = A_0e_0. \quad (\text{A.4})$$

The product D is squarefree. The hypothesis on A/C gives $D > 1$.

In the real quadratic field $K_D = \mathbb{Q}(\sqrt{D})$, using the standard ideal-and-unit description of quadratic orders [11], the element

$$\alpha = X + w\sqrt{D}$$

belongs to the ring of integers $\mathcal{O}_{K_D}$, since $\sqrt{D}$ is an algebraic integer. Its norm is m_0 . Consequently

$$(\alpha)(\bar{\alpha}) = (m_0),$$

and the integral ideal (α) divides the ideal (m_0) . For each rational prime $p^a \parallel m_0$, the number of choices for the part above p is at most $(a+1)^2$: it equals $(a+1)^2$ if p splits, $a+1$ if it is

inert, and $2a + 1 \leq (a + 1)^2$ if it ramifies, since then $(p) = \mathfrak{p}^2$. The number of ideal divisors of (m_0) is therefore at most

$$\prod_{p^a \parallel m_0} (a + 1)^2 = \tau(|m_0|)^2.$$

Only the principal ideals $\mathfrak{a} = (\alpha)$ of absolute norm $|m_0|$ that admit a generator of norm exactly m_0 can come from a solution of (A.4). Fix such an ideal and a generator α_0 of norm m_0 . Then

$$|\sigma_1(\alpha_0)\sigma_2(\alpha_0)| = |m_0|.$$

All relevant generators are among

$$\pm \alpha_0 \varepsilon_D^n, \quad n \in \mathbb{Z},$$

where $\varepsilon_D > 1$ generates the free part of the unit group. If a unit of norm -1 exists, the norm- m_0 condition selects only one parity class of n , which cannot increase the number of solutions.

The bounds on A, C, z, w give, for the two real embeddings σ_1, σ_2 ,

$$|\sigma_j(\alpha)| \leq H := M^{O_{K_0}(1)}.$$

The inequalities

$$|\sigma_1(\alpha_0)|\varepsilon_D^n \leq H, \quad |\sigma_2(\alpha_0)|\varepsilon_D^{-n} \leq H$$

define an interval of integers n of length at most

$$\frac{2 \log H - \log |m_0|}{\log \varepsilon_D} + O(1) = O_{K_0}(\log M),$$

since every real quadratic unit $\varepsilon > 1$ satisfies $\varepsilon \geq (1 + \sqrt{5})/2$. Indeed its conjugate is $\pm \varepsilon^{-1}$ and its trace is an integer. For norm -1 , this gives $\varepsilon - \varepsilon^{-1} \geq 1$; for norm $+1$, $\varepsilon + \varepsilon^{-1} \geq 3$. Both give the asserted lower bound. Thus the number of solutions is at most

$$O_{K_0}(\tau(|m_0|)^2 \log M).$$

Since $|m_0| \leq M^{O_{K_0}(1)}$, the standard explicit divisor bound [14] gives

$$\tau(|m_0|)^2 \log M = \exp\left(O_{K_0}\left(\frac{\log M}{\log \log M}\right)\right),$$

which proves (A.3); the sign of m_0 does not affect the argument. $\square$

The exponent K_0 is fixed in every use. There is no appeal to a uniform lower bound for a regulator that deteriorates with the discriminant: the elementary absolute lower bound on a real quadratic fundamental unit already controls the number of generators in each ideal class.

A.3 Square-class localization in bounded height

In Lemma 3.12, the decisive set is $S = \{p : p \mid e \prod_{i < j} (h_i - h_j)\}$. For a solution, a prime outside S can divide only one shifted factor; its valuation there is therefore even. We enumerate the squarefree parts of only the first two factors, not of all d factors. There are at most $4^{|S|}$ pairs, and the two-factor equation has a nonzero right-hand side because $h_1 \neq h_2$.

Here is an explicit check on the order of parameters. Fix d, K first. For $M \geq 2$,

$$\log |e\Delta| \leq K \log M + \binom{d}{2} (K \log M + \log 2) = O_{d,K}(\log M).$$

Thus $4^{|S|} = \exp(O_{d,K}(\log M / \log \log M))$. Independently, every retained squarefree coefficient is at most $2M^K$, not an unrestricted product of primes in S . The coefficients, right-hand side and square variables of its Pell equation are therefore all covered by the fixed exponent $K + 1$ (increasing the ambient threshold if needed). The Pell constant does not depend on S , the offsets or the chosen pair. When the squarefree coefficients agree, divisor factorization treats the case that the nonsquare Pell lemma deliberately excludes.

For a component, $d = |J| \leq K_0$ and the normalizing coefficient is the squarefree part of $d_C \prod_{i \in I} (x + i)$. The smooth coefficient $d_C \leq \exp(O(B))$ and $|I| \leq K_0$ put this integer in one fixed polynomial range. This supplies exactly the uniformity consumed by Lemma 3.14, even when the two starts have different scales.

A.4 Why a single rational unit is not contracted

Two units force the reduced channel height to be at most L . One unit does not. For example, take $N = 2^{100}$, $L = 100$, $a = 103$, $b = 102$ and $t = 12427947061061072563693168687$, with $x = bt$, $y = at$. Both starts lie in $[N, 2N)$, and $103i = 102j$ has only $(i, j) = (0, 0)$ in $\{-1, \dots, 99\}^2$. Yet $103 > B = 101$, and $t \equiv 100 \pmod{103}$ gives $v_{103}(x) = 0$, $v_{103}(y) = 1$. No vertex of the first window is divisible by 103, so the second occurrence is pinned rather than isolated with the first. This is why the article defines $\sigma = 0$, $D^\# = D$ and $c^\# = c$ when $m \leq 1$. The aligned deep-core proof removes the at most two components meeting a single unit directly; it never assigns that unit a false private-kernel identity.

A.5 Uniformity across the master interval

All vertices are $O(M)$ and at least $M^\delta - 1$. Squarefree B -smooth coefficients have size $M^{O(1)}$, while a fixed number of offsets and smooth choices contributes $M^{o(1)}$. On a larger-start terminal slice $X \leq \max(x, y) < 2X$, $B/\log X$ remains in a fixed compact interval depending only on $(\delta, \beta_-, \beta_+)$. Thus the Pell estimates can be applied at scale X there, even when the other start is smaller. The proof never identifies a macroscopic cross-block pair with an intra-block pair, and never applies a fixed-ratio theorem with a growing ratio.

B Prime cutoffs and scalar smoothing

This appendix verifies the free cutoff before specializing it to either saddle. The distinction matters: the hard route deletes all exceptional starts, whereas the soft scalar route leaves them in the Stein equation.

B.1 A uniform weighted prime sum

Proposition B.1 (Free Rankin calculation and both saddles). *The bounds and expansions in proposition 4.2 hold uniformly for w in a fixed multiplicative neighbourhood of $\sqrt{H \log H}$, where $H = \log N$.*

Proof. Every Y -defective integer has the unique form $n = a^2 s$ with s squarefree and Y -smooth. Therefore, for $0 < \zeta < 1/2$,

$$\frac{A(X, Y)}{X} \leq X^{-\zeta} \prod_{p \leq Y} (1 + p^{-1+\zeta}). \quad (\text{B.1})$$

Indeed sum $\sqrt{X/s}$ over $s \leq X$ and use $s^{-1/2} \leq X^{1/2-\zeta} s^{-1+\zeta}$. Put $\nu = H/w$, choose the larger solution $u \geq 1$ of $e^u/u = \nu$, and set $\zeta = u/w$. Uniformly in the stated range,

$$u \sim \frac{1}{2} \log H, \quad \nu \asymp \sqrt{H/\log H}, \quad \zeta \asymp \sqrt{\log H/H}, \quad \log H = o(\nu).$$

For X/N in a fixed compact subset of $(0, \infty)$, $\zeta \log X = u\nu + o(\nu)$.

We require the prime number theorem inside the weighted Euler product, not merely an upper bound with an unspecified constant. Let $R(t) = \pi(t) - \text{Li}(t)$. Given $\epsilon > 0$, choose a fixed lower split point beyond which $|R(t)| \leq \epsilon t / \log t$. Stieltjes integration by parts gives, uniformly in w ,

$$\begin{aligned} \sum_{p \leq e^w} p^{-1+\zeta} &= \int_2^{e^w} \frac{t^{-1+\zeta}}{\log t} dt + O_\epsilon(1) + O\left(\epsilon \frac{e^u}{w} + \epsilon \int_2^{e^w} \frac{t^{-1+\zeta}}{\log t} dt\right) \\ &= \text{Ei}(u) + o(\nu). \end{aligned}$$

For the second equality first let $N \rightarrow \infty$ and then $\epsilon \downarrow 0$. The substitution $z = \zeta \log t$ costs only $O(\log(1/\zeta)) = o(\nu)$ at the lower endpoint, and $\text{Ei}(u) \asymp e^u/u = \nu$. Also $\sum_p p^{-2+2\zeta} = O(1)$, so logarithms of the Euler factors give

$$\sum_{p \leq e^w} \log(1 + p^{-1+\zeta}) = \text{Ei}(u) + o(\nu). \quad (\text{B.2})$$

Floors in $Y = \lfloor e^w \rfloor$ alter these estimates by a negligible quantity. Equations (B.1)–(B.2) yield

$$\log(A(X, Y)/X) \leq -\nu u + \text{Ei}(u) + o(\nu) = -D(\nu) + o(\nu).$$

A bad integer contaminates at most B windows; $\log B = o(\nu)$ absorbs this factor. The graph-edge estimate in the article gives $E_Y/N^2 \leq e^{-w+o(\nu)} + N^{-1+o(1)}$. A window uses at most CBH/w large primes, and each prime meets at most $CB(N/Y + 1)$ starts. Hence

$$\Delta_Y \ll B^2(H/w)(N/e^w + 1), \quad \Delta_Y/N \leq e^{-w+o(\nu)} + N^{-1+o(1)}. \quad (\text{B.3})$$

This maximum-degree bound applies to the graph on all starts, including exceptional ones.

For a unified expansion write $V_a = aD(H/V_a)$ for $a = 1$ or 2 , $\nu_a = H/V_a$, and $u_a = u(\nu_a)$. Since $D' = u > 0$, the defining function decreases strictly in V_a . At zero its sign is positive; at H/e its sign is negative for sufficiently large H . Thus the root is unique. The parametric identities are

$$\nu_a = e^{u_a}/u_a, \quad V_a = a(e^{u_a} - \text{Ei}(u_a)).$$

Using $\text{Ei}(u) = e^u u^{-1}(1 + u^{-1} + O(u^{-2}))$ gives

$$\log H = 2u_a - \log u_a + \log a + O(u_a^{-1}), \quad u_a = \frac{1}{2}(\log H + \log \log H - \log(2a)) + o(1),$$

and $V_a^2/H = a(u_a - 1 + O(u_a^{-1}))$. Substitution gives (4.6) and (4.7). $\square$

The second-order constants belong to this PNT main-term optimization. They are not matching lower bounds for the probability distance.

B.2 A finite soft-exception Stein lemma

Lemma B.2 (Retention of exceptional indicators). *Let a finite indicator family $(I_\alpha)_{\alpha \in \mathcal{I}}$ have an exact dependency graph conditionally on $\mathcal{F}$. Split $\mathcal{I} = \mathcal{G} \sqcup \mathcal{X}$ deterministically, and assume $\mathbb{E}(I_\alpha \mid \mathcal{F}) = p$ for every $\alpha \in \mathcal{G}$. Put $W = \sum I_\alpha$, $\lambda = |\mathcal{I}|p$, $\lambda_{\mathcal{X}} = |\mathcal{X}|p$ and $W_{\mathcal{X}} = \sum_{\alpha \in \mathcal{X}} I_\alpha$. Let $p_\beta(\omega) = \mathbb{E}(I_\beta \mid \mathcal{F})(\omega)$. The mean conditional Poisson distance is at most a constant times*

$$(1 \wedge \lambda^{-1}) \left[|\mathcal{G}|p^2 + p \sum_{\alpha \in \mathcal{G}} \sum_{\beta \in B_\alpha \setminus \{\alpha\}} \mathbb{E} I_\beta + \sum_{\alpha \in \mathcal{G}} \sum_{\beta \in B_\alpha \setminus \{\alpha\}} \mathbb{E}(I_\alpha I_\beta) \right] + (1 \wedge \lambda^{-1/2})(\lambda_{\mathcal{X}} + \mathbb{E} W_{\mathcal{X}}). \quad (\text{B.4})$$

The two minimum factors are interpreted as one if $\lambda = 0$. No marginal assumption is made for exceptional neighbours.

Proof. If $\lambda = 0$, all good indicators vanish almost surely and $d_{\text{TV}}(\mathcal{L}(W \mid \mathcal{F}), \delta_0) \leq \mathbb{E}(W_{\mathcal{X}} \mid \mathcal{F})$; averaging proves the claim. For $\lambda > 0$ and a target test set $A \subset \mathbb{N}_0$, let $f = f_{\lambda, A}$ solve $\lambda f(k+1) - kf(k) = \mathbf{1}_A(k) - \text{Pois}(\lambda)(A)$. The standard factors are

$$\|f\|_\infty \leq 1 \wedge \sqrt{2/(e\lambda)}, \quad \|\Delta f\|_\infty \leq (1 - e^{-\lambda})/\lambda \leq 1 \wedge \lambda^{-1}$$

[12, (2.15)]; see also [15, (2.4)]. For $\alpha \in \mathcal{G}$ set $U_\alpha = \sum_{\beta \in B_\alpha \setminus \{\alpha\}} I_\beta$ and $V_\alpha = W - I_\alpha - U_\alpha$. On each fibre I_α is independent of V_α . Add and subtract $p\mathbb{E}[f(V_\alpha + 1) \mid \mathcal{F}]$. Telescoping gives

$$\begin{aligned} & |p\mathbb{E}[f(W+1) \mid \mathcal{F}] - \mathbb{E}[I_\alpha f(W) \mid \mathcal{F}]| \\ & \leq \|\Delta f\|_\infty \left[p^2 + p \sum_{\beta \in B_\alpha \setminus \{\alpha\}} p_\beta + \sum_{\beta \in B_\alpha \setminus \{\alpha\}} \mathbb{E}(I_\alpha I_\beta \mid \mathcal{F}) \right]. \end{aligned}$$

The remaining term is bounded directly by

$$|\mathbb{E}[\lambda_{\mathcal{X}} f(W+1) - W_{\mathcal{X}} f(W) \mid \mathcal{F}]| \leq \|f\|_\infty (\lambda_{\mathcal{X}} + \mathbb{E}[W_{\mathcal{X}} \mid \mathcal{F}]).$$

The bound is uniform in the test set, so taking the supremum on each fibre and then averaging proves (B.4). $\square$

B.3 Applying the lemma at a movable cutoff

Take $\mathcal{X} = D_Y$, $p = 2^{-L}$ and $\lambda = Np > 0$. Set $\ell = \log^+ \lambda$ and $K = \max(1, \lambda)$. The elementary inequalities

$$(1 \wedge \lambda^{-1})\lambda^2 \leq K, \quad (1 \wedge \lambda^{-1})(\lambda^2 + \lambda) \leq 2K, \quad (1 \wedge \lambda^{-1/2})\lambda \leq \sqrt{K}$$

hold on both sides of intensity one. For neighbour marginals use the maximum degree of the graph on all sites:

$$p \sum_{x \in \mathcal{G}} \sum_{y \in B_x \setminus \{x\}} \mathbb{E} J_{y,L} \leq p \Delta_Y \mathbb{E} Z_{N,L} \leq p \Delta_Y (\lambda + p \mathcal{M}_B).$$

The first-difference factor makes this at most $K(\Delta_Y/N)(1 + N^{-1/2+o(1)})$. The true marginals of exceptional neighbours have not been replaced by p . For the joint sum, overlapping starts contribute zero and the local bound treats touching pairs; the separated pairs give

$$\sum_{x \in \mathcal{G}} \sum_{y \in B_x \setminus \{x\}} \mathbb{P}(J_{x,L} J_{y,L} = 1) \ll p^2(N^{1+o(1)} + N\Delta_Y + R_2(N, L)).$$

Since $R_2/N^2 \ll N^{-1/3+\varepsilon/4}(1 + \lambda^{-1})$, its contribution, after the Stein factor, is at most $KN^{-1/3+\varepsilon/2}$ after enlarging the threshold. The diagonal and local polynomial costs are smaller. Finally the exceptional zeroth-order term is at most

$$C\sqrt{K} \left(\frac{\mathcal{M}_B}{N} + \frac{|D_Y|}{N} \right) \leq Ce^{\ell/2 - D(H/w) + o(H/w)} + KN^{-1/3+\varepsilon/2}.$$

Together with (B.3), these estimates give (4.10) for all intensities.

At $w = \widehat{V}_N$, let $\nu = H/w$ and choose a common $\eta_N \geq 0$, $\eta_N \rightarrow 0$, dominating the displayed remainder terms. Put

$$q = e^{-(w-\ell)/2 + \eta_N \nu}, \quad r = e^{\ell - w + \eta_N \nu}, \quad P = N^{-1/3+\varepsilon/2}.$$

Up to an absolute constant, the error is at most $q + r + KP$. If $q + N^{-1/3+\varepsilon} \geq 1$, the trivial total-variation bound suffices. Otherwise $q < 1$ implies $\ell \leq w$ and $r \leq q$. Since $w = o(\log N)$, eventually $K \leq e^w \leq N^{\varepsilon/2}$, so $KP \leq N^{-1/3+\varepsilon}$. This proves the soft minimum bound, including $\lambda < 1$, at the same cutoff and on the full $\mathcal{F}_Y$. All thresholds depend only on the fixed band and the error margin, not on its length parameter. The numerical constants in a coarser ledger do not affect the claimed asymptotic bound.

All terms admit their conditional versions: integrate the fibrewise bound over $C \in \mathcal{F}_Y$ and divide by $\mathbb{P}(C)$. A single factor e^I multiplies the ledger, where $I = -\log \mathbb{P}(C)$; the leading soft exponent consequently has sufficient budget $2I + \log^+ \lambda < \widehat{V}$. This argument applies only to events measurable in the chosen prime field. It does not turn arbitrary post-selection on a run statistic into an arithmetic conditioning event.

C Marked fields and spatial aggregation

The distinction between a labelled field and its aggregate is already visible in a finite family. This appendix gives the finite inequalities first; all passages to countably many marks are then controlled by a single tail event.

C.1 Uniform and averaged dictionary comparisons

The marginal-ceiling calculation in the proof of Theorem 5.1 gives the finite inequality

$$\mathbb{E}\Delta_{\mathcal{W}} \ll a(\mathcal{M}_B^{\text{val}} + |D_Y^{\text{val}}|) + a^2(NB + E_Y + R_2^{\text{val}, [1/a]}(N, B)) + \Lambda\Omega(\mathcal{W}),$$

where $a = m2^{-B}$ and $\Lambda = Na$. The comparison retains the full site–word index set; only the error terms have been summed over words. We verify here the substitutions and the uniform absorption of secondary terms. With $T = N/\Lambda$ and $Q_B = Nm/\Lambda \geq T$, the three capped-profile terms become, respectively,

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

These substitutions are exact because $a = \Lambda/N$ and $Q_B = Nm/\Lambda$. The restrictions $1 \leq m \leq 2^B$ also give $0 < a \leq 1$, so the arithmetic cap $T = 1/a$ is always at least one. The secondary edge contribution is absorbed in every admissible intensity regime, not only at criticality: $m \geq \Lambda/N$ implies

$$\frac{\Lambda^2}{N} \leq N^{-1/3} \Lambda^{4/3} m^{2/3}.$$

Factors $N^{o(1)}$ in these smaller terms are absorbed by fixing a smaller arithmetic error exponent first. At bounded positive Λ and $m \leq N^{1/2-\eta}$, the powers of N in the three terms, with $\varepsilon = \eta/6$, are at most

$$-\frac{1}{4}, \quad -\frac{\eta}{2}, \quad -\frac{1}{3} + \frac{\eta}{6} \leq -\frac{\eta}{2} \quad (0 < \eta < 1/2).$$

Thus all three are $O(N^{-\eta/2})$, uniformly in the dictionary. No independent random marks are added to the multiplicative sequence, and no word identity is lost when the error terms are summed.

Averaging the finite dictionary comparison

We verify (5.14) before using any asymptotic profile. Fix $N \geq 2$, $B \geq 1$, $Y > 2B$, a deterministic start mask $A \subset I_N$, and a random size- m dictionary, $1 \leq m \leq 2^B$, independent of the prime signs. Only the inclusion probabilities (5.11) are needed; both ensembles in the article have them. For each realization of the dictionary, apply the process inequality on the finite coordinate set $A \times \mathcal{W}$.

Deletion and the exact graph. Let $D = D_Y^{\text{val}} \cap A$. Removing real indicators has mean cost, after averaging the dictionary and the prime environment,

$$\sum_{x \in D} \mathbb{E}_f \mathbb{P}_{\mathcal{W}}(\mathbf{f}_x \in \mathcal{W}) = a|D|.$$

The Poisson target has total deleted mean $a|D|$. This uses no uniformity of a bad window's law. At a retained site every coordinate has conditional mean $p_B = 2^{-B}$. Its neighbours include all words at its own site, at each overlapping support and at each support sharing an odd prime coordinate above Y . Families with no edge depend on disjoint remaining prime signs, so their conditional independence is exact. The site graph and its good set do not depend on $\mathcal{W}$. The first Chen–Stein sum, including each diagonal coordinate, is bounded by $Ca^2(|A|B + E_Y(A))$.

The second sum. At the same site its off-diagonal terms vanish by mutual exclusivity; a coordinate's own diagonal belongs to the first sum only. For distinct sites the sum over word labels is exactly $\mathbb{P}_f(\mathbf{f}_x, \mathbf{f}_y \in \mathcal{W} \mid \mathcal{F}_Y)$. After both averages it is $b + (a - b)\mathbb{P}_f(\mathbf{f}_x = \mathbf{f}_y)$, as in (5.12). At a good local pair with displacement $d < B$, write the union values as $z_0, \dots, z_{B+d-1}$. The equations $z_j = z_{j+d}$, $0 \leq j < B$, form the d paths with residue classes modulo d . There are $B + d$ vertices and B edges. A prime above $Y > 2B$ can occur at only one union vertex, and each vertex has such an odd coordinate, so the conditional vector is uniform. Its collision probability is exactly 2^{-B} on every fibre.

For disjoint supports, let $v(n)$ denote the complete binary valuation vector at the finite set of primes dividing the union. The equality matrix has rows $v(x - 1 + j) + v(y - 1 + j)$. Its

coefficient kernel maps injectively by

$$t \mapsto (t, t), \quad \sum_j t_j (v(x-1+j) + v(y-1+j)) = 0,$$

into the coefficient kernel of the stacked value matrix. The equality right-hand side is zero, so its probability is exactly $2^{-B+\kappa}$, regardless of whether either value window is individually uniform. This proves (5.13) with the full, not the large-prime-truncated, nullity. It is used only after averaging the prime fibres. Consequently a separated edge costs at most

$$a^2 + ap_B + ap_B(2^{\rho_B^{\text{val}}} - 1) \leq 2a^2 + ap_B(2^{\rho_B^{\text{val}}} - 1).$$

The local sum is $O(a^2|A|B)$ and the separated sum is $O(a^2 E_Y(A) + ap_B R_2^{\text{val}}(A, B))$. All site pairs here are ordered; the two orientations are included. Adding the two deletion costs proves (5.14) with an absolute implied constant. This finite inequality also covers $m = 1$, $B = 1$ and the empty mask.

Normalization and the order of averaging. With $a = \Lambda/N$ and $p_B = \Lambda/(Nm)$, the two terms of the coarse profile give, separately,

$$ap_B N^{5/3} = \frac{\Lambda^2}{m} N^{-1/3}, \quad ap_B N^{2/3} 2^B = \Lambda N^{-1/3}.$$

The factor N^ε is inserted only after these exact identities. For an independent-letter word field, distinct disjoint supports need no edge, and every local collision again has probability p_B . Its averaged Poisson error is therefore $O(a^2|A|B)$, without a restriction on m . This is a fresh finite calculation, not an application of the uniform dictionary theorem outside its range.

The supremum over target tests is taken for each fixed dictionary and prime fibre *before* averaging their distances. Finite sums over the auxiliary dictionary law then give the claimed mean bound. The class of good dictionaries supplied by Markov controls all coordinate restrictions simultaneously by contraction. No mask may use the observed prime signs. A dictionary-dependent deterministic restriction is harmless once that dictionary has been selected. No coupling over different scales, or almost-sure choice of a good dictionary at every scale, is inferred.

C.2 Local pairs and the role of the maximal support

From here on, $p = 2^{-L}$, $q_e = 2^{-e-1}$, $p_e = pq_e$, $\lambda = Np$ and $0 \leq e \leq E$. The largest row length is $Q = L + E + 1$; its support has $Q + 1$ vertices. An unsigned mark of excess e imposes $L + e + 1$ relative equations; prescribing its sign fixes all $L + e + 2$ values on the same support.

A good maximal support contains no Y -defective vertex and has diameter $Q < Y$. Every exact-mark system on that support consequently has full rank above Y , with conditional marginal p_e . If two actual supports intersect, their union has diameter at most $2Q < Y$; every vertex of the union then has a prime coordinate private to that union. For two compatible exact runs the only possible intersecting cases are a common terminal/border vertex or a common terminal edge. Equivalently, for $x < y$, they satisfy $y - x = L + e + 1$ or $y - x = L + e$. The union of the two constraint trees has cycle rank at most one, so its probability is at most $2p_e p_f$.

The actual supports can also be disjoint while the *maximal* supports intersect. In that case all their private coordinates are still distinct, and the probability is $p_e p_f$. Thus the bound

$$\mathbb{P}(K_{x,e} K_{y,f} = 1 \mid \mathcal{F}_Y) \leq 2p_e p_f \quad (0 < |x - y| \leq Q) \quad (\text{C.1})$$

holds for every good pair; it does not assert that only two separations are possible among pairs of intersecting maximal supports. At the same site, different exact marks are mutually exclusive.

For $|x - y| > Q$, zero-extension of the coefficient vectors embeds the mixed relation space into the two-window space at row length Q . Therefore

$$\mathbb{P}(K_{x,e}K_{y,f} = 1) \leq p_e p_f 2^{\rho_Q(x,y)}. \quad (\text{C.2})$$

This is an averaged, not necessarily fibrewise, estimate: the small-prime phase in the affine formula can vary with the environment.

For signed marks, put $p_{e,s} = p_e/2$ and let $a = L + e$, $b = L + f$. On a good local union the raw value vector is uniform. For $x < y$, $d = y - x$, the exact joint probabilities are

$$\mathbb{P}(K_{x,e}^s K_{y,f}^t = 1 \mid \mathcal{F}_Y) = \begin{cases} 0, & d < a, \\ 4p_{e,s}p_{f,t}\mathbf{1}_{\{t=-s\}}, & d = a, \\ 2p_{e,s}p_{f,t}\mathbf{1}_{\{t=s\}}, & d = a + 1, \\ p_{e,s}p_{f,t}, & a + 1 < d \leq Q. \end{cases}$$

The final case concerns disjoint actual supports inside intersecting maximal supports. The factor four, rather than two, is needed when signs are prescribed. After summing signs one recovers (C.1). For separated signed systems, zero-extension instead gives the bound $p_{e,s}p_{f,t}2^{\rho_{Q+1}^{\text{val}}(x,y)}$. Proposition 3.24 controls its weighted sum.

C.3 The directional multivariate Stein factor

Let $\pi = \bigotimes_{i=1}^d \text{Pois}(\lambda_i)$, with $d \geq 2$, $\lambda_i > 0$, $\lambda' = \sum_i \lambda_i$ and $\mu_i = \lambda_i/\lambda'$. For a test set $A \subseteq \mathbb{N}_0^d$, use Barbour's immigration-death solution, with its sign chosen so that

$$\mathcal{A}g(z) = \mathbf{1}_A(z) - \pi(A), \quad \mathcal{A}g(z) = \sum_i \{\lambda_i \Delta_i g(z) + z_i [g(z - \mathbf{e}_i) - g(z)]\}.$$

At $z_i = 0$ the corresponding death term is zero. Write $H(z) = (\Delta_{ij}g(z))_{i,j}$ and $c(\lambda') = (1 + 2 \log^+(2\lambda'))/2$. Barbour's Lemmas 2 and 3 [4, p. 179] give, for this same solution,

$$|\Delta_{ij}g(z)| \leq 1, \quad |\alpha^T H(z) \alpha| \leq c(\lambda') \sum_i \frac{\alpha_i^2}{\lambda_i}.$$

The weighted bound, stated there for integer directions, extends to rational directions by homogeneity and to real directions by continuity. The literature input consists of the solution and these analytic bounds, not an arithmetic or aggregated total-variation estimate.

For clarity, the constant entrywise bound also follows directly from the solution's semigroup representation. Couple the processes started at z , $z + \mathbf{e}_i$, $z + \mathbf{e}_j$ and $z + \mathbf{e}_i + \mathbf{e}_j$ by adding two independent particles of rate-one exponential lifetimes. The second difference vanishes once either added particle dies; before that time the second difference of an indicator is at most two in absolute value. Its time integral is thus bounded by $\int_0^\infty 2e^{-2t} dt = 1$. The argument includes $i = j$.

To obtain the weighted entrywise bound, conjugate the symmetric matrix $H(z)$ by $D = \text{diag}(\sqrt{\lambda_i})$. The quadratic inequality gives $\|DH(z)D\|_{2 \rightarrow 2} \leq c(\lambda')$. Bounding each entry by

this norm, and using the constant bound above, yields

$$|\Delta_{ij}g(z)| \leq \min \left\{ 1, \frac{1 + 2 \log^+(2\lambda')}{2\lambda' \sqrt{\mu_i \mu_j}} \right\} \leq \min \left\{ 1, \frac{C(1 + \log^+(2\lambda'))}{\lambda' \sqrt{\mu_i \mu_j}} \right\}. \quad (\text{C.3})$$

The logarithmic term in multivariate Stein factors is discussed in [15]; the two bounds used above are taken directly from Barbour. For the truncated geometric marks, $S_E = \sum_{e=0}^E q_e \in [1/2, 1]$, $\lambda' = \lambda S_E$ and $\mu_e = q_e/S_E$. The second expression is thus bounded by $C(1 + \log^+(2\lambda))/(\lambda \sqrt{q_e q_f})$. The summability that removes dependence on the number of marks is

$$\sum_{e \geq 0} \sqrt{q_e} = 1 + \sqrt{2}. \quad (\text{C.4})$$

Theorem C.1 (Finite aggregated comparison). *Let $X_e = \sum_{x \in I_N} K_{x,e}$, and let $D_Y(Q)$ be the bad maximal supports. Put*

$$s(\lambda) = 1 \wedge \frac{C(1 + \log^+(2\lambda))}{\lambda}.$$

Uniformly while $L + 1$ and $Q + 1$ remain in a fixed positive logarithmic band and $2Q < Y$, one has

$$\begin{aligned} \mathbb{E} d_{\text{TV}} \left(\mathcal{L}((X_e)_{0 \leq e \leq E} \mid \mathcal{F}_Y), \bigotimes_{e=0}^E \text{Pois}(\lambda q_e) \right) \\ \ll p\mathcal{M}_B + p|D_Y(Q)| + s(\lambda)p^2(NQ + E_Y(Q) + R_2(N, Q)). \end{aligned} \quad (\text{C.5})$$

Here $\mathcal{M}_B$ is the one-window defect mass at the base length, and $E_Y(Q)$ is the ordered large-prime edge count for maximal supports.

Proof. Write $D = D_Y(Q)$ and retain the good-site counts $G_e = \sum_{x \notin D} K_{x,e}$. Independently of the arithmetic variables, introduce independent $Z_e \sim \text{Pois}(\zeta_e)$ with $\zeta_e = |D|pq_e$, and put $W = G + Z$. Removing the actual bad-site marks and adding this Poisson filling changes the law by at most

$$\sum_{x \in D} \mathbb{P}(J_{x,L} = 1) + \sum_{e=0}^E \zeta_e \ll p\mathcal{M}_B + p|D|.$$

This is an averaged conditional coupling bound. It has no factor E , since the actual marks are disjoint refinements of $J_{x,L}$ and $\sum_{e \leq E} q_e \leq 1$.

Use the Stein solution for the *full* target means λq_e . For the coordinate vector $\mathbf{e}_e$ and $\Delta_e g(w) = g(w + \mathbf{e}_e) - g(w)$, Poisson summation gives, conditionally on the good indicators and $\mathcal{F}_Y$,

$$\mathbb{E}[Z_e \Delta_e g(W - \mathbf{e}_e) \mid G, \mathcal{F}_Y] = \zeta_e \mathbb{E}[\Delta_e g(W) \mid G, \mathcal{F}_Y].$$

Indeed $z\mathbb{P}(Z_e = z) = \zeta_e \mathbb{P}(Z_e = z - 1)$ for $z \geq 1$; the term with $Z_e = 0$ is zero. Thus the filling terms cancel the ζ_e part of the birth rates in

$$\sum_{e=0}^E \{\lambda q_e \Delta_e g(W) - W_e \Delta_e g(W - \mathbf{e}_e)\}.$$

Only the good indicators remain to be telescoped. The target parameters are still $\lambda' = \lambda S_E$ and $\mu_e = q_e/S_E$, so $\lambda' \sqrt{\mu_e \mu_f} = \lambda \sqrt{q_e q_f}$ exactly. No comparison of the good-site intensity with λ is needed, even if every site was removed. The independent vector Z stays in the outside count during each telescoping step, so the good-site dependency graph is unchanged. For good sites, apply the multivariate dependency-graph Stein argument [4,15] to the *counts by mark*, not to

counts by site and mark. On the first-difference telescoping terms each pair of types contributes its joint or product mass multiplied by $|\Delta_{ef}g|$. The product masses are $p_e p_f$. For local pairs use (C.1); for separated pairs use (C.2) after averaging. Decompose $2^{\rho_Q} = 1 + (2^{\rho_Q} - 1)$ and sum the second part by $R_2(N, Q)$. The diagonal and all local pairs are bounded by $O(NQp^2)$.

Using the first bound in (C.3) sums the mark weights to at most one. Using the second sums them to $\sum_{e,f} \sqrt{q_e q_f} \leq (1 + \sqrt{2})^2$. Take the better of these two estimates. This proves (C.5) when $E \geq 1$. For $E = 0$, retain both signs during the same comparison, so there are two positive target coordinates. At fixed excesses the factor does not depend on the chosen signs; summing their probabilities before the relative-pair bound again gives $R_2(N, Q)$. The signed local factor is at most four, an absolute constant. Thus the result also covers $E = 0$ without extending the literature statement to dimension one. Notice that the arithmetic argument is used at Q , rather than merely at L : after insertion of the directional weights, the unweighted folding identity for marks is not the inequality being summed. $\square$

At $Y = Y_N^\sharp$, the first two arithmetic error sources in (C.5) are bounded by

$$\lambda e^{-V_N + o(\nu_N)} + N^{-1/2 + o(1)} \lambda, \quad \lambda(1 + \log^+(2\lambda)) e^{-V_N + o(\nu_N)}.$$

For clarity, before absorbing subpolynomial factors the profile term is

$$p^2 R_2(N, Q) \ll N^{-1/3 + \varepsilon} (\lambda^2 + \lambda 2^{E+2}).$$

If $I + \log^+ \lambda = O(V_N)$ and $E = O(V_N)$, multiplication by $e^I s(\lambda)$ leaves only an $N^{o(1)}$ factor. Start with a smaller arithmetic ε to obtain the final remainder $N^{-1/3 + \varepsilon}$ at any prescribed positive ε . This is the step that prevents an unrecorded exponential mark-cutoff loss.

For aggregation with signs retained, replace q_e by $q_{e,s} = q_e/2$ and $R_2(N, Q)$ by $R_2^{\text{val}}(N, Q+1)$ in (C.5). The local factor is at most four and $\sum_{e,s} \sqrt{q_{e,s}} = \sqrt{2}(1 + \sqrt{2})$, both fixed constants. The full-value profile at $Q+1$ vertices has the same envelope as the relative profile at Q rows. It therefore leaves the same polynomial remainder and the same one-factor budget. By contrast, the unweighted labelled argument sums signs and marks first and continues to use $R_2(N, L)$ at the base length.

C.4 Removing the mark cutoff

Assume $I + \log^+ \lambda \leq V_N - c\nu_N$, and choose $0 < c' < c'' < c$. Set

$$E_N = \left\lceil \frac{\log^+ \lambda + c'' \nu_N}{\log 2} \right\rceil.$$

Then $E_N = O(V_N) = o(\log N)$, and the target tail has mass $\mu_N = \lambda 2^{-E_N - 1} \leq e^{-c'' \nu_N}$. The actual tail event is exactly a hit at length $L + E_N + 1$. The scalar hard-conditioning inequality at that shifted length gives

$$\mathbb{P}(\exists x, e > E_N : K_{x,e} = 1 \mid C) \leq \mu_N + O\left(e^I \mu_N (e^{-V_N + o(\nu_N)} + N^{-1/3 + \varepsilon})\right).$$

It is $O(\mu_N)$ under the displayed budget. Couple each tail to zero, apply the finite theorem, and absorb $\log(1 + \log^+ \lambda) = o(\nu_N)$. This gives

$$d_{\text{TV}} \left(\mathcal{L}((X_e)_{e \geq 0} \mid C), \bigotimes_{e \geq 0} \text{Pois}(\lambda q_e) \right) \ll e^{-c' \nu_N} + N^{-1/3 + \varepsilon}.$$

For finite-mass sequences, tail summation is a measurable bijection, with inverse $X_e = T_e - T_{e+1}$. The entire threshold path has exactly the same total-variation distance.

C.5 Keeping the spatial labels

For the labelled comparison no directional Stein factor is inserted. The identities

$$\sum_e p_e \leq p, \quad \sum_{e,f} \mathbb{P}(K_{x,e} K_{y,f} = 1) \leq \mathbb{P}(J_{x,L} J_{y,L} = 1)$$

fold the separated-pair contribution at the base length. Their signed versions hold after summing both signs. Thus the relation profile is $R_2(N, L)$, whereas the graph and deleted set use the maximal support.

Free-cutoff bound and tails. Let $H = \log N$, $w \in [V_N, \widehat{V}_N]$, $Y = \lfloor e^w \rfloor$ and $\nu_w = H/w$. Suppose $C \in \mathcal{F}_Y$ has $I = -\log \mathbb{P}(C) \geq 0$, and $\lambda = N2^{-L} \geq 1$ with $\ell = \log \lambda$ and $I + \ell = O(V_N)$. Choose

$$E = \left\lceil \frac{I + \ell + w}{\log 2} \right\rceil, \quad Q = L + E + 1.$$

Since $L = (H - \ell)/\log 2$, the maximal number of values is

$$Q + 1 = \frac{H + I + w}{\log 2} + O(1) = \log_2 N + O(V_N).$$

Both the base and maximal lengths remain in a fixed positive logarithmic band, and $Y > 2(Q + 1)$ eventually, uniformly in these parameters. By Theorem 5.8 and the uniform free-cutoff estimates, the finite field distance, multiplied by the information cost, is at most

$$C \left\{ e^{I + \ell - D(H/w) + o(\nu_w)} + e^{I + 2\ell - w + o(\nu_w)} + e^I \lambda(1 + \lambda) N^{-1/3 + \varepsilon/2 + o(1)} \right\}. \quad (\text{C.6})$$

The weighted one-window term and the $N^{-1+o(1)}$ part of the edge bound fit the polynomial term. Constants in the little-oh terms are uniform for w in the displayed interval and for the enlarged lengths. A fixed multiplicative enlargement of this interval is also permitted by proposition B.1.

An excess greater than E is exactly a threshold hit at $L + E + 1$. The unconditional first moment gives actual conditional tail cost at most

$$e^I \lambda 2^{-E-1} (1 + N^{-1/2+o(1)}) = O(e^{-w}),$$

and the target tail has no larger cost. This step uses conditioning only as the bound $\mathbb{P}(H \mid C) \leq e^I \mathbb{P}(H)$; no conditioning theorem at a different cutoff is invoked. Coupling both tails to zero adds $O(e^{-w})$ to (C.6). The full site-sign-excess space is countable, with finite expected total mass, so these tail couplings exhaust its coordinates.

In the applications below, $I + \ell \leq w \leq \widehat{V}_N$. One may instead use the common truncation $E = 3\lceil \widehat{V}_N / \log 2 \rceil$, for which $e^I \lambda 2^{-E-1} \leq \frac{1}{2} e^{-2\widehat{V}_N}$. It still has $Q + 1 = \log_2 N + O(V_N)$, so the finite comparison is unchanged and both tails are $O(e^{-2\widehat{V}_N})$. This is the truncation used in the proof of proposition 6.2; the sharper choice above is not needed for that conclusion.

The information-dependent crossing. For $0 \leq I \leq V_N$, set $F_I(w) = 2D(H/w) - w - I$. The function D is continuous and strictly increasing on the larger saddle branch. Since H/w strictly decreases with w , F_I is continuous and strictly decreasing. Its endpoint values on

$[V_N, \widehat{V}_N]$ are $V_N - I$ and $-I$. There is therefore exactly one root $w_N(I)$, including the endpoints for $I = V_N$ and $I = 0$. The first branch of $g_{N,I}(w) = \min\{D(H/w) - I, (w - I)/2\}$ decreases strictly; the second increases strictly. The root is their unique maximum on the interval. It obeys

$$D(H/w_N(I)) = \frac{w_N(I) + I}{2}, \quad \ell_{\text{lab},N}(I) = D(H/w_N(I)) - I.$$

If $\ell \leq \ell_{\text{lab},N}(I) - c\nu_w$, the first two terms in (C.6) are respectively $e^{-c\nu_w + o(\nu_w)}$ and $e^{-2c\nu_w + o(\nu_w)}$. The polynomial multiplier is $N^{o(1)}$ because $I + 2\ell \leq w_N(I) = O(V_N)$. Also $w/\nu_w = w^2/H \asymp \log H \rightarrow \infty$, so e^{-w} is smaller than any fixed $e^{-c'\nu_w}$. This proves proposition 6.2. The hypotheses are vacuous at $I = V_N$ for $\lambda \geq 1$, since then the allowed leading logarithmic intensity is zero and the margin is strictly negative.

For $I > 0$, the root satisfies $w_N(I) < \widehat{V}_N$. Hence $2D(H/w_N(I)) > 2D(H/\widehat{V}_N) = \widehat{V}_N$, and its defining equation gives $w_N(I) + I > \widehat{V}_N$. Thus $w_N(I) > \widehat{V}_N - I$, giving the claimed strict improvement of the leading budget over $\widehat{V}_N/2 - I$ without differentiating the implicitly defined root. Uniformly for $w = tV_N$, $1 \leq t \leq \widehat{V}_N/V_N$, the expansion $D(\nu) = \nu u(\nu)(1 + O(1/u(\nu)))$ gives $D(H/w)/V_N = t^{-1} + o(1)$. The saddle equation consequently yields $t^2 + \theta t - 2 = 0$ when $I/V_N \rightarrow \theta$. The positive root is $(\sqrt{\theta^2 + 8} - \theta)/2$; subtracting θ and dividing by two proves (6.8).

A fixed event and a constrained cutoff. Suppose $C \in \mathcal{F}_{\lfloor e^{w_0} \rfloor}$ for a given $w_0 \in [V_N, \widehat{V}_N]$. The prime sigma-fields increase with the cutoff. On the certified admissible interval $[w_0, \widehat{V}_N]$, $g_{N,I}$ increases up to $w_N(I)$ and decreases afterwards; hence its maximum is at $\max\{w_0, w_N(I)\}$. For any admissible w the two inequalities

$$I + \ell \leq D(H/w) - cH/w, \quad I + 2\ell \leq w - 2cH/w$$

are sufficient. In particular no event measurable only above the proposed cutoff is retained by an algebraic optimization. All cutoffs here are deterministic functions of the event and its probability, not of the realized prime environment. This is an optimization of the proved labelled bounds, not an optimality assertion about the law.

C.6 Macroscopic transport and containment

Let $L_M = \lfloor \log_2 M \rfloor - a_M$, $a_M = o(\log M)$, and $\lambda = M2^{-L_M}$. The base-contained interior starts are $2 \leq x \leq M - L_M + 1$; their number is $M - L_M$. First retain $\lceil M^{1/2} \rceil \leq x \leq M - L_M + 1$. The conditional ranks and graph construction are local and unchanged. The one-window defect masses, summed on dyadic slices, are $M^{1/2+o(1)}$; the edge estimate uses the upper height M ; and the separated-pair mass is bounded by the macroscopic profile (3.4). Thus the labelled finite comparison is

$$\ll \lambda(1 + \lambda)(e^{-V_M + o(\nu_M)} + M^{-1/3+\epsilon}).$$

Restoring the deep starts costs the bound (7.7),

$$O\left(\lambda M^{-1/2} + \exp(-c_0 \log M / \log \log M)\right).$$

The target deletion costs only $O(\lambda M^{-1/2})$. Conditioning on a compatible event in the chosen prime field multiplies the averaged error by at most e^I .

Choose $E = \lceil (I + V_M + \log(2 + \lambda)) / \log 2 \rceil$ in the hard labelled regime. The exact mark-tail identity and the global scalar bound control the real tail. Besides the intensity-linear tail term, the scalar bound has the separate deep-start error just displayed; that error is already present

in the full-field inequality. In particular it must not be silently dropped from a finite-prefix tail calculation. Both tails are absorbed by (7.14). The free-cutoff proof is identical with V_M replaced by the optimized cutoff in its two leading terms.

For aggregated counts apply Theorem C.1 on the retained macroscopic set, with $R_{2,\delta}(M, Q)$ in place of $R_2(N, Q)$. The same geometric square-root sum applies. Under $I + \log^+ \lambda \leq V_M - c\nu_M$, the deep-start correction times e^I is smaller than every fixed $e^{-c'\nu_M}$, and the aggregate mark cutoff above removes both tails. This proves the macroscopic one-factor staircase without an additional marked arithmetic input.

At exact centered level r the target mean is precisely $(M - L_M)2^{-b_M-r-1}$, not $M2^{-b_M-r-1}$. All these fields retain the full rightward mark of a base-contained start. They are not the list of blocks censored at the right end of the finite sequence. The containment diagram in the article distinguishes the two objects.

C.7 Deterministic masks and diffuse limits

A deterministic restriction of a lattice field contracts total variation. At bounded intensity, if $N^{-1} \sum_{x \in A_N} \delta_{x/N} \Rightarrow \mu$, the restricted marked field therefore converges weakly to $\text{PPP}(\lambda\mu \otimes \nu_{\text{geo}})$, along a subsequence $N2^{-L} \rightarrow \lambda$. The same statement holds for the macroscopic field, with scaling by M . It follows directly from the target Laplace functional and weak convergence of the deterministic empirical measures. There is no universal numerical rate for this last deterministic step without a quantitative assumption on the mask or the test function.

Nor is the diffuse limit a total-variation target. The actual field is supported on its lattice, while the diffuse target has positive probability $1 - e^{-\lambda\mu(\mathbb{R})}$ of a nonempty configuration outside that lattice. The approximation in total variation always takes place *before* passing from the grid to the continuum.

D Conditional laws and the Poisson target

The arithmetic theorems provide total-variation errors. This appendix records exactly when those errors are small enough to transfer a local probability, a conditional state, or a Gaussian limit. Properties of the independent Poisson target are not additional arithmetic results.

D.1 Scalar Gaussian and local estimates

Let $P_\lambda \sim \text{Pois}(\lambda)$ and let ϵ_N denote any proved conditional scalar total-variation bound in the article. If $\lambda \rightarrow \infty$,

$$\sup_t \left| \mathbb{P} \left(\frac{Z - \lambda}{\sqrt{\lambda}} \leq t \mid C \right) - \Phi(t) \right| \leq \epsilon_N + C\lambda^{-1/2}.$$

For example, write P_λ as a sum of $\lceil \lambda \rceil$ independent Poisson variables of mean at most one and use the classical Berry bound [5]. Their third absolute centered moments are $O(\lambda/\lceil \lambda \rceil)$, which gives the displayed scale. The normalization uses the deterministic comparison mean λ ; replacing it by an actual conditional mean requires a separate moment estimate and is not inferred from total variation alone.

For every integer $n \geq 1$ and every $\lambda > 0$, one has the exact form

$$\mathbb{P}(P_\lambda = n) = \frac{e^{-\delta_n}}{\sqrt{2\pi n}} \exp[-\lambda h(n/\lambda)], \quad 0 \leq \delta_n \leq \frac{1}{12n}. \quad (\text{D.1})$$

Here $h(u) = u \log u - u + 1$, and δ_n is independent of λ . To check the remainder, set

$$\delta_n = \log(n!) - (n + \tfrac{1}{2}) \log n + n - \tfrac{1}{2} \log(2\pi).$$

Stirling's limit gives $\delta_n \rightarrow 0$. With $t = (2n + 1)^{-1}$,

$$\begin{aligned} \delta_n - \delta_{n+1} &= (n + \tfrac{1}{2}) \log(1 + 1/n) - 1 = \sum_{j \geq 1} \frac{t^{2j}}{2j + 1}, \\ 0 \leq \delta_n - \delta_{n+1} &\leq \frac{t^2}{3(1 - t^2)} = \frac{1}{12n(n + 1)}. \end{aligned}$$

Telescoping to infinity proves the stated bounds. Substitution of the factorial identity into $e^{-\lambda} \lambda^n / n!$ proves (D.1); moreover $0 \leq 1 - e^{-\delta_n} \leq 1/(12n)$, uniformly in the intensity. For $n = 0$ the mass is exactly $e^{-\lambda}$ and the square-root formula is not used.

Thus relative transfer at n requires $\epsilon_N / \mathbb{P}(P_\lambda = n) \rightarrow 0$. For the hard scalar rate a sufficient condition is

$$I + \log^+ \lambda + \tfrac{1}{2} \log n + \lambda h(n/\lambda) \leq V_N - c\nu_N,$$

and for the soft rate a sufficient condition is

$$2I + \log^+ \lambda + \log n + 2\lambda h(n/\lambda) \leq \hat{V}_N - c\hat{\nu}_N.$$

In both cases the event must be measurable at the corresponding prime cutoff. The logarithms of the reciprocal target masses are then $O(V_N)$ or $O(\hat{V}_N)$, so division by that mass does not destroy the polynomial remainder.

In the central range $n = \lambda + t\sqrt{\lambda} + O(1)$, $|t| = o(\lambda^{1/6})$, Taylor expansion gives $\lambda h(n/\lambda) = t^2/2 + o(1)$ and $\log n = \log \lambda + o(1)$. The respective central local budgets are therefore

$$I + \tfrac{3}{2} \log \lambda + \tfrac{1}{2} t^2 \leq V_N - c\nu_N, \quad 2I + 2 \log \lambda + t^2 \leq \hat{V}_N - c\hat{\nu}_N.$$

A global scalar approximation at its upper intensity endpoint does not by itself imply a relative local approximation there.

For completeness, the analogous relative upper-tail transfer has the additional denominator $\mathbb{P}(P_\lambda \geq \lceil \lambda + t\sqrt{\lambda} \rceil)$. Stirling's formula and summation of successive Poisson masses yield the usual Gaussian moderate-tail comparison for $0 \leq t = o(\lambda^{1/6})$. Equivalently, expanding the exponent uniformly on the contributing $O(\sqrt{\lambda}/(1+t))$ scale gives relative error $O((1+t^3)/\sqrt{\lambda})$; the exponentially smaller outer part of the sum is bounded by the ratio of successive masses. Since $1 - \Phi(t) \asymp e^{-t^2/2}/(1+t)$ for $t \geq 0$, a sufficient hard arithmetic condition is

$$I + \log^+ \lambda + \tfrac{1}{2} t^2 + \log(1+t) \leq V_N - c\nu_N.$$

This condition must be combined with the intrinsic Poisson restriction $t = o(\lambda^{1/6})$. It is not a claim about larger deviation scales.

D.2 Resolving an entire conditional path

Let ϵ^{path} be the aggregate staircase error based at the present threshold. The sharpened conditioning inequality (6.9) gives a full future-path error at most $\epsilon^{\text{path}}/p_\lambda(n)$, where $p_\lambda(n) = \mathbb{P}(P_\lambda = n) > \epsilon^{\text{path}}$. The same assertion holds jointly with any fixed available lower segment. Forward transitions are binomial thinning; backward transitions add independent Poisson increments.

Before division, the aggregate proof gives

$$\epsilon^{\text{path}} \ll e^I \lambda (1 + \log^+(2\lambda)) e^{-V_N + o(\nu_N)} + N^{-1/3+\varepsilon},$$

provided the information, log-intensity and mark cutoff are $O(V_N)$. Combining this with (D.1) proves the budget (6.12). Its hypotheses also validate that use of the common logarithmic band. Indeed, for $n \geq 2$, if $\lambda \geq 1$ the budget bounds $\log \lambda$ by $O(V_N)$. If $\lambda < 1$, use

$$\lambda h(n/\lambda) = n \log n - n + \lambda - n \log \lambda$$

to conclude $|\log \lambda| = O(V_N)$. Thus the base offset, information cost and chosen mark cutoff are all $O(V_N) = o(\log N)$. The atom's reciprocal is $\exp(O(V_N)) = N^{o(1)}$, which can be absorbed by first choosing a smaller arithmetic ε . For a fixed present state at a fixed positive limiting intensity the atom stays bounded away from zero, and no extra logarithmic cost is needed.

The condition $p_\lambda(n) > \epsilon^{\text{path}}$ only makes the conditional law well-defined and gives a bound. Convergence requires $\epsilon^{\text{path}}/p_\lambda(n) \rightarrow 0$. Nothing here supplies a single finite-intensity comparison for the entire reverse path: its target mass diverges toward thresholds $m \rightarrow -\infty$.

D.3 Quenched transfer along geometric scales

The proof of Corollary 6.5 gives the single-length Markov–Borel–Cantelli argument. Its uniform extension uses the bound

$$e^{-(c-\beta)\nu_{N_k} + o(\nu_{N_k})} + e^{\beta\nu_{N_k}} N_k^{-1/3+\varepsilon}, \quad 0 < \beta < c,$$

for the exceptional probability on $N_k \asymp q^k$, $q > 1$. Since $\nu_{N_k} \asymp \sqrt{k/\log k}$, multiplication by $O(\log N_k) = O(k)$ preserves summability. A single union bound therefore covers every integer length in a fixed logarithmic band, provided the error margin is uniform. The resulting almost-sure statement concerns these conditional laws at individual scales; it requires no independence between scales and asserts no canonical coupling over all scales.

D.4 Exact target consequences

For a phase $\vartheta \in [0, 1]$, let $P_r \sim \text{Pois}(2^{\vartheta-r-1})$ independently for $r \in \mathbb{Z}$, and $S_m = \sum_{r \geq m} P_r$. Each tail sum is finite almost surely. Then

$$S_m \sim \text{Pois}(2^{\vartheta-m}), \quad \text{Cov}(S_m, S_n) = 2^{\vartheta-\max(m,n)}.$$

Splitting $S_m = P_m + S_{m+1}$ gives the exact forward thinning and backward immigration kernels in the article. For $s \geq 0$ the joint probability generating function is

$$\mathbb{E} \prod_{j=0}^s z_j^{S_{m+j}} = \exp \left[\sum_{r=m}^{m+s-1} 2^{\vartheta-r-1} \left(\prod_{j=0}^{r-m} z_j - 1 \right) + 2^{\vartheta-m-s} \left(\prod_{j=0}^s z_j - 1 \right) \right].$$

Conditionally on $S_m = n$, the n points have independent uniform positions and independent geometric overshoots $\mathbb{P}(E = e) = 2^{-e-1}$. For lattice targets the positions are uniform on the corresponding finite grid. Conditioning the critical prelimit field on a fixed count transfers this assertion by the sharp conditioning lemma; only the final grid-to-uniform passage is weak convergence.

The complete centered target is locally finite on $[1, 2] \times \mathbb{Z}$, although it has infinite total mass toward $r = -\infty$. For every nonnegative continuous compactly supported g put

$$\Psi_{\vartheta}(g) = \exp \left[- \int_1^2 \sum_{r \in \mathbb{Z}} 2^{\vartheta-r-1} (1 - e^{-g(t,r)}) dt \right].$$

Compact support meets finitely many levels. At those levels the lattice comparison and a Riemann sum give

$$\mathbb{E}(e^{-\langle g, \mathcal{E}_N \rangle} \mid C_N) = \Psi_{\vartheta_N}(g) + o_g(1)$$

whenever $I_N \leq V_N - c\nu_N$. A convergent phase subsequence is needed only to replace this moving deterministic target by a single limit law.

For a base intensity $\lambda_N \rightarrow \infty$, the exact maximum identity in the article, applied at $m_N = \lfloor \log_2 \lambda_N \rfloor + j$, gives, along $\{\log_2 \lambda_N\} \rightarrow \vartheta$,

$$\mathbb{P}(M_{N,L} - \lfloor \log_2 \lambda_N \rfloor \leq j) \longrightarrow e^{-2^{\vartheta-j-1}}.$$

This is a lattice extreme law. The identity uses the scalar comparison at the shifted critical length, not an unproved uniform law for every large marked field.

Finally, the Poisson–Gaussian bridge follows by splitting the independent levels at a fixed critical level r_0 . Below it, the normalized Poisson sums satisfy a multivariate central limit theorem with covariance $2^{-\max(j,k)}$. The centered contribution of the levels at or above r_0 has variance $O(\Lambda_N^{-1})$ after normalization, whereas those levels retain their Poisson law and are independent of the lower part. This proves joint independence in the limit. No continuous-time embedding or phase-removing auxiliary randomization is used in this arithmetic argument.

E The microscopic boundary and relative factorization

The exact border cylinder and microscopic non-vacancy are different objects. The first has probability $2^{-\pi(L)}$ without an error term; the second also allows an interior start. This appendix proves that the entire interior first moment is exponentially smaller than the border probability. A union bound is then sufficient; no microscopic overlap expansion is needed.

E.1 Valuation matrices and the transition range

Put $B = L + 1$, $V_x = \{x - 1, \dots, x + L - 1\}$, and let M_x have rows $v(n) = (v_p(n) \bmod 2)_p$ for $n \in V_x$. The start matrix is $A_x = D_x M_x$, where D_x is the edge-difference matrix of a tree on V_x . Its kernel consists exactly of the constant vectors. Consequently

$$\text{rank } A_x \geq \text{rank } M_x - 1, \quad \mathbb{P}(J_{x,L} = 1) \leq 2^{-\text{rank } A_x}. \quad (\text{E.1})$$

For $2 \leq x \leq L - 1$ a start is impossible: both $2x$ and $2x - 2$ lie inside the putative constant run, and complete multiplicativity would force $f(2) = 1$ and $f(2) = -1$ simultaneously.

Lemma E.1 (Transition identity minor). *Uniformly for $x \in \{B - 1, B, B + 1\}$, $\text{rank } A_x \geq (2 - o(1))\pi(B)$.*

Proof. Write $a = x - 1 \in \{B - 2, B - 1, B\}$ and $Y_0 = 2\sqrt{B}$. For each prime $Y_0 < p \leq B$ choose $m_p = p \lceil a/p \rceil \in [a, a + B - 1]$. The cofactor satisfies $\lceil a/p \rceil \leq B/p + 1 < \sqrt{B}/2 + 1 < Y_0 < p$ for large B . Thus m_p has valuation one at p and no other prime factor above Y_0 . Each prime $B < q \leq a + B - 1$ occurs as its own vertex. On these rows and columns the valuation matrix has an identity minor of size $\pi(a + B - 1) - \pi(Y_0) = (2 - o(1))\pi(B)$. Apply (E.1). $\square$

E.2 The uniform prime-divisor input

The precise literature input is

$$\forall \varepsilon > 0 \exists B_0(\varepsilon) \forall B \geq B_0(\varepsilon) \forall n > B : \quad \omega \left(\prod_{j=0}^{B-1} (n+j) \right) \geq (2 - \varepsilon) \pi(B). \quad (\text{E.2})$$

It is equation (14) of Laishram–Shorey [13]. Its uniformity is important. For $n > 17B/12$, their Theorem 2 gives at least $\pi(2B)$ prime divisors, apart from a fixed exceptional pair. For $B < n \leq 17B/12$, every prime at most B divides the block product, and each prime in $[n, n+B-1]$ is an additional divisor. Hence the number is at least

$$\pi(B) + \pi(n+B-1) - \pi(n-1) = (2 - o(1))\pi(B),$$

uniformly in that bounded ratio, by the prime number theorem. This verifies the quantifier order in (E.2).

Suppose now that $B+2 \leq x \leq B^2+2$. Every vertex is at most B^2+B . Equation (E.2) supplies $t_x \geq (1 - o(1))\pi(B)$ primes $q > B$ in the block. Each has valuation one, since $q^2 \geq (B+1)^2 > B^2+B$; it occurs at one vertex only, since $q > B-1$; and two such primes cannot occupy the same vertex, since their product exceeds B^2+B . Their columns span a coordinate subspace U_x of dimension t_x . Call its rows the private rows.

E.3 A global count of even valuations

Fix an integer $K \geq 2$ and use the medium primes $B/K < p \leq B$. Let E_x be their total odd-valuation incidence mass in V_x . We claim

$$E_x \geq (H_K - 1 - o(1))\pi(B), \quad H_K = \sum_{j=1}^K \frac{1}{j}. \quad (\text{E.3})$$

The block contains at least $\lfloor B/p \rfloor$ multiples of p . An incidence of positive even valuation is among those with $n = ap^2 \leq B^2+B$. Thus $a \leq K^2+1$ for sufficiently large B . For fixed a , successive distinct primes $p, q > B/K$ give values ap^2, aq^2 separated by more than $2aB/K$; only $1 + K/(2a)$ such values can lie in an interval of length B . Summing over the fixed number of possible a shows that the total number of square divisibilities is $O_K(1)$, uniformly in x . Subtracting all of them is a permissible overcount of even valuations.

The main incidence count can be evaluated without rounding every summand:

$$\begin{aligned} \sum_{B/K < p \leq B} \left\lfloor \frac{B}{p} \right\rfloor &= \sum_{j=1}^{K-1} (\pi(B/j) - \pi(B/K)) \\ &= (H_{K-1} - (K-1)/K + o(1))\pi(B) \\ &= (H_K - 1 + o(1))\pi(B). \end{aligned}$$

This proves (E.3). The improvement over a prime-by-prime loss is essential: subtracting one for each prime would lose a positive proportion of $\pi(B)$.

E.4 Rank of a simple incidence matrix

Lemma E.2 (Simple incidence rank). *Let H be a finite binary matrix. Suppose every row has weight at most two, every column has weight at most K , and no pair of columns occurs as the*

support of two weight-two rows. If E is the number of entries equal to one, then

$$\text{rank } H \geq \frac{E}{K+1}.$$

Proof. Remove zero columns and view a weight-two row as an edge between its columns, and a weight-one row as a pin. In a connected component with r active columns, an unpinned component has rank $r - 1$ and a pinned component has rank r . The first assertion follows because the edge vectors of a connected graph span the even-parity subspace; one pin then spans its missing direction.

A pinned component has incidence mass at most Kr , which is at most $(K+1)$ times its rank. An unpinned component has $r \geq 2$; simplicity and the degree bound give incidence mass at most $\min\{r(r-1), Kr\}$. If $r \leq K+1$ use the first bound, and if $r \geq K+1$ use the second. In either case the mass is at most $(K+1)(r-1)$. Sum over the components. $\square$

Delete the private rows from the medium-prime columns of M_x , and call the resulting matrix H_x . For fixed K and sufficiently large B , a private row carries at most one medium prime: otherwise a vertex would exceed $B(B/K)^2 > B^2 + B$. Thus deletion removes at most t_x incidences. A remaining row carries at most two medium primes, since $(B/K)^3 > B^2 + B$. A column has weight at most K . Finally two weight-two rows with the same support p, q would be two multiples of $pq > (B/K)^2 > B$ inside one block, which is impossible. All hypotheses of Lemma E.2 are therefore met. Modulo U_x , a medium column is represented by its restriction to the remaining rows, so

$$\begin{aligned} \text{rank } A_x &\geq t_x + \frac{(E_x - t_x)_+}{K+1} - 1 \\ &\geq \left(1 + \frac{H_K - 2}{K+1} - o(1)\right) \pi(B). \end{aligned} \tag{E.4}$$

For the last step, if $E_x \leq t_x$ use $t_x \geq E_x$; otherwise use $t_x \geq (1 - o(1))\pi(B)$ and (E.3). The parameter can be chosen exactly within this family. Set $\delta_K = (H_K - 2)/(K+1)$. Then

$$\delta_{K+1} - \delta_K = \frac{3 - H_K}{(K+1)(K+2)}.$$

Since $H_{10} < 3 < H_{11}$ and H_K is strictly increasing, the unique maximizer over integers $K \geq 2$ is $K = 11$. Consequently we use

$$H_{11} = \frac{83711}{27720}, \quad \delta_* := \frac{H_{11} - 2}{12} = \frac{28271}{332640} > \frac{1}{12}.$$

This optimizes only the displayed fixed-cutoff family of bounds, not the rank surplus among all possible methods.

The local product comparisons hold for $B \geq 11^3 + 1 = 1332$: $(B/11)^3 > B^2 + B$, $B^3/121 > B^2 + B$ and $(B/11)^2 > B$. In the square-divisibility count one may then use $a \leq 122$ and at most seven values per a , giving the safe bound 854. These elementary checks do *not* make 1332 a threshold for the complete asymptotic theorem: the prime-divisor and prime number theorem thresholds still depend on the desired error margin.

E.5 Post-quadratic starts and summation

Lemma E.3 (Post-quadratic defect maximum). *There is an absolute constant θ_0 such that, for all sufficiently large B and every $x > B^2 + 2$, the number m_x of B -defective vertices in V_x*

is at most $B - g_L$, where

$$g_L = \frac{B}{\log B} (\log \log B - \log \log \log B - \theta_0), \quad g_L / \pi(B) \longrightarrow \infty.$$

In particular $\mathbb{P}(J_{x,L} = 1) \leq 2^{-g_L}$.

Proof. We use the Balasubramanian–Shorey exclusion [3] in the form recorded by Shorey [16, pp. 55–56, equation (15) with exponent 2 and the subsequent square-case statement]. It has an absolute constant θ_0 and, for every fixed $\varepsilon > 0$, a threshold $k_0(\varepsilon)$ such that, for $k \geq k_0(\varepsilon)$,

$$\prod_{i=1}^t (z + d'_i) = by^2, \quad 0 \leq d'_1 < \cdots < d'_t < k, \quad b > 0, \quad P^+(b) \leq k, \quad z > e^{1-\theta_0+\varepsilon} \frac{k \log k}{\log \log k}$$

forces $t < \mu_k(\theta_0)$, where

$$\mu_k(\theta) = k \left(1 - \frac{\log \log k}{\log k} + \frac{\log \log \log k}{\log k} + \frac{\theta}{\log k} \right).$$

The smoothness condition on b is part of (15); it is not a size bound on b . All thresholds are uniform in z, b, y and the shifts. To obtain the form used here, first fix $\varepsilon = 1$. Increase the absolute threshold until

$$k^2 > e^{2-\theta_0} \frac{k \log k}{\log \log k}.$$

For $m > k^2$ and $1 \leq d_1 < \cdots < d_t \leq k$, set $z = m + 1$ and $d'_i = d_i - 1$. These substitutions preserve every factor $m + d_i = z + d'_i$, put all shifts in $[0, k)$, and satisfy the required height bound. Hence the product equation forces $t < \mu_k(\theta_0)$. Write $V_x = \{m + 1, \dots, m + B\}$ with $m = x - 2 > B^2$. The product of its B -defective vertices is a square times a B -smooth integer. There is no extra upper bound on that smooth coefficient in the imported statement. Its contrapositive gives $m_x < \mu_B(\theta_0) = B - g_L$ beyond the fixed absolute threshold. The affine defect bound is $\mathbb{P}(J_{x,L} = 1) \leq 2^{-L} 2^{\max(m_x-1, 0)}$, which is at most 2^{-g_L} for large B . $\square$

Partition the starts $2 \leq x \leq 2L^2$ into the forbidden early range, the three transition starts, $B + 2 \leq x \leq B^2 + 2$, and the remaining post-quadratic starts. Their total probabilities are respectively

$$0, \quad 3 \cdot 2^{-(2-o(1))\pi(B)}, \quad O(B^2) 2^{-(1+\delta_*-o(1))\pi(B)}, \quad O(L^2) 2^{-g_L}.$$

Since $\log B = o(\pi(B))$ and $\pi(B) = \pi(L) + O(1)$, their sum is at most $2^{-(1+\delta_*-o(1))\pi(L)}$. Adding the exact border event proves

$$q_L = 2^{-\pi(L)} (1 + O_{\delta_0}(2^{-\delta_0 \pi(L)})) \quad (0 < \delta_0 < 28271/332640),$$

including the displayed choice $\delta_0 = 1/12$. The same interior sum divided by q_L bounds the conditional probability that the exact border is not the unique microscopic start.

The cutoff-invariance and mesoscopic estimates use one further observation. At heights above $2L^2$, two B -defective vertices at distance at most L have distinct squarefree B -smooth parts; equal parts would place consecutive squares too far apart. The number of choices of these parts is $\exp(O(L/\log L))$, and the uniform Pell bound gives the same type of bound on the number of corresponding windows in the entire admissible population up to twice the ambient height. Multiplication by 2^{-g_L} remains $o(2^{-\pi(L)})$, without an exterior height-slice factor. Starts with at most one defect cost only their number times 2^{-L} . This proves the

bounds in the article with cutoff $K \leq 2^L$, using 2^L as an artificial upper ambient scale; it does not require a dyadic lower endpoint comparable to that artificial scale.

E.6 The prime clock

Section 7.6 of the article proves the exact prime-clock law and explains why integer overshoots need not be tight. We record here the conventions needed for the conditional comparisons. Let $J_\partial = \min\{j : f(p_j) = -1\}$. On $C_L = 1$ and the full-measure event $J_\partial < \infty$, put $G_L = J_\partial - \pi(L) - 1$; assign $G_L = 0$ elsewhere. Then

$$\mathbb{P}(G_L = k \mid C_L = 1) = 2^{-k-1}, \quad \mathbb{P}(\ell_\partial - L \geq h \mid C_L = 1) = 2^{-[\pi(L+h) - \pi(L)]}.$$

Since $\{C_L = 1\} \subseteq \{B_L^{\text{mic}} > 0\}$, the conditional law of G_L given microscopic non-vacancy is a mixture whose $C_L = 1$ component is exactly geometric. Its distance from that law is at most $\mathbb{P}(C_L = 0 \mid B_L^{\text{mic}} > 0) = O(2^{-\delta_0 \pi(L)})$. The next subsection constructs the finite-cylinder representative used to carry this mark through the stable product comparison.

E.7 Relative bulk estimates and affine conditioning

Let $\lambda = M2^{-L} \rightarrow 0$, $L = O(\log M)$ and $U = [\lceil M^\delta \rceil, M - L + 1] \cap \mathbb{Z}$. Fix δ and an upper bound for $L/\log M$. With $Y = Y_M^\sharp$, truncate the signed exact marks at E and put $Q = L + E + 1$. The averaged conditional process error is at most

$$C \left[p\mathcal{M}_B + p|D_Y(Q)| + p^2(MQ + E_Y(Q) + R_{2,\delta}(M, L)) \right].$$

The local ranks are those in Appendix C. Summing signs and exact marks before the separated-pair estimate leaves the relation sum at the base length L ; only the bad supports and shared-prime graph use Q . After division by $Mp = \lambda$, the contributions are

Source	Normalized contribution
One-window defects	$M^{-1/2+o(1)}$
Bad maximal supports	$e^{-V_M+o(\nu_M)}$
Diagonal and local pairs	pQ
Shared-prime edges	$\lambda(e^{-V_M+o(\nu_M)} + M^{-1+o(1)})$
Base-length relations	$M^{-1/3+\varepsilon}(1 + \lambda)$
Actual and target mark tails	$O(2^{-E-1})(1 + M^{-1/2+o(1)})$

For the last row, the exact tail event is a start at length $L + E + 1$; the masked first-moment estimate at that length gives the displayed bound after averaging the conditional tail probability. Choose $E = \lceil V_M / \log 2 \rceil$. Then $Q = O(\log M)$, $pQ = O(\log M/M)$ for large M , and every row tends to zero. This proves the signed, marked relative error in (7.23), not merely its unmarked projection.

The microscopic vector is measurable at the same prime cutoff because $2L^2 + L < Y$ for large M . For a fixed K , define $G_{L,K}$ from the first K prime signs after L : count consecutive positive signs, capped at K , and put the value zero off $C_L = 1$. This finite-cylinder variable agrees almost surely with $G_L \wedge K$; this makes measurability on the finite prime sigma-field explicit, independently of the convention on the null nonterminating event. Bertrand's postulate bounds the K th prime after L by $2^K L < Y$ eventually, for fixed K . The stable lift retains $(\mathbf{B}_L, G_{L,K})$ jointly with the full signed marked bulk field, at cost $\Delta_M = o(\lambda)$.

Write $\alpha = 2^{-\pi(L)}$, $b = |U|p$, and let u_M and v_M denote the interior microscopic first moment and the mesoscopic hit probability. The source bounds give $u_M + v_M = o(\alpha + b)$. The one-hit calculation in the proof of Theorem 7.11 gives sparse-target mass $e^{-b}[\alpha + (1 - \alpha)b]$ and a change

of mixture weights at most b relative to (7.30). The conditional bulk mark, sign and position retain their full laws; only the boundary clock is truncated. Simultaneous sources and multiple bulk points have probabilities at most αb and $b^2/2$. Since the sparse-event mass is asymptotic to $\alpha + b$, normalizing the source subprobability measures gives

$$d_{\text{TV}}(\mathcal{L}(\Gamma_{M,L}^{[K]} \mid R_M \geq L), Q_{M,L}^{[K]}) \ll \frac{\Delta_M + u_M + v_M + \alpha b + b^2}{\alpha + b} + b.$$

The additional b is already absorbed by the last two numerator terms. For the actual first-hit law the discarded boundary tail is at most 2^{-K-1} : its unnormalized mass is at most $\alpha 2^{-K-1}$ and the hit probability is at least α . The target tail is also at most 2^{-K-1} . First let $M \rightarrow \infty$, then let $K \rightarrow \infty$. No lower bound on either mixture weight, convergence of their ratio, or exact finite-size independence is required. The grid-to-Lebesgue passage is a separate weak limit.

Now let $A = \{GX = b\}$ be a compatible affine cylinder on the exposed prime signs. Write $r = \text{rank } G$, $I = r \log 2$, and $U_L = \text{span}\{e_p : p \leq L\}$. If the stack with C_L is compatible, its rank is $r + \pi(L) - \kappa$, where $\kappa = \dim(\text{row}(G) \cap U_L)$. Consequently

$$\mathbb{P}(C_L = 1 \mid A) = 2^{-(\pi(L) - \kappa)}$$

exactly, whenever the exposed cutoff contains the border coordinates. This finite identity needs no rare-intensity regime. If the stack is incompatible, that probability is zero. The parameter κ counts consequences entirely supported on the border coordinates, not input equations that merely mention them.

For the asymptotic conclusions assume $M, L \rightarrow \infty$, $L \leq A_0 \log M$, $M 2^{-L} \rightarrow 0$, and eventually $I \leq V_M - c\nu_M$ for fixed $c > 0$. The matrices, their dimensions and right-hand sides may vary with M . Multiply the quantitative normalized bulk ledger by e^I before taking limits. The cutoff term is at most $e^{-c\nu_M + o(\nu_M)}$; all polynomial terms still vanish because $I = O(V_M) = o(\log M)$. Thus the amplified bulk error is $o(\lambda)$; amplification of an unspecified little-oh alone would not justify this conclusion. Also $I = O(V_M) = o(\pi(L))$ and $\kappa \leq r$. The conditioned microscopic interior mass, relative to $2^{-\pi(L) + \kappa}$, is at most

$$\exp(I) 2^{-\delta_0 \pi(L) - \kappa} = o(1).$$

The mesoscopic first term is $\lambda M^{\delta-1}$ and remains $o(\lambda)$ after multiplication by e^I ; its other term is dominated by 2^{-g_L} and remains negligible relative to the border mass. The same one-hit calculation, normalized by $2^{-\pi(L) + \kappa} + b$, therefore proves the affine-conditioned crossover with $\pi(L)$ replaced by $\pi(L) - \kappa$. No assumption on future prime coordinates is used for the rare tail, border concentration or location laws.

Finally let $q_1 < \dots < q_K$ be the first primes after L , and put $Q_{L,K} = \text{span}\{e_{q_1}, \dots, e_{q_K}\}$ and $H_L = \text{row}(G) + U_L$. If the extended stack is compatible,

$$\mathbb{P}(G_L \geq K \mid A, C_L = 1) = 2^{-[K - \dim(H_L \cap Q_{L,K})]},$$

otherwise this tail probability is zero. Thus the ordinary geometric mark is preserved when, for every fixed K , the condition $H_L \cap Q_{L,K} = \{0\}$ holds beyond a threshold that may depend on K . This is separate from the parameter κ controlling the border mass. Fix K , use the finite-cylinder representative and joint stable lift, and take the size limit before $K \rightarrow \infty$. No common threshold in K and no neutrality for a growing truncation $K(M)$ is assumed. The exact rank formula remains the alternative when neutrality fails; no relative estimate for a disappearing mixture component is inferred.

F Microscopic comparison by largest-odd-prime pivots

We prove Theorem 7.7. Throughout this appendix the hypotheses are (7.15) and (7.17), with $\Lambda \rightarrow \infty$. Suppress the index M on I, V, ν, Y , and put $H = \log M$, $u = u(\nu)$. Write $\mathbb{P}_{\mathcal{A}}, \mathbb{E}_{\mathcal{A}}$ for conditioning on $\mathcal{A}_M$.

All arithmetic inputs in this proof are established in the present article–companion pair and are identified in the local register below. The final reciprocal bound also follows from the classical largest-prime-factor asymptotics of [10]. We retain a direct uniform Rankin proof using the prime number theorem [18], then spell out the alternative deduction. Neither the reciprocal bound nor the general Palm–Stein principle is claimed as a new analytic tool. Palm–Stein comparison is classical [4,8]; the finite total-variation inequality is included explicitly.

F.1 The local signed input register

Put $E_* = E_M^*$ from (7.19), $Q = L + E_* + 1$, and $U = \{x \in \mathbb{Z} : \lceil \sqrt{M} \rceil \leq x \leq n+1\}$. A maximal support at $x = j+1$ is $\{j, \dots, j+Q\}$, with $Q+1$ raw vertices. Its base support has $L+1$ vertices; these two lengths are not interchanged. Let $\mathcal{D}_Y(Q)$ be the indices j with $j+1 \in U$ whose maximal support contains a Y -defective integer. Put

$$\mathcal{M}_{L+1} = \sum_{x \in U} (2^{m_{L+1}(x)} - 1), \quad R_{2,U}^{\text{val}}(M, Q+1) = \sum_{\substack{x, y \in U \\ |x-y| > Q}} (2^{\rho_{Q+1}^{\text{val}}(x,y)} - 1).$$

The full-value kernel is the one defined in section 3.7, with no block-parity restriction.

Lemma F.1 (Signed inputs and an independent shifted scalar tail). *Under (7.15) and (7.17), $Q = \log_2 M + O(V)$ and $Y > 2(Q+1)$ eventually. Uniformly in this band,*

$$\mathcal{M}_{L+1} \leq M^{1/2+o(1)}, \quad |\mathcal{D}_Y(Q)|/M \leq e^{-V+o(\nu)}, \quad (\text{F.1})$$

$$R_{2,U}^{\text{val}}(M, Q+1) \ll_{\varepsilon} M^{\varepsilon} (M^{5/3} + M^{2/3} 2^{Q+1}). \quad (\text{F.2})$$

On every retained support the signed low-type marginal is exactly $p\gamma_{e,\sigma}$, on every environment in $\mathcal{F}_Y$. For two retained sites with $0 < |j-k| \leq Q$, the conditional joint mass is at most $4p^2\gamma_{e,\sigma}\gamma_{f,\tau}$. For separated sites its unconditional counterpart is at most

$$p^2\gamma_{e,\sigma}\gamma_{f,\tau} 2^{\rho_{Q+1}^{\text{val}}(j+1,k+1)}. \quad (\text{F.3})$$

Finally, for $L' = L + E_* + 1$ and $\lambda' = M2^{-L'}$,

$$\mathbb{P}(\exists 2 \leq x \leq n+1 : J_{x,L'} = 1) \ll \lambda' + e^{-c_0 \log M / \log \log M}, \quad e^I \lambda' \leq e^{-V}. \quad (\text{F.4})$$

This scalar tail is established without the microscopic comparison.

Proof. Write $\lambda = Mp$ for the ambient intensity in the cutoff, whereas $\Lambda = np$ is the exact target intensity in the hypothesis. Since $L = O(\log M)$, eventually $\Lambda \leq \lambda \leq 2\Lambda$. For any fixed $0 < c_1 < c$, the cost $\log 2$ is absorbed by $(c - c_1)\nu$, giving $I + \log \lambda \leq V - c_1\nu$ eventually. Thus no stronger conditioning hypothesis is introduced. The cutoff gives $2^{E_*} \geq e^{I+V}(2 + \lambda)$ and, in particular, $e^I \lambda 2^{-E_*-1} \leq e^{-V}/2$. The budget also gives $I, \log \Lambda, a_M, E_* = O(V) = o(\log M)$. Proposition 2.4 and Corollary 2.5, including the macroscopic form (2.4), give the weighted mass and, for a deterministic mask $A \subset U$, the bound

$$\sum_{x \in A} \mathbb{P}(J_{x,L} = 1) \leq p|A| + p\mathcal{M}_{L+1}.$$

Propositions B.1 and 4.2 apply to raw heights $M + O(\log M)$ and give the bad-support count. The factor $Q + 1$ and changes of this size in the ambient height have logarithm $o(\nu)$. Proposition 3.24 restricted to U gives the displayed full-value envelope. It is specialized to $M^{5/3+o(1)}$ only after $Q = \log_2 M + O(V)$ has been checked.

For the signed marginals and local pairs, apply section C.2. A prime above $Y > 2(Q + 1)$ occurring oddly at one vertex is private throughout a local union. Its raw values are thus independent fair signs conditionally on the small-prime environment. For $j < k$, $a = L + e$ and $d = k - j$, the ratios to $p^2 \gamma_{e,\sigma} \gamma_{f,\tau}$ are

$$0 \ (d < a), \quad 4\mathbf{1}_{\{\tau=-\sigma\}} \ (d = a), \quad 2\mathbf{1}_{\{\tau=\sigma\}} \ (d = a + 1), \quad 1 \ (a + 1 < d \leq Q).$$

Zero extension to the maximal supports gives (F.3); it is an unconditional, not a fibrewise, profile bound. For a general $\mathcal{A}_M$, nonnegative averaged errors may be multiplied by e^I as in Lemma 6.1. The local bound by four times the rate product already holds on every small-prime fibre and therefore also after mixing under $\mathcal{A}_M$, without that factor. The later common factor e^I is a valid, coarser way of writing the combined ledger.

Proposition 7.3 gives the independent deletion of $x < \lceil \sqrt{M} \rceil$, of mass $O(\Lambda M^{-1/2} + e^{-c_0 \log M / \log \log M})$. For the shifted tail, L' remains in a fixed logarithmic band. Apply Theorem 7.4 to its count over $2 \leq x < M$ and then test non-vacancy. Since $1 - e^{-\lambda'} \leq \lambda'$, this gives (F.4); the retained indices are a subset of that range. Its proof uses the scalar graph estimate, the two-window profile and proposition 7.3, not the theorem being proved here. Lastly (7.19) gives $e^I M p 2^{-E_*-1} \leq e^{-V}$ directly. Summing signs does not enlarge the tail event. $\square$

F.2 A finite conditional coupling inequality

Let $\mathcal{J}$ be finite, let $W = (W_a)_{a \in \mathcal{J}}$ be indicators with positive means λ_a , and for each a couple W^a to W with $\mathcal{L}(W^a) = \mathcal{L}(W \mid W_a = 1)$. Denote a coordinate unit vector by e_a .

Lemma F.2 (Dimension-free Palm–Stein bound). *With the supremum convention for total variation,*

$$d_{\text{TV}} \left(\mathcal{L}(W), \bigotimes_a \text{Pois}(\lambda_a) \right) \leq \sum_a \lambda_a \mathbb{E} \|W - (W^a - e_a)\|_1. \quad (\text{F.5})$$

No dependency graph is assumed.

Proof. The immigration–death generator with invariant product-Poisson law $\mathbf{Q}$ is

$$\mathcal{A}g(w) = \sum_a \lambda_a \Delta_a g(w) + \sum_a w_a \{g(w - e_a) - g(w)\}.$$

For $0 \leq h \leq 1$, its semigroup solution $g_h(w) = -\int_0^\infty (\mathbb{E}h(Z_t^w) - \mathbf{Q}h) dt$ satisfies $\mathcal{A}g_h = h - \mathbf{Q}h$. The integral is finite by coupling the initial particles to a stationary initial configuration; the expected number of unmatched survivors is at most $(|w| + \sum_a \lambda_a)e^{-t}$. Couple two extra initial particles independently, each with death rate one. Their simultaneous survival probability is e^{-2t} and the absolute second difference of an indicator is at most two. Hence

$$|\Delta_a \Delta_b g_h(w)| \leq \int_0^\infty 2e^{-2t} dt = 1, \quad |\Delta_a g_h(w) - \Delta_a g_h(v)| \leq \|w - v\|_1.$$

The first inequality includes $a = b$. For the second, put $u = w \wedge v$, the coordinatewise minimum. Telescope from u to w and from u to v along coordinate steps. Both paths stay in the nonnegative lattice and their total length is $\|w - u\|_1 + \|v - u\|_1 = \|w - v\|_1$. Thus

comparability of w and v is not required. The conditional-law identity gives

$$\mathbb{E} \mathcal{A} g_h(W) = \sum_a \lambda_a \mathbb{E} \{ \Delta_a g_h(W) - \Delta_a g_h(W^a - e_a) \}.$$

Here $W^a - e_a \geq 0$. Bound the differences and take the supremum over indicator tests. All sums are finite. $\square$

For categorical indicators $W_{j,\alpha}$, suppose forcing (j, α) makes it the unique occupied type at j and leaves every site outside a deterministic set $\mathcal{D}_j$ unchanged. Put $s_j = \sum_\alpha \lambda_{j,\alpha}$. The same-site part of (F.5) is $\sum_j s_j^2$. The off-site part is at most

$$\sum_j \sum_{k \in \mathcal{D}_j \setminus \{j\}} \left\{ s_j s_k + \sum_{\alpha, \alpha'} \mathbb{P}(W_{j,\alpha} = W_{k,\alpha'} = 1) \right\}. \quad (\text{F.6})$$

Indeed $|W_{k,\alpha'} - W_{k,\alpha'}^{j,\alpha}| \leq W_{k,\alpha'} + W_{k,\alpha'}^{j,\alpha}$, and $\lambda_{j,\alpha} \mathbb{E} W_{k,\alpha'}^{j,\alpha} = \mathbb{P}(W_{j,\alpha} = W_{k,\alpha'} = 1)$. Thus a very small type probability cancels, rather than appearing as an inverse minimum probability. The analytic solution and its Hessian bound are the explicit input used by the associated Lean proof of Lemma F.2; the semigroup construction printed above is not reconstructed by that formal proof. This distinction concerns formal coverage, not an extra hypothesis in the written coupling argument.

F.3 A uniform reciprocal-pivot estimate

For $m \geq 1$ define $\kappa_{\text{odd}}(m)$ as in the article and put $C(X, z) = \#\{1 \leq m \leq X : \kappa_{\text{odd}}(m) \leq z\}$. For $z \geq 1$, the condition is exactly that every valuation at a prime above z is even, namely the earlier defective-integer predicate. The integer zero is never included. The pivot is not the ordinary largest prime factor: for instance $\kappa_{\text{odd}}(2 \cdot 13^2) = 2$. Equivalently $\kappa_{\text{odd}}(m) = P^+(\text{core}(m))$, with $P^+(1) = 1$; this is $P^*(m)$ in [6, eq. (5)]. Writing $m = a^2 v$ with v squarefree gives, for $s > 1/2$,

$$\sum_{\kappa_{\text{odd}}(m) \leq z} m^{-s} = \zeta(2s) \prod_{q \leq z} (1 + q^{-s}), \quad C(X, z) \leq X^s \zeta(2s) \prod_{q \leq z} (1 + q^{-s}). \quad (\text{F.7})$$

The square part includes primes larger than z with even exponent and cannot be omitted.

Lemma F.3 (Uniform Rankin estimate). *For $X = M + O(\log M)$, uniformly for $V - 2 \leq w \leq 3V + 2$,*

$$C(X, e^w) \leq X \exp\{-D(H/w) + o(\nu)\}. \quad (\text{F.8})$$

The remainder is deterministic and uniform in this band.

Proof. Put $v = u(H/w)$ and $s = 1 - v/w$. Throughout the band, $H/w \asymp \nu$, $v = u + O(1)$, and $s \rightarrow 1$, so eventually $s \geq 3/4$. Therefore

$$\log \left\{ \zeta(2s) \prod_{q \leq e^w} (1 + q^{-s}) \right\} = \sum_{q \leq e^w} q^{-s} + O(1).$$

The sum is $\text{Ei}(v) + o(\nu)$ uniformly. To check the required precision, split at $e^{w/2}$. Chebyshev's bound and partial summation bound the lower part by $O(\log w + e^{v/2}) = o(\nu)$. On the upper part the relative error in $\pi(t) \sim t/\log t$ is uniform for $t \geq e^{(V-2)/2}$ and tends to zero. Partial

summation compares the sum to

$$\int_{e^{w/2}}^{e^w} \frac{t^{-s}}{\log t} dt = \text{Ei}(v) - \text{Ei}(v/2).$$

This integral is $O(\nu)$. The boundary terms $O(e^v/w + e^{v/2}/w)$, the replacement of a partial-summation factor s by one, and $\text{Ei}(v/2)$ are all $o(\nu)$. Since $\log X = H + o(1)$, the logarithm of the normalized bound in (F.7) is

$$(s-1)\log X + \text{Ei}(v) + o(\nu) = -(H/w)v + \text{Ei}(v) + o(\nu) = -D(H/w) + o(\nu).$$

□

Lemma F.4 (Convex saddle envelope). *In the preceding band,*

$$w + D(H/w) \geq 2V - V \left(\sqrt{\nu u/V} - 1 \right)^2 = 2V - O(\nu/u) = 2V - o(\nu). \quad (\text{F.9})$$

Proof. A parameter calculation gives the tangent inequality directly. Put $\nu(t) = e^t/t$ and $D_{\text{par}}(t) = e^t - \text{Ei}(t)$ for $t > 1$. At the fixed saddle parameter u ,

$$\frac{d}{dt} \{D_{\text{par}}(t) - u\nu(t)\} = \frac{e^t(t-1)(t-u)}{t^2}.$$

This expression is nonpositive up to u and nonnegative afterwards, so the bracket has its minimum at u . Since $D_{\text{par}}(u) = V$ and $\nu(u) = \nu$, evaluation at the parameter corresponding to H/w gives $D(H/w) \geq V + u(H/w - \nu)$, without differentiating an inverse function. With $a = \nu u/V$ and $t = w/V$, add w and minimize $t + a/t$:

$$w + D(H/w) \geq V(t + 1 - a + a/t) \geq V(1 - a + 2\sqrt{a}) = 2V - V(\sqrt{a} - 1)^2.$$

Now $\nu u - V = \text{Ei}(u) = \nu(1 + O(1/u))$ and $V \sim \nu u$, so

$$V(\sqrt{a} - 1)^2 = \frac{\text{Ei}(u)^2}{(\sqrt{\nu u} + \sqrt{V})^2} = O(\nu/u).$$

The minimizer need not be exactly $w = V$; the displayed precision, not exact optimality of that choice, preserves the fixed margin $c\nu$. □

Proposition F.5 (Reciprocal-pivot sum). *For $X = M + O(\log M)$ and $Y = \lfloor e^V \rfloor$,*

$$S_M(X) := \sum_{\substack{m \leq X \\ \kappa_{\text{odd}}(m) > Y}} \frac{1}{\kappa_{\text{odd}}(m)} \leq X \exp\{-2V + o(\nu)\}. \quad (\text{F.10})$$

Proof. Terms with $\kappa_{\text{odd}}(m) > e^{3V}$ contribute at most Xe^{-3V} . Divide the remaining logarithmic range into intervals $[b, b+1)$, starting at $\log Y = V + o(1)$. Each contributes at most

$$e^{-b}C(X, e^{b+1}) \leq X \exp\{-b - D(H/(b+1)) + o(\nu)\} \leq X \exp\{-2V + o(\nu)\},$$

by Lemmas F.3 and F.4. Indeed $-b - D(H/(b+1)) = 1 - \{b+1 + D(H/(b+1))\}$: the next-slice shift costs the fixed factor e before absorption in $o(\nu)$. For an integer pivot, $\kappa_{\text{odd}} > \lfloor e^V \rfloor$ is exactly $\kappa_{\text{odd}} > e^V$; no threshold mass is lost. There are $O(V)$ intervals, with $\log V = o(\nu)$. Their uniform remainders can be summed. The omitted tail is smaller since $V/\nu \rightarrow \infty$. □

Classical antecedent and square-part reduction. The bound in proposition F.5 is not an autonomous analytic novelty. With $P^+(1) = 1$, write $S_+(x) = \sum_{n \leq x} 1/P^+(n)$. Equation (1.3) of [10], at $r = 1$, gives

$$S_+(x) = x \exp\{-G(\log x) + o(\nu_x)\}, \quad G(H) = \sqrt{2H(\log H + \log \log H - \log 2 - 2)}.$$

The expansion of V_x in proposition 4.2 gives $G(\log x) = 2V_x + o(\nu_x)$. In the classical formula, the omitted terms after the first correction have size $O(\sqrt{H \log H}(\log \log H)^2/(\log H)^2) = o(\nu_x)$. This identifies the required second-scale precision, not just the leading square-root exponent.

Here is the deduction for odd-prime pivots. Uniquely write $n = a^2v$ with v squarefree. Positivity gives

$$\sum_{n \leq x} \frac{1}{\kappa_{\text{odd}}(n)} = \sum_{a \leq \sqrt{x}} \sum_{\substack{v \leq x/a^2 \\ v \text{ squarefree}}} \frac{1}{P^+(v)} \leq \sum_{a \leq \sqrt{x}} S_+(x/a^2).$$

The terms $a > \exp(3V_x)$ are $O(xe^{-3V_x})$ by $S_+(t) \leq t$. For the other terms, put $H = \log x$ and $t = x/a^2$. Uniformly here, $\log t = H - O(V_x)$, $\nu_t \sim \nu_x$, and the remainder in the classical asymptotic is $o(\nu_x)$. The explicit smooth function G satisfies $G'(H) = O(\sqrt{\log H/H}) = o(1)$ throughout this range. The mean-value theorem therefore yields

$$S_+(x/a^2) \leq x \exp\{-2V_x + o(\nu_x)\} a^{-2+o(1)}.$$

Only the explicit function G has been differentiated, not an unspecified asymptotic remainder. The weights sum to $O(1)$, so

$$\sum_{n \leq x} \frac{1}{\kappa_{\text{odd}}(n)} \leq x \exp\{-2V_x + o(\nu_x)\}.$$

Restricting to $\kappa_{\text{odd}}(n) > Y$ only decreases this sum. At $x = M + O(\log M)$, the hard scales agree at the stated precision, recovering proposition F.5. This is a deduction from the cited asymptotic, not an assertion that its statement uses the present pivot notation. The direct Rankin argument above remains a complete proof of the estimate used below.

F.4 Directed change sets and exact forcing

Take $E_* = E_M^*$ from (7.19), $Q = L + E_* + 1$, and $\mathcal{T}_* = \{0, \dots, E_*\} \times \{-1, +1\}$. Use the original deterministic good set

$$G_0 = \{j : 1 \leq j \leq n, j+1 \geq \lceil \sqrt{M} \rceil, \kappa_{\text{odd}}(j+a) > Y \ (0 \leq a \leq Q)\}. \quad (\text{F.11})$$

For $j \in G_0$, put $q_{j,a} = \kappa_{\text{odd}}(j+a)$. These primes are distinct and greater than $Y > 2(Q+1)$: a selected prime divides its own vertex to an odd exponent and no other vertex of that maximal support. For this privacy assertion alone, exceeding Q suffices: a common prime divisor of two distinct vertices would divide their nonzero difference of size at most Q . Odd valuation, not valuation exactly one, supplies the parity coordinate. The stronger $Y > 2(Q+1)$ is retained for the separate local-pair argument.

The finite cylinder must include every prime up to an integer $R \geq \max\{Y, n+Q\}$. This covers all marked vertices and the entire conditioning event, not only the terminal prefix M . Restricting the infinite prime-sign source to this cylinder preserves the conditional laws of all

the finite events below. Define

$$\mathcal{D}_j = \{k \in G_0 : q_{j,a} \mid k + b \text{ for some } 0 \leq a, b \leq Q\}, \quad \mathcal{B}_M = \sum_{j \in G_0} |\mathcal{D}_j|. \quad (\text{F.12})$$

These are the possible changes for a specified forcing, not neighbourhoods asserted to be a new dependency graph. Divisibility, including even exponents, is a harmless overestimate of possible change.

Proposition F.6 (The total directed footprint). *With $X = n + Q = M + O(\log M)$,*

$$|\mathcal{D}_j| \leq (Q + 1) \sum_{a=0}^Q (n/q_{j,a} + 1), \quad (\text{F.13})$$

$$\mathcal{B}_M \leq n(Q + 1)^2(1 + S_M(X)) \ll n^2 e^{-2V+o(\nu)} + n(Q + 1)^2. \quad (\text{F.14})$$

Proof. For a fixed prime and target offset b , at most $n/q + 1$ grid indices satisfy the divisibility; no bound on b/q is required. Sum over offsets and selected primes. When summing reciprocals over j, a , each raw integer $m = j + a$ occurs at most $Q + 1$ times. Apply Proposition F.5, use $X \asymp n$, and absorb $\log(Q + 1) = O(\log H) = o(\nu)$. $\square$

For example, let $Q = 3$, $Y = 11$, $j = 246$ and $k = 101$. The source vertices 246, 247, 248, 249 have pivots 41, 19, 31, 83. The all-large-prime graph connects them to 101, 102, 103, 104 through 13, since $247 = 13 \cdot 19$ and $104 = 8 \cdot 13$. None of the four source pivots divides a target vertex, so $k \notin \mathcal{D}_j$. Conversely the pivot 13 of 104 divides 247; the direction matters.

Proposition F.7 (Exact one-word conditional law). *Let $W = (I_{j,\alpha})_{j \in G_0, \alpha \in \mathcal{T}_*}$. For every $a = (j, e, \sigma)$, there is a deterministic modification of prime signs under $\mathbb{P}_{\mathcal{A}}$ whose resulting vector W^a has law*

$$\mathcal{L}_{\mathcal{A}}(W^a) = \mathcal{L}_{\mathcal{A}}(W \mid I_{j,e,\sigma} = 1).$$

Only $q_{j,0}, \dots, q_{j,L+e+1}$ are modified. The small-prime environment and every coordinate at a site outside $\mathcal{D}_j$ remain unchanged. If the word already holds, the modification is the identity.

Proof. Write $f(q) = (-1)^{\eta_q}$ and $\sigma = (-1)^\tau$. The required raw word has bits $b_0 = b_{L+e+1} = 1 - \tau$ and $b_a = \tau$ for $1 \leq a \leq L + e$. By privacy, the raw bit at $j + a$ is $\eta_{q_{j,a}} + \ell_{j,a}(\eta_{\text{nonpivots}})$, where $\ell_{j,a}$ uses none of the pivot bits being forced. Keep the other bits fixed and set

$$\eta'_{q_{j,a}} = b_a + \ell_{j,a}(\eta_{\text{nonpivots}}) \pmod{2}, \quad 0 \leq a \leq L + e + 1. \quad (\text{F.15})$$

Every other bit is left unchanged, including the pivots with $L + e + 2 \leq a \leq Q$ in the maximal support. Conditional on all the nonpivot bits and on $\mathcal{A}_M$, the distinct pivot bits are independent and fair. Thus the event has probability $2^{-(L+e+2)} = p\gamma_{e,\sigma}$ for every nonpivot configuration. Conditioning on it leaves the complete nonpivot law unchanged and fixes exactly the values in (F.15). This proves the joint conditional law of the entire finite vector. All forced primes exceed Y . Outside $\mathcal{D}_j$, none divides any of the raw vertices of a low-type event, so that event is unchanged. Categorical exclusivity gives the unique forced atom at the source site. $\square$

This is the step that deals with higher-order dependence: a complete conditional law is constructed, not inferred from pair covariances. The estimates below use pairs only after that law and its possible changes have been established.

F.5 Finite signed ledger and completion

At each $j \in G_0$, every low-type conditional mean is $p\gamma_\alpha$ on every small-prime environment and hence under $\mathbb{P}_\mathcal{A}$. Apply Lemma F.2 and (F.6) to the preceding forcing. The same-site sum is at most np^2 and the product-of-means sum is at most $p^2\mathcal{B}_M$. For $0 < |j - k| \leq Q$, the signed local calculation in Lemma F.1 bounds the joint mass by $4p^2\gamma_\alpha\gamma_{\alpha'}$, already uniformly over environments. There are $O(nQ)$ such pairs. For separated sites the unconditional bound is

$$\mathbb{P}(I_{j,\alpha} = I_{k,\alpha'} = 1) \leq p^2\gamma_\alpha\gamma_{\alpha'} 2^{\rho_{Q+1}^{\text{val}}(j+1,k+1)}.$$

Bound its conditional version by e^I times this nonnegative quantity, split $2^\rho = 1 + (2^\rho - 1)$, and enlarge only the excess to all separated pairs. The constant term stays on the directed change sets; extending it to all pairs would add an unnecessary quadratic baseline. Since $\sum_{\alpha \in \mathcal{T}_*} \gamma_\alpha \leq 1$, this proves

$$d_{\text{TV}} \left(\mathcal{L}_\mathcal{A}(W), \bigotimes_{j \in G_0, \alpha \in \mathcal{T}_*} \text{Pois}(p\gamma_\alpha) \right) \ll np^2 + e^I p^2 \{nQ + \mathcal{B}_M + R_{2,U}^{\text{val}}(M, Q+1)\}. \quad (\text{F.16})$$

Neither the number of exact types nor their smallest probability appears.

Embed the finite field by zero elsewhere. The total actual and target deletion-and-tail cost is bounded by a fixed multiple of

$$d_{0,M} = e^I \left[p\mathcal{M}_{L+1} + p|\mathcal{D}_Y(Q)| + \Lambda M^{-1/2} + e^{-c_0 H / \log H} + Mp 2^{-E_*-1} \right]. \quad (\text{F.17})$$

This follows from the masked first moment, independent deep-start bound and shifted scalar tail in Lemma F.1. The actual upper-tail event is a base hit at $L + E_* + 1$; signs do not enlarge it. The independent target cost is its omitted intensity. Filling is unnecessary for (F.16), because the Stein bound is dimension free for the reduced deterministic target. Thus

$$\Delta_M^{\text{mic}} \ll d_{0,M} + np^2 + e^I p^2 \{nQ + \mathcal{B}_M + R_{2,U}^{\text{val}}(M, Q+1)\}. \quad (\text{F.18})$$

By (F.14), the footprint term is at most

$$e^{I+2\log \Lambda - 2V + o(\nu)} + M^{-1+o(1)} \leq e^{-I-2c\nu+o(\nu)} + M^{-1+o(1)}.$$

The deletion term is bounded by

$$e^{I+\log \Lambda - V + o(\nu)} + M^{-1/2+o(1)} + e^{I-c_0 H / \log H} + O(e^{-V}).$$

Here $H / \log H \gg V$ makes the deep-start term negligible. Since $I, \log \Lambda, E_* = O(V) = o(H)$ and $Q = \log_2 M + O(V)$, the full-value input gives

$$e^I p^2 R_{2,U}^{\text{val}}(M, Q+1) \ll_{\varepsilon_0} M^{-1/3+\varepsilon_0+o(1)}.$$

The local and same-site terms are $M^{-1+o(1)}$. Choose $\varepsilon_0 < \varepsilon$ and absorb the uniform $o(\nu)$ terms with $c' < c$. This proves (7.18).

Dependency boundary. This proof uses a single forced word at a time. It does not condition on a Poisson-size plant, estimate an exponentially small planted void, or bound a global absolute cumulant activity. The constructions in Appendix G are separate consequences and comparisons. The selected-pivot footprint replaces the all-large-prime baseline $e^{I+2\log \Lambda - V + o(\nu)}$; there is no additional inverse-intensity Stein factor.

G Private kernels, signed Palm voids, and the absolute-cumulant obstruction

This appendix is a methodological complement for the same binary model. It retains three distinct conclusions: regular configurations drawn from the target, finite signed Palm identities, and a counterexample to a global absolute-cumulant smallness condition. All arguments are given below. The one-sided Palm deficit follows *from* Theorem 7.7, never the reverse. None of these constructions is an input to section F. Keep the hard scales, cutoff E_* , maximal support diameter Q , original good set G_0 and conditional law $\mathbb{P}_{\mathcal{A}}$ used there, and put $\beta = \log 2$.

G.1 Odd rough kernels and a stronger good set

Define

$$r_Y(m) = \prod_{\substack{q > Y \\ v_q(m) \text{ odd}}} q, \quad Z_Y(s) = \zeta(2s) \prod_{q \leq Y} (1 + q^{-s}), \quad B(X, Y; s) = X^{s-1} Z_Y(s), \quad \frac{1}{2} < s < 1.$$

The product $r_Y(m)$ and the single pivot $\kappa_{\text{odd}}(m)$ have different roles. Every integer has a unique form $m = a^2 v r$, with v squarefree and Y -smooth and $r = r_Y(m)$ squarefree with primes above Y . If $D_Y(x) = \#\{m \leq x : r_Y(m) = 1\}$, Rankin gives $D_Y(x) \leq x^s Z_Y(s)$. The exact convolutions over squarefree rough r imply

$$\frac{1}{X} \#\{m \leq X : r_Y(m) \leq T\} \leq B(X, Y; s) \left(1 + \frac{T^{1-s}}{1-s}\right), \quad (\text{G.1})$$

$$\begin{aligned} \frac{1}{X} \sum_{\substack{m \leq X \\ r_Y(m) > 1}} \frac{z^{\omega(r_Y(m))}}{r_Y(m)} &\leq B(X, Y; s) \left\{ \prod_{q > Y} (1 + z q^{-1-s}) - 1 \right\} \\ &\leq B(X, Y; s) \{\exp(z/(sY^s)) - 1\}, \quad z \geq 1. \end{aligned} \quad (\text{G.2})$$

For the first inequality sum $D_Y(X/r)$ over rough $r \leq T$ and enlarge to all positive integers. For the second sum $z^{\omega(r)} D_Y(X/r)/r$; the empty kernel is removed before forming the Euler product. Finally $\sum_{q > Y} q^{-1-s} \leq \sum_{m > Y} m^{-1-s} \leq Y^{-s}/s$.

At $s = 1 - u/V$ and $X = M + O(Q)$, the calculation of Lemma F.3 at $w = V$ gives

$$B(X, Y; s) \leq e^{-V+o(\nu)}, \quad Y^{-s} = e^{-V+u+o(1)}. \quad (\text{G.3})$$

Fix $0 < \vartheta < c$, and put

$$T_M = \exp(\vartheta H/u), \quad G_{\vartheta} = \{j \in G_0 : r_Y(j+a) > T_M \ (0 \leq a \leq Q)\}. \quad (\text{G.4})$$

This is a threshold on a product of primes, not a change of the conditioning cutoff Y . Eventually $T_M > Y$ and $T_M = M^{o(1)}$.

Proposition G.1 (Cost of stronger deletion). *One has*

$$p|G_0 \setminus G_{\vartheta}| \leq \exp\{\log \Lambda - V + \vartheta \nu + o(\nu)\} = o(1), \quad (\text{G.5})$$

and replacing these extra deleted low-type coordinates by independent target coordinates costs at most $2p|G_0 \setminus G_{\vartheta}|$ under $\mathbb{P}_{\mathcal{A}}$. Moreover

$$|G_{\vartheta}^c| \leq O(\sqrt{M}) + M e^{-V+\vartheta \nu + o(\nu)} = o(M/H). \quad (\text{G.6})$$

Completing the original deletions and mark tails has cost $O(d_{0,M} + p|G_0 \setminus G_\vartheta|) = o(1)$.

Proof. Each integer is used in at most $Q + 1$ maximal supports. Use (G.1)–(G.3) and $(1 - s) \log T_M = \vartheta \nu$; the logarithms of $Q + 1$ and $1/(1 - s)$ are $O(\log H) = o(\nu)$. This proves the count and the first bound, whose exponent is at most $-I - (c - \vartheta)\nu + o(\nu)$. At a site of G_0 , each low-type mean is exactly $p\gamma_\alpha$ on every environment. Thus the added real and target masses are each at most p per site without an extra factor e^I . The original deletion and high marks are covered by (F.17). The last equality in (G.6) follows from $V/\nu \sim u \rightarrow \infty$ and $\log H = o(\nu)$. $\square$

Call a collection of maximal supports *private-pivot regular* if each raw vertex occurrence has a prime above Y , occurring there to an odd exponent, that divides no other vertex occurrence of the collection. Repeated vertices preclude regularity.

Lemma G.2 (CRT allocation of all odd prime factors). *Assume $n \geq 1$, $Q \leq n$ and $Y > Q$, as holds eventually in the ambient regime. Let $J_1, \dots, J_k$ be independent uniform indices in $\{1, \dots, n\}$. Conditional on $J_i = j$ and on a source vertex $m = j + a$ with $r = r_Y(m) > 1$, the probability that every prime of r is hosted in one of the other $k - 1$ supports is at most*

$$\frac{\{3(k - 1)(Q + 1)\}^{\omega(r)}}{r}. \quad (\text{G.7})$$

For $k \geq 2$ and $z = 3k(Q + 1) < Y$, the probability that all indices lie in G_ϑ but their supports are not regular is at most $k(Q + 1)T_M^{-1 + \log z / \log Y}$. A single good support is regular.

Proof. Assign each prime of r to a target block and offset where it is hosted. There are at most $((k - 1)(Q + 1))^{\omega(r)}$ assignments. For a fixed assignment, CRT fixes each target index modulo the product r_h of its assigned primes. A target block receiving no prime has probability exactly one. For a nonempty block, $r_h \leq r \leq m \leq n + Q \leq 2n$, so that residue has probability at most $1/r_h + 1/n \leq 3/r_h$. There are at most $\omega(r)$ nonempty blocks; their nontrivial moduli have product r . Independence therefore bounds the assignment probability by $3^{\omega(r)}/r$, not by a factor three for every empty block as well. Summing assignments gives (G.7). If $r > T_M$, then $\omega(r) \leq \log r / \log Y$ and $z^{\omega(r)}/r \leq T_M^{-1 + \log z / \log Y}$. Union-bound over the $k(Q + 1)$ source occurrences. A prime above $Y > Q$ cannot be hosted elsewhere inside its own support. $\square$

Let $\mathcal{R}_M$ be the target configurations of mass at most $K_M = \lceil 2\Lambda \rceil$, with low types $e \leq E_*$, sites in G_ϑ , and regular maximal supports. Equivalently, enumerate all its atoms, with their multiplicities, and impose regularity on every raw vertex occurrence of that enumeration. This definition is invariant under permutation, so it passes from the ordered sample to the point configuration without losing repeated occurrences. The empty configuration is included.

Theorem G.3 (Private pivots for a Poisson-size target cloud). *For the full same-grid target law $\mathbf{Q}_M$,*

$$\begin{aligned} \mathbf{Q}_M(\mathcal{R}_M^c) &\leq \varepsilon_{\text{reg},M} \ll \Lambda M^{-1/2} + e^{\log \Lambda - V + \vartheta \nu + o(\nu)} + \Lambda 2^{-E_* - 1} \\ &\quad + e^{-(2 \log 2 - 1)\Lambda} + e^{-c\vartheta \nu^2 / (4u)} = o(1). \end{aligned} \quad (\text{G.8})$$

For every regular $z = \{(j_i, \alpha_i) : 1 \leq i \leq k\}$,

$$\mathbb{P}(\text{Occ}(z) \mid \mathcal{F}_Y) = \prod_{i=1}^k p\gamma_{\alpha_i} \quad (\text{G.9})$$

on every small-prime environment, and therefore under $\mathbb{P}_A$.

Proof. Conditional on target mass $N = k$, sites are independent uniform grid indices and types have law γ . Exponential Markov bounds give $\mathbb{P}(N > K_M) \leq e^{-(2 \log 2 - 1)\Lambda}$. The bad-site and high-mark expected counts give the first three terms of (G.8). For $k \leq K_M$, eventually

$$1 - \frac{\log(3K_M(Q+1))}{\log Y} \geq \frac{c\nu}{2V}.$$

Apply Lemma G.2, obtaining a remaining bound $K_M(Q+1)e^{-c\partial\nu^2/(2u)}$. Since $\nu^2/(uV) \rightarrow \infty$ and $\log(K_M(Q+1)) \leq V + O(\log H)$, this is at most the final term of (G.8). For a regular configuration choose one private odd prime per raw vertex. The selected columns of the valuation matrix form an identity minor over $\mathbb{F}_2$ even conditional on all unchosen signs. The raw values are therefore independent and fair. Each exact signed word fixes $L + e_i + 2$ of them; integrating unused maximal-support vertices proves (G.9). $\square$

The first assertion is about configurations drawn from the target. The second prescribes presence, not absence of further occurrences. For instance, if $X_{2i-1} = X_{2i} = B_i$ with independent Bernoulli- p variables, then target configurations using at most one site per pair have asymptotic probability one when $np^2 \rightarrow 0$ and their prescribed probabilities equal p^k . Yet when $np \rightarrow \infty$, the actual field lies in that regular set only when empty, of probability $(1-p)^{n/2} \rightarrow 0$. Its microscopic distance to the Poisson target tends to one. This surrogate is not the arithmetic model.

G.2 Exact configuration masses and ordinary Palm deletion

Put $G = G_\vartheta$, $g = |G|$, $\mu = gp$, and restrict the full arithmetic and target fields to G , without truncating exact lengths or filling. Denote their laws by $\mathbf{P}_M^G, \mathbf{Q}_M^G$, and their distance by Δ_M^G . Let $d_A = \mathbb{P}_A(\text{a base start occurs outside } G)$ and $\delta_G = p(n-g)$; both tend to zero by Proposition G.1 and the original tail estimates. Projection and the deletion coupling give

$$0 \leq \Delta_M^{\text{mic}} - \Delta_M^G \leq d_A + \delta_G. \quad (\text{G.10})$$

Let $\mathcal{R}_M^G$ be the same regularity conditions inside G , with mass bounded by K_M . Its target exceptional probability $\varepsilon_{\text{reg},G} = o(1)$ follows by restriction from Theorem G.3. For regular z put $\lambda_z = \prod_{(j,\alpha) \in z} p\gamma_\alpha$ and

$$v(z) = \mathbb{P}_A(\mathbf{I}_M = z \mid \text{Occ}(z)), \quad v_G(z) = \mathbb{P}_A(\mathbf{I}_M|_G = z \mid \text{Occ}(z)).$$

The regular configurations have distinct sites, so $\mathbf{Q}_M^G(\{z\}) = e^{-\mu}\lambda_z$ and $\mathbf{P}_M^G(\{z\}) = \lambda_z v_G(z)$. The normalized retained void is

$$R_z(t) = \frac{\mathbb{E}_A\left[\prod_{j \in G \setminus \text{supp } z} (1 - tJ_{j+1,L}) \mid \text{Occ}(z)\right]}{(1 - tp)^{g-|z|}}, \quad 0 \leq t \leq 1. \quad (\text{G.11})$$

It is nonnegative and finite; $p < 1$ and the plant is regular.

Proposition G.4 (Averaged deletion and the normalized void deficit). *One has*

$$\begin{aligned} & \mathbb{E}_{\mathbf{Q}_M^G} \mathbf{1}_{\mathcal{R}_M^G}(z) e^\mu (v_G(z) - v(z)) \\ &= \mathbb{P}_A(\mathbf{I}_M|_G \in \mathcal{R}_M^G, \text{a start occurs outside } G) \leq d_A. \end{aligned} \quad (\text{G.12})$$

For the normalized void R_z in (G.11),

$$\left| \Delta_M^{\text{mic}} - \mathbb{E}_{\mathbf{Q}_M^G} \mathbf{1}_{\mathcal{R}_M^G}(z)(1 - R_z(1))_+ \right| \leq d_A + \delta_G + \varepsilon_{\text{reg},G} + O(\Lambda^2/n). \quad (\text{G.13})$$

Proof. Presence factorization implies $\lambda_z(v_G - v) = \mathbb{P}_{\mathcal{A}}(\mathbf{I}_M|_G = z, \mathbf{I}_M \neq z)$. Multiplication by the target mass and summation over disjoint events proves (G.12). For any countable laws, $d_{\text{TV}}(\mathbf{P}, \mathbf{Q}) = \sum_z (\mathbf{Q}(z) - \mathbf{P}(z))_+$. Thus, with $D_G = \mathbb{E}_{\mathbf{Q}_M^G} \mathbf{1}_{\mathcal{R}_M^G}(1 - e^\mu v_G)_+$,

$$0 \leq \Delta_M^{\text{mic}} - D_G \leq d_A + \delta_G + \varepsilon_{\text{reg},G}.$$

Once the exact planted words are fixed, no further retained atom is, almost surely, the event that every remaining base-start indicator is zero. Simultaneous finiteness of the runs gives this equivalence, and both positive-probability conditionings preserve that full-measure event. Hence $v_G(z) = (1 - p)^{g-k} R_z(1)$, where $k = |z| \leq K_M$. Uniformly on regular configurations,

$$|\mu + (g - k) \log(1 - p)| \leq kp + \frac{gp^2}{2(1 - p)} = O(\Lambda^2/n).$$

For $b > 0, x \geq 0$, $|(1 - bx)_+ - (1 - x)_+| \leq 1 - e^{-|\log b|} \leq |\log b|$; no bound on $x = R_z(1)$ is needed. This proves (G.13). $\square$

The argument pays d_A , not $e^\Lambda d_A$. For the full target, the same mass identity gives directly

$$0 \leq \Delta_M^{\text{mic}} - \mathbb{E}_{\mathbf{Q}_M} \mathbf{1}_{\mathcal{R}_M}(z)(1 - e^\Lambda v(z))_+ \leq \varepsilon_{\text{reg},M}.$$

Consequently Theorem 7.7 makes both one-sided deficits vanish. In the full-target form, setting $L_z = (-\log v(z) - \Lambda)_+$ on regular configurations when $v(z) > 0$, $L_z = +\infty$ there when $v(z) = 0$, and $L_z = 0$ otherwise, gives $(1 - e^\Lambda v(z))_+ = 1 - e^{-L_z}$ on the regular set, with $e^{-\infty} = 0$. The extended convention is essential at a zero void. Its expectation vanishes exactly when $L_z \rightarrow 0$ in target probability. This is not uniform control of every Palm void, and it does not result from dividing an absolute approximation by $e^{-\Lambda}$.

Corollary G.5 (Completion of the one-sided Palm deficit). *Under Theorem 7.7,*

$$\mathbb{E}_{\mathbf{Q}_M^G} \left[\mathbf{1}_{\mathcal{R}_M^G}(z)(1 - R_z(1))_+ \right] \longrightarrow 0. \quad (\text{G.14})$$

Equivalently the lower relative-void defect tends to zero in target-configuration probability. This is an averaged, one-sided statement, not a uniform relative estimate over every plant.

Proof. Proposition G.4 identifies the left side with $\Delta_M^{\text{mic}} + o(1)$, using exact configuration masses and ordinary deletion costs. Apply Theorem 7.7, whose proof does not use that identity. $\square$

G.3 The finite signed polynomial, including the environment

Fix a regular plant z and write $\mathcal{B}_z = G \setminus \text{supp } z$, $m = g - |z|$. Let ϵ be the bits of primes at most Y , and ξ those above Y among the finitely many queried integers. Successive relative signs give matrices

$$J_{j+1,L} = \mathbf{1}_{\{A_j \epsilon + B_j \xi = b\}}, \quad b = (1, 0, \dots, 0)^\top \in \mathbb{F}_2^L, \quad A_z \epsilon + B_z \xi = b_z$$

for the base event and the planted raw equations, respectively. Regularity makes B_z full row rank $q_z = \sum_{(j,e,\sigma) \in z} (L + e + 2)$. Planting has conditional probability 2^{-q_z} on every environment,

and so leaves the law of $\epsilon \mid \mathcal{A}_M$ unchanged. Given that environment, ξ is uniform on the displayed affine solution space. Put $\hat{\mu}_A(s) = \mathbb{E}[(-1)^{s \cdot \epsilon} \mid \mathcal{A}_M]$. For $S \subseteq \mathcal{B}_z$, $S \neq \emptyset$, define

$$C_z(S) = \sum_{\substack{u_j \in \mathbb{F}_2^L \setminus \{0\}, j \in S \\ v \in \mathbb{F}_2^{d_z}, \sum_j u_j B_j = v B_z}} (-1)^{\sum_j u_j \cdot b + v \cdot b_z} \hat{\mu}_A \left(\sum_j u_j A_j + v A_z \right), \quad C_z(\emptyset) = 1. \quad (\text{G.15})$$

For a given collection (u_j) there is at most one v , by row independence of B_z .

Proposition G.6 (Signed avoidance and a target-sign bound). *Writing $\mathbb{E}_z$ for the planted conditional expectation, for $0 \leq t \leq 1$,*

$$\mathbb{E}_z \prod_{j \in S} (J_{j+1,L} - p) = p^{|S|} C_z(S), \quad (\text{G.16})$$

$$R_z(t) = 1 + \sum_{\emptyset \neq S \subseteq \mathcal{B}_z} \left(-\frac{tp}{1-tp} \right)^{|S|} C_z(S). \quad (\text{G.17})$$

These finite identities include singleton corrections and all overlaps. For fixed unsigned regular geometry ζ , expand the sign dependence as

$$R_{\zeta, \sigma}(1) = \sum_{D \subseteq [k]} a_D(\zeta) \prod_{i \in D} \sigma_i, \quad \mathcal{V}(\zeta) = \sum_{D \neq \emptyset} a_D(\zeta)^2.$$

Under the independent fair target signs,

$$\mathbb{E}_{\sigma}(1 - R_{\zeta, \sigma}(1))_+ \leq \min \left\{ 1, (1 - a_{\emptyset}(\zeta))_+ + \frac{1}{2} \sqrt{\mathcal{V}(\zeta)} \right\}. \quad (\text{G.18})$$

Proof. The character expansion is $J_{j+1,L} - p = p \sum_{u_j \neq 0} (-1)^{u_j \cdot b} (-1)^{(u_j A_j) \cdot \epsilon + (u_j B_j) \cdot \xi}$. On the planted affine space a character $w \cdot \xi$ has zero mean unless $w = v B_z$, in which case its mean is $(-1)^{v \cdot b_z + (v A_z) \cdot \epsilon}$. Multiplication and environment averaging prove (G.16). Expand $\prod_j (1 - t J_{j+1,L}) / (1 - tp)^m = \prod_j \{1 - t(J_{j+1,L} - p) / (1 - tp)\}$ to obtain (G.17).

In planted block i , $b_{z,i} = c_i + \tau_i \mathbf{1}$ with $\sigma_i = (-1)^{\tau_i}$, where c_i is one at the two borders and zero inside. Thus grouping terms by the block parities of v gives the Walsh expansion. Orthogonality gives mean $a_{\emptyset}$ and variance $\mathcal{V}$. For any integrable R of mean a ,

$$\mathbb{E}(1 - R)_+ \leq (1 - a)_+ + \mathbb{E}(a - R)_+ = (1 - a)_+ + \frac{1}{2} \mathbb{E}|R - a| \leq (1 - a)_+ + \frac{1}{2} \sqrt{\text{Var } R}.$$

Here $R \geq 0$, so the left side is also at most one. This proves (G.18). $\square$

For the unconditioned model $\hat{\mu}_A(s) = \mathbf{1}_{\{s=0\}}$; for a general admissible event it is not that indicator. Nor are the planted marginal means assumed to equal p . The coefficients and their signs must be retained before any norm is taken. Theorem 7.7 controls the deficit in (G.13), but does not establish that $(1 - a_{\emptyset})_+$ or $\mathcal{V}$ vanish separately. No such stronger assertion is needed.

Two other exact formulations are useful for separating sufficient conditions from the concluded deficit. Cumulants below are defined by finite moment-partition identities. Deterministic recentering leaves orders at least two unchanged, but $\text{cum}_z(J_{j+1,L} - p) = \mathbb{E}_z J_{j+1,L} - p$ at order one. Since p is a reference rate, not necessarily the planted mean, these singleton terms

must be kept. If $W_j = J_{j+1,L} - p$ and

$$A_z(t) = \sum_{\emptyset \neq S \subseteq \mathcal{B}_z} \left(\frac{t}{1-tp} \right)^{|S|} |\text{cum}_z(W_j : j \in S)|,$$

then the finite moment-partition expansion gives $|R_z(t) - 1| \leq e^{A_z(t)} - 1$, including singleton Palm corrections. Its per-site weight at $t = 1$ is $(1-p)^{-1}$, not the global weight 2 discussed below. Also, with $N_z = \sum_{j \in \mathcal{B}_z} J_{j+1,L}$ and the planted law tilted by $(1-t)^{N_z} / \mathbb{E}_z(1-t)^{N_z}$ for $t < 1$,

$$-\frac{d}{dt} \log R_z(t) = \frac{\mathbb{E}_{z,t} N_z}{1-t} - \frac{mp}{1-tp}.$$

Integration to one is an improper identity, with value $+\infty$ when the void is zero. We assert no smallness estimate for either right-hand side.

G.4 The absolute conditional pair layer

There is a useful pair estimate independent of the microscopic proof. For two good disjoint prescribed raw words of total length $q_1 + q_2 = q$, write $A\epsilon + B\xi = b$, $r_0 = 2^{-q}$, $\rho_{>} = q - \text{rank } B$ and $\rho_{\text{all}} = q - \text{rank}[A \ B]$. If the full system is compatible, a fraction $a = 2^{\rho_{\text{all}} - \rho_{>}}$ of environments is admissible, and on those environments its conditional probability is $r_0 2^{\rho_{>}}$, otherwise zero. Thus

$$\mathbb{E}_\epsilon |\mathbb{P}(B\xi = b - A\epsilon \mid \epsilon) - r_0| = r_0 \{2^{\rho_{\text{all}}} - 1 + 2(1 - 2^{\rho_{\text{all}} - \rho_{>}})\}. \quad (\text{G.19})$$

If the full system is incompatible the left side is r_0 , and in both cases it is at most $r_0 \{(2^{\rho_{\text{all}}} - 1) + 2\mathbf{1}_{\{\rho_{>} > \rho_{\text{all}}\}}\}$. In the incompatible case $\rho_{\text{all}} \geq 1$.

For a concrete distinction, take $L = 3$, $e = f = 0$, both signs positive, $Y = 11$, and starts $x = 92, y = 114$. The ten raw vertices are $91, \dots, 95$ and $113, \dots, 117$. Each individual rough row rank is five, their combined rough rank is seven, and their full rank is ten. Thus the unconditioned joint mass is 2^{-10} , but its conditional value is 2^{-7} on one eighth of environments and zero elsewhere. The mean absolute conditional covariance is $7/4096$.

Let $H_2^>(M, Q; Y)$ count ordered separated good maximal-support pairs with nonzero rough nullity. Such a relation uses a vertex m in one support, and every prime of $r_Y(m)$ must be hosted by the other support. Assign those primes to target offsets and use CRT. With $X = M + Q + 1$ this gives

$$H_2^>(M, Q; Y) \ll (Q+1)X \sum_{\substack{m \leq X \\ r_Y(m) > 1}} \frac{(Q+1)^{\omega(r_Y(m))}}{r_Y(m)}.$$

There is only one target block, whose residue count is $O(X/r_Y(m))$ since $r_Y(m) \leq X$. Equations (G.2)–(G.3) at $z = Q + 1$ imply

$$H_2^>(M, Q; Y) \leq M^2 e^{-2V + o(\nu)}. \quad (\text{G.20})$$

Indeed $(Q+1)Y^{-s} \rightarrow 0$ and $u + \log Q = o(\nu)$.

Proposition G.7 (The full signed pair activity). *Uniformly for deterministic $G \subseteq G_0$, the*

low-type pair activity obeys

$$\begin{aligned}
 \mathcal{K}_{2,M} &:= 2 \sum_{\substack{j < k \\ j, k \in G}} \sum_{\alpha, \alpha' \in \mathcal{T}_*} |\text{Cov}_{\mathcal{A}}(I_{j,\alpha}, I_{k,\alpha'})| \\
 &\ll e^I p^2 \{nQ + R_{2,U}^{\text{val}}(M, Q+1) + H_2^>(M, Q; Y)\} \\
 &\ll_{\varepsilon} e^{I+2 \log \Lambda - 2V + o(\nu)} + M^{-1/3+\varepsilon} = o(1).
 \end{aligned} \tag{G.21}$$

Proof. Good-word marginal means are constant on environments. Hence absolute covariance under $\mathcal{A}_M$ is at most e^I times mean absolute covariance over the unconditioned small-prime environments. Apply (G.19). Zero extension bounds the full-kernel excess by the maximal full-value profile; the second term requires a rough relation host. Local pairs cost $O(e^I p^2 nQ)$ with signed factor at most four. Sum $\gamma_{\alpha} \gamma_{\alpha'}$ before applying (G.20) and the profile. The budget bounds the first exponential by $e^{-I-2c\nu+o(\nu)}$. $\square$

This is an absolute conditional covariance result, not a graph theorem: zero pair covariance does not imply independence from families of other sites. In particular it cannot replace the exact one-word coupling in Appendix F.

G.5 The finite cumulant bound and its arithmetic obstruction

Let $G \subseteq G_0$ be a deterministic retained set and let $\mathcal{T}_* = \{0, \dots, E_M^*\} \times \{-1, +1\}$. Define the global distinct-site activity

$$\mathcal{K}_{\text{abs},M} = \sum_{\substack{S \subseteq G \\ |S| \geq 2}} 2^{|S|-1} \sum_{\alpha \in \mathcal{T}_*^S} |\text{cum}(I_{j,\alpha_j} : j \in S)|. \tag{G.22}$$

Here the cumulants are under the stated arithmetic law, not under a product target.

For a finite family of categories $Y_j \in \{\partial\} \cup \mathcal{T}_j$ with marginals ν_j , let $I_{j,a} = \mathbf{1}_{\{Y_j=a\}}$ and let $\mathcal{K}$ be the activity (G.22) with these types and their actual law.

Lemma G.8 (A finite categorical cumulant bound). *One has*

$$d_{\text{TV}} \left(\mathcal{L}((Y_j)_j), \bigotimes_j \nu_j \right) \leq \min\{1, (e^{2\mathcal{K}} - 1)/2\}. \tag{G.23}$$

For $t \geq 0$, if B_j are Bernoulli variables of means r_j , and $F_t = \mathbb{E} \prod_j \{1 + t(B_j - r_j)\}$, then

$$|F_t| \leq \exp \left\{ \sum_{\substack{S \\ |S| \geq 2}} t^{|S|} |\text{cum}(B_j : j \in S)| \right\}. \tag{G.24}$$

For categorical occupancies $B_j = \sum_a I_{j,a}$ with $2 \max_j r_j < 1$, $\mathcal{K} \geq \frac{1}{2} \log F_2$.

Proof. The exact identity $\delta_{Y_j} = \nu_j + \sum_a (I_{j,a} - \mathbb{E} I_{j,a})(\delta_a - \delta_{\partial})$ gives, after tensor expansion, centered moment coefficients on subsets S . The signed tensor has zero-mass supremum norm $2^{|S|-1}$. Singletons vanish. Expand each centered moment into cumulant partitions, with blocks of size at least two. If $b_B = 2^{|B|-1} \sum_a |\text{cum}(I_{j,a_j} : j \in B)|$, a partition into m blocks contributes at most $2^{m-1} \prod_B b_B$. Sum disjoint block collections, then drop disjointness to obtain $\frac{1}{2} \sum_{m \geq 1} (2 \sum_B b_B)^m / m! = (e^{2\mathcal{K}} - 1)/2$. The identical finite partition expansion with weights $t^{|B|}$ proves (G.24). Multilinearity gives $2\mathcal{K} \geq \sum_{|S| \geq 2} 2^{|S|} |\text{cum}(B_j : j \in S)|$. When $2 \max_j r_j < 1$,

all factors in F_2 are positive, so taking logarithms proves the last assertion. No analytic infinite cumulant series is assumed. $\square$

The law on the right of (G.23) is the product of categorical marginals, not yet independent Poisson coordinates. Poissonizing occupied mass $s_j \leq p$ costs at most $\sum_j s_j^2 \leq np^2$ by coupling the total count and using the same type kernel. The finite inequality remains true, but its global-smallness hypothesis fails as follows.

Theorem G.9 (Small microscopic distance with divergent absolute activity). *Let q tend to infinity through odd primes and set*

$$a_q = \lfloor \log_2 q \rfloor, \quad M_q = 2^{q-1+a_q}, \quad L_q = q - 1.$$

Take $\mathcal{A}_{M_q}$ to be the whole space. If the retained sets have the exact low-type marginals and omit $o(M_q / \log M_q)$ sites, then

$$\liminf_{q \rightarrow \infty} \frac{\log M_q}{M_q} \mathcal{K}_{\text{abs}, M_q} \geq \frac{\log 2}{2} (\log 3 - 1) > 0. \quad (\text{G.25})$$

This applies to G_0 of Appendix F and to the stronger retained sets in Appendix G. The two-site contribution tends to zero, so the activity of orders at least three diverges. Nevertheless $\Delta_{M_q}^{\text{mic}} \rightarrow 0$, already within the earlier sufficient labelled range.

Proof of Theorem G.9. For $M = M_q$, $L = q - 1$, $a_q = \lfloor \log_2 q \rfloor$,

$$H \sim \beta q, \quad \Lambda = 2^{a_q} - (q - 1)2^{-(q-1)} \asymp q.$$

Thus $a_q = o(\log M)$ and both $\log \Lambda$ and $\log \Lambda + \log(1 + \Lambda)$ are $O(\log H) = o(V - c\nu)$ for every fixed c . Let $g = |G|$, $S_* = 1 - 2^{-E_*-1}$ and $B_j = \sum_{\alpha \in \mathcal{T}_*} I_{j,\alpha}$. Their common exact mean is $r = pS_* \leq p$, and $gr \leq \Lambda$.

Consider the event

$$\mathcal{E}_q = \{f(q) = -1, f(\ell) = 1 \text{ for every prime } \ell \leq M + Q, \ell \neq q\}.$$

Its probability is $2^{-\pi(M+Q)}$ in the original unconditioned model; it is not a new admissible conditioning event. The cutoff $M + Q$ covers the queried vertices and has $Q = o(M)$. A cylinder enlarged to $2M$ would change the witness cost and cannot be substituted here. On this event every queried value is $(-1)^{v_q(m)}$. For $(t + 1)q \leq M$ and $q \nmid t(t + 1)$, the endpoints $tq, (t + 1)q$ are negative and the $q - 1$ interior values are positive. Therefore the site $j = tq$ carries the low type $(0, +1)$. With $T = \lfloor M/q \rfloor - 1$, the number of certified indices is exactly

$$T - \lfloor T/q \rfloor - \lfloor (T + 1)/q \rfloor = (1 - 2/q)M/q + O(1).$$

The omissions are counted in $[1, M]$, including the L terminal indices outside $G \subseteq \{1, \dots, M - L\}$. Their additional cost $L = O(\log M)$ is negligible. Omitting $o(M/H) = o(M/q)$ sites leaves at least $N_0 = (1 - o(1))M/q$ retained low-type atoms. For $N = \sum_{j \in G} B_j$ and sufficiently large q ,

$$F_2 = (1 - 2r)^g \mathbb{E} \left(\frac{3 - 2r}{1 - 2r} \right)^N \geq 2^{-\pi(M+Q)} (1 - 2r)^g \left(\frac{3 - 2r}{1 - 2r} \right)^{N_0}.$$

The prime number theorem gives $\pi(M + Q) = (1 + o(1))M/H$; $g \log(1 - 2r) \geq -O(\Lambda)$; and $\log((3 - 2r)/(1 - 2r)) = \log 3 + O(r)$. Since $\Lambda \asymp H = o(M/H)$, it follows that

$$\log F_2 \geq \{\beta(\log 3 - 1) + o(1)\}M/H.$$

Lemma G.8 proves (G.25). The omission condition holds for G_ϑ by (G.6) and for G_0 by its smaller deletion. Proposition G.7 gives the vanishing pair layer, and hence divergence of the remaining orders. These parameters lie even in the graph-based labelled budget. In particular Theorem 7.7 gives the asserted vanishing microscopic distance, independently of the activity calculation. $\square$

For a fixed $t > 0$, the same witness has leading lower exponent $\beta\{\log(1+t) - 1\}M/H$. It diverges when $t > e - 1$; this is a threshold for the witness, not an optimal boundary of the model. Restricting the occupancy to $N \leq b\Lambda$, $b > 1$, can have vanishing ordinary total-variation cost, but it changes its means, cumulants, and conditional forcing identities. No such nonlocal truncation is used to retain an invalid absolute-smallness claim.

Availability and declaration

The scope, explicit literature inputs and limitations of the Lean development, together with its seven Palomar registrations, are described in the article’s section “Formalization and Palomar registrations”. The article also records the use of AI assistance and the author’s responsibility. This companion and the article form one proof package, with their sources and bibliography available together.

References

- [1] R. Arratia, L. Goldstein, and L. Gordon. “Two moments suffice for Poisson approximations: The Chen–Stein method”. In: *Annals of Probability* 17.1 (1989), pp. 9–25. doi: [10.1214/aop/1176991491](https://doi.org/10.1214/aop/1176991491).
- [2] R. Arratia, L. Goldstein, and L. Gordon. “Poisson approximation and the Chen–Stein method”. In: *Statistical Science* 5.4 (1990), pp. 403–424. doi: [10.1214/ss/1177012015](https://doi.org/10.1214/ss/1177012015).
- [3] R. Balasubramanian and T. N. Shorey. “Squares in products from a block of consecutive integers”. In: *Acta Arithmetica* 65.3 (1993), pp. 213–220. doi: [10.4064/aa-65-3-213-220](https://doi.org/10.4064/aa-65-3-213-220).
- [4] A. D. Barbour. “Stein’s method and Poisson process convergence”. In: *Journal of Applied Probability* 25A (1988), pp. 175–184. doi: [10.2307/3214155](https://doi.org/10.2307/3214155).
- [5] A. C. Berry. “The accuracy of the Gaussian approximation to the sum of independent variates”. In: *Transactions of the American Mathematical Society* 49.1 (1941), pp. 122–136. doi: [10.1090/S0002-9947-1941-0003498-3](https://doi.org/10.1090/S0002-9947-1941-0003498-3).
- [6] R. de la Bretèche, V. Y. Wang, and M. W. Xu. *Random multiplicative functions and making squares from polynomial values*. 2026. arXiv: [2607.06398v1](https://arxiv.org/abs/2607.06398). url: <https://arxiv.org/abs/2607.06398v1>.
- [7] Y. M. Chee, H. M. Kiah, P. Purkayastha, and C. Wang. “Cross-bifix-free codes within a constant factor of optimality”. In: *IEEE Transactions on Information Theory* 59.7 (2013), pp. 4668–4674. doi: [10.1109/TIT.2013.2252952](https://doi.org/10.1109/TIT.2013.2252952).
- [8] L. H. Y. Chen and A. Xia. “Stein’s method, Palm theory and Poisson process approximation”. In: *Annals of Probability* 32.3B (2004), pp. 2545–2569. doi: [10.1214/009117904000000027](https://doi.org/10.1214/009117904000000027).
- [9] J. Cilleruelo and M. Z. Garaev. “Concentration of points on two and three dimensional modular hyperbolas and applications”. In: *Geometric and Functional Analysis* 21 (2011), pp. 892–904. doi: [10.1007/s00039-011-0127-6](https://doi.org/10.1007/s00039-011-0127-6). arXiv: [1007.1526](https://arxiv.org/abs/1007.1526).
- [10] P. Erdős, A. Ivić, and C. Pomerance. “On sums involving reciprocals of the largest prime factor of an integer”. In: *Glasnik Matematički, Series III* 21(41) (1986), pp. 283–300. url: <https://math.dartmouth.edu/~carlp/PDF/54.pdf>.

- [11] F. Halter-Koch. *Quadratic Irrationals: An Introduction to Classical Number Theory*. Monographs and Textbooks in Pure and Applied Mathematics 306. Boca Raton: CRC Press, 2013. doi: [10.1201/b14968](https://doi.org/10.1201/b14968).
- [12] K. Krokowski. “Poisson approximation of Rademacher functionals by the Chen–Stein method and Malliavin calculus”. In: *Communications on Stochastic Analysis* 11.2 (2017), pp. 195–222. doi: [10.31390/cosa.11.2.05](https://doi.org/10.31390/cosa.11.2.05). arXiv: [1505.01417](https://arxiv.org/abs/1505.01417).
- [13] S. Laishram and T. N. Shorey. “Number of prime divisors in a product of consecutive integers”. In: *Acta Arithmetica* 113.4 (2004), pp. 327–341. doi: [10.4064/aa113-4-3](https://doi.org/10.4064/aa113-4-3).
- [14] J.-L. Nicolas and G. Robin. “Majorations explicites pour le nombre de diviseurs de N ”. In: *Canadian Mathematical Bulletin* 26.4 (1983), pp. 485–492. doi: [10.4153/CMB-1983-078-5](https://doi.org/10.4153/CMB-1983-078-5).
- [15] A. Röllin. “On the optimality of Stein factors”. In: *Probability Approximations and Beyond*. Ed. by A. Barbour, H. P. Chan, and D. Siegmund. Vol. 205. Lecture Notes in Statistics. Springer, 2012, pp. 61–72. doi: [10.1007/978-1-4614-1966-2_5](https://doi.org/10.1007/978-1-4614-1966-2_5). arXiv: [0706.0879](https://arxiv.org/abs/0706.0879).
- [16] T. N. Shorey. “Applications of Baker’s theory of linear forms in logarithms to exponential diophantine equations”. In: *RIMS Kôkyûroku* 886 (1994), pp. 48–60. url: <https://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/pdf/0886-05.pdf>.
- [17] T. Tao. *Products of consecutive integers with unusual anatomy*. 2026. arXiv: [2603.27990v2](https://arxiv.org/abs/2603.27990v2). url: <https://arxiv.org/abs/2603.27990v2>.
- [18] G. Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. 3rd ed. Graduate Studies in Mathematics 163. Providence, RI: American Mathematical Society, 2015. doi: [10.1090/gsm/163](https://doi.org/10.1090/gsm/163).