Sharp Poisson approximation and second-order correction for long multiplicative runs

Brice Pouly

With mathematical and editorial assistance from ChatGPT (GPT-6 Astra)

hexagon:2611.00447v1

Abstract

Let f be completely multiplicative with independent fair signs at primes. We count starts of constant runs of length at least L in $[N, 2N)$, with $L = \log_2 N + O(1)$ and both signs included. The total variation distance to the Poisson law with nominal or exact mean has an explicit equivalent of order $LN^{-2/3}$, with a coefficient depending on $N2^{-L}$ and $L \bmod 3$. After retaining the exact mean and second factorial cumulant, the quadratic Poisson–Charlier approximation has optimal error power $4/3$: for every fixed $\varepsilon > 0$, its ℓ^1 error lies between constant multiples of $N^{-4/3-\varepsilon}$ and $N^{-4/3+\varepsilon}$, uniformly in each fixed critical band.

The leading correction comes from rational pairs. Keeping all channels up to a fixed height h gives an explicit distance equivalent with power $h/(h+1)$, determined by the first omitted height. The third factorial cumulant has optimal power $41/30$. The proof expands around the exact Bernoulli marginals and stops at the first excluded dependent configuration. Arithmetic bounds for fixed configurations and a separate uniform tail estimate control the resulting positive boundary mass. Short square relations and a common algebraic curve bound the triple interactions; a fourth Charlier test detects the two-pair contribution that makes the quadratic exponent optimal.

Contents

1 Introduction 3

2 Exact events and their relation spaces 7

3 Pair and triple scales 10

4 Stopping the expansion at its first excluded prefix 14

5 Reconstructing the law with every marginal retained 16

6 Why the quadratic approximation stops at the fourth-order scale 20

7 Completing the approximation and its consequences 22

8 Explicit references and the first omitted channel 26

9 What the proof makes reusable 29

Formal proofs and reproducibility 31

Statement on AI assistance 32

References 32

1 Introduction

A long constant run in independent fair signs starts with probability 2^{-L}. Thus N possible starts have a bounded expected count when $L = \log_2 N + O(1)$. A completely multiplicative sign sequence has the same rare-event scale, but its distant windows share prime signs. The resulting dependencies leave an explicit signature in the Poisson approximation.

The first signature comes from pairs of windows in the ratio $2 : 3$. Their total contribution has size $LN^{-2/3}$. We determine its coefficient and then ask how accurately the whole counting law can be recovered after the exact pair correction is retained. The answer has optimal power $4/3$: two pair interactions survive at fourth order, whereas the total triple activity decays with power $41/30$. The gap $41/30 - 4/3 = 1/30$ is enough, provided that the expansion also controls configurations whose cardinality grows with N.

1.1 Model and principal results

Choose independent fair signs $f(\ell)$, indexed by primes, and set

$$f(n) = \prod_{\ell} f(\ell)^{v_\ell(n)} \qquad (n \geq 1).$$

The function f is completely multiplicative; in particular, $f(n^2) = 1$. Fix $C > 0$. Throughout, $N \geq 2$ and $L \geq 1$ are integers satisfying

$$I_N = [N, 2N) \cap \mathbb{Z}, \qquad |L - \log_2 N| \leq C, \qquad B = L + 1, \qquad p = 2^{-L}, \qquad \lambda_N = Np. \tag{1.1}$$

All logarithms without a subscript are natural. The nominal mean λ_N remains in a compact subinterval of $(0, \infty)$, but need not converge. Define

$$J_x = \mathbf{1}_{\{f(x-1)=-f(x),\ f(x)=f(x+1)=\cdots=f(x+L-1)\}}, \qquad Z_N = \sum_{x \in I_N} J_x. \tag{1.2}$$

Both signs of the run are included. The left boundary condition counts a long run once, at its start. A retained run may continue beyond $2N$, and a run starting before N is not counted.

Write

$$p_x = \mathbb{E}J_x, \qquad \mu_N = \mathbb{E}Z_N, \qquad \alpha_N = \mathrm{Var}(Z_N) - \mu_N.$$

The exact marginals p_x need not equal p. For a signed measure ν on $\mathbb{N} = \{0, 1, \ldots\}$, let $\|\nu\|_1 = \sum_{k \geq 0} |\nu\{k\}|$; total variation between probabilities is $d_{\mathrm{TV}}(P, Q) = \frac{1}{2}\|P - Q\|_1$. Set $\Pi_\mu = \mathrm{Pois}(\mu)$ and $\Delta = \delta_1 - \delta_0$. Products of measures are convolutions. For $r \geq 1$, write $(z)_r = z(z-1)\cdots(z-r+1)$, with $(z)_0 = 1$, and set $\Delta^{*0} = \delta_0$. The quadratic reference is

$$Q_{\mu,\alpha} = \Pi_\mu + \frac{\alpha}{2}\Pi_\mu * \Delta^{*2}. \tag{1.3}$$

It has total mass one and, in the range of the theorem below, is also nonnegative.

Theorem 1.1 (Optimal quadratic approximation with exact parameters)**.** *Fix $C > 0$. For every $\varepsilon > 0$, there are positive constants $c_{C,\varepsilon}$, $C_{C,\varepsilon}$ and a threshold $N_0(C, \varepsilon)$ such that, for all $N \geq N_0$ and every L satisfying (1.1),*

$$c_{C,\varepsilon} N^{-4/3-\varepsilon} \leq \|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \leq C_{C,\varepsilon} N^{-4/3+\varepsilon}. \tag{1.4}$$

Moreover, $\mu_N - \lambda_N = O_{C,\varepsilon}(N^{-1+\varepsilon})$, and Q_{μ_N,α_N} is a probability measure for all sufficiently large N, uniformly in the same band.

The theorem determines the optimal power of the remainder, with an arbitrary fixed power margin. It does not determine its logarithmic factor or leading constant. The exact mean matters: replacing it by λ_N can have a larger cost of order $N^{-1+o(1)}$. Proposition 7.4 makes this distinction precise.

The pair calculation gives a full leading equivalent. If $r \in \{0, 1, 2\}$ is the residue of L modulo 3, then

$$\alpha_N \sim \frac{r+2}{18\, 2^{r/3}} \lambda_N^{5/3} L N^{-2/3}. \tag{1.5}$$

Here and below, an equivalent with moving parameters means that the ratio tends to one uniformly over the fixed critical band.

Corollary 1.2 (Exact Poisson total variation scale)**.** *For every $\varepsilon > 0$, uniformly under (1.1),*

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), \Pi_{\lambda_N}) = \frac{\alpha_N}{4} \|\Pi_{\lambda_N} * \Delta^{*2}\|_1 + O_{C,\varepsilon}(N^{-1+\varepsilon}). \tag{1.6}$$

Consequently,

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), \Pi_{\lambda_N}) \sim \frac{r+2}{72\, 2^{r/3}} \lambda_N^{5/3} \|\Pi_{\lambda_N} * \Delta^{*2}\|_1 \, L N^{-2/3}. \tag{1.7}$$

With Π_{μ_N} as the reference, the first formula holds with μ_N in the difference norm and error $O_{C,\varepsilon}(N^{-4/3+\varepsilon})$; the equivalent is unchanged.

Thus the power 2/3, the factor L, and the leading coefficient are determined for both specified Poisson means. No minimization over all possible means is asserted.

Example 1.3 (A fixed phase). Take $N = 2^{3k}$ and $L = 3k$. Then $\lambda_N = 1$, $r = 0$, and $\|\Pi_1 * \Delta^{*2}\|_1 = 3/e$. Hence

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), \Pi_1) \sim \frac{1}{12e} \, L N^{-2/3}. \tag{1.8}$$

The constant combines the arithmetic coefficient $1/9$, the second-difference norm $3/e$, and the factor $1/4$ from the cumulant and total variation conventions.

A rational channel $a : b$ has coprime integers $a < b < 2a$ and height b. An explicit reference can retain the complete contributions of all channels up to a fixed integer $h \geq 3$. The following table distinguishes the resulting approximation statements. All errors are measured in total variation, uniformly in the fixed critical band; h and every margin $\varepsilon > 0$ are fixed before $N \to \infty$.

Reference	Parameters retained	Precision established here
Poisson	Nominal mean λ_N or exact mean μ_N	Explicit equivalent of order $LN^{-2/3}$ (Corollary 1.2).
Exact quadratic	μ_N, α_N	Optimal power 4/3, with arbitrary fixed power margins (Theorem 1.1).
Explicit quadratic	Nominal mean and complete channel sums of heights $\leq h$	Explicit equivalent of order $LN^{-h/(h+1)}$ (Theorem 8.1).
Exact Hermite	μ_N, α_N	Upper bound $O_{C,\varepsilon}(N^{-4/3+\varepsilon})$; optimality is not established (Corollary 8.4).

For $h = 4$, the explicit quadratic reference retains the $2 : 3$ and $3 : 4$ channels and has optimal power 4/5. Theorem 3.3 also determines the optimal power 41/30 of the third factorial cumulant; its positive witness consists of triples in the ratio $10 : 12 : 15$.

1.2 How the first channel produces its scale

The interval selects the first interior ratio. If $x \in I_N$, then $2x \notin I_N$, so the ratio $1:2$ has no macroscopic family of interior placements. The next ratio, $2:3$, does. For $L = 3m$, the windows at $2t$ and $3t$ contain the matched values $2(t+j)$ and $3(t+j)$, $0 \le j < m$. For $j \neq k$,

$$f(2(t+j))f(2(t+k)) = f(3(t+j))f(3(t+k)).$$

The canonical matching has rank $m-1$ when there are no further square relations. Its prescriptions are compatible because the predecessor signs do not occur in these matches. Such a pair has

$$\mathbb{P}(J_{2t} = J_{3t} = 1) = 2^{-2L+m-1},$$

rather than the independent weight 2^{-2L}.

Allowing the offset $2y-3x$ to vary over a range of length $O(L)$ gives $O(NL)$ placements. Their typical joint weight is $N^{-5/3}$, which explains the product

$$(NL)N^{-5/3} = LN^{-2/3}.$$

This calculation identifies the mechanism. To obtain the coefficient, one must sum the exact endpoint and parity conditions and bound pairs with extra relations. Theorem 3.1 performs that calculation. The complete finite $2:3$ contribution matters: its next term, of order $N^{-2/3}$, is larger than the leading $3:4$ contribution of order $LN^{-3/4}$. Successive channels are therefore isolated by subtracting complete finite expressions.

1.3 From arithmetic interactions to the whole law

The reference is first a sum of independent Bernoulli variables with the exact probabilities p_x. This retains every one-window effect, including repeated indices. Each event is an affine condition on the prime signs over $\mathbb{F}_2$. Effective square-class spaces record its constraints after local relations have been removed. A homogeneous joint probability then dominates the exact rank weights, actual joint probabilities and absolute centered activities, so a single geometric estimate can control several terms in the expansion.

The finite stopping identity in Section 4 retains a hereditary class of configurations with a computable probability. It stops at the first excluded prefix and leaves the remaining indicators inside a positive conditional measure. The resulting error is controlled by the mass of that boundary. Fixed configurations are handled by conditional arithmetic counts. A separate certificate argument sums the large-cardinality tail. Bounded moments at every fixed order do not replace this uniform estimate.

For

$$A_3 = \sum_{\substack{T \subseteq I_N \\ |T|=3}} \left| \mathbb{E} \prod_{x \in T} (J_x - p_x) \right|,$$

Theorem 5.3 gives

$$\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \ll_{C,\varepsilon} N^{\varepsilon} A_3 + N^{-4/3+\varepsilon}.$$

The arithmetic problem is therefore to place A_3 below the power $4/3$. For a pair with sufficiently large shared rank, short relations produce a common curve and force many integer points in a logarithmic offset grid. A degree-uniform point estimate rules out this rank for nonaligned pairs. Combining this restriction with the explicit geometry of aligned triples gives $A_3 \ll N^{-41/30+\varepsilon}$. Taking a sufficiently small fixed margin yields

$$N^{\varepsilon} A_3 \ll N^{-4/3-(1/30-2\varepsilon)}, \qquad 0 < \varepsilon < 1/60.$$

This is the arithmetic gain that closes the upper bound.

For the lower bound, two pair interactions have size $\alpha_N^2 = N^{-4/3+o(1)}$. A fourth Charlier polynomial detects the corresponding term in the true law. The proof keeps a nonnegative compatible contribution and bounds the incompatible part before applying independently proved fixed-moment estimates. This establishes noncancellation for the specified quadratic reference. The exact Hermite reference already contains a fourth-order pair-pair term, so the same lower bound does not apply to it.

1.4 Context and organization

For independent signs, long runs and rare patterns have a classical Poisson theory; see Arratia, Gordon, and Waterman [2] and Godbole [8]. In the multiplicative model, Najnudel [10, Lemma 3.1 and Proposition 6.3] describes consecutive values through valuation relations and proves almost-sure frequencies of fixed words. Fixed-length frequencies do not supply the uniform rare-event estimates needed when $L \asymp \log N$. The binary completely multiplicative model is also distinct from the squarefree-supported convention for a Rademacher multiplicative function; the former is called extended Rademacher in de la Bretèche, Wang, and Xu [6].

Basak, Nath, and Zaharescu [16, Theorem 1] prove related Poisson behavior for chains of quadratic non-residues. Their statistic counts successful disjoint blocks in moving windows of block indices, for almost all primes in a short interval. The block length tends to infinity at most as fast as four iterated logarithms of the prime scale. This differs from the random run starts and critical length regime considered here.

Pouly [11, Theorem 1.1] establishes a Poisson approximation for the sign-resolved run field and already uses the $2:3$ interaction in a lower bound [11, Section 3.6, (3.22)]. Here the observable is the scalar count with both signs included. Its finer approximation requires the complete pair coefficient, control of the law below that coefficient, and a lower bound for the residual after correction.

Poisson approximation for dependent indicators has a general Chen–Stein framework [1]; asymptotic refinements include [3,4]. The reference $Q_{\mu,\alpha}$ is explicit in Barbour and Eagleson [4, Theorem 1, p. 588]. The Poisson–Charlier organization by factorial cumulants is classical [5, Sections 5 and 8.1]. Čekanavičius and Vellaisamy [13, Theorems 3.1–3.2] distinguish quadratic and exponential corrections in a locally dependent setting. Their later lower bound [15, Theorem 2.1, (2.3)] exhibits the squared second-factorial-cumulant obstruction. These local-dependence hypotheses do not cover the separated arithmetic windows here. Our contribution is the arithmetic asymptotics of the parameters and the uniform control and noncancellation needed for the optimal remainder power.

The distance-equivalent theorem of Chhaibi, Delbaen, Méliot, and Nikeghbali [14, Theorem 3.11] uses a diverging parameter and regular Fourier residues; our Poisson means stay in a compact subset of $(0,\infty)$. This restriction belongs to that theorem: the finite bounds of [5, Theorem 5.1] also treat bounded Poisson means once the required residue estimates are available.

The finite stopping identity is a specialization of the abstract lace expansion [12]. Its application here depends on a retained class adapted to rational pairs and on arithmetic bounds for its boundary. Those bounds, the explicit channel coefficients and the noncancellation argument give the stated precision for multiplicative runs.

This article and its *Technical companion* form one work. References beginning with T belong to the companion. The article develops the exact finite expansion, states the arithmetic estimates it consumes, and carries out reconstruction and optimality. The companion proves the arithmetic and uniform counting estimates. Every error margin, fixed order and channel height is chosen before $N \to \infty$; any growing range has a separate uniform estimate.

2 Exact events and their relation spaces

The optimal remainder is smaller than N^{-1}, the cost of deleting all irregular windows. We therefore keep every event with nonzero probability from the outset. Its effective square-class space records all local relations once, and its actual probability is the correct centering parameter.

Let $\mathcal{I}_N = \{x \in I_N : p_x > 0\}$, where $p_x = \mathbb{E}J_x$. An event outside $\mathcal{I}_N$ is identically zero and can be removed without changing Z_N. For every $x \in I_N$, let S_x be the image of the even selections of $V_x = \{x-1, \dots, x+L-1\}$ in $\mathbb{Q}^\times/\mathbb{Q}^{\times 2}$, regarded as a vector space over $\mathbb{F}_2$. Write $d_x = \dim S_x$. The even-selection space has dimension L. Its kernel consists of local even square relations. The following bound makes the possible effective ranks explicit.

Lemma T1.18 in the companion proves that the local even square-relation code has dimension at most one in the critical band. Its proof combines a short-interval square-product bound with the binary Griesmer bound.

For an admissible window the affine prescription is consistent. Its probability is 2^{-d_x}, as the character expansion in Section 2.1 also shows. The lemma therefore gives

$$d_x \in \{L-1, L\}, \qquad p_x = 2^{-d_x} \in \{p, 2p\}, \qquad p_* := \max_{x \in \mathcal{I}_N} p_x \ll_C N^{-1}. \tag{2.1}$$

The assertion about d_x holds for all sufficiently large N in a fixed critical band. A local square relation may instead contradict the run prescription; that is precisely the case already removed by $p_x = 0$.

For a finite $I \subseteq \mathcal{I}_N$, put

$$\rho_{\mathrm{rel}}(I) = \sum_{x \in I} d_x - \dim \sum_{x \in I} S_x, \qquad d_{xy} = \dim(S_x \cap S_y), \tag{2.2}$$

$$q(I) = \mathbb{E} \prod_{x \in I} J_x, \qquad w(I) = 2^{\rho_{\mathrm{rel}}(I)} \prod_{x \in I} p_x, \tag{2.3}$$

$$c(I) = \mathbb{E} \prod_{x \in I} (J_x - p_x). \tag{2.4}$$

The empty products are one. We have $q(I) = w(I)$ when the affine prescriptions are jointly compatible, and $q(I) = 0$ otherwise. Write $G_N = \{x \in I_N : d_x = L\}$ for the regular set. On G_N, these definitions agree with the homogeneous-marginal notation used in the companion. Outside G_N, ρ_{rel} counts dependencies between effective spaces; it does not count local relations a second time.

The notions of directness, connectedness, and primitive family refer to the direct-sum decomposition of the effective spaces proved below. A family is direct when its effective spaces are in direct sum, and connected when it has just one canonical direct component. In particular, a primitive family has a nonzero relation and every proper subfamily is direct. Every nonzero relation of a primitive family has a nonzero projection in each of its blocks. A connected family need not have a relation with this property on its entire support. This distinction matters in the conditional count below.

For fixed $r \geq 2$, define

$$A_r = \sum_{\substack{I \subseteq \mathcal{I}_N \\ |I| = r}} |c(I)|, \qquad K_3^+ = \sum_{\substack{|T|=3,\ T \text{ connected} \\ q(T) > 0}} c(T). \tag{2.5}$$

The nonnegativity in the definition of K_3^+ will be proved in Proposition T9.3. The exact second and third factorial cumulants satisfy

$$\alpha_N = 2 \sum_{|e|=2} c(e) - \sum_x p_x^2, \tag{2.6}$$

$$\kappa_{3,N} = 6 \sum_{|T|=3} c(T) - 6 \sum_{e=\{x,y\}} (p_x + p_y) c(e) + 2 \sum_x p_x^3. \tag{2.7}$$

These identities follow by expanding the first three factorial moments. They display the repeated-index terms that an independent Bernoulli reference will retain exactly.

Fixed sources on effective spaces. The fixed-source estimates stated in Propositions T4.2 and T4.3 apply on $\mathcal{I}_N$, with the new conditional rank

$$\delta_z = \dim\bigl(S_z \cap \textstyle\sum_{x\in I} S_x\bigr)$$

and weight $p_z 2^{\delta_z}$. To see this, choose in each even-selection space a linear complement to its local kernel. The valuation map is injective on that complement and maps it onto S_x. Every label regularization, prime row, and disjoint square equation in the proofs is applied to this linear realization. Restriction to a complement restricts the allowed binary selections within the same physical coordinates; it does not enlarge any prime support. The factor p_z is at most $2p$, and the number of sources and targets is fixed in those propositions. Thus their constants change by a fixed factor. The same reasoning applies to homogeneous prescriptions. In particular, for fixed s, m and $\varepsilon > 0$,

$$\sum_{\substack{T\subseteq\mathcal{I}_N\setminus I\\ |T|=m}} q(I\cup T) \ll_{s,m,C,\varepsilon} N^{\varepsilon} q(I) \quad (|I| = s), \tag{2.8}$$

and the analogous homogeneous bound holds with q replaced by w. If $q(I) = 0$, every summand on the left is zero. No estimate here is asserted for s or m growing with N.

2.1 Affine structure and centered coefficients

The arithmetic estimates naturally count relations, whereas the approximation problem concerns a signed difference of laws. We introduce the exact identities linking these two objects. They will be used with a stopping rule, so that large connected families outside the required nullity range never need to be summed in absolute value.

Throughout this section $I \subseteq \mathcal{I}_N$, and $\rho_{\mathrm{rel}}(I)$ is the effective relative nullity. Write $b_x = \delta_0 + p_x\Delta$. The identities below hold for heterogeneous marginals. On the regular set G_N, one has $p_x = p$ and $d_x = L$. The same identities therefore cover both regular and irregular windows.

2.1.1 An exact expansion around independent Bernoulli variables

Set

$$Y_x = J_x - p_x, \qquad c(I) = \mathbb{E}\prod_{x\in I} Y_x, \qquad c(\varnothing) = 1.$$

Proposition 2.1 (Centered Bernoulli expansion)**.** *For every N,*

$$\mathcal{L}(Z_N) = \sum_{I\subseteq\mathcal{I}_N} c(I)\Delta^{*|I|} * \mathop{*}_{x\notin I} b_x. \tag{2.9}$$

The complementary convolution products are deterministic probabilities. The terms of sizes zero and two are

$$\mathop{*}_{x} b_x + \sum_{|e|=2} c(e)\Delta^{*2} * \mathop{*}_{x\notin e} b_x.$$

Terms of size one vanish.

Proof. For an indicator, $\delta_{J_x} = b_x + Y_x\Delta$. Expand the finite convolution product over $\mathcal{I}_N$ and take expectations. No independence assumption is used in this identity. □

Each $s \in S_x$ is a square-class vector. Let χ_s be its character on the prime signs, and let $\eta_x(s) \in \{\pm 1\}$ be the affine sign prescribed by the run-start event. The local prescription is consistent on $\mathcal{I}_N$, so its affine character η_x is well defined on the effective space. Thus

$$J_x = p_x \sum_{s \in S_x} \eta_x(s)\chi_s, \qquad \eta_x(0) = 1.$$

Lemma 2.2 (Full-support Fourier formula)**.** *For nonempty $I \subseteq \mathcal{I}_N$,*

$$c(I) = \Big(\prod_{x \in I} p_x\Big) \sum_{\substack{s_x \in S_x \setminus \{0\} \ (x \in I) \\ \sum_x s_x = 0}} \prod_{x \in I} \eta_x(s_x). \tag{2.10}$$

Consequently,

$$|c(I)| \le \Big(\prod_{x \in I} p_x\Big)(2^{\rho_{\mathrm{rel}}(I)} - 1) \le \Big(\prod_{x \in I} p_x\Big) 2^{\rho_{\mathrm{rel}}(I)}.$$

If $c(I) \neq 0$, there is a relative relation meeting every block of I.

Proof. Subtracting p_x removes the zero character from each factor. The expectation of a character product is zero unless the sum of its indices vanishes. The kernel of $\bigoplus_{x \in I} S_x \to \sum_{x \in I} S_x$ has $2^{\rho_{\mathrm{rel}}(I)}$ elements. The displayed sum omits its zero element and possibly others. Every affine coefficient has modulus one. □

We extend the same formulas for $q(I)$, $c(I)$ and relative nullity to arbitrary original window families when stating the next lemma.

Lemma 2.3 (A common homogeneous envelope)**.** *For a nonempty family $I \subseteq I_N$ of original windows, set*

$$q_0(I) = 2^{-\dim(\sum_{x \in I} S_x)}, \qquad w(I) = \Big(\prod_{x \in I} p_x\Big) 2^{\rho_{\mathrm{rel}}(I)}.$$

Then

$$q_0(I) = 2^{\rho_{\mathrm{rel}}(I)} \prod_{x \in I} 2^{-d_x}, \qquad w(I),\ q(I),\ |c(I)| \le q_0(I). \tag{2.11}$$

These assertions allow locally inconsistent start prescriptions and zero marginals. In the critical band, for all sufficiently large N,

$$q_0(I) \le 2^{|I|} p^{|I|} 2^{\rho_{\mathrm{rel}}(I)}. \tag{2.12}$$

Proof. Rank–nullity gives the displayed factorization. A consistent affine prescription has the same probability as the corresponding homogeneous system; an inconsistent one has probability zero. Thus $p_x \le 2^{-d_x}$ and $q(I) \le q_0(I)$, proving the assertions for $w(I)$ and $q(I)$. If some $p_x = 0$, then $J_x = 0$ almost surely and $c(I) = 0$. Otherwise the full-support Fourier formula in Lemma 2.2 gives $|c(I)| \le w(I)$. Finally, local rigidity gives $d_x \ge L - 1$ uniformly on I_N in the critical band, so $2^{-d_x} \le 2p$ and the last inequality follows. □

The homogeneous quantity is an actual probability: it prescribes zero for every relative bit. It can be positive even when an original start is impossible. Its finite domination is valid at every cardinality; absorbing the factor $2^{|I|}$ into a constant requires a fixed cardinality.

2.1.2 The nullity that survives affine compatibility

Define

$$\rho_*(I) = \max\{\rho_{\mathrm{rel}}(A) : A \subseteq I,\ q(A) > 0\}, \qquad q(\varnothing) = 1, \quad \rho_{\mathrm{rel}}(\varnothing) = 0. \tag{2.13}$$

This parameter records the largest rank loss in a compatible subfamily. It can be smaller than the nullity of the whole family.

Lemma 2.4 (Compatible-nullity majorant)**.** *For every $I \subseteq \mathcal{I}_N$,*

$$|c(I)| \le 2^{|I|}\Big(\prod_{x\in I} p_x\Big)2^{\rho_*(I)}, \qquad \rho_*(I) \le \rho_{\mathrm{rel}}(I). \tag{2.14}$$

Equality holds in the second inequality when $q(I) > 0$.

Proof. The exact inclusion-exclusion formula is

$$c(I) = \sum_{A\subseteq I} (-1)^{|I|-|A|} q(A) \prod_{x\in I\setminus A} p_x.$$

A compatible term has absolute value $(\prod_{x\in I} p_x)2^{\rho_{\mathrm{rel}}(A)}$, and an incompatible term is zero. There are at most $2^{|I|}$ terms. Nullity increases under inclusion, because adding the block S_x increases the total rank by at most d_x. This proves the other assertions. □

2.1.3 Canonical direct components

A partition $\mathcal{P}$ of I is called direct if the spaces $V_C = \sum_{x\in C} S_x$, $C \in \mathcal{P}$, are in direct sum. Its finest such partition is the canonical direct decomposition of I. We call a nonempty family *connected* if this partition has only one part.

Lemma 2.5 (Existence, uniqueness, and factorization)**.** *Every finite family I has a unique finest direct partition $\mathcal{P}(I)$. For it,*

$$q(I) = \prod_{C\in\mathcal{P}(I)} q(C), \qquad c(I) = \prod_{C\in\mathcal{P}(I)} c(C), \qquad \rho_{\mathrm{rel}}(I) = \sum_{C\in\mathcal{P}(I)} \rho_{\mathrm{rel}}(C). \tag{2.15}$$

Proof. If $\mathcal{P}, \mathcal{Q}$ are direct partitions, their nonempty intersections form a direct partition. Indeed, inside each $P \in \mathcal{P}$, the spaces $\sum_{x\in P\cap Q} S_x$ lie in the mutually direct spaces V_Q, and sum to V_P. The direct sum over P proves the claim. Intersecting the finitely many direct partitions gives the unique finest one.

Characters indexed by mutually direct subspaces of the finite prime-sign vector space are independent uniform character families. Events, and centered variables, depending on different V_C are therefore independent. This proves both probability identities; rank additivity proves the last one. □

Connectedness here is defined by all relative relations. A minimal dependent family of original blocks is primitive, but primitive block supports alone do not determine the canonical components: an additional relation can join a third block to an already dependent pair. Exact pair compression will retain this possibility.

3 Pair and triple scales

The first deviation from Poisson is determined by rational coincidences between two separated windows. We keep the complete finite coefficient: its lower-order terms matter before the next

channel becomes visible. The counting estimates behind this calculation are proved in Sections T1 and T2. For the pair profiles write $\rho(I) = |I|L - \dim \sum_{x\in I} S_x$ and let $\eta(I)$ indicate affine compatibility, so that $q(I) = p^{|I|}\eta(I)2^{\rho(I)}$. This raw nullity includes local kernels and is distinct from the effective relative nullity used in Section 2.

3.1 The two leading rational channels

An exact unit of the reduced channel $a : b$, $a < b$, is a pair of values (at, bt) in the two windows. A relative rational relation selects an even number of these units. If the first and last units are t and $t + d$, the number of selections is

$$A_1 = 1, \qquad A_d = 2^{d-2} \quad (d \geq 2).$$

Summing over positions in one window gives $L + 1 - D$ for a selection of diameter D. Retaining the predecessor sign changes this to

$$g_L(D) = (L - 1 - D)\mathbf{1}_{\{0\leq D\leq L\}}. \tag{3.1}$$

Indeed exactly one of the $L + 1 - D$ placements has its first selected vertex at the predecessor; its sign changes from $+1$ to -1. The value $g_L(L) = -1$ is part of the formula.

Define

$$U_{a,b}(L) = \sum_{d=1}^{\lfloor L/b\rfloor} A_d(L + 1 - ad)(L + 1 - bd),$$
$$S_{a,b}(L) = \sum_{d=1}^{\lfloor L/b\rfloor} A_d(L - 1 - ad)(L - 1 - bd).$$

For $L = bm + r$, $0 \leq r < b$ and $m \geq 1$, both sums are evaluated by

$$\begin{aligned}\sum_{d=1}^{m} & A_d(L + s - ad)(L + s - bd)\\ &= 2^{m-1}\big[(b - a)(b + r + s)m + (r + s)^2 + (a + b)(r + s) + 3ab\big]\\ &\quad - (a + b)(L + s) - ab, \qquad s \in \{-1, 1\}.\end{aligned} \tag{3.2}$$

When $m = 0$ the defining sums are empty and equal to zero. To verify the identity for $m \geq 1$, expand the product and use

$$\sum A_d = 2^{m-1}, \quad \sum dA_d = (m - 1)2^{m-1} + 1, \quad \sum d^2A_d = (m^2 - 2m + 3)2^{m-1} - 1.$$

In particular, for $m \geq 1$,

$$S_{2,3}(L) = 2^{m-1}\big[(r + 2)m + r^2 + 3r + 14\big] - 5L - 1, \qquad L = 3m + r, \quad 0 \leq r < 3, \tag{3.3}$$
$$S_{3,4}(L) = 2^{m-1}\big[(r + 3)m + r^2 + 5r + 30\big] - 7L - 5, \qquad L = 4m + r, \quad 0 \leq r < 4. \tag{3.4}$$

Theorem 3.1 (Complete pair profile)**.** *The sums below are over ordered separated pairs in I_N:*

$$\sum_{|x-y|>L} (2^{\rho(x,y)} - 1) = \frac{N}{3}\{U_{2,3}(L) + U_{3,4}(L)\} + O_{C,\varepsilon}(N^{6/5+\varepsilon}), \tag{3.5}$$

$$\sum_{|x-y|>L} (\eta(x, y)2^{\rho(x,y)} - 1) = \frac{N}{3}\{S_{2,3}(L) + S_{3,4}(L)\} + O_{C,\varepsilon}(N^{6/5+\varepsilon}). \tag{3.6}$$

Writing $\mu_N = \mathbb{E}Z_N$ and $\alpha_N = \mathrm{Var}(Z_N) - \mu_N$, one has

$$\alpha_N = \frac{\lambda_N^2}{3N}\{S_{2,3}(L) + S_{3,4}(L)\} + O_{C,\varepsilon}(N^{-4/5+\varepsilon}). \tag{3.7}$$

Thus, with $r_3 = L \bmod 3$,

$$\alpha_N \sim \frac{r_3+2}{18\,2^{r_3/3}}\lambda_N^{5/3} L N^{-2/3}. \tag{3.8}$$

After the entire first-channel term is removed, the next term is

$$\alpha_N - \frac{\lambda_N^2}{3N} S_{2,3}(L) \sim \frac{r_4+3}{24\,2^{r_4/4}}\lambda_N^{7/4} L N^{-3/4}, \qquad r_4 = L \bmod 4. \tag{3.9}$$

The exact-unit calculation gives the displayed channel sums. Both orientations of each of the channels $2:3$ and $3:4$ have a translation interval of total length $N/3 + O(B)$. Private prime coordinates identify their rational kernels; the exceptional normalizations and all remaining channels have the stated residual mass. Section T2.4.3 gives the complete proof, including the endpoint, contact, and mean-transfer estimates.

The subtraction in (3.9) must retain the whole finite expression $S_{2,3}(L)$. Its own term of size $N^{-2/3}$ without the factor L is larger than $LN^{-3/4}$. The two exact channel polynomials together depend on $L \bmod 12$.

3.2 A computable two-channel reference

An entirely explicit reference has a different optimal rate. Retain the complete $2:3$ and $3:4$ channel sums, and set

$$\widehat{\alpha}_N = \frac{\lambda_N^2}{3N}\{S_{2,3}(L) + S_{3,4}(L)\}. \tag{3.10}$$

The sums have been defined above and are evaluated by (3.3)–(3.4).

Corollary 3.2 (An explicit probability reference)**.** *For every* $\varepsilon > 0$*, uniformly in the critical band,*

$$\left\| \mathcal{L}(Z_N) - \Pi_{\lambda_N} - \frac{\widehat{\alpha}_N}{2}\Pi_{\lambda_N} * \Delta^{*2} \right\|_1 \ll_{C,\varepsilon} N^{-4/5+\varepsilon}. \tag{3.11}$$

For all sufficiently large N*, the reference* $Q_{\lambda_N,\widehat{\alpha}_N}$ *is a probability measure. Its mass at* $k \geq 0$ *is*

$$e^{-\lambda_N}\frac{\lambda_N^k}{k!}\left[1 + \frac{\widehat{\alpha}_N}{2}\left\{\frac{k(k-1)}{\lambda_N^2} - \frac{2k}{\lambda_N} + 1\right\}\right]. \tag{3.12}$$

All parameters in this reference are determined by N, L*.*

Proof. Theorem 3.1 gives $\alpha_N - \widehat{\alpha}_N = O_{C,\varepsilon}(N^{-4/5+\varepsilon})$. Since $\|\Pi_{\lambda_N} * \Delta^{*2}\|_1 \leq 4$, replacing the coefficient costs at most $2|\alpha_N - \widehat{\alpha}_N|$. Apply Theorem 1.1. The replacement of μ_N by λ_N costs $O(N^{-1+\varepsilon})$, which is absorbed here. The finite formulas give $\widehat{\alpha}_N \sim \alpha_N > 0$ and $\widehat{\alpha}_N \to 0$, uniformly in the band. The positivity criterion in Lemma 7.1 therefore applies, since λ_N remains in a compact subinterval of $(0,\infty)$. □

The exponent $4/5$ in (3.11) is optimal for this two-channel reference. Its leading error is half the height-five contribution times $\|\Pi_{\lambda_N} * \Delta^{*2}\|_1$, as proved in Corollary 8.3. The explicit parameters do not supply a numerical value of the constant in the remainder or a numerical threshold for N. The same two-channel coefficient also gives a positive Hermite reference; see Corollary 8.4.

The probability of no retained start. Evaluating (3.11) at zero gives the immediate consequence

$$\mathbb{P}(Z_N = 0) = e^{-\lambda_N}\left(1 + \frac{\widehat{\alpha}_N}{2}\right) + O_{C,\varepsilon}(N^{-4/5+\varepsilon}), \tag{3.13}$$

because $(\Pi_{\lambda_N} * \Delta^{*2})\{0\} = e^{-\lambda_N}$. The positive correction shows an effect of the arithmetic pairs that is absent from the mean alone. Here $Z_N = 0$ means that no retained run start lies in I_N. A run starting before N may still intersect the interval, and a run starting in I_N may continue beyond $2N$. The exact-parameter refinement is Corollary 7.3.

3.3 Why triple interactions are smaller

The third factorial cumulant is

$$\kappa_{3,N} = \mathbb{E}(Z_N)_3 - 3\mu_N \mathbb{E}(Z_N)_2 + 2\mu_N^3. \tag{3.14}$$

It is distinct from the ordinary third cumulant, which equals the mean for a Poisson variable.

Theorem 3.3 (Optimal power of the third factorial cumulant)**.** *For each fixed critical band $C > 0$, there is a constant $c_C > 0$ such that, for all sufficiently large N,*

$$c_C N^{-41/30} \le \kappa_{3,N} \ll_{C,\varepsilon} N^{-41/30+\varepsilon} \qquad (\varepsilon > 0). \tag{3.15}$$

Both bounds are unconditional. Thus $41/30$ is the optimal power exponent for this factorial cumulant. The theorem does not assert an asymptotic equivalent or identify a logarithmic factor for the full cumulant.

The complete proof is given in Section T10. It bounds all admissible triples, with the exact marginals retained. The following form is the input to the reconstruction.

Corollary 3.4 (Absolute third-order activity)**.** *For every $\varepsilon > 0$,*

$$A_3 := \sum_{|T|=3} |c(T)| \ll_{C,\varepsilon} N^{-41/30+\varepsilon}. \tag{3.16}$$

Proof. Equation (T9.1) gives $A_3 = K_3^+ + O(N^{-3/2+\varepsilon'})$. Use the upper bound of Theorem 3.3 and reduce the internal margin. □

The geometric step in the upper bound changes the scale at which points are counted. Write $d_{xy} = \dim(S_x \cap S_y)$ and suppose $d_{xy} > ((2/3)\log 2 + \delta)B$ for a fixed $\delta > 0$. Short-relation extraction yields a space of dimension $\gg_\delta B$ and a common curve through (x, y), of degree $D \le B^{a+o(1)}$ for some fixed $a < 1/2$. For the relations in a subspace of codimension at most one, the associated products are squares in the curve's function field. At a selected noncritical offset u, the simple zeros of $X + u$ must therefore be matched by zeros of factors $Y + v$ from the other window. After $O(D^2)$ exceptional offsets are removed, these fibres give $\gg_\delta B$ integer points in $[-B, B]^2$. Their coordinates are logarithmic-size offsets, although the original centres have size N. For $D \ge 2$, the degree-uniform point bound gives only $o(B)$ points, a contradiction. For $D = 1$, matching the roots gives a small rational slope and an alignment. Thus nonaligned pairs obey the rank bound of Theorem T8.6; the full argument is in Section T8.

The configuration $10t, 12t, 15t$ explains the third-order power. Multiplication by $6, 5, 4$ gives a common center $60t$ and offset steps $6, 5, 4$. The canonical relative nullity is

$$\rho_L = \left\lfloor \frac{L-1}{6} \right\rfloor + \left\lfloor \frac{L-1}{3} \right\rfloor + \left\lfloor \frac{L-1}{5} \right\rfloor - \left\lfloor \frac{L-1}{15} \right\rfloor = \frac{19}{30}L + O(1).$$

A compatible triple with no additional square relations therefore has weight $N^{-71/30+o(1)}$. For a concrete phase, take $L = 30m$ and $N = 2^L$. Then

$$\rho_L = (5m-1) + (10m-1) + (6m-1) - (2m-1) = 19m - 2,$$
$$q(\{10t, 12t, 15t\}) = p^3 2^{\rho_L} = \frac{1}{4}N^{-71/30}.$$

There are $N/30 + O(1)$ translations in the dyadic interval. Proposition T10.4 shows that all but $N^{1/2+o(1)}$ of them have private prime coordinates and compatible predecessor signs. On this population, the largest pair exponent subtracted in the centered coefficient is $10m - 1$, so $c(T) = q(T)(1 + o(1))$. This family thus contributes $(1/120 + o(1))N^{-41/30}$ to K_3^+; the coefficient

is specific to this family, not to the full cumulant. The upper bound also needs pair rigidity and an exhaustive partition of the other triples. It is not a consequence of this example alone.

The distinction between pairs and triples is decisive: $\alpha_N^2 = N^{-4/3+o(1)}$, whereas $A_3 \ll N^{-41/30+\varepsilon}$ and $41/30 - 4/3 = 1/30$. The reconstruction will retain the first pair term and allow its square in the error. The fourth-order test will prove that this remaining scale cannot be improved in power for the quadratic reference.

4 Stopping the expansion at its first excluded prefix

The leading correction comes from retaining one dependent pair. An expansion of the entire counting law would also retain every possible interaction after that pair. Taking absolute values at that stage can be expensive, even when the interactions that first leave the retained class have small total probability. We therefore stop each term at its first excluded prefix and leave all subsequent indicators inside a probability kernel.

The first identity is finite and does not require independence. Its exact-marginal application will reduce the approximation problem to families whose effective nullity is at most $2L$. The first-excluded-prefix identity is a specialization of Zeilberger's abstract lace expansion [12, formula (*)]. The broader inclusion–exclusion context includes the abstract-tube identities [9,7]. We give the prefix proof and its convolution form; the model-specific estimates concern the retained arithmetic class and its full boundary.

4.1 An identity for an arbitrary hereditary class

Let G be a finite totally ordered set. A family $\mathscr{D} \subseteq 2^G$ is hereditary if $I \in \mathscr{D}$ and $J \subseteq I$ imply $J \in \mathscr{D}$. Assume throughout that $\varnothing \in \mathscr{D}$. Its *ordered boundary* is

$$\partial\mathscr{D} = \{I \subseteq G : I \notin \mathscr{D},\ I \setminus \{\max I\} \in \mathscr{D}\}. \tag{4.1}$$

This definition records the first excluded prefix in the prescribed order. It is not the family of all inclusion-minimal excluded sets: an element of $\partial\mathscr{D}$ can contain a smaller excluded subset which is not a prefix.

Lemma 4.1 (Expansion at the first excluded prefix)**.** *Let $(u_x)_{x\in G}$ belong to a commutative unital algebra. Then, as a polynomial identity,*

$$\prod_{x\in G}(1+tu_x) = \sum_{I\in\mathscr{D}} t^{|I|}\prod_{x\in I}u_x + \sum_{I\in\partial\mathscr{D}} t^{|I|}\prod_{x\in I}u_x \prod_{\substack{y\in G\\ y>\max I}}(1+tu_y). \tag{4.2}$$

Proof. Expand the product on the left and index its monomials by $S \subseteq G$. If $S \notin \mathscr{D}$, read the elements of S in their prescribed order. There is a unique first prefix $I \notin \mathscr{D}$. Its preceding prefix belongs to $\mathscr{D}$, so $I \in \partial\mathscr{D}$, and every element of $S \setminus I$ is larger than $\max I$.

Conversely, take $I \in \partial\mathscr{D}$ and $T \subseteq \{y \in G : y > \max I\}$. Every proper prefix of I belongs to $\mathscr{D}$, by heredity. The set $I \cup T$ is excluded, since it contains I, and its first excluded prefix is I. Thus the pairs (I, T) partition the monomials indexed by $S \notin \mathscr{D}$. Summing these monomials proves the identity. □

To identify the specialization explicitly, define $\ell(S) = S$ for $S \in \mathscr{D}$, and otherwise let $\ell(S)$ be its first excluded prefix. This is a lace map: $\ell(S) \subseteq S$; $\ell(T) = \ell(S)$ whenever $\ell(S) \subseteq T \subseteq S$; and equal images of two sets remain unchanged under their union. Its fixed points are $\mathscr{D} \cup \partial\mathscr{D}$. The compatible-extension set is empty for $I \in \mathscr{D}$, and equals $\{y \in G : y > \max I\}$ for $I \in \partial\mathscr{D}$. Zeilberger's formula (*) with $Y_x = tu_x$ is therefore exactly (4.2).

For indicators $(J_x)_{x\in G}$, write

$$Z=\sum_{x\in G}J_x,\qquad q(I)=\mathbb{E}\prod_{x\in I}J_x,\qquad q(\varnothing)=1.$$

Let $\Delta=\delta_1-\delta_0$, and use the convention $\Delta^{*0}=\delta_0$. The measure retained by $\mathscr{D}$ is

$$T_{\mathscr{D}}=\sum_{I\in\mathscr{D}}q(I)\Delta^{*|I|}. \tag{4.3}$$

It has total mass one, although its individual masses need not be nonnegative.

Proposition 4.2 (A positive remainder kernel)**.** *For arbitrary dependence among the indicators,*

$$\mathcal{L}Z-T_{\mathscr{D}}=\sum_{I\in\partial\mathscr{D}}\Delta^{*|I|}*\nu_I, \tag{4.4}$$

$$\nu_I=\mathbb{E}\left[\left(\prod_{x\in I}J_x\right)\delta_{\sum_{y>\max I}J_y}\right].$$

The measure ν_I is positive and has mass $q(I)$. Consequently,

$$\|\mathcal{L}Z-T_{\mathscr{D}}\|_1\le\sum_{I\in\partial\mathscr{D}}2^{|I|}q(I). \tag{4.5}$$

Proof. Apply Lemma 4.1 in the convolution algebra of finite signed measures, with unit δ_0 and $t=\Delta$. For an indicator,

$$\delta_0+J_x\Delta=\delta_{J_x},\qquad \underset{y>\max I}{*}(\delta_0+J_y\Delta)=\delta_{\sum_{y>\max I}J_y}.$$

The expectation of the complete product is $\mathcal{L}Z$. Taking expectations in (4.2) therefore gives (4.4). Positivity and the mass formula for ν_I follow directly from its definition. Finally, $\|\Delta^{*|I|}\|_1=2^{|I|}$ and the convolution norm is submultiplicative, which proves (4.5). □

If $q(I)>0$, the normalized measure $\nu_I/q(I)$ is the law of the suffix count conditioned on the activity of every member of I. This conditional law can differ substantially from its unconditional version. The proof uses only its total mass, so no independence between the prefix and suffix is required. In particular, it incurs no factor exponential in the number of suffix indicators. The use of the ordinary ℓ^1 norm matters: an exponential weight would also charge the suffix count.

Example 4.3 (A suffix that remains fully dependent). Take $G=\{1,2,3\}$ in its natural order, and retain only the empty set and the singletons. Then

$$\partial\mathscr{D}=\{\{1,2\},\{1,3\},\{2,3\}\},\qquad T_{\mathscr{D}}=\delta_0+(q_1+q_2+q_3)\Delta,$$

and the exact remainder is

$$\mathcal{L}Z-T_{\mathscr{D}}=\Delta^{*2}*\left(\mathbb{E}[J_1J_2\delta_{J_3}]+(q_{13}+q_{23})\delta_0\right).$$

For $J_1=J_2=J_3=H$, where $\mathbb{P}(H=1)=\theta$, the conditional suffix following $\{1,2\}$ is identically one. The formula becomes

$$\mathcal{L}(3H)-(\delta_0+3\theta\Delta)=\theta\Delta^{*2}*(2\delta_0+\delta_1).$$

The remainder has norm 6θ, while (4.5) gives the valid bound 12θ. The mechanism is the positive kernel of mass 3θ, not a comparison of the conditional suffix with an independent one. This finite example explains the stopping identity; the remaining task in our model is to prove that the appropriate boundary has small weighted mass.

5 Reconstructing the law with every marginal retained

At a precision finer than N^{-1}, removing the exceptional windows is too expensive. We retain their exact marginals instead. This also keeps the repeated occurrence of a single Bernoulli variable inside the reference law, where it can be evaluated exactly. The remaining task is to control interactions between distinct windows and to sum the expansion without applying a fixed-order estimate at an order growing with N.

We use the full admissible set $\mathcal{I}_N = \{x \in I_N : p_x > 0\}$, the exact marginals $p_x = \mathbb{E}J_x$, the centered coefficients $c(I) = \mathbb{E}\prod_{x\in I}(J_x - p_x)$, and the relative nullity defined in Section 2. Throughout this section we abbreviate $\rho_{\mathrm{rel}}(I)$ to $\rho(I)$; this convention uses the effective local ranks and does not add the local rank defects again. Put

$$A_2 = \sum_{|e|=2} |c(e)|, \qquad A_3 = \sum_{\substack{|T|=3\\ T \text{ connected}}} |c(T)|, \qquad p_* = \max_{x\in\mathcal{I}_N} p_x.$$

The marginal and pair estimates give $p_* \le 2p \ll_C N^{-1}$, $\mu = \sum_x p_x = O_C(1)$, and $A_2 \ll_{C,\varepsilon} N^{-2/3+\varepsilon}$. The connection between the arithmetic estimates and the distribution is the following bound:

$$\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \ll_{C,\varepsilon} N^{\varepsilon}A_3 + N^{-4/3+\varepsilon}. \tag{5.1}$$

We prove it in Theorem 5.3. The two terms have different origins: the first records a genuinely connected three-window interaction, while the second permits two pair interactions.

Three errors in the reconstruction. For $e = \{x, y\}$ write $S_e = S_x + S_y$. The retained class $\mathscr{D}_2$ consists of direct families and families $e\cup U$ with a dependent pair e for which S_e and the spaces $(S_z)_{z\in U}$ form a direct sum. The dependent pair is then unique. Write $T_{\mathscr{D}_2} = \sum_{I\in\mathscr{D}_2} q(I)\Delta^{*|I|}$. With the covariance-corrected Bernoulli measure B_2 defined in (5.2), the proof separates

$$\begin{aligned}\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \le \|\mathcal{L}(Z_N) - T_{\mathscr{D}_2}\|_1 + \|T_{\mathscr{D}_2} - B_2\|_1 \\ + \|B_2 - Q_{\mu_N,\alpha_N}\|_1.\end{aligned}$$

The first term is the stopping boundary, with the unexpanded suffix kept in a positive conditional measure. The second consists of the independent and pair deficits in (5.8); the covariance signs are retained before taking norms. In particular, an incompatible pair is allowed in $\mathscr{D}_2$: its joint probability is zero, but its covariance is still present in B_2. The last term costs $O(N^{-5/3+\varepsilon})$ by Lemma 5.1; that rate concerns the full measure B_2, including every covariance. Theorem 5.2 assembles the three estimates, using fixed orders and the separately uniform tail of Proposition T12.11 for the boundary.

5.1 The exact Bernoulli reference and its conversion

For any finite family of Bernoulli variables, without an independence assumption, define $b_i = \delta_0 + p_i\Delta$ and

$$B_2 = \mathop{*}_{i\in\mathcal{I}_N} b_i + \sum_{|e|=2} c(e)\Delta^{*2} * \mathop{*}_{i\notin e} b_i. \tag{5.2}$$

This is a signed measure. Its first term is the law of a sum of independent Bernoulli variables with the exact individual marginals; its second term restores every covariance once.

Lemma 5.1 (Conversion with heterogeneous marginals)**.** *Let $s_j = \sum_i p_i^j$, $C_2 = \sum_{|e|=2} c(e)$, and $\alpha = 2C_2 - s_2$. If $\mu \le M$ and $p_* \le 1/4$, then*

$$\|B_2 - Q_{\mu,\alpha}\|_1 \le C_M\{s_3 + s_2^2 + A_2(s_2 + p_*)\}. \tag{5.3}$$

For the present model this is $O_{C,\varepsilon}(N^{-5/3+\varepsilon})$.

Proof. In the convolution algebra of finite signed measures, the series for $\log_*(\delta_0 + p_i\Delta)$ converges absolutely, since $\|p_i\Delta\|_1 = 2p_i \le 1/2$. Summing it gives

$$\log_*\Big(\underset{i}{*}\, b_i\Big) = \mu\Delta - \frac{s_2}{2}\Delta^{*2} + E, \qquad \|E\|_1 \le \sum_{j\ge 3} \frac{2^j s_j}{j} \le \frac{16}{3} s_3.$$

Exponentiation and $\exp_*(\mu\Delta) = \Pi_\mu$ imply

$$\Big\| \underset{i}{*}\, b_i - \Pi_\mu + \frac{s_2}{2}\Pi_\mu * \Delta^{*2} \Big\|_1 \le C_M(s_3 + s_2^2).$$

The product omitting a pair $e = \{i, j\}$ differs from $\Pi_{\mu - p_i - p_j}$ by $O_M(s_2)$ in norm, and

$$\|\Pi_{\mu - p_i - p_j} - \Pi_\mu\|_1 \le 2(p_i + p_j).$$

Replacing each such product in (5.2) by Π_μ therefore costs $O_M(A_2(s_2 + p_*))$. The exact identity $\operatorname{Var} Z - \mathbb{E}Z = 2C_2 - s_2$ proves the lemma. Here $s_2 \le p_*\mu = O_C(N^{-1})$ and $s_3 \le p_*^2\mu = O_C(N^{-2})$, giving the stated application. □

Algebraic form of the conversion. The estimate in Lemma 5.1 also holds for arbitrary real pair weights $c(e)$ in (5.2), with $C_2 = \sum_e c(e)$, $A_2 = \sum_e |c(e)|$ and $\alpha = 2C_2 - s_2$. The mean and marginal bounds are unchanged. The two approximations in its proof and the triangle inequality give this algebraic form; only the interpretation $\alpha = \operatorname{Var} Z - \mathbb{E}Z$ uses actual covariances.

5.2 Arithmetic estimates used in the assembly

The proof below uses three estimates proved in Section T12 of the companion. They separate fixed geometry from the growing number of windows. For each fixed $r \ge 4$, the total connected activity is

$$\sum_{\substack{|I|=r \\ I \text{ connected}}} |c(I)| \ll_{r,C,\varepsilon} N^{\varepsilon} A_3 + N^{-4/3+\varepsilon},$$

and its incompatible part is $O(N^{-3/2+\varepsilon})$. The independent deficit of all pair extensions is $O(N^{-3/2+\varepsilon})$ at every cardinality. Finally, the connected activity beyond the fixed size 256, up to the logarithmic boundary cutoff and with compatible nullity at most $2L$, is $O(N^{-4/3+\varepsilon})$; the corresponding primitive tail has the same bound without a cardinality cutoff. These are Proposition T12.9, Proposition T12.8 and Proposition T12.11, respectively.

The last estimate has a separate proof. It uses small-kernel counts, moderate congruence certificates and a graph with small incoming and outgoing degrees. It does not let the order in the first estimate grow with N. The choice 256 is sufficient for these estimates; no optimality is attached to that cutoff.

Dependencies of the reconstruction. The table separates the arithmetic inputs by the quantity they control. The fixed-order estimates and the uniform tails enter different parts of the stopped expansion; the lower bound uses the independently proved moment estimate.

Quantity	Main inputs in the companion	Use in the article
Complete pair coefficient	Sections T1–T3	Theorem 3.1
Rank of a nonaligned pair	T8.1–T8.6; norm estimate in T8.13	Theorem 3.3
Absolute triple activity	Section T9; T10.1–T10.4	Corollary 3.4
Connected activity at fixed orders	T12.1, T12.4, T12.6, T12.9	Theorem 5.3
Pair deficit at every cardinality	T12.5–T12.8	Equation (5.9)
Uniform tails	Section T11; T12.5, T12.7–T12.11	Theorems 5.2–5.3
Quadratic lower bound	Theorem T4.4; incompatible part of Proposition T12.9	6.1–6.2

5.3 The stopped expansion and its finite input

Let $\mathscr{D}_2$ consist of the direct families and the families $e \cup U$ for which e is any dependent pair and $S_e, (S_z)_{z\in U}$ are direct. Affine incompatibility of e is allowed. This class is hereditary; every non-direct member contains exactly one dependent pair. For a fixed ordering of $\mathcal{I}_N$, its boundary contains the first excluded prefixes. A boundary family has relative nullity at most $2L$: its retained predecessor has at most L, and adding one block adds at most L. In particular every subfamily of a boundary family has compatible nullity at most $2L$.

Theorem 5.2 (Finite transfer to the complete law)**.** *Fix $1 < \beta \le 4/3$. Suppose that, for each of the fixed integers $3 \le r \le 256$ and every $\varepsilon > 0$,*

$$\sum_{\substack{|I|=r,\ I \text{ connected}\\ \rho_*(I)\le 2L}} |c(I)| \ll_{r,C,\varepsilon} N^{-\beta+\varepsilon}. \tag{5.4}$$

Then

$$\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \ll_{C,\varepsilon} N^{-\beta+\varepsilon}. \tag{5.5}$$

Only these 254 fixed orders are assumed in (5.4); the larger orders are supplied by Proposition T12.11.

Proof. Fix a weight $w > 0$. The hypothesis at sizes at most 256 and (T12.17) give, for $K_N = o(L)$,

$$T := \sum_{\substack{I \text{ connected},\ 3\le|I|\le K_N\\ \rho_*(I)\le 2L}} w^{|I|}|c(I)| \ll N^{-\beta+\eta}. \tag{5.6}$$

The finitely many factors w^r at $r \le 256$ only change constants. For a primitive P, $\rho(P) \le L$ and $\prod_{x\in P} p_x \le |c(P)|$, because $c(P) = q(P) - \prod_P p_x$ and $q(P)$ is either zero or at least twice the independent product. Together with (T12.18), this gives

$$\sum_{\substack{P \text{ primitive}\\ |P|\ge 3}} w^{|P|}|c(P)| \ll N^{-\beta+\eta} \tag{5.7}$$

over all cardinalities.

Put $T_{\mathscr{D}_2} = \sum_{I\in\mathscr{D}_2} q(I)\Delta^{*|I|}$. The finite expansion of B_2 yields the identity

$$T_{\mathscr{D}_2} - B_2 = -\sum_{I\notin\mathscr{D}_2} \Delta^{*|I|}\left[\prod_{i\in I} p_i + \sum_{\substack{e\subset I\\ |e|=2}} c(e) \prod_{i\in I\setminus e} p_i\right]. \tag{5.8}$$

For the independent term, choose a primitive subfamily of a discarded family. If it has at least three members, use (5.7); if it is a pair e, the complement must be a non-direct extension of S_e, and $\prod_{i\in e} p_i \le |c(e)|$. The pair terms in (5.8) are bounded by the same all-cardinality pair deficit. Free complements have total weight at most $e^{2\mu}$. Consequently

$$\|T_{\mathscr{D}_2} - B_2\|_1 \ll N^{-\beta+\eta}. \tag{5.9}$$

Take $K_N = (5+o(1))\log N/\log\log N$. The exact stopping identity from Section 4 gives

$$\|\mathcal{L}(Z_N) - T_{\mathscr{D}_2}\|_1 \le \sum_{I\in\partial\mathscr{D}_2} 2^{|I|} q(I).$$

For $|I| \le K_N$, expand $q(I)$ in centered coefficients. The terms of size zero and two are the independent and pair deficits already bounded. Every remaining nonzero centered term has either a connected component of size at least three or at least two pair components. Canonical factorization, (5.6), and the free independent complements bound their sum by a constant times

$$e^{4A_2+T} - 1 - 4A_2 = O(T + A_2^2).$$

Here $T = o(1)$ and $A_2^2 \ll N^{-4/3+\eta}$; the centered coefficient of a singleton is zero. This is where the linear contribution of one pair is subtracted before taking an absolute bound.

On the larger boundary, $q(I) \le 2^{2L}\prod_I p_x$, so

$$\sum_{\substack{I\in\partial\mathscr{D}_2\\ |I|>K_N}} 2^{|I|} q(I) \ll_C N^2 \sum_{r>K_N} \frac{(2\mu)^r}{r!} = N^{-3+o(1)}.$$

Combining these bounds, (5.9), and Lemma 5.1 proves (5.5). All internal margins are chosen after the final ε; none is made to vary with N. □

Theorem 5.3 (Reconstruction from triple activity)**.** *With the exact mean and second factorial cumulant,*

$$\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \ll_{C,\varepsilon} N^{\varepsilon} A_3 + N^{-4/3+\varepsilon}. \tag{5.10}$$

In particular the arithmetic triple estimate gives the upper bound $O_{C,\varepsilon}(N^{-4/3+\varepsilon})$ for the full model.

Proof. In the proof of Theorem 5.2, replace the finite activity input by Proposition T12.9. At sizes $3,\dots,256$ its total cost is $N^{\varepsilon'}A_3 + N^{-4/3+\varepsilon'}$; the larger sizes have the second cost by Proposition T12.11. The primitive deficits, pair deficits, and boundary expansion obey the same bounds. In the component generator,

$$T \ll N^{\varepsilon'}A_3 + N^{-4/3+\varepsilon'}, \qquad e^{4A_2+T} - 1 - 4A_2 = O(T + A_2^2).$$

The established arithmetic estimate makes $T = o(1)$, so this passage loses no new fixed power of N. The large boundary costs $N^{-3+o(1)}$ and conversion costs $N^{-5/3+\varepsilon'}$. To make the choices explicit, denote the requested final margin by $\varepsilon_{\mathrm{out}} = \varepsilon$ and choose $0 < \tau < \min(\varepsilon_{\mathrm{out}}/4, 1/960)$. Invoke each completed fixed-order estimate with sufficiently small fixed input margins that its total displayed loss is at most τ. For example, use the pair estimate with loss $\tau/2$ before squaring A_2. Only the fixed orders $3,\dots,256$ occur in this choice; the larger orders use the uniform certificate bounds, not constants from a fixed-order theorem. With $\varepsilon' = \tau$ the preceding argument proves (5.10), since $\tau < \varepsilon_{\mathrm{out}}$.

Finally invoke Corollary 3.4 with loss τ. The triple term is then at most $N^{-41/30+2\tau}$, and

$$\frac{41}{30} - 2\tau - \frac{4}{3} = \frac{1}{30} - 2\tau > 0.$$

The conversion and large-boundary exponents exceed $4/3$ by $1/3$ and $5/3$, respectively. These comparisons give the announced upper bound. All accuracy choices, source counts and fixed activity orders used here are fixed before N; the growing cardinality cutoff K_N is controlled separately. □

Power margins. The strict power savings used in the last stages are recorded below. Each margin compares complete contributions, after coefficient choices, propagation and rank factors have been included.

Comparison	Exact margin	Location
Nonaligned target above a rational pair versus dominant triples	$11/8 - 41/30 = 1/120$	(T10.11)
Incompatible triples versus the third factorial cumulant	$3/2 - 41/30 = 2/15$	(T9.1)
Third activity versus the quadratic remainder	$41/30 - 4/3 = 1/30$	(5.10)
Low-nullity uniform tail versus the quadratic remainder	$27/20 - 4/3 = 1/60$	(T12.20)
Quartic signed error versus α_N^2	$3/2 - 4/3 = 1/6$	(6.2)

The smallest displayed margin is $1/120$. In contrast, two pair activities have size $A_2^2 \ll N^{-4/3+\eta}$ and are retained at the target scale. They have no strict power saving to spend. For a prescribed final accuracy $\varepsilon_0 > 0$, one may choose $0 < \tau < \min(\varepsilon_0/4, 1/960)$ and invoke completed auxiliary estimates with total loss at most τ each. Their own input margins are chosen smaller as needed; in a product, divide the allotted loss between the factors. Thus the same symbol ε in successive statements never requires the same numerical value at every stage. Rank thresholds and fixed geometric degrees have separate roles. The fixed connected-activity estimates in reconstruction are used only through order 256; larger cardinalities have the uniform certificate estimate and a factorial boundary bound.

6 Why the quadratic approximation stops at the fourth-order scale

The quadratic reference retains the second factorial cumulant linearly. Two independent pair contributions therefore leave a fourth-order term of size α_N^2. To turn this observation into a lower bound, one must check that incompatible arithmetic configurations cannot cancel that term. The half-exponent count supplies this missing sign information. The formal term $\alpha_N^2 \Pi_{\mu_N} * \Delta^{*4}/8$ is classical in Poisson–Charlier expansions [5,13]. Lower bounds with a squared second-cumulant contribution are known for locally dependent sums [15, Theorem 2.1, (2.3)]. The issue here is its arithmetic size, the sign of the other fourth-order contributions, and the passage from an unbounded polynomial test to the full norm. Throughout this section $\mu = \mu_N$, $\alpha = \alpha_N$, and all marginals and centered coefficients are those of the full admissible set.

6.1 Cancellation of the disconnected pair terms

Write

$$C_2^{\mathrm{s}} = \sum_{|e|=2} c(e), \qquad K_4 = \sum_{\substack{|I|=4 \\ I \text{ connected}}} c(I), \qquad K_4^+ = \sum_{\substack{|I|=4 \\ I \text{ connected},\ q(I)>0}} c(I).$$

Compatibility makes every character on the relation space positive, so $K_4^+ \geq 0$.

Lemma 6.1 (The fourth factorial cumulant)**.** *For every fixed critical band and every $\varepsilon > 0$,*

$$\kappa_4 = 24K_4 + O_{C,\varepsilon}(N^{-5/3+\varepsilon}), \qquad K_4 = K_4^+ + O_{C,\varepsilon}(N^{-3/2+\varepsilon}). \tag{6.1}$$

Consequently

$$\kappa_4 + 3\alpha^2 = 3\alpha^2 + 24K_4^+ + O_{C,\varepsilon}(N^{-3/2+\varepsilon}). \tag{6.2}$$

Proof. Expanding the first four factorial moments with the individual p_i gives the exact identity

$$\kappa_4 = 24 \sum_{|I|=4} c(I) - 12(C_2^{\mathrm{s}})^2 - 24 \sum_{|T|=3} \Big(\sum_{i\in T} p_i\Big) c(T) + 24 \sum_{e=\{i,j\}} (p_i^2 + p_i p_j + p_j^2) c(e) - 6 \sum_i p_i^4. \tag{6.3}$$

For example, first write

$$\mathbb{E}(1+t)^Z = \prod_i (1+p_i t) \left[1 + \sum_{2\le |I|\le 4} \frac{c(I) t^{|I|}}{\prod_{i\in I}(1+p_i t)} + O(t^5) \right],$$

then take 24 times the coefficient of t^4 in its formal logarithm. This is a finite coefficient calculation; no analytic logarithm of the true generating function is required.

Canonical direct decomposition gives $\sum_{|I|=4} c(I) = K_4 + D_{22}$, where D_{22} sums $c(e)c(f)$ over unordered disjoint pairs of edges whose spaces are direct. In $(C_2^{\mathrm{s}})^2/2 - D_{22}$ only repeated edges, edges sharing a vertex, and disjoint edges with a nonzero intersection remain. The positive propagation estimate, together with the product of the marginals, implies

$$\sum_{y\neq x} |c(x,y)| \ll_{C,\eta} p_x N^{\eta}.$$

Repeated edges and shared vertices consequently have total absolute weight $O(p_* A_2 N^{\eta}) = O(N^{-5/3+2\eta})$.

Apply the pair partition of Lemma T9.4 with a small fixed parameter η. A product involving a weak or exceptional edge has total weight at most $A_2(A_{\mathrm{w}} + A_{\mathrm{s}}) \ll N^{-5/3+O(\eta)}$. For two good edges from the remaining finite family of channels, fix the first edge and the offset pattern of the second compression. There are $NB^{O(1)}$ choices for the first edge. A nonzero intersection imposes an exceptional target of fixed width, giving $N^{1/2+o(1)}$ choices for the second compressed center by the elementary exceptional-target count in Proposition T1.15. The proof of that count uses only a fixed width and applies to this compression as well. Each product has weight at most $N^{-10/3+o(1)}$. Hence this last cost is $N^{-11/6+o(1)}$, and

$$D_{22} = (C_2^{\mathrm{s}})^2/2 + O(N^{-5/3+\varepsilon}).$$

The remaining terms of (6.3) are bounded by $O(p_* A_3 + p_*^2 A_2 + p_*^3 \mu)$. The triple estimate in Corollary 3.4 gives $A_3 \ll N^{-41/30+\eta}$, so each is smaller than the stated error. This proves the first identity.

The incompatible part of Proposition T12.9 at the fixed order $r = 4$ gives

$$|K_4 - K_4^{+}| \le \sum_{\substack{|I|=4,\ I \text{ connected} \\ q(I)=0}} |c(I)| \ll_{C,\varepsilon} N^{-3/2+\varepsilon}.$$

Compatible coefficients are nonnegative. This proves the second identity and then (6.2). □

6.2 A Charlier polynomial detects the remaining mass

Theorem 6.2 (Optimal power for the exact quadratic reference)**.** *For every $C > 0$ and $\varepsilon > 0$, there are $c_{C,\varepsilon} > 0$ and $N_0(C,\varepsilon)$ such that*

$$\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \ge c_{C,\varepsilon} N^{-4/3-\varepsilon} \tag{6.4}$$

whenever $N \ge N_0$ and $|L - \log_2 N| \le C$. Together with Theorem 5.3 and the triple activity bound, this identifies the optimal power as $4/3$. It does not assert an asymptotic constant or a logarithmic factor for the remainder.

Proof. The arithmetic coefficient estimate gives $\alpha \asymp_C LN^{-2/3} > 0$ uniformly in the critical band. The sign argument has the margin $3/2 - 4/3 = 1/6$. Choose a fixed total error loss $\tau < 1/6$ in (6.2); then

$$\frac{N^{-3/2+\tau}}{\alpha^2} \ll_C L^{-2}N^{-1/6+\tau} \longrightarrow 0.$$

Positivity therefore yields

$$\kappa_4 + 3\alpha^2 \geq (3 - o_C(1))\alpha^2.$$

The fourth Charlier polynomial is

$$H_4(z;\mu) = (z)_4 - 4\mu(z)_3 + 6\mu^2(z)_2 - 4\mu^3 z + \mu^4 = 4![t^4]e^{-\mu t}(1+t)^z.$$

Its integral against $Q_{\mu,\alpha}$ vanishes, because the factorial generating function of that reference is $e^{\mu t}(1 + \alpha t^2/2)$. On the true law its expectation equals $\kappa_4 + 3\alpha^2$. Thus, with $R = \mathcal{L}(Z_N) - Q_{\mu,\alpha}$,

$$\left|\int H_4\, dR\right| \gg_C L^2N^{-4/3}.$$

The test is unbounded, so a moment argument is needed before it gives a norm estimate. For every fixed integer $k > 1$, the fixed moments from Theorem T4.4 give $\int |H_4(z;\mu)|^k\, d\mathcal{L}(Z_N)(z) \ll_{C,k} 1$. The same bound holds for $|Q_{\mu,\alpha}|$, since μ stays in a compact subset of $(0,\infty)$, $\alpha = O_C(1)$, and

$$|Q_{\mu,\alpha}| \leq \Pi_\mu + \frac{|\alpha|}{2}\Pi_\mu * (\delta_2 + 2\delta_1 + \delta_0).$$

Consequently Hölder's inequality for the positive measure $|R|$ gives

$$\left|\int H_4\, dR\right| \ll_{C,k} \|R\|_1^{1-1/k}, \qquad \|R\|_1 \gg_{C,k} (L^2N^{-4/3})^{k/(k-1)}.$$

After ε has been prescribed, choose the integer k so that $4/[3(k-1)] \leq \varepsilon$. It remains fixed as N tends to infinity, and proves (6.4). The upper bound follows from the reconstruction and $A_3 \ll N^{-41/30+\eta}$, since $41/30 > 4/3$. □

The obstruction is specific to the quadratic reference. The Hermite law with the same first two factorial cumulants already contains the term $\alpha^2\Pi_\mu * \Delta^{*4}/8$. The fourth Charlier test against that law sees κ_4 instead of $\kappa_4 + 3\alpha^2$, so the lower bound above does not establish an optimal Hermite approximation rate.

7 Completing the approximation and its consequences

The reconstruction is carried out in the algebra of signed measures. For the parameters of the model its final reference is a probability. The following positivity and continuity facts are elementary and do not depend on the law approximation. The deduction of Corollary 1.2 from the main theorem is given below.

7.1 Continuity of the reference

For nonnegative μ, μ',

$$\|\Pi_\mu - \Pi_{\mu'}\|_1 \leq 2|\mu - \mu'|. \tag{7.1}$$

For example, if $\mu' = \mu + h$, couple by adding an independent Poisson variable of mean h. Its probability of being nonzero is $1 - e^{-h} \leq h$, proving the inequality. It follows that

$$\|Q_{\mu,\alpha} - Q_{\mu',\alpha'}\|_1 \leq 2(1 + 2|\alpha|)|\mu - \mu'| + 2|\alpha - \alpha'|. \tag{7.2}$$

Here we have expanded the difference with coefficient α on $\Pi_\mu - \Pi_{\mu'}$ and used $\|\Pi_{\mu'} * \Delta^{*2}\|_1 \leq 4$.

7.2 Positivity and the coefficient of total variation

The signed representation of the reference is useful in the proof, but for the parameters of the theorem it is eventually a probability. The following exact criterion also specifies its domain for general parameters.

Lemma 7.1 (Positivity of the quadratic Poisson correction)**.** *Let $\mu > 0$, $\alpha \in \mathbb{R}$, and $\theta = \mu - \lfloor \mu \rfloor$. Then $Q_{\mu,\alpha}$ is a probability measure if and only if*

$$0 \le \alpha \le \frac{2\mu^2}{\mu + \theta(1-\theta)}. \tag{7.3}$$

Its masses are

$$Q_{\mu,\alpha}\{k\} = e^{-\mu}\frac{\mu^k}{k!}\left[1 + \frac{\alpha}{2\mu^2}\big((k-\mu)^2 - k\big)\right], \qquad k \ge 0. \tag{7.4}$$

Proof. Formula (7.4) follows by expressing the Poisson masses at $k-1$ and $k-2$ relative to the mass at k; negative-index masses are taken to be zero. The total mass is one because Δ^{*2} has total mass zero.

If $\alpha < 0$, the bracket is negative for all sufficiently large k, so positivity is impossible. For $\alpha \ge 0$, minimize the quadratic $f(k) = (k-\mu)^2 - k$ over nonnegative integers. Its successive difference is $f(k+1) - f(k) = 2(k-\mu)$. If μ is not an integer, the minimum is attained at $k = \lceil \mu \rceil$; if μ is an integer, it is attained at both $k = \mu$ and $k = \mu + 1$. In either case its value is $-\mu - \theta(1-\theta)$. Nonnegativity of the bracket is therefore exactly (7.3). □

In the theorem, $\alpha_N > 0$ for large N and $\alpha_N \to 0$, while μ_N and λ_N stay in a compact subinterval of $(0, \infty)$. Both corrected references are therefore probabilities for all sufficiently large N. The norm estimate simultaneously controls all their masses; it is not only a pointwise assertion at each fixed count.

Define, for $\lambda > 0$,

$$\mathcal{C}(\lambda) = \frac{1}{4}\|\Pi_\lambda * \Delta^{*2}\|_1 = \frac{1}{4}\mathbb{E}\left|1 - \frac{2P_\lambda}{\lambda} + \frac{P_\lambda(P_\lambda - 1)}{\lambda^2}\right|, \qquad P_\lambda \sim \mathrm{Pois}(\lambda). \tag{7.5}$$

The coefficient is explicit even when λ_N does not converge. A summable signed sequence of total mass zero has ℓ^1 norm at least twice any positive coefficient. The mass $e^{-\lambda}$ at zero of $\Pi_\lambda * \Delta^{*2}$ and its convolution norm therefore give

$$\frac{e^{-\lambda}}{2} \le \mathcal{C}(\lambda) \le 1. \tag{7.6}$$

7.3 Completion of the main theorem

The arithmetic and probabilistic estimates now meet at a single quantity, the absolute third-order activity. We complete the main theorem, then record what its precision allows for unbounded scalar observables and for the choice of mean.

Proof of Theorem 1.1. Write $\varepsilon_{\mathrm{out}} = \varepsilon$ for the requested final margin and choose $0 < \tau < \min(\varepsilon_{\mathrm{out}}/4, 1/960)$. Invoke Theorem 5.3 with accuracy τ and Corollary 3.4 with accuracy τ. The triple contribution is

$$N^{\tau} A_3 \ll N^{-41/30+2\tau} = N^{-4/3-(1/30-2\tau)}.$$

Since $1/30 - 2\tau > 0$ and the other reconstruction term is $N^{-4/3+\tau}$, this proves the upper bound. Theorem 6.2 gives the lower bound for the same exact parameters, using its own fixed Charlier interpolation order.

The mean estimate is Proposition T2.1. The pair asymptotic (1.5) shows that $\alpha_N > 0$, $\alpha_N \to 0$, and μ_N stays in a compact subset of $(0, \infty)$. Lemma 7.1 therefore makes the reference a probability measure for sufficiently large N. Every auxiliary accuracy parameter and fixed order was chosen after C, ε and before N. The conclusion is uniform on the entire fixed critical band. □

7.4 Polynomial weights and fixed moments

Corollary 7.2 (Every fixed polynomial weight)**.** *For every fixed integer $m \geq 0$ and $\varepsilon > 0$,*

$$\sum_{k\geq 0}(1+k)^m \left|\mathbb{P}(Z_N = k) - Q_{\mu_N,\alpha_N}\{k\}\right| \ll_{C,m,\varepsilon} N^{-4/3+\varepsilon}. \tag{7.7}$$

For every fixed $r \geq 2$, consequently,

$$\mathbb{E}(Z_N)_r = \mu_N^r + \binom{r}{2}\mu_N^{r-2}\alpha_N + O_{C,r,\varepsilon}(N^{-4/3+\varepsilon}). \tag{7.8}$$

The equalities for $r = 0, 1, 2$ hold without a remainder when the right side is interpreted in its natural form.

Proof. Let $R = \mathcal{L}Z_N - Q_{\mu_N,\alpha_N}$. For every fixed integer $v > 1$, Hölder gives

$$\sum_k (1+k)^m |R\{k\}| \leq \|R\|_1^{1-1/v} \left(\sum_k (1+k)^{mv}|R\{k\}|\right)^{1/v}.$$

The second factor is uniformly bounded. For the true law this follows from the independent fixed-moment proof in Theorem T4.4; for the reference it follows by dominating its variation by a fixed sum of shifted Poisson laws, as in Theorem 6.2. Choose v sufficiently large after ε and then a sufficiently small margin in the main theorem. This proves (7.7). The factorial generating function of $Q_{\mu,\alpha}$ is $e^{\mu t}(1 + \alpha t^2/2)$; its coefficient of t^r, multiplied by $r!$, gives the moment formula. □

This interpolation concerns the scalar count and fixed polynomial degrees.

Corollary 7.3 (No retained start)**.** *Uniformly in a fixed critical band,*

$$\mathbb{P}(Z_N = 0) = e^{-\mu_N}\left(1 + \frac{\alpha_N}{2}\right) + O_{C,\varepsilon}(N^{-4/3+\varepsilon}).$$

Proof. Evaluate the approximation at zero, using $(\Pi_\mu * \Delta^{*2})\{0\} = e^{-\mu}$. □

The event is the absence of a retained start in I_N. It does not exclude a long run that started before N.

7.5 Why the mean must be specified at this precision

Proposition 7.4 (The quantized mean bias)**.** *Let η_x indicate one-window compatibility and put $h_x = L - \dim S_x$. Then*

$$\mu_N - \lambda_N = pM_{N,L}, \qquad M_{N,L} = \sum_{x\in I_N}(\eta_x 2^{h_x} - 1) \in \mathbb{Z}. \tag{7.9}$$

Thus either $\mu_N = \lambda_N$ or $|\mu_N - \lambda_N| \geq p \asymp_C N^{-1}$. Keeping the same exact α_N, one has

$$\|\mathcal{L}Z_N - Q_{\lambda_N,\alpha_N}\|_1 \ll_{C,\varepsilon} N^{-1+\varepsilon}.$$

On every sufficiently large admissible pair (N, L) for which $M_{N,L} \neq 0$,

$$\|\mathcal{L}Z_N - Q_{\lambda_N,\alpha_N}\|_1 \asymp_C |\mu_N - \lambda_N|. \tag{7.10}$$

No assertion that $M_{N,L} \neq 0$ at infinitely many scales is included.

Proof. Sum the one-window formula $\mathbb{E}J_x = \eta_x p 2^{h_x}$ over the exactly N centres to obtain (7.9). For two means in the compact range in question, (7.2) gives

$$\|Q_{\mu,\alpha_N} - Q_{\lambda,\alpha_N}\|_1 \ll_C |\mu - \lambda|.$$

The reverse bound follows from the mass at zero, $e^{-\theta}(1+\alpha_N/2)$, by the mean value theorem. The upper bound for the nominal reference follows from the main theorem and $|\mu_N - \lambda_N| \ll N^{-1+\varepsilon}$.

If $M_{N,L} \neq 0$, the mean difference is at least p. Choose a fixed margin less than $1/3$ in the main theorem; its remainder is then $o(p)$. The triangle and reverse triangle inequalities transfer the two-sided reference comparison to (7.10). □

Irregular compatible and incompatible windows contribute with opposite signs to $M_{N,L}$. Their presence alone does not prove a nonzero total bias.

7.6 The total-variation coefficient

Proof of Corollary 1.2. Theorem 1.1 and the reverse triangle inequality first give

$$d_{\mathrm{TV}}(\mathcal{L}Z_N, \Pi_{\mu_N}) = \alpha_N \mathcal{C}(\mu_N) + O_{C,\varepsilon}(N^{-4/3+\varepsilon}).$$

Use $|\mu_N - \lambda_N| \ll N^{-1+\varepsilon}$ and (7.2) to obtain

$$\mathcal{L}Z_N - \Pi_{\lambda_N} = \frac{\alpha_N}{2}\Pi_{\lambda_N} * \Delta^{*2} + O_{\ell^1,C,\varepsilon}(N^{-1+\varepsilon}).$$

The reverse triangle inequality, followed by the factor $1/2$ converting the ℓ^1 norm to total variation, yields

$$d_{\mathrm{TV}}(\mathcal{L}Z_N, \Pi_{\lambda_N}) = \alpha_N \mathcal{C}(\lambda_N) + O_{C,\varepsilon}(N^{-1+\varepsilon}). \tag{7.11}$$

The pair asymptotic from Theorem 3.1 now gives

$$d_{\mathrm{TV}}(\mathcal{L}Z_N, \Pi_{\lambda_N}) \sim \mathcal{C}(\lambda_N)\frac{r_N+2}{18\,2^{r_N/3}}\,\lambda_N^{5/3} L N^{-2/3}, \qquad r_N = L \bmod 3. \tag{7.12}$$

The lower bound in (7.6) is uniform on the fixed critical band, and the remainder is $o(LN^{-2/3})$. This proves the sharp order asserted in the corollary. The equivalence is relative to the displayed oscillating expression; no convergence of its phase or intensity factors is required. □

Replacing α_N by its leading equivalent proves (7.12). It does not by itself preserve the error $O(N^{-1+\varepsilon})$ in (7.11), which uses the exact second factorial cumulant.

Proposition T14.1 gives $\|\Pi_1 * \Delta^{*2}\|_1 = 3/e$, hence $\mathcal{C}(1) = 3/(4e)$. Along $N = 2^L$ with $3 \mid L$, the pair coefficient therefore gives

$$d_{\mathrm{TV}}(\mathcal{L}Z_N, \Pi_1) \sim \frac{1}{12e}\, L N^{-2/3}. \tag{7.13}$$

This is the coefficient in the example announced in the introduction.

The total-variation asymptotic uses the exact-marginal reconstruction and the explicit pair coefficient. The independent positivity and reference-continuity estimates do not use the reconstruction.

8 Explicit references and the first omitted channel

The exact second factorial cumulant retains every pair interaction. An explicit reference can instead retain the complete contributions of all rational channels up to a chosen height. The first omitted height then determines its error, including the leading coefficient. This gives a different optimal exponent for each finite truncation, even though all these references approximate the same random variable.

For coprime integers $b/2 < a < b$, $b \ge 3$, retain the finite sums $S_{a,b}(L)$ from (3.2) and put

$$w_{a,b} = \frac{2(2a-b)}{ab}, \qquad T_b(L) = \sum_{\substack{b/2<a<b \\ (a,b)=1}} w_{a,b} S_{a,b}(L),$$
$$\widehat{\alpha}_N^{(\le h)} = \frac{\lambda_N^2}{N} \sum_{b=3}^{h} T_b(L), \qquad h \ge 3. \tag{8.1}$$

The factor two in $w_{a,b}$ counts the two orientations of a pair. Here and throughout this section, the integer h is fixed before N. In particular, $\widehat{\alpha}_N^{(\le 4)} = \widehat{\alpha}_N$ from (3.10).

Theorem 8.1 (The optimal error of every fixed height truncation)**.** *Fix $h \ge 3$ and $C > 0$, and write $b = h+1$ and $r = L \bmod b$. For every $\varepsilon > 0$, uniformly in the critical band,*

$$\alpha_N - \widehat{\alpha}_N^{(\le h)} = \frac{\lambda_N^2}{N} T_b(L) + O_{h,C,\varepsilon}(N^{-b/(b+1)+\varepsilon}), \tag{8.2}$$

$$\mathcal{L}(Z_N) - Q_{\lambda_N, \widehat{\alpha}_N^{(\le h)}} = \frac{\lambda_N^2 T_b(L)}{2N} \Pi_{\lambda_N} * \Delta^{*2} + O_{\ell^1,h,C,\varepsilon}(N^{-b/(b+1)+\varepsilon}). \tag{8.3}$$

Define the positive number

$$c_b(r) = \frac{b+r-1}{b^2\, 2^{r/b}} \sum_{\substack{b/2<a<b \\ (a,b)=1}} \frac{(2a-b)(b-a)}{a}. \tag{8.4}$$

The reference is a probability measure for all sufficiently large N, and

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), Q_{\lambda_N, \widehat{\alpha}_N^{(\le h)}}) \sim \frac{c_b(r)}{4} \lambda_N^{2-1/b} \|\Pi_{\lambda_N} * \Delta^{*2}\|_1 \, L N^{-1+1/b}. \tag{8.5}$$

The ratio tends to one uniformly over the critical band with the displayed moving parameters retained. Thus the optimal power exponent of this specified reference is $h/(h+1)$.

Before proving the theorem, we account for the nonaligned pairs. The unweighted count alone would not suffice: a pair carrying many relations has a correspondingly larger joint probability.

Lemma 8.2 (Weighted nonaligned pairs)**.** *For every $\varepsilon > 0$,*

$$\sum_{\substack{x,y \in G_N,\ |x-y|>L \\ (x,y)\ \text{nonaligned}}} (2^{\rho(x,y)} - 1) \ll_{C,\varepsilon} N^{1+\varepsilon}. \tag{8.6}$$

Consequently the sum of the absolute covariances on these pairs is $O_{C,\varepsilon}(N^{-1+\varepsilon})$.

Proof. Use the large-prime component decomposition (T2.5), so that $\rho(x,y) \le D(x,y) + c(x,y)$ and $D(x,y) = o(B)$ uniformly. Fix $\eta > 0$ temporarily. On $c(x,y) < \eta B$, the weight is at most $2^{\eta B + o(B)}$. A pair in G_N with $\rho(x,y) > 0$ has a nonzero intersection between its two effective spaces. The pair count of Corollary T6.3 therefore bounds the contribution of this sector by $N^{1+\eta+\varepsilon'+o(1)}$. On $c(x,y) \ge \eta B$, Lemma T1.8 already bounds the entire weighted sum by $N^{1+\varepsilon'}$. Choose η and then ε' inside the prescribed margin.

On G_N both marginals are p. The affine identity (T2.1) gives $|\operatorname{Cov}(J_x, J_y)| \le p^2(2^{\rho(x,y)} - 1)$, including incompatible pairs. Since $p \asymp_C N^{-1}$, this proves the last assertion. □

Proof of Theorem 8.1. We first prove, for every fixed integer $m \geq 3$,

$$\alpha_N - \widehat{\alpha}_N^{(\leq m)} = O_{m,C,\varepsilon}(N^{-m/(m+1)+\varepsilon}). \tag{8.7}$$

The regular-set transfer in Proposition T2.1, justified by the independent fixed-moment bound of Theorem T4.4, changes the second factorial cumulant by $O(N^{-1+\varepsilon})$. On G_N, that cumulant is the sum of covariances over ordered distinct pairs minus $|G_N|p^2$. The last term is $O(N^{-1})$.

Strict contacts are incompatible and number $O(NB)$, so their covariance mass is $O(B/N)$. At separation exactly L, the union of the two windows lies in an interval of length $O(B)$ and has formal rank $2L - o(L)$, by Lemma T1.3; this band has the same $N^{-1+o(1)}$ cost. Lemma 8.2 treats the separated nonaligned pairs.

For aligned separated pairs, reduced slopes are unique for sufficiently large N: their separation is at least B^{-6}, whereas the allowed offset changes the ratio of the starts by only $B^{O(1)}/N$. There are $NB^{O(1)}$ aligned hosts altogether. A slope of height one has no exact unit for separated windows and hence only residual nullity $o(L)$ by Lemma T1.9. The slope $1:2$ is confined to $B^{O(1)}$ pairs at the dyadic endpoints; its total mass is at most $N^{-3/2+o(1)}$. Every remaining reduced slope is an interior channel $b/2 < a < b$, $b \geq 3$.

For each fixed channel of height at most m, normalize to $bV_x \cup aV_y$. Lemma T3.5 gives its ordered contribution $p^2 N w_{a,b} S_{a,b}(L) + O_{m,C,\varepsilon}(N^{-1+\varepsilon})$. Changing from the good-pair joint mass to the covariance costs at most $O_m(p^2 NB)$. The omitted pairs of bad normalization have total weight $N^{-3/2+1/b+o(1)}$, as in that lemma's proof, and are smaller still. Only finitely many channels have been retained.

For all remaining heights $b \geq m+1$, including heights varying up to B^3, the uniform aligned rank bound gives $\rho(x,y) \leq L/(m+1) + o(L)$. Their total contribution, in absolute value, is at most

$$p^2 N B^{O(1)} 2^{L/(m+1)+o(L)} = N^{-m/(m+1)+o(1)}.$$

This proves (8.7). Apply it at $m = h+1$ and separate the last retained term to obtain (8.2).

The mean estimate of Proposition T2.1 gives $|\mu_N - \lambda_N| \ll N^{-1+\varepsilon'}$. Coupling Poisson variables shows $\|\Pi_u - \Pi_v\|_1 \leq 2|u-v|$. Since $\alpha_N = o(1)$,

$$\|Q_{\mu_N,\alpha_N} - Q_{\lambda_N,\alpha_N}\|_1 \ll |\mu_N - \lambda_N|.$$

Theorem 1.1, this estimate, and the exact linearity of $Q_{\lambda,a}$ in a give (8.3). Both the mean error of power one and the exact-reference error of power $4/3$ are smaller than $N^{-b/(b+1)+\varepsilon}$ for this fixed b.

To identify the leading term, write $L = bq + r$, $0 \leq r < b$. The finite identity (3.2), with $s = -1$, yields

$$S_{a,b}(L) = 2^{q-1}\big[(b-a)(b+r-1)q + O_b(1)\big] + O_b(L).$$

As $q \sim L/b$ and $2^q = 2^{-r/b}(N/\lambda_N)^{1/b}$, summation gives

$$\frac{\lambda_N^2}{N} T_b(L) \sim c_b(r) \lambda_N^{2-1/b} L N^{-1+1/b}. \tag{8.8}$$

Every summand in (8.4) is positive, and $a = b-1$ is always present. The phase ranges over a fixed finite set and λ_N over a compact subset of $(0,\infty)$. Taking $\varepsilon < 1/(b(b+1))$ makes the error in (8.3) negligible relative to its leading term. The reverse triangle inequality then gives the norm equivalent; division by two gives (8.5). The factor $1/4$ combines this division with the quadratic coefficient $1/2$.

Finally, $\widehat{\alpha}_N^{(\leq h)} \sim \alpha_N > 0$ and tends to zero. The positivity criterion of Lemma 7.1 therefore applies uniformly on the compact range of λ_N. The leading coefficient in (8.5) is bounded above and away from zero, since $\|\Pi_{\lambda_N} * \Delta^{*2}\|_1 \geq e^{-\lambda_N}$. This proves optimality of the stated power for the specified reference. □

8.1 What the first omitted height contributes

The four channels of heights at most five are $2:3$, $3:4$, $3:5$, and $4:5$. Thus the next explicit correction after (3.10) has a short closed formula.

Corollary 8.3 (The error of the two-channel reference)**.** *Write $L = 5q + r$, $0 \le r < 5$, and, for $q \ge 1$, set*

$$\begin{aligned} M_5(N,L) &= \frac{\lambda_N^2}{N}\left\{\frac{2}{15}S_{3,5}(L) + \frac{3}{10}S_{4,5}(L)\right\} \\ &= \frac{\lambda_N^2}{30N}\left\{2^{q-1}\big[17(r+4)q + 13r^2 + 87r + 620\big] - 113L - 127\right\}. \end{aligned} \tag{8.9}$$

Then

$$\mathcal{L}(Z_N) - Q_{\lambda_N,\widehat{\alpha}_N} = \frac{M_5(N,L)}{2}\Pi_{\lambda_N} * \Delta^{*2} + O_{\ell^1,C,\varepsilon}(N^{-5/6+\varepsilon}), \tag{8.10}$$

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), Q_{\lambda_N,\widehat{\alpha}_N}) \sim \frac{17(r+4)}{1200\,2^{r/5}}\lambda_N^{9/5}\|\Pi_{\lambda_N} * \Delta^{*2}\|_1\, LN^{-4/5}. \tag{8.11}$$

The four-channel coefficient is $\widehat{\alpha}_N^{(\le 5)} = \widehat{\alpha}_N + M_5(N,L)$. Its reference has optimal exponent $5/6$, with leading phase constant $c_6(r) = (r+5)/(45\,2^{r/6})$, where now $r = L \bmod 6$.

Proof. The finite identity gives

$$\begin{aligned} S_{3,5}(L) &= 2^{q-1}\big[2(r+4)q + r^2 + 6r + 38\big] - 8L - 7, \\ S_{4,5}(L) &= 2^{q-1}\big[(r+4)q + r^2 + 7r + 52\big] - 9L - 11. \end{aligned}$$

Their weighted sum proves (8.9). Theorem 8.1 at $h = 4$ gives the first two conclusions. At height six the only reduced interior channel is $5:6$; substitution in (8.4) gives c_6 and the assertion for the four-channel reference. □

For example, if $N = 2^L$ and $5 \mid L$, the error of the two-channel reference is asymptotic in total variation to $17LN^{-4/5}/(100e)$, because $\|\Pi_1 * \Delta^{*2}\|_1 = 3/e$. The table records how successive complete channels change the error.

Height h	Retained channels	Optimal power	First omitted channels
3	$2:3$	$3/4$	$3:4$
4	$2:3,\ 3:4$	$4/5$	$3:5,\ 4:5$
5	$2:3,\ 3:4,\ 3:5,\ 4:5$	$5/6$	$5:6$
h fixed	all reduced interior heights $\le h$	$h/(h+1)$	height $h+1$

The full coefficient in (8.1) has a phase dividing $\mathrm{lcm}(3,\dots,h)$, whereas its leading omitted term uses only $L \bmod (h+1)$, together with λ_N. The theorem permits any fixed height, with constants depending on that height. It makes no assertion for a height growing with N. In particular it does not identify an obstruction of power one for other explicit references: the stated optimality concerns precisely the finite truncations defined in (8.1).

8.2 Positive Hermite references

The classical Hermite distribution [18] has generating function

$$\exp\{(\mu-\alpha)(z-1) + (\alpha/2)(z^2-1)\}.$$

For $0 \le \alpha \le \mu$, it is the law of a Poisson variable of mean $\mu-\alpha$ plus twice an independent Poisson variable of mean $\alpha/2$.

Corollary 8.4 (Hermite approximation with single and double jumps)**.** *For all sufficiently large N, let U_N, V_N be independent Poisson variables with means $\mu_N - \alpha_N$ and $\alpha_N/2$, respectively. Then*

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), \mathcal{L}(U_N + 2V_N)) \ll_{C,\varepsilon} N^{-4/3+\varepsilon}.$$

For every fixed $h \geq 3$, the entirely explicit reference

$$\widehat{U}_N + 2\widehat{V}_N, \qquad \widehat{U}_N \sim \mathrm{Pois}(\lambda_N - \widehat{\alpha}_N^{(\leq h)}), \qquad \widehat{V}_N \sim \mathrm{Pois}(\widehat{\alpha}_N^{(\leq h)}/2),$$

with independent summands has optimal power $h/(h+1)$ and the same leading distance equivalent as the corresponding quadratic reference in Theorem 8.1. In particular the two-channel choice $h = 4$ has error $O(N^{-4/5+\varepsilon})$. No matching lower bound is asserted here for the exact-parameter Hermite reference.

Proof. Both means are nonnegative for large N. In the convolution algebra, the law of $U_N + 2V_N$ is

$$\exp_*\left(\mu_N \Delta + \frac{\alpha_N}{2}\Delta^{*2}\right).$$

Expanding the second exponential to first order gives Q_{μ_N,α_N} with an ℓ^1 error $O_C(\alpha_N^2)$. Indeed $\|\Delta^{*2}\|_1 = 4$, and the exponential remainder is bounded by $\frac{1}{2}(2\alpha_N)^2 e^{2\alpha_N}$. Since $\alpha_N^2 = N^{-4/3+o(1)}$, the main theorem proves the claim. For each fixed-height explicit reference, the coefficient is positive and tends to zero uniformly in the critical band, while λ_N stays bounded away from zero. Thus both proposed Poisson means are nonnegative for all sufficiently large N. Repeat the same expansion with parameters $(\lambda_N, \widehat{\alpha}_N^{(\leq h)})$, and apply Theorem 8.1. The additional norm error is $O_C((\widehat{\alpha}_N^{(\leq h)})^2) = N^{-4/3+o(1)}$, smaller than the positive first omitted channel term of order $LN^{-h/(h+1)}$. The reverse triangle inequality preserves its equivalent and hence its optimal exponent. □

The double jump represents the leading pair interaction at the level of the scalar law. The contribution here is the precision of the Hermite approximation for Z_N. Its double jump intensity uses either the exact second factorial cumulant or a complete finite channel sum. This is a representation of the reference law; it does not decompose the actual run starts into independent components. Keeping only the leading equivalent of that coefficient does not preserve the precision of a fixed-height truncation. The exact Hermite law already contains the quartic term $\alpha_N^2 \Pi_{\mu_N} * \Delta^{*4}/8$, so the lower bound for the quadratic reference does not transfer to it.

This distinction is already present in the locally dependent setting of Čekanavičius and Vellaisamy [13, Theorem 3.1, (3.5), and Theorem 3.2, (3.6)]. Their quadratic reference is $\Pi_{\Gamma_1} * (\delta_0 + \Gamma_2 \Delta^{*2})$, whereas the exponential reference is $\exp_*(\Gamma_1 \Delta + \Gamma_2 \Delta^{*2})$, with $\Gamma_2 = (\mathrm{Var}\, Z - \mathbb{E}Z)/2$. The quadratic estimate contains an additional squared error term that is absent from the exponential estimate. These results concern sums reduced to one-dependence under specified moment conditions. We use this comparison to place the references in context; the arithmetic estimates here provide their approximation to the multiplicative run count.

9 What the proof makes reusable

The exponent $4/3$ results from the relation between two scales: the leading pair mass is $N^{-2/3+o(1)}$, and the quadratic reference omits a contribution formed from two such interactions. The arithmetic work shows that all other terms relevant to the reconstruction can be controlled at this precision. The proof also separates several choices that can be made in other rare-event problems.

9.1 Choose a reference before estimating the error

The independent Bernoulli reference retains every exact marginal. Its centered expansion has no one-site term, and repeated indices are already accounted for. The retained pair correction then removes the complete second-order contribution. This makes the quantity to be bounded a residual interaction, with a specified sign and normalization.

A homogeneous envelope provides a second useful reduction. The joint probability of the homogeneous constraints bounds the actual joint probability, the exact marginal rank weight and the absolute centered activity. Geometric counting can therefore be performed once and used for each of these weights. At a fixed arity r, the comparison with the nominal weight costs at most 2^r. That fixed factor can be absorbed into the constant of a fixed-order estimate. The estimates that sum over unbounded arities keep their own weights and quantifiers.

Resultants and degree-uniform curve bounds also appear in Shparlinski and Sleiman's count of consecutive multiplicatively dependent triples [17, Theorem 1.1 and Section 3]. The additional feature used here is the simultaneous placement of a large space of square-class relations on one common curve, together with the integer fibres that those relations force in the offset grid. This converts geometric information about logarithmic windows into a bound for their shared rank.

9.2 Stop where the remaining dependence is still positive

The specialization of the abstract lace expansion [12] in Section 4 retains a hereditary class of subsets with computable joint probabilities. At the first excluded prefix, all later variables remain in a positive conditional measure. Its total mass is the joint probability of the prefix. The boundary therefore has a direct probabilistic weight even though the retained expansion is signed. Incompatible pair and triple terms must still be combined before absolute values are taken.

To use this identity in another model, one needs a retained class that matches the chosen reference and quantitative control of the first excluded configurations. Here compression makes a dependent pair one affine block. Conditional arithmetic counts control a fixed number of blocks, and congruence certificates control the large-cardinality part. The reconstruction estimate records exactly what these inputs accomplish:

$$\|\mathcal{L}(Z_N) - Q_{\mu_N,\alpha_N}\|_1 \ll N^{\varepsilon} A_3 + N^{-4/3+\varepsilon}.$$

The finite identity itself does not give the arithmetic bounds needed on its right-hand side.

9.3 Separate fixed orders from a uniform tail

A fixed-order theorem allows its constant and scale threshold to depend on the order. It cannot be summed up to an order growing with N without further information. The proof instead chooses a finite cutoff after the desired accuracy, treats the finitely many lower orders by their individual estimates, and uses one uniform weighted tail above that cutoff.

Proposition T13.1 illustrates this arrangement. Finite layers give a half power for each effective block, while the independent tail gives every power strictly below 2. Their combination yields the exponent $\min(s/2, 2)$, with an arbitrary fixed margin and a common constant for all $s \geq 2$. This is a bound on a deterministic sum of independent counting weights. It is not an exponential-tail statement for the law of Z_N.

The independent fixed-moment proof has a different use: it turns a polynomial test into a norm lower bound and gives fixed polynomially weighted consequences of the main theorem. The finite-difference example in Section T15 shows why even very strong information at each fixed moment order does not by itself provide a polynomial rate for the complete law.

9.4 Test the approximation that was actually retained

The fourth Charlier polynomial detects $\kappa_{4,N} + 3\alpha_N^2$ after subtraction of the quadratic reference. The nonnegative compatible part of the quartic identity prevents cancellation at the scale of α_N^2. For the exact Hermite reference the same test instead detects $\kappa_{4,N}$, because its exponential already contains the two-pair term. A lower bound is therefore tied to a specified reference, not just to the underlying dependent count.

The full equivalent for the exact quadratic remainder, the full constant for the third factorial cumulant, and the sharp error of the exact Hermite reference remain open in this work. So do the errors after cubic or quartic correction and the optimization of total variation over all Poisson means. The explicit channel coefficients determine reference laws, but the constants in the asymptotic remainders and the thresholds in N are not supplied numerically. Applying the method elsewhere requires counterparts of the conditional counts, compatibility cancellations and uniform boundary estimate; the stopping identity alone does not impose a universal approximation exponent.

Formal proofs and reproducibility

The accompanying Lean development provides proofs for the 118 numbered results in the article and companion (26 and 92, respectively). Its statement index records the correspondence of clauses, definitions and inherited hypotheses, including the conditional premise of Theorem 5.2. The documentation of the recorded replay is fixed at commit `a1c8092f7e25`. The linked source manifest and validation record identify the proof source used for the recorded full replay as `aa12e797ce89`. That record specifies Lean 4.36.0-rc2 and Mathlib commit `9d581dfff2e4`, and includes compilation results, extracted statement types and the kernel-axiom audit. The documentation snapshot and the replayed proof source have distinct roles.

The correspondence between the written and formal statements was examined with AI assistance. It is not a machine translation of the written proofs or a substitute for independent human review. The conditional-sum and propagation clauses of Proposition T4.3 allow homogeneous prescriptions; its small-contact clause uses the original run-start events. The formal interfaces distinguish these scopes as well.

The final five-dossier allocation selects 42 formal statements covering every clause of the article's 26 numbered results. It preserves the four issued registrations and merges the two previously unregistered remainder scopes. The allocation is as follows, with numbers referring to the article:

Dossier	Main results	Statements
Quadratic approximation and Poisson scales	1.1–1.2, 7.2–7.4	10
Cumulants and optimality	3.3–3.4, 6.1–6.2	8
Pair profiles and channel hierarchy	3.1–3.2, 8.1–8.3	13
Stopping, conversion and reconstruction	4.1–4.2, 5.1–5.3	8
Uniform support and positive reference limits	2.1–2.5, 7.1, 8.4	3

The last dossier uses three compound statements, retaining all thirteen remaining main-result clauses and directly selecting companion Propositions T13.1 and T15.1, with the latter's displayed sequence corollary. The former controls the weighted full-support sum at all cardinalities with one threshold before every subset and minimum size. The latter shows that fixed factorial moments and uniform exponential-moment control cannot replace growing-order estimates for a polynomial total-variation rate; no multiplicative realization of that counterexample is claimed. The other companion results supply proof dependencies, without being claimed as all directly registered. The exact clause allocation fixes these scopes.

As of 13 November 2026, all five entries are published, in table order:

PALOMAR-2026-11-12-000013; PALOMAR-2026-11-13-000001;
PALOMAR-2026-11-13-000002; PALOMAR-2026-11-13-000008;
PALOMAR-2026-11-13-000020.

The first entry fixes source f9a2d213265a; entries 2–4 fix d6a0f6c51098; the final dossier fixes eeda0e055bc1. The final interface has a separate audit of 111 declarations and fresh statement comparison under the same Lean and Mathlib pins. Individual public entries provide the actual mechanical reports, rendered statements, preserved sources and published warnings.

Statement on AI assistance

ChatGPT (GPT-6 Astra) was used to organize the research material, develop and rewrite mathematical arguments, inspect references, and prepare the article and its technical companion. Parallel AI-assisted reviews examined selected identities, constants, proof dependencies and the exposition. The associated Lean development also used AI assistance for proof construction and statement comparison. AI-assisted review and machine checking are distinct from independent human verification. The author remains responsible for the mathematical claims, their correspondence with the formal statements, and any submission.

References

[1] R. Arratia, L. Goldstein, and L. Gordon, *Two moments suffice for Poisson approximations: the Chen–Stein method*, Annals of Probability **17** (1989), no. 1, 9–25. doi:10.1214/aop/1176991491.

[2] R. Arratia, L. Gordon, and M. S. Waterman, *The Erdős–Rényi law in distribution, for coin tossing and sequence matching*, Annals of Statistics **18** (1990), no. 2, 539–570. doi:10.1214/aos/1176347615. Article text.

[3] A. D. Barbour, *Asymptotic expansions in the Poisson limit theorem*, Annals of Probability **15** (1987), no. 2, 748–766. doi:10.1214/aop/1176992169. Article text.

[4] A. D. Barbour and G. K. Eagleson, *An improved Poisson limit theorem for sums of dissociated random variables*, Journal of Applied Probability **24** (1987), no. 3, 586–599. doi:10.2307/3214091. Article text, Theorem 1, p. 588.

[5] A. D. Barbour, E. Kowalski, and A. Nikeghbali, *Mod-discrete expansions*, Probability Theory and Related Fields **158** (2014), 859–893. doi:10.1007/s00440-013-0498-8. Section numbers refer to the 26-page author text, Sections 5 and 8.1: https://people.math.ethz.ch/~kowalski/mod-discrete.pdf.

[6] R. de la Bretèche, V. Y. Wang, and M. W. Xu, *Random multiplicative functions and making squares from polynomial values*, arXiv:2607.06398v1 (2026), Theorem 1.2 and Lemmas 6.2–6.3.

[7] K. Dohmen, *Improved inclusion-exclusion identities and inequalities based on a particular class of abstract tubes*, Electronic Journal of Probability **4** (1999), paper 5, 1–12. doi:10.1214/EJP.v4-42. Sections 2–3.

[8] A. P. Godbole, *Poisson approximations for runs and patterns of rare events*, Advances in Applied Probability **23** (1991), no. 4, 851–865. doi:10.2307/1427680.

[9] D. Q. Naiman and H. P. Wynn, *Abstract tubes, improved inclusion-exclusion identities and inequalities and importance sampling*, Annals of Statistics **25** (1997), no. 5, 1954–1983. doi:10.1214/aos/1069362380.

[10] J. Najnudel, *On consecutive values of random completely multiplicative functions*, Electronic Journal of Probability **25** (2020), paper 59, 1–28. doi:10.1214/20-EJP456. arXiv:1702.01470v2, Lemma 3.1 and Proposition 6.3.

[11] B. Pouly, *Long runs and rare patterns of a random completely multiplicative function*, with technical companion, Zenodo (2026). doi:10.5281/zenodo.22872154. Theorem 1.1 and Section 3.6, (3.22).

[12] D. Zeilberger, *The abstract lace expansion*, Advances in Applied Mathematics **19** (1997), no. 3, 355–359. doi:10.1006/aama.1997.0552. Author version, formula (*).

[13] V. Čekanavičius and P. Vellaisamy, *Discrete approximations for sums of m-dependent random variables*, ALEA, Latin American Journal of Probability and Mathematical Statistics **12** (2015), 765–792. Journal text, Theorem 3.1, equation (3.5), and Theorem 3.2, equation (3.6). arXiv:1301.7196v3.

[14] R. Chhaibi, F. Delbaen, P.-L. Méliot, and A. Nikeghbali, *Mod-ϕ convergence: Approximation of discrete measures and harmonic analysis on the torus*, Annales de l'Institut Fourier **70** (2020), no. 3, 1115–1197. doi:10.5802/aif.3332. Journal text, Theorem 3.11.

[15] V. Čekanavičius and P. Vellaisamy, *Lower bounds for discrete approximations to sums of m-dependent random variables*, Probability and Mathematical Statistics **40** (2020), no. 1, 23–35. doi:10.37190/0208-4147.40.1.2. Journal text, Theorem 2.1, equation (2.3).

[16] D. Basak, K. Nath, and A. Zaharescu, *Poisson behavior for chains of small quadratic non-residues*, International Mathematics Research Notices **2024** (2024), no. 4, 3356–3390. Published online 24 July 2023. doi:10.1093/imrn/rnad168. Theorem 1.

[17] I. E. Shparlinski and N. Sleiman, *Counting consecutive multiplicatively dependent triples*, arXiv:2609.29408v1 (24 September 2026), Theorem 1.1, Lemmas 2.1–2.2, and Section 3.

[18] C. D. Kemp and A. W. Kemp, *Some properties of the 'Hermite' distribution*, Biometrika **52** (1965), nos. 3–4, 381–394. doi:10.1093/biomet/52.3-4.381.

Sharp Poisson approximation and second-order correction for long multiplicative runs

Technical companion

Brice Pouly

With mathematical and editorial assistance from ChatGPT (GPT-6 Astra)

Abstract

This companion supplies the arithmetic and uniform counting estimates for the article *Sharp Poisson approximation and second-order correction for long multiplicative runs*. The two texts prove the explicit Poisson distance equivalent, the optimal quadratic error power 4/3, the third factorial cumulant power 41/30, and the hierarchy of errors for rational channels of fixed height.

The proofs are organized by the estimates they supply: exact pair profiles, fixed-source and half-power counts, confinement and pair rigidity, positive triple masses, and independent certificate tails. A common homogeneous probability bounds the several weights used in these arguments. The resulting consequences include positive Palm costs for arbitrary target windows and a support count uniform in its lower cardinality. Fixed-order and growing-cardinality estimates retain distinct hypotheses throughout.

Contents

How the estimates fit together 3

T1 Uniform arithmetic estimates 3

T2 Regular windows, pair profiles and compression 19

T3 Coefficient budgets, coordinate restrictions, and fixed channels 23

T4 Fixed sources, primitive families, and fixed moments 27

T5 Square products with short selections 33

T6 A half power for each new window 36

T7 Confining large common divisors 40

T8 Short square relations and the rank of a pair 45

T9 Separating compatible triples from cancellations 55

T10 The third factorial cumulant and rational triples 57

T11 Uniform certificate tools 65

T12 Arithmetic estimates for the exact reconstruction 68

T13 Uniform independent support counts 79

T14 Poisson differences and reference coefficients 80

T15 Fixed moments do not give a polynomial rate 82

Formal proofs and reproducibility 82

Statement on AI assistance 83

References 84

How the estimates fit together

The model, exact marginals, effective relation spaces and centered coefficients are defined in Sections 1.1 and 2 of the article. References without a T prefix refer to that article. Both documents are needed for the complete proof.

Pair coefficients and independent fixed moments. Sections T1–T3 give the arithmetic pair estimates. The elementary residual power 6/5 suffices for the first pair profile; its later quasi-linear strengthening is not used to prove itself. Section T4 establishes the fixed moments independently of the approximation of the law. The good-pair compression estimate remains necessary in this moment proof.

The gain below the quadratic scale. Short-selection counts, conditional half-counts, confinement and pair rigidity culminate in the triple analysis of Section T10. Its positive and signed estimates give $A_3 \ll N^{-41/30+\varepsilon}$. The all-target Palm estimate is a consequence of the positive connected-triple mass and the absolute marginal bias. It is not used as a premise for that mass.

Summing the entire expansion. The elementary certificates of Section T11 are used in Section T12. The latter proves separate finite-order estimates and a uniform tail beyond order 256. Its homogeneous primitive estimate rests on the geometric dense-host construction and the half-count, not on a law approximation. The article applies these estimates to its finite stopped expansion.

Two useful consequences. The uniform support estimate in Section T13 combines finitely many half-power layers with one independent tail. Its threshold and constant are uniform in the lower support size. The final finite-difference example explains why estimates at all fixed moment orders alone cannot supply a polynomial distance rate.

In every asymptotic assertion, the critical-band constant, accuracy parameters, fixed source counts and fixed channel heights precede the scale threshold. Growing degrees and cardinalities occur only in estimates that state their uniform range explicitly.

T1 Uniform arithmetic estimates

This section supplies the uniformity needed when the window length grows. The degree parameters in the determinant arguments are fixed before N tends to infinity. Dependence on such fixed parameters is permitted; dependence on a squarefree coefficient of size $N^{O(1)}$ is not.

T1.1 Split products, defects, and the regular set

Products in short intervals have a long history; see Erdős and Turk [4]. The polynomial square-root argument for a fixed even number of distinct factors appears in the proof of Munsch, Shparlinski, Sun, and Xiao [6, Lemma 3.3]. We record an explicit bound uniform in the degree, as required when the number of selected factors grows. No constant depending implicitly on that degree is used in the following lemma.

Lemma T1.1 (A split-product bound uniform in the degree)**.** *Let $d = 2k \geq 2$ and let $\gamma_1, \dots, \gamma_d$ be distinct integers with $|\gamma_i| \leq R$, $R \geq 1$. If $U \geq 2R$ is an integer and $\prod_i (U + \gamma_i)$ is a square, then*

$$U \leq (64R)^d.$$

Proof. Write $G(T) = \prod_i (T + \gamma_i)$ and

$$\phi(z) = \prod_i (1 + \gamma_i z)^{1/2} = \sum_{j \geq 0} c_j z^j, \qquad P(T) = \sum_{j=0}^{k} c_j T^{k-j}.$$

The branches equal one at the origin. Cauchy's bound on $|z| = (2R)^{-1}$ gives $|c_j| \leq (3/2)^{d/2}(2R)^j$ and, for $j \leq k$, $|c_j| \leq 2^d R^j$. Comparing coefficients in ϕ^2 inductively gives $2^{2j-1}c_j \in \mathbb{Z}$ for $j \geq 1$, so $2^d P \in \mathbb{Z}[T]$. For $U > 4R$ the tail satisfies

$$|\sqrt{G(U)} - P(U)| \leq 4\, 3^{d/2} R^{k+1}/U.$$

If this difference is nonzero, it is at least 2^{-d}, whence $U \leq (8R)^d$. If it vanishes, the nonzero integer polynomial $Q = 2^{2d}(P^2 - G)$ has root U. It is nonzero because G has simple roots. Keeping the homogeneous powers of R in the coefficient products gives

$$|[T^{d-j}](P^2 - G)| \leq (d+2)4^d R^j \leq 8^d R^j, \qquad H(Q) \leq (32R)^d.$$

The integral-root bound gives $U \leq 1 + H(Q) \leq (64R)^d$. The range $2R \leq U \leq 4R$ already satisfies the conclusion. □

The following coefficient-uniform count is classical; see Cilleruelo and Garaev [12, Section 3, Lemma 4 and Proposition 1] and the directly comparable formulation of Tao [13, Lemma 2.10]. When ab is not a square, the substitution $(au)^2 - abv^2 = ah$ reduces it to the former lemma; the square case is a divisor count. We include a proof to keep the polynomial coefficient range explicit.

Lemma T1.2 (Uniform quadratic equation count)**.** *Fix $A_0 > 0$. Uniformly for positive integers a, b and nonzero integers h with $a, b, |h| \leq N^{A_0}$, the number of positive solutions to*

$$au^2 - bv^2 = h, \qquad u, v \leq N^{A_0},$$

is $N^{o_{A_0}(1)}$.

Proof. If ab is a square, multiply the equation by a and factor the resulting difference of squares; the divisor bound suffices. Otherwise write $ab = Ds^2$ with D squarefree and nonsquare. Each solution gives the algebraic integer $\xi = au + sv\sqrt{D}$ of norm ah in the maximal order of $\mathbb{Q}(\sqrt{D})$. Its principal ideal has norm $|ah|$. The number of ideals of this norm is at most $\tau(|ah|)$: at a split prime there are at most $e + 1$ ideals of norm ℓ^e, and at an inert or ramified prime there are no more. Elements generating the same ideal differ by a unit. A real quadratic unit greater than one is at least $(1 + \sqrt{5})/2$, as follows from its integral trace and norm ± 1. The positive values of ξ arising here lie in $[1, N^{O(A_0)}]$; there are therefore $O_{A_0}(\log N)$ possibilities in each ideal class of generators. Finally $\tau(|ah|) = N^{o_{A_0}(1)}$. This argument makes the coefficient uniformity explicit and requires no bound for a fundamental unit or a class number. □

Lemma T1.3 (Local and global defect counts)**.** *Suppose $B \asymp \log N$ and $H \asymp B$, with fixed comparison constants. Any interval of diameter $O(B)$ in $[cN, CN]$ contains $O(B/\log B)$ H-defective integers. Moreover*

$$\#\{n \leq CN : K_H(n) = 1\} \leq N^{1/2 + o(1)}, \tag{T1.1}$$

$$\#\{x \in I_N : V_x \text{ contains two } H\text{-defective integers}\} \leq N^{o(1)}. \tag{T1.2}$$

All statements remain valid after fixed dilations and translations of the intervals.

Proof. For D defective integers in a short interval, form the binary matrix with their small-prime valuation columns and an additional parity row. Its kernel has redundancy at most $\pi(H)+1$. A nonzero word selects an even square product, so Lemma T1.1 gives minimum distance $\delta \gg \log N/\log B$. Choose $t \asymp \log N/\log B$ with $2t < \delta$. The Hamming bound gives

$$\sum_{j=0}^{t} \binom{D}{j} \le 2^{\pi(H)+1}.$$

If $D \ge 2t$, use $\binom{D}{t} \ge (D/t)^t$. Since $(\pi(H)+1)/t = O(1)$, this proves $D = O(B/\log B)$; the other case is immediate.

Write a defective integer uniquely as $n = a^2c$, where c is squarefree and H-smooth. Then

$$\#\{n \le CN : K_H(n) = 1\} \le \sqrt{CN} \prod_{\ell \le H} (1 + \ell^{-1/2}) \le \sqrt{CN}\exp(O(\sqrt{H})).$$

For two distinct defective values in one window, write $x+i = c_1a^2$, $x+j = c_2b^2$. The offsets and the squarefree smooth parts have $B^2 4^{\pi(H)} = N^{o(1)}$ choices. All their heights are polynomial in N, and $c_1a^2 - c_2b^2 = i - j \ne 0$. Lemma T1.2 supplies $N^{o(1)}$ possibilities. This proves (T1.2) and the stated uniform versions. □

T1.2 Two small kernels with their smooth coefficients retained

The following form of the approximate determinant method is needed because the small-prime parts are fixed separately and may be polynomially large. It is an extension of the proof of [8, Sections 2.1–2.3], rather than an application of its fixed-coefficient statement.

Lemma T1.4 (Uniform determinant estimate)**.** *Fix $A_0, \delta, \varepsilon > 0$. Suppose $1 \le c_1, c_2 \le 3N$, $0 < |h| \le (\log N)^{A_0}$, and*

$$c_1D^2U \asymp c_2E^2V \asymp N, \quad D, E \ll \sqrt{N}, \quad c_1c_2(DE)^2 \ge N, \quad \log(DE), \log(UV) \ge \delta \log N.$$

The comparison constants are fixed. The number of positive solutions $d \asymp D, e \asymp E, u \asymp U, v \asymp V$ to $c_2e^2v - c_1d^2u = h$ is

$$\ll_{A_0,\delta,\varepsilon} N^{\varepsilon} \min\{(DE\mathcal{M})^{1/2} + D + E, (UV\mathcal{M})^{1/2} + U + V\}, \qquad \log \mathcal{M} = \frac{9}{8}\frac{\log(DE)\log(UV)}{\log N}. \tag{T1.3}$$

Fixed or polylogarithmic factors in the lower bound for $c_1c_2(DE)^2$ can be absorbed by decreasing the error margin.

Parameters in the determinant argument. Here and in the proof, height of a polynomial means the maximum absolute value of its integer coefficients. The ledger below distinguishes the fixed auxiliary degrees from the coefficients that vary with N. The symbols A, B_0, κ, L_1, T_1 are specified in Steps 2–3.

Object and dimension	Degree	Point or box scale	Coefficient height
Input hypersurface in $\mathbb{A}^4$, dimension 3	Total degree 3	(D, E, U, V), each $O(N)$	$c_1, c_2 \le 3N$; $\lvert h\rvert \le (\log N)^{A_0}$
Auxiliary curve in $\mathbb{P}^1 \times \mathbb{P}^1$, dimension 1	Bidegree $(a, b) \le (A, B_0)$, fixed	Primitive pairs of heights $L_1, T_1 \le N^C$	$\lVert G\rVert \le N^{\kappa}$ after the lattice changes
Eliminant curve in $\mathbb{A}^2$, when $a = 1$ or $b = 1$	Bounded in terms of A, B_0	The same primitive boxes	Resultants and nonzero right-hand sides $N^{O(\kappa)}$

All displayed height exponents, including C, are fixed, uniformly over the input boxes. The relative chart has errors $\alpha = O(M^{-1})$ and $\beta = O(|h|/N)$, with $M = N^{\zeta}\mathcal{M}$; no coefficient ratio is charged to the latter error. Choose the final ε first, then a sufficiently small ζ, then A, B_0 and κ, and finally the curve-count exponent ϵ' so that $(C+\kappa)\epsilon' \le \zeta$. Only after these choices does N tend to infinity. In particular, the factor $\|G\|^{\epsilon'}$ in Heath-Brown's bound is explicitly included in this budget. The coefficient uniformity in this ledger is established below; Reuss's stated theorem allows its constant to depend on the fixed shift h and does not itself assert this extension.

Proof. The proof has four steps. They identify separately the determinant scale, the uniform count on each auxiliary curve, and the two lattice summations. Only the last step sums over the short intervals.

Step 1: primitive pairs and relative normalization. If $g = \gcd(d, e)$ and $q = \gcd(u, v)$, then $g^2 q \mid h$. Divide by $g^2 q$ and replace the four dyadic scales accordingly. Both pairs are now primitive. The effective size $N/(g^2 q)$ and the hypotheses on the scales change only by polylogarithmic factors; in particular the lower bound for $c_1 c_2 (DE)^2$ has the permitted polylogarithmic loss. There are only polylogarithmically many reductions, which will be summed at the end. We keep the notation N, D, E, U, V, h for a reduced box.

Set $s = d/e$, $t = v/u$. The exact curve approximation is

$$t = (c_1/c_2)s^2 + h/(c_2 e^2 u).$$

On a relative interval of length M^{-1} write

$$s = s_0(1+\alpha), \qquad t = t_0(1 + 2\alpha + \alpha^2 + \beta), \qquad t_0 = (c_1/c_2)s_0^2 \asymp V/U.$$

Then $|\alpha| \ll M^{-1}$ and $|\beta| \ll |h|/N$, uniformly in c_1, c_2. The relevant vertical error is relative to t_0; its bound therefore contains no additional factor c_1/c_2.

Step 2: the determinant with variable coefficients. For the auxiliary monomials $s^a t^b$, $0 \le a \le A$, $0 \le b \le B_0$, the row denominator is $e^A u^{B_0}$. Extracting $s_0^a t_0^b$ from the columns leaves the same logarithmic denominator cost as in the normalized calculation:

$$A \log E + B_0 \log U + \frac{A}{2}\log(D/E) + \frac{B_0}{2}\log(V/U) = \frac{A}{2}\log(DE) + \frac{B_0}{2}\log(UV).$$

Take $M = N^{\zeta}\mathcal{M}$, with a small fixed $\zeta > 0$, and use $N/|h|$ for the vertical approximation scale. The strict determinant inequality in [8, Section 2.1, Lemma 7] then holds, with degrees fixed in terms of ζ, δ. The resulting integer polynomial has coefficient height $N^{\kappa(\zeta,\delta)}$. The exponent is uniform in c_1, c_2: their sizes, and the rational column scalings, are polynomial in N, while the degrees are fixed.

To check the allowed scale, put $z = \log(DE)/\log N$ and $w = 1 - \log(c_1 c_2)/(2\log N)$. Then

$$\frac{\log(UV)}{\log N} = 2w - 2z + O(1/\log N), \qquad \frac{\log \mathcal{M}}{\log N} = \frac{9}{4}z(w-z) + o(1) \le \frac{9}{16} + o(1).$$

Thus $M \le N/|h|$ for sufficiently small ζ. The two logarithmic size ratios stay in a compact subset of the positive quadrant, so the degree choices are uniform over all boxes. The parameter order is: choose ζ from the final error margin, then the auxiliary degrees and the height exponent κ, and only then the internal exponent in the curve count below.

Step 3: fibres and the count on one auxiliary curve. Apply the two unimodular lattice changes of [8, Section 2.2] to (d, e) and (v, u). Write the resulting primitive pairs as $\lambda = (\lambda_1, \lambda_2)$ and $\tau = (\tau_1, \tau_2)$, in boxes of side lengths $L_1 \ge L_2$ and $T_1 \ge T_2$, respectively. The changes of coordinates have polynomial height; enlarging κ absorbs their contribution to the heights of the original and auxiliary equations. The original equation still has bidegree $(2, 1)$.

An auxiliary factor of degree zero in (d,e) fixes only boundedly many primitive (u,v); Lemma T1.2 counts the remaining quadratic equation uniformly. For degree zero in (u,v), fix (d,e) and set $g_f = \gcd(c_1d^2, c_2e^2)$. Solvability requires $g_f \mid h$. The step of the resulting one-dimensional lattice in u is c_2e^2/g_f; the step in v is c_1d^2/g_f. Its intersection with the box has at most

$$O\left(1+\frac{Ng_f}{c_1c_2(DE)^2}\right) \le O\left(1+\frac{N|h|}{c_1c_2(DE)^2}\right)$$

points. This is polylogarithmic by the stated lower bound.

For an absolutely irreducible factor of bidegree (a,b), $a,b \ge 1$, [5, Lemma 2] bounds primitive solutions by $O_{a,b,\epsilon'}(X^{2/b+\epsilon'}\|G\|^{\epsilon'})$ when the first pair has height at most X; the second pair has no height restriction. Thus $b \ge 2$ gives the required bound in L_1, and $a \ge 2$ gives the bound in T_1 after exchanging the pairs. Choose ϵ' after κ so that the coefficient-height factor costs at most the chosen power $N^{O(\zeta)}$.

The missing bound when one bidegree is one follows directly by elimination. In bidegree $(1,b)$ write $\lambda_1F_1(\tau)+\lambda_2F_2(\tau)=0$. Primitivity gives $q\lambda_1 = F_2(\tau)$, $q\lambda_2 = -F_1(\tau)$ for a nonzero integer q. The two dehomogenizations of the coprime forms F_1, F_2 give nonzero resultants R, R' of height $N^{O(\kappa)}$ and

$$q \mid R\tau_1^m, \qquad q \mid R'\tau_2^{m'}.$$

Since $(\tau_1,\tau_2)=1$, prime by prime q divides RR'. There are $N^{o(1)}$ possibilities, uniformly in the coefficients. Substitution into the original equation gives $F_3(\tau_1,\tau_2) = hq^2$, where F_3 is homogeneous of fixed degree $2b+1$. If F_3 is zero there is no solution. Otherwise $F_3 - hq^2$ is a nonzero polynomial. Fixing the first coordinate leaves boundedly many second coordinates, except for boundedly many full fibres, each of length $O(T_2)$. As $T_2 \ll T_1$, the count is $O(T_1)$. This includes powers of a linear form and needs no uniform height-free bound for a Thue equation. Exchanging the pairs handles bidegree $(a,1)$ with count $O(L_1)$ and right-hand side hq.

These cases prove, with arbitrarily small internal losses, the local bound of [8, Lemma 8] in both orientations. Reducible auxiliary polynomials have boundedly many factors. Rational points on an irreducible factor that is not geometrically integral lie in its bounded-degree conjugate intersections. The resulting local count is therefore uniform in all the variable coefficients.

Step 4: the two short-vector summations. We give the scale check behind [8, Section 2.3]. For the first lattice, use intervals with $s_0 = (D/E)x_3/M$, $x_3 \asymp M$. A shortest primitive vector (x_1,x_2) in a box with $L < L_1 \le 2L$ satisfies

$$x_2x_3 = (ME/D)x_1 + O(E/L), \qquad |x_1| \ll D/L, \quad |x_2| \ll E/L, \quad L \gg (DE/M)^{1/2}.$$

If $x_1x_2 \ne 0$, the integer product x_2x_3 has $O(E/L)$ choices for each of $O(D/L)$ choices of x_1. The divisor bound gives $N^{o(1)}DE/L^2$ intervals. Their local contributions sum to $N^{O(\zeta)}DE/L$; summing the dyadic L gives $N^{O(\zeta)}(DEM)^{1/2}$. The primitive axis vectors have one coordinate equal to ±1 and give the additional costs $D+E$.

For the other lattice choose $U = N/(c_1D^2)$ and $V = N/(c_2E^2)$ within the dyadic comparison constants. Then

$$t_0 = (c_1/c_2)(D/E)^2(x_3/M)^2 = (V/U)(x_3/M)^2.$$

Replace this by the grid point $(V/U)x_4/M$, where $x_4 = \lfloor x_3^2/M \rfloor$. The change is within the allowed interval width. Since $x_3 \asymp M$, this discretization has bounded multiplicity and $x_4 \asymp M$. In particular no factor c_1/c_2 remains. The same divisor summation gives $N^{O(\zeta)}\{(UVM)^{1/2}+U+V\}$. Finally take ζ and the internal margins sufficiently small, replace M by $N^{\zeta}\mathcal{M}$, and sum the primitive reductions. This proves (T1.3). □

Lemma T1.5 (Two small odd kernels)**.** *Uniformly for $B \asymp \log N$, fixed $0<\beta<1$, and $\varepsilon>0$,*

$$G_2(N,N^{\beta}) := \#\{x \in I_N : \#\{n \in V_x : K_B(n) \le N^{\beta}\} \ge 2\} \ll_{\beta,C,\varepsilon} N^{\max(\beta,7/12)+\varepsilon}.$$

For $7/12 \le \beta < 1$ this exponent is sharp in the sense $G_2(N, N^\beta) = N^{\beta+o(1)}$.

Proof. Two selected integers have the unique forms $n = c_1 d^2 u$ and $n+h = c_2 e^2 v$, where $1 \le h < B$, $u, v \le T = N^\beta$, and c_1, c_2 are squarefree and B-smooth. There are $N^{o(1)}$ choices for c_1, c_2, h. Fix them and dyadically subdivide d, e. With D, E fixed, $U \asymp N/(c_1 D^2)$ and $V \asymp N/(c_2 E^2)$. The elementary congruence count, for fixed d, e, gives in this box

$$\ll \frac{N|h|}{c_1 c_2 DE} + DE \ll (\log N)^{O(1)} T + DE. \tag{T1.4}$$

Indeed the common divisor of the two moduli divides h, and $DE \gg N/(\sqrt{c_1 c_2}\, T)$. If $c_1 c_2 (DE)^2 < N$, then $DE < \sqrt{N}$ and (T1.4) suffices. It also suffices for $DE \le N^\delta$, with a small fixed δ. For $UV \le N^\delta$ fix u, v and apply Lemma T1.2. These edge regimes cost $N^{\max(\beta,1/2)+O(\delta)}$.

In the other boxes apply Lemma T1.4. In its notation the main exponents are

$$f_1(z, w) = z/2 + \tfrac{9}{8} z(w - z), \qquad f_2(z, w) = w - z + \tfrac{9}{8} z(w - z).$$

They meet at $z = 2w/3$; f_1 is increasing to the left of that point and f_2 decreasing to its right. Since $0 \le w \le 1 + o(1)$,

$$\min(f_1, f_2) \le w/3 + w^2/4 \le 7/12 + o(1).$$

The remaining terms satisfy $D + E \ll N^{1/2}$ and $U + V \ll T$. Sum the polylogarithmic boxes and reduce the internal margins.

For the lower bound let $D \asymp \sqrt{N/T}$ with a fixed sufficiently large comparison constant, and take coprime $d, e \in [D, 2D]$. The conditions $n \equiv 0 \pmod{d^2}$, $n + 1 \equiv 0 \pmod{e^2}$ have $\gg N/D^4$ solutions in a fixed interior subinterval of $[N, 2N)$, because $\beta > 1/2$. Both remaining cofactors are at most T. The total number of incidences is $\gg D^2 N/D^4 \asymp T$. A fixed n is counted at most $\tau(n)\tau(n+1) = N^{o(1)}$ times. The two neighboring integers lie in a common window, proving the claim. □

Lemma T1.6 (A large selected kernel)**.** *For any fixed $0 < \tau < 1/2$ and $1 \le T \le 3N$, the number of separated pairs possessing a mixed square relation with a selected vertex whose B-kernel exceeds T is at most*

$$N^{3/2+o(1)} T^{-1/2+\tau}.$$

Also the number of windows containing one vertex with kernel at most T is at most $N^{1/2+o(1)} T^{1/2}$.

Proof. Write $n = a^2 cr$, with c squarefree and B-smooth and $r = K_B(n)$. The one-kernel assertion follows from

$$\sum_{\substack{n \le 3N \\ K_B(n) \le T}} 1 \le \sqrt{3N} \prod_{\ell \le B} (1 + \ell^{-1/2}) \sum_{r \le T} r^{-1/2} \le N^{1/2+o(1)} T^{1/2},$$

with a factor B for its placements. For the first assertion anchor a selected value $n = x + i$ with $K_B(n) > T$. Every prime of its kernel must occur at one of the B offsets of the other window. There are at most $B^{\omega(r)}$ assignments; each specifies a residue class modulo r. Since $r \le 3N$, there are $O(N/r)$ possible other starts, including $r > N$. Summation over anchors is bounded by

$$BN^{3/2} \prod_{\ell \le B} (1 + \ell^{-1/2}) \sum_{\substack{r > T \text{ squarefree} \\ \ell | r \Rightarrow \ell > B}} B^{\omega(r)} r^{-3/2}.$$

Rankin's inequality bounds the last sum by

$$T^{-1/2+\tau} \prod_{\ell > B} (1 + B\ell^{-1-\tau}) = T^{-1/2+\tau} \exp(O_\tau(B^{1-\tau})).$$

Both products and the placement factor are $N^{o(1)}$. □

T1.3 Multiquadratic surfaces and the curves that can survive

We use two uniform determinant estimates. A geometrically integral affine surface of degree d in $\mathbb{A}^3$ has, outside $O_{d,\varepsilon}(H^{2/\sqrt{d}+\varepsilon})$ integral points of height at most H, a covering by $O_{d,\varepsilon}(H^{1/\sqrt{d}+\varepsilon})$ curves of bounded degree [9, Lemma 2]. An integral curve of degree e in fixed affine space has $O_{e,n,\varepsilon}(H^{1/e+\varepsilon})$ integral points in the same box [7, Theorem A]. These constants do not depend on the defining coefficients. For a curve irreducible over $\mathbb{Q}$ but not geometrically integral, its rational points lie in the intersection of distinct conjugate components and have bounded cardinality.

Consider equations

$$w_\nu^2 = d_\nu P_\nu(X) Q_\nu(Y), \quad P_\nu(X) = \prod_{i\in I_\nu}(X+i), \quad Q_\nu(Y) = \prod_{j\in J_\nu}(Y+j), \quad 1 \le \nu \le k, \tag{T1.5}$$

where I_ν, J_ν are nonempty, disjoint within each list, and $|I_\nu| + |J_\nu| \le K$. The offsets are $O(B)$ and the nonzero coefficients d_ν are squarefree and $O(B)$-smooth. Constants k, K are fixed.

Lemma T1.7 (The uniform geometric counts)**.** *Fix $A_0, K, \varepsilon > 0$. For sufficiently large fixed k, a system of the form (T1.5), with arbitrary nonzero integer coefficients $|d_\nu| \le N^{A_0}$, has $O_{A_0,K,\varepsilon}(N^{1/2+\varepsilon})$ integral hosts $X, Y \asymp N$ at which the even-selection square-class maps on the two actual sets of selected values are injective and their images are in direct sum. If every Q_ν has degree one, the bound is $O_{A_0,K,\varepsilon}(N^{\varepsilon})$. If the degrees of the P_ν vary, select sufficiently many equations of one of their finitely many possible degrees before applying the assertion. These estimates are uniform in the coefficients and do not require their smoothness; any enumeration of such coefficients is a separate cost. They remain valid after other fixed centers have been specialized into the d_ν.*

For one variable, a bounded-degree system of split-product square equations with polynomial-height coefficients has $N^{o(1)}$ possible integer values as soon as its disjoint root sets include at least two distinct selected offsets. This includes either one equation of degree at least two or two equations with disjoint nonempty root sets.

The following graph consequences hold, with the number of component equations chosen sufficiently large after the stated fixed parameters.

1. *For any $\eta, \varepsilon > 0$, pairs of separated nonaligned logarithmic windows with at least ηB nontrivial unpinned components number $O_{\eta,C,\varepsilon}(N^{1/2+\varepsilon})$.*
2. *For every fixed integer $a \ge 1$ and $\varepsilon > 0$, there is k_a such that k_a components having a vertices in one window and one in the other have $O_{a,C,\varepsilon}(N^{\varepsilon})$ dyadic hosts. For $a = 1$ aligned hosts are excluded; for $a \ge 2$ no such exclusion is required.*
3. *After one start is fixed, a component with at least two vertices in the other window has $N^{o(1)}$ possible partners, uniformly for a fixed bound on the component size and polynomial coefficient heights.*

Parameters in the geometric counts. The external estimates are applied in the following embeddings. Write $m = \binom{D+2}{2} + k - 1$ for the ambient dimension of the affine Veronese chart, and $q = (a + 1 + A_0)/2$ in the degree-$(a, 1)$ case.

External count	Dimension and degree	Point height	Coefficient height
Surface covering, Steps 1–4	Surface in $\mathbb{A}^3$; $2^k \le d_{\mathrm{im}}$, $d_{\mathrm{im}} = O_{k,K}(1)$	$H \ll N^D$	Polynomial in N; no coefficient-height dependence in the covering constant
Lifted curves, Steps 2–4	Curves in $\mathbb{A}^m$; $2D \le e = O_{D,k,K}(1)$	$H \ll N^D$	No bound on the coefficients of the covering curves is required
Degree-$(a,1)$ surface covering, Steps 5–6	Surface in $\mathbb{A}^3$; $2^k \le d_{\mathrm{im}}$, $d_{\mathrm{im}} = O_{k,a}(1)$	$H \ll N^q$	Polynomial in N; the same coefficient-uniform covering estimate
Degree-$(a,1)$ lifted curves, Steps 5–6	Curves in $\mathbb{A}^{k+2}$; $E+1 \le e = O_{k,a}(1)$	$H \ll N^q$	No coefficient-height bound is required by the curve estimate

The curve rows concern components that can contain the hosts being counted; Steps 3 and 5–6 exclude the lower degrees. The input coefficients satisfy $|d_\nu| \le N^{A_0}$, and the offsets are $O(\log N)$. Thus the defining equations and the projected surface equations have polynomial height with fixed exponents, while the external constants depend only on the fixed dimensions, degrees and internal error margins. For the first bound choose A_0, K, ε, then $D \ge (K + A_0)/2$, then k, and finally the internal margin in the surface and curve counts. For the degree-$(a,1)$ bound choose q, then E with $q/(E+1) < \varepsilon/8$, then k satisfying the two conditions in Step 6, and finally that internal margin. All these choices precede $N \to \infty$. The count for one system is uniform over arbitrary coefficients in the stated height range; the $N^{o(1)}$ enumeration in the graph application uses smoothness separately.

Proof. We first count one fixed system, keeping the coefficient height visible; only afterwards do we sum the patterns arising from the graph. We will retain a sufficiently large fixed number k of equations, chosen after the height parameter D below. All degrees and ambient dimensions below depend on that number and on the fixed parameters, not on the values of the coefficients.

Step 1: geometric integrality and the finite affine projection. Over an algebraically closed field of characteristic zero, the square classes on the right of (T1.5) are independent: the valuation along a factor $X + i$ detects exactly its own equation. The coordinate ring is free of rank 2^k over the polynomial ring in X, Y, and its generic localization is a field. It is therefore integral.

The affine projection $(X, Y, w) \mapsto (X, Y, \sum_\nu w_\nu)$ is finite and birational onto its image. Finiteness follows from integrality over $\mathbb{Q}[X, Y]$. Birationality follows because no nontrivial sign change fixes the sum: the individual radicals are independent basis elements of the multiquadratic extension. The image surface in $\mathbb{A}^3$ has degree at least 2^k and degree bounded in terms of k, K. The projection need not extend across the projective boundary; only its affine finiteness is used when curves are lifted.

Step 2: projective degrees and the height of integral points. Choose an integer $D \ge (K + A_0)/2$ before choosing k, and homogenize in the weighted ring

$$\mathbb{Q}[X, Y, Z, W_1, \dots, W_k] \Big/ \left(W_\nu^2 - d_\nu Z^{2D-|I_\nu|-|J_\nu|} \prod_{i \in I_\nu} (X + iZ) \prod_{j \in J_\nu} (Y + jZ) \right),$$

with $\deg X = \deg Y = \deg Z = 1$, $\deg W_\nu = D$. Its Dth Veronese ring is generated in degree one by the degree-D monomials in X, Y, Z and the W_ν. This embeds an integral projective surface finite of degree 2^k over $\mathbb{P}^2$. The hyperplane bundle is the pullback of $\mathcal{O}_{\mathbb{P}^2}(D)$, so its degree is $2^k D^2$ and every integral curve C satisfies

$$\deg C = D \deg \sigma(C)\, [\mathbb{Q}(C) : \mathbb{Q}(\sigma(C))]. \tag{T1.6}$$

Here σ is the finite map to $\mathbb{P}^2$. At an integral point in the dyadic region, $|w_\nu| \ll N^{(K+A_0)/2}$. In the chart $Z = 1$, the Veronese coordinates are integers of height $H \ll N^D$, including the W_ν. The same bound holds for the affine projection in Step 1. Thus both applications of the determinant estimates use one common height bound.

Step 3: excluding the curves of smallest possible degree. A curve of degree D maps with degree one to a line. A vertical or horizontal line meeting the positive region is impossible: a nonconstant product of distinct linear factors in the other variable would have to be a rational square. On an oblique line, square products force every root on one side to match exactly one root on the other. Two component equations give two distinct exact units, hence a positive rational alignment of height $O(B)$. Thus a nonaligned point is on no curve of degree D; by (T1.6), its curves have degree at least $2D$.

In the direct-sum formulation the exceptional degree-D curves are excluded for a more intrinsic reason: their two exact units give nonempty even selections, two values on each side, with equal square classes. Injectivity makes both classes nonzero, contradicting the direct sum. Thus the argument does not require absence of geometric alignment when the actual spaces are already in direct sum.

Step 4: the surface covering and the first graph consequence. Apply the surface covering, lift its curves by the finite affine map, and take their closures in this embedding. Their numbers of components and degrees are bounded in terms of the fixed parameters by Bézout. Pila's estimate and the degree lower bound $2D$ give, per pattern,

$$O(N^{2D2^{-k/2}+\varepsilon'} + N^{1/2+D2^{-k/2}+\varepsilon'}).$$

Choose k large enough after D, ε and reduce the internal margin ε'. This proves the per-system bound.

For the graph consequence, smoothness gives $|d_\nu| \le N^{A_0}$ for a fixed A_0. At least half the components in a population $c \ge \eta B$ have size at most $K = \lceil 4/\eta \rceil$. Select k of them. The offsets and smooth coefficients have $B^{O_K(k)} 2^{k\pi(O(B))} = N^{o(1)}$ patterns. Choosing k large after D, ε, and summing these patterns, proves the first assertion.

Step 5: reduced fibres for equations of type $(a, 1)$. For the stronger degree-one assertion we exclude all curves of a fixed degree at most E, before making k large. For this part use the ordinary affine embedding with coordinates $X, Y, w_1, \dots, w_k$. Write the equations as $w_\nu^2 = d_\nu P_\nu(X)(Y + j_\nu)$ with $\deg P_\nu = a$. On the normalization of an integral curve of degree $e \le E$, the coordinate functions X, Y have degrees $m_X, m_Y \le e$ and the genus is at most $(e-1)(e-2)/2$. Neither function can be constant: otherwise the many distinct roots in the other variable would all be branch values, contradicting Riemann–Hurwitz. Discard equations having a branch value among their selected roots, or a finite fibre meeting the pole divisor of the other coordinate. At most

$$(2g - 2 + 2m_X) + (2g - 2 + 2m_Y) + m_X + m_Y \le 2e^2$$

equations are lost. For each remaining equation all finite fibres are reduced and avoid the poles. Parity of divisors gives

$$F_Y(-j_\nu) = \sum_{i \in I_\nu} F_X(-i), \qquad m_Y = a m_X.$$

In fact

$$\operatorname{div}\left(\frac{P_\nu(X)}{Y + j_\nu}\right) = F_Y(\infty) - a F_X(\infty),$$

independently of ν. For two good equations the quotients therefore differ by a nonzero constant. Eliminating this common quotient gives $Y = R(X)$ with $\deg R \le a$. Since $m_Y = (\deg R) m_X = a m_X$, the degree is exactly a.

Step 6: interpolation on the offsets and exclusion of low-degree curves. If $k \geq 4(E+1)^2$, there are at least three good equations. Their $3a$ distinct grid points $(-i, -j_\nu)$ lie on the graph of R: the reduced-fibre identity shows that R is finite and equals $-j_\nu$ at each selected $-i$. These grid points determine R. Choose $2a+1$ of them. Rational interpolation, using maximal minors of the integer evaluation matrix, gives coprime integer polynomials P, Q of degrees at most a, coefficients $B^{O_a(1)}$, and $R = P/Q$. The kernel of this interpolation system is one-dimensional: two solutions would have cross-product of degree at most $2a$ vanishing at $2a+1$ distinct abscissae. Since the reduced rational function has degree exactly a, no nonconstant polynomial multiple of its coprime numerator and denominator still has both degrees at most a. This makes the coefficient uniformity explicit: the interpolation data are the offsets, so their height bound is independent of the d_ν.

If Q is nonconstant, the nonzero integer resultant $\mathrm{Res}(P, Q)$ has size $B^{O_a(1)}$ and $Q(X)$ divides it at an integral point. Thus $|X| \leq B^{O_a(1)}$. If Q is constant and $a \geq 2$, then $|Y| \gg |X|^a / B^{O_a(1)}$, impossible for $X, Y \asymp N$. If $a = 1$ and Q is constant, the graph is a rational line; its grid points give an exact positive alignment. Thus every remaining dyadic nonaligned point lies on a curve of degree greater than E.

In the ordinary affine embedding the height is $H \ll N^q$ for a fixed $q = (a+1+A_0)/2$. Choose E so that $q/(E+1) < \varepsilon/8$, then $k \geq 4(E+1)^2$ so that $2q2^{-k/2} < \varepsilon/8$. The same finite projection, surface covering and curve estimate now give N^ε per fixed coefficient pattern. For $a = 1$, its exceptional line is also excluded under the direct-sum hypothesis: two grid matches give the forbidden even selections on the two actual blocks. Hence the degree-one per-system estimate is proved with arbitrary polynomial-height coefficients. In the graph application one sums the $N^{o(1)}$ smooth patterns, proving the second graph assertion, including $a = 1$.

Step 7: fixing one center and reducing to a quadratic equation. For the last assertion fix the first start and the component data. The other start satisfies $\prod_{j \in J}(Y + j) = ev^2$, with fixed squarefree $e \leq N^{O(1)}$ and $|J| \geq 2$. Outside primes dividing $e \prod_{i \neq j}(i - j)$, every selected factor has even valuation. Fixing the squarefree parts of two factors costs $2^{O(\omega(e \prod_{i \neq j}(i-j)))} = N^{o(1)}$ for fixed $|J|$. Their difference is a nonzero fixed offset, so Lemma T1.2 gives $N^{o(1)}$ starts. This is a bounded-height argument; no height-free theorem for all integral points is being invoked. The one-variable assertion with two separate equations follows in exactly the same localization: take one selected factor from each equation and put the prime divisors of both coefficients and both within-equation discriminants into the exceptional prime set. Outside it each chosen factor has even valuation. The two distinct offsets again give one quadratic equation counted by Lemma T1.2. □

Lemma T1.8 (The dense weighted sector)**.** *For every fixed $\eta, \varepsilon > 0$,*

$$\sum_{\substack{x,y \in I_N,\ |x-y|>L \\ (x,y)\ \text{nonaligned},\ c(x,y) \geq \eta B}} (2^{\rho(x,y)} - 1) \ll_{C,\eta,\varepsilon} N^{1+\varepsilon}.$$

The same bound holds for the full value kernel, without block-parity rows.

Proof. Let t_2, t_3 count nontrivial unpinned components with two and three vertices. If either exceeds a sufficiently large fixed threshold, Lemma T1.7(2), in both orientations for size three, gives N^ε hosts. Their weight is at most $2^{B+o(B)} = N^{1+o(1)}$, since each other component needs two vertices and the number of defective vertices is $o(B)$. Otherwise $t_2, t_3 = O(1)$ and the vertex budget gives $4c - 2t_2 - t_3 \leq 2B$, hence $c \leq B/2 + O(1)$. Lemma T1.7(1) gives $N^{1/2+\varepsilon}$ hosts, each of weight at most $2^{B/2+o(B)} = N^{1/2+o(1)}$. Reduce the internal margins in both estimates. □

Lemma T1.9 (Uniform residual rank on an aligned pair)**.** *For separated aligned windows with reduced height $b = \max(a, b) \leq B^3$, let u be the number of exact units. Then*

$$\rho(x, y) \leq (u - 1)_+ + o(B) \leq L/b + o(B),$$

uniformly over the channels. For fixed channel height b and fixed $D \geq 6b$, an enlarged prime cutoff $6bB \leq H \leq DB$ leaves residual $O_{C,b,D}(B/\log B)$.

Proof. Normalize the two windows to bV_x and aV_y, and merge equal values. All remaining distinct integers lie in an interval of diameter $O(B^4)$ centered at an integer at least N. Remove prime rows $\ell \leq B$ and rows with $\ell \mid ab$. Every other prime row meets at most two distinct normalized vertices, since it occurred at most once in each original window. Its equations again leave one variable per unpinned component. Let c_K count these components of size at most a fixed K.

For these components, the removed prime rows and one total-parity row form a binary matrix of redundancy at most $s = \pi(B) + \omega(ab) + 1 = O(B/\log B)$. A nonzero kernel word gives an even square product of distinct normalized integers. By Lemma T1.1, it uses at least $\delta \gg \log N/\log B$ vertices, and hence at least δ/K components. The Hamming argument in Lemma T1.3 yields $c_K = O_{C,K}(B/\log B)$. Components larger than K number at most $2B/K$. Since K can be chosen arbitrarily large before N, the dimension of the value kernel on the distinct normalized set is $o(B)$.

The kernel of the merging map on the original relative selections is exactly the even-selection space of the u exact units, of dimension $(u-1)_+$. Its image lies in the value kernel just bounded, proving the first assertion. For fixed height the normalized interval has diameter $O(B)$. Private primes above a sufficiently large fixed multiple of B leave only $O(B/\log B)$ defective normalized vertices, directly by Lemma T1.3; this gives the last assertion. □

T1.4 Asymmetric extensions and homogeneous seeds

Lemma T1.10 (An asymmetric pair profile)**.** *Fix integers $D, Q \geq 1$ and $D_{\mathrm{cut}} \geq 2D+1$. Let $B = L+1$ and $|L - \log N/\log 2| \leq C$. For each $N \leq X < N+QN$, let U_X be any subset of $X + [-DB, DB]$, and let V_z be a new window with $z \in I_N$. Choose an integer cutoff $(2D+1)B \leq H \leq D_{\mathrm{cut}}B$. For this asymmetric comparison call (X, z) aligned at radius D if there are coprime positive integers $r, s \leq B^3$ such that $|rz - sX| \leq D(r+s)B$; use its complement in the sum below. Let $\delta^*(X, z)$ be the dimension of the even selections of V_z whose prime valuations above H belong to the value space of U_X. Then*

$$\sum_{\substack{N \leq X < N+QN,\ z \in I_N \\ (X,z) \text{ nonaligned at radius } D}} (2^{\delta^*(X,z)} - 1) \ll_{C,D,D_{\mathrm{cut}},Q,\varepsilon} N^{31/24+\varepsilon}.$$

The choice $Q = 1$ gives the two-dyadic-window version. Adding prime coordinates at most H to the old data leaves δ^* unchanged.

Proof. All alignments in this proof use the radius DB of the old enclosing interval. The cutoff lower bound is larger than both interval diameters, so a prime above H occurs at most once in each block. The fixed expanded old-centre range changes only constants in the biform support estimates and the bounded-degree graph counts. For the graph above H, let D_z be the number of new defective vertices and let c count nontrivial unpinned components. Projection to the new block gives $\delta^* \leq \min(L, D_z + c)$, with $D_z = o(B)$. A nonzero even selection uses at least two new vertices. If it uses no old vertex, the new window has two defects and contributes $N^{1+o(1)}$ hosts after the old parameter is chosen. Otherwise it uses at least one old and two new vertices. Set $T = N^{5/12}$. If all their kernels are at most T, Lemmas T1.5 and T1.6 give

$$N^{7/12+\varepsilon'} N^{1/2+5/24+o(1)} = N^{31/24+\varepsilon'+o(1)}.$$

If one kernel exceeds T, the same CRT anchor proof gives $N^{3/2+o(1)} T^{-1/2+\tau} = N^{31/24+(5/12)\tau+o(1)}$. All these arguments concern the large-prime valuations only, so they do not impose parity on the old selection. They are unchanged at the fixed enlarged threshold H.

For $c < \eta B$, the weight is at most $2^{\eta B+o(B)}$, and the support bound suffices. For $c \ge \eta B$, the proof of Lemma T1.7(1) applies to the unequal fixed widths and gives $N^{1/2+\varepsilon'}$ hosts. Fix a large integer K. Components with one new vertex and more than K old vertices number $O_D(B/K)$. A sufficiently large fixed population of any other type $(a,1)$, $1 \le a \le K$, has only $N^{\varepsilon'}$ hosts by Lemma T1.7(2). On this exceptional set use $2^{\delta^*} \le 2^L \ll N$. Else the number c_1 of one-new-vertex components is $O_D(B/K) + O_K(1)$, and $2c - c_1 \le B$. Thus $\delta^* \le B/2 + O_D(B/K) + o(B)$ and the remaining dense mass is $N^{1+O_D(1/K)+\varepsilon'}$. Choose K, then η, τ, ε' after the required margin. Finally shrinking the old set only decreases δ^*, and prime bits at most H have already been forgotten. □

Alignment convention for asymmetric applications. The enlarged radius in this lemma follows the actual old interval. Original pairs of equal windows retain the alignment convention in Section T2. Whenever a source is compressed or an enclosing interval is enlarged below, its aligned sector is enlarged at the same time; that sector still has $NB^{O(1)}$ hosts and is treated by the separate aligned-rank estimates. The radius-one assertion with arbitrary D is false, as the following family shows.

Necessity of the enlarged alignment radius. In Lemma T1.10, the alignment radius follows the old enclosing interval $X + [-DB, DB]$: the exclusion is $|rz - sX| \le D(r+s)B$. The centres satisfy $N \le X < N + QN$, $z \in I_N$, with fixed $Q \ge 1$, and the cutoff satisfies $(2D+1)B \le H \le D_{\mathrm{cut}}B$. The weighted bound for $2^{\delta^*} - 1$ has power $31/24$. A shifted $2:3$ family with $L = 3k$, $B = 3k+1$, $X = 2t$, $z = 3t + 3B$ and $U_X = [X, X+3B] \cap \mathbb{Z}$ has $\delta^* \ge k - 1$. It is nonaligned at radius one for $t > 5B^4$ but aligned at radius two; its dyadic mass of order $N^{4/3}$ refutes a radius-one estimate for arbitrary D. Compressed-source applications use the same enlarged radius on both sides of the alignment partition.

Lemma T1.11 (Aligned loss above a rational pair)**.** *Fix a geometrically good pair in channel $a : b = 2:3$ or $3:4$, regardless of its affine compatibility. Among targets z with $|z-x|, |z-y| \ge L$ and distinct from the two sources, there are $B^{O(1)}$ aligned targets per source. Except for $B^{O(1)}$ source pairs at dyadic endpoints, their additional relative rank loss modulo the old value space is at most*

$$L/3 + o(L) \quad (a:b = 2:3), \qquad L/2 + o(L) \quad (a:b = 3:4).$$

The result allows the prime bits of a, b in the old information.

Proof. Represent the source by $x = at - i_0$, $y = bt - j_0$ and put $X = abt$. The true old normalized values form a subset of $X + [-bB, bB]$. Applying the residual argument of Lemma T1.9 to a fixed-width enclosing interval and the new interval leaves $o(B)$ dimensions after exact units are removed. If an exact unit of the enlarged graph uses an old value absent from the true set, its new nondefective vertex is pinned in the true graph. Consequently the remaining loss is at most the number of new vertices matched to actual old values, plus $o(B)$.

Write an aligned target as $z = (ct + h)/d$, with c, d positive coprime, $c, d \le B^{O(1)}$ and $h/d = O(B)$. A match to the first old family requires $h + dk \equiv 0 \pmod{c/(c,a)}$; the second family requires the same condition modulo $c/(c,b)$. Their intersection is the class modulo c, because $(a,b) = 1$. Thus the union has at most

$$B\frac{(c,a) + (c,b) - 1}{c} + O(1) \tag{T1.7}$$

new offsets. Conditions on the true intervals can only reduce this count.

For $2:3$, the source range is $N/2 - O(B) \le t \le 2N/3 + O(B)$, so $3/2 \le c/d \le 4$ apart from endpoint errors of order B/N. A distinct rational of polynomial height cannot cross either endpoint by such a small amount. For $c \ge 7$ the density in (T1.7) is at most $1/3$ (check $7 \le c \le 12$,

then bound the numerator by four). For $c = 5$ it is $1/5$. The cases $c/d = 2, 3$ physically reproduce one old family. Their supports intersect the new support in at most one integer, because the starts are at distance at least L. The other family supplies at most $B/3 + O(1)$ units, counted in its actual window of width B. The only remaining possibilities with $c \le 6$ are $c/d = 4, 3/2$; these force $t = N/2 + O(B)$ or $t = 2N/3 + O(B)$. They account for only $B^{O(1)}$ source pairs.

For $3 : 4$, one has $N/3 - O(B) \le t \le N/2 + O(B)$ and $2 \le c/d \le 6$. The density is at most $1/2$ for $c \ge 9$ (check $9 \le c \le 12$, then use numerator at most six), and also for $c = 5, 7, 8$. The physical cases $c/d = 3, 4$ leave at most one unit in one old family and $B/4 + O(1)$ in the other. The remaining fractions $2, 6$ force $t = N/2 + O(B)$ or $N/3 + O(B)$ and are endpoint sources. Finally the canonical slopes and displacements have polynomially many choices in B, proving the target count. □

Proposition T1.12 (Homogeneous seeds for fixed extensions)**.** *Let $\mathcal{A}_{\mathrm{raw}}$ be the geometric good-edge family obtained by omitting the compatibility condition in the definition of $\mathcal{E}_N$. For a finite set $I \subset G_N$ put*

$$W(I) = p^{|I|} 2^{\rho(I)}.$$

All configurations in the following sums have distinct, pairwise separated starts. For every $\varepsilon > 0$:

1. *The sum of $W(\{x, y\})$ over dependent pairs outside $\mathcal{A}_{\mathrm{raw}}$ is $O_{C,\varepsilon}(N^{-4/5+\varepsilon})$. Moreover $\sum_{e \in \mathcal{A}_{\mathrm{raw}}} W(e) \ll_C LN^{-2/3}$.*
2. *For a geometric good source $e \in \mathcal{A}_{\mathrm{raw}}$ in channel $a : b \in \{2 : 3, 3 : 4\}$, sum over targets z for which $S_z \cap \sum_{x \in e} S_x \ne \{0\}$. Then*

$$\sum_{e,z} W(e \cup \{z\}) \ll_{C,\varepsilon} N^{-41/30+\varepsilon}.$$

3. *For two distinct geometric good edges, whose union has three or four starts,*

$$\sum_{e \ne f} W(e \cup f) \ll_C L^2 N^{-4/3}.$$

These are rank bounds without any affine compatibility restriction.

Proof. The homogeneous indicators $H_x = \mathbf{1}_{\{f(x-1)=\cdots=f(x+L-1)\}}$ have zero affine target, so $\mathbb{E} \prod_{x \in I} H_x = W(I)$. The first assertion follows from the positive profile and the exceptional-channel estimates in Theorem 3.1; on a dependent pair, $2^\rho \le 2(2^\rho - 1)$.

For (2), the good source is dependent, so positive crossing gives a connected triple. The separated homogeneous variant in Section T10.5.1, with at most three marked source pairs per triple, gives $O(N^{-41/30+\varepsilon})$.

Two interacting raw edges. Write $W_e = \sum_{x \in e} S_x$ and sum over ordered pairs:

$$\mathfrak{I}_N = \sum_{\substack{e,f \in \mathcal{A}_{\mathrm{raw}},\ e \ne f \\ e \cup f \text{ separated},\ W_e \cap W_f \ne \{0\}}} W(e \cup f).$$

We prove $\mathfrak{I}_N \ll_{C,\varepsilon} N^{-41/30+\varepsilon}$. If the two edges share a vertex, their union is a connected triple; the same homogeneous estimate, with bounded marking multiplicity, applies.

For disjoint edges first treat nonaligned compressed blocks, using a fixed radius enclosing both true value sets in the alignment test. Its complementary enlarged aligned sector is retained below. Write $r_f = |U_f| - 1 \ge 5L/3 - O(1)$ and let δ^* be the loss of the even-selection rank on U_f after the old normalized values and small primes have been exposed. The excess above $W(e)W(f)$ is

at most $W(e)W(f)(2^{\delta^*}-1)$. The weak-component support proof of Lemma T1.10 still applies to two subsets of fixed-width logarithmic intervals: it needs two selected new values and one selected old value. It gives $N^{31/24+\varepsilon}$ hosts. In this sector $W(e)W(f) \le N^{-10/3+o(1)}$, so its mass is $N^{-49/24+\varepsilon}$.

In the dense sector, the multiquadratic argument gives $N^{1/2+\varepsilon}$ hosts. A fixed large population of components with one new vertex and boundedly many old vertices has only N^{ε} hosts, by Lemma T1.7(2). On that exceptional population use $W(f)2^{\delta^*} \le 1$ and $W(e) \le N^{-5/3+o(1)}$. Otherwise all but $O(B/K)+O_K(1)$ components use two new vertices. Hence

$$\delta^* \le |U_f|/2 + O(B/K) + o(B), \qquad W(f)2^{\delta^*} \le N^{-5/6+O(1/K)+o(1)}.$$

The remaining dense mass is $N^{-2+\varepsilon}$. The entire nonaligned excess is therefore $N^{-5/3+\varepsilon}$. On an interacting pair let $\delta = \dim(W_e \cap W_f) > 0$. Since $2^{\delta} \le 2(2^{\delta}-1) \le 2(2^{\delta^*}-1)$, the same nonaligned excess bound controls its full homogeneous mass.

For disjoint edges with aligned compressed centres, the fixed channel, offset and alignment codes leave $NB^{O(1)}$ quadruples. Choose $z \in f$. If $S_z \cap W_e \neq 0$, the triple $e \cup \{z\}$ is connected. For each such triple the mate of z in f has $O(B)$ choices, because the raw good-edge graph has degree $O(B)$. Since $W(e \cup f) \le W(e \cup \{z\})$, the homogeneous triple bound applies. If $S_z \cap W_e = 0$, instead $W(e \cup f) \le W(e \cup \{z\}) = W(e)p \le N^{-\kappa_b-1+o(1)}$. The host count gives $N^{-\kappa_b+o(1)}$ in this case. This includes a positive intersection between the two pair spaces even when both individual target crossings vanish. Thus all sectors satisfy the asserted bound for $\mathfrak{I}_N$.

Finally, pairs with disjoint effective spaces factor exactly. Hence

$$\sum_{\substack{e,f \in \mathcal{A}_{\mathrm{raw}},\ e \neq f \\ e \cup f \text{ separated}}} W(e \cup f) \le \Big(\sum_{e \in \mathcal{A}_{\mathrm{raw}}} W(e) \Big)^2 + \mathfrak{I}_N.$$

The first assertion bounds the squared sum by $O_C(L^2 N^{-4/3})$. Fix an internal margin below $1/30$; since $41/30 > 4/3$, the interaction term is absorbed into that scale with a fixed constant. This proves (3). □

T1.5 Support and branching at a fixed number of blocks

Lemma T1.13 (Small-kernel counts for fixed-width blocks)**.** *Let $U_X \subset X + [-DB, DB]$ with D fixed and $X \asymp N$, and let $H \asymp B$ exceed the diameter. Uniformly for $1 \le T \le CN$, the number of translations containing one value with $K_H(n) \le T$ is at most $N^{1/2+o(1)}T^{1/2}$. For fixed $0 < \theta < 1$, the number containing two such distinct values with $T = N^{\theta}$ is $O_{C,D,\theta,\varepsilon}(N^{\max(\theta,7/12)+\varepsilon})$. In particular it is $O(N^{\theta+\varepsilon})$ for $\theta \ge 7/12$.*

Proof. Enlarge the selected set to its interval. The proofs of Lemmas T1.6 and T1.5 use only a fixed logarithmic width and a prime threshold of that order. In particular $2^{\pi(H)} = N^{o(1)}$, the smooth parts have fixed polynomial height, and all changes of offset count are powers of B. □

Here a projection is an element of the effective even-selection image in the square-class space, rather than a formal coordinate vector.

Proposition T1.14 (A relative relation meeting several blocks)**.** *Fix integers $r \ge s \ge 2$. For r separated logarithmic windows, the number of tuples having a relative square relation whose projection is nonzero in at least s blocks is*

$$O_{C,r,\varepsilon}\left(N^{r-s/2+\varepsilon}\right).$$

The same assertion holds for effective even-selection images on subsets of fixed-width logarithmic intervals, including the compressed good edges, with only $B^{O_r(1)}$ choices of fixed offsets and fixed dilation factors.

Proof. Choose s blocks on which a witnessing relation has nonzero effective projection, and fix the other $r-s$ centres. Their values generate the old space in Theorem T6.1. The selected s nonzero projections sum into that old space; they need not remain nonzero modulo it. The theorem therefore gives $O(N^{s/2+\varepsilon'})$ choices for the selected centres. Summing the $O_r(1)$ choices of blocks and $O(N^{r-s})$ other centres gives $O(N^{r-s/2+\varepsilon'})$.

The same theorem permits the effective images of all fixed-width blocks in its stated height range. Fixed dilations preserve that range, and the $B^{O_r(1)}$ offset and motif choices are absorbed by reducing ε'. Injectivity of the local selection maps is unnecessary: the projections must be nonzero in their own effective images, even when they vanish modulo the old space. Formal vectors with zero local image are not counted. □

Proposition T1.15 (A fixed number of exposed source windows)**.** *Fix $j \geq 1$ and source starts $x_1, \dots, x_j \in I_N$. Outside a deterministic set of $N^{1/2+o_{C,j}(1)}$ possible starts, the new relative vector is uniform and independent of every value in the old windows. An explicit bound for the exceptional set is*

$$j(2L+1) + B\sqrt{3N}\exp(2\sqrt{m_{j,N}}), \qquad m_{j,N} = \left\lceil \pi(B) + \frac{jB\log(3N)}{\log B} \right\rceil.$$

Fixed dilations, subsets, and fixed enlarged logarithmic windows satisfy the same $N^{1/2+o(1)}$ conclusion.

Proof. Let $\mathcal{P}$ contain all primes at most B and all primes dividing an old value. It has at most $m_{j,N}$ members. The number of integers at most $3N$ whose squarefree part is supported on $\mathcal{P}$ is at most

$$\sqrt{3N}\prod_{\ell\in\mathcal{P}}(1+\ell^{-1/2}) \leq \sqrt{3N}\exp(2\sqrt{|\mathcal{P}|}),$$

because the ith distinct prime is at least i. Discard overlapping starts and windows containing two such values; the stated estimate follows even by charging each such window to one of its values and multiplying by B. If a nonzero new relative combination belonged to the old value space, its even selection would contain at least two integers. Every large odd prime of each selected integer is private inside the new window, so it would have to divide an old value. Both selected integers would have squarefree part supported on $\mathcal{P}$, contrary to the retained set. The new rank therefore increases by exactly L over the old value space, which proves the asserted conditional uniformity. For fixed j, $m_{j,N} = O_{C,j}((\log N)^2/\log\log N)$. □

T1.6 Products grouped by their ternary support

Proposition T1.16 (A ternary group budget)**.** *Let $U_i \subset X_i + [-DB, DB]$, $X_i \asymp N$, be three sets with injective even-selection square-class maps whose images are pairwise in direct sum. The width constant and the number of offset patterns are fixed. Choose H greater than the three diameters. For an offset triple $u = (u_1, u_2, u_3)$ let*

$$G_u = \prod_{\substack{\ell > H \text{ prime} \\ v_\ell(X_i+u_i) \text{ odd } (i=1,2,3)}} \ell.$$

Retain the groups greater than one and write their decreasing values $G_{(1)} > G_{(2)} > \cdots$. Put $\mathcal{G}_k = \prod_{i=1}^{k} G_{(i)}$, or zero if fewer than k exist. For every fixed $k \geq 3$, $\varepsilon > 0$, and $1 \leq T \leq N$,

$$\#\{\textit{tuples} : \mathcal{G}_k > T\} \ll_{C,D,k,\varepsilon} N^{\varepsilon}(N^3/T^2 + N^{6/5}). \tag{T1.8}$$

Proof. A prime larger than H has only one occurrence in each block, so the groups are pairwise coprime and $G_u \mid \gcd(X_1+u_1, X_2+u_2, X_3+u_3)$. Two groups satisfy $G_u G_v \ll NB$. If they share

an occurrence this follows from divisibility of that integer. Otherwise choose two blocks in which the offsets differ. Their product divides a determinant $nm' - n'm$ of size $O(NB)$. Its vanishing would give a nonzero relative relation on those two blocks, contrary to their direct sum.

Three groups with product greater than a sufficiently large constant times NB^2 force their three value vectors to have zero determinant. The corresponding offset triples cannot be collinear. Indeed, writing $u_j = u_1 + t_j d$, with d a primitive integer vector and t_j integers, in the collinear case, all three coprime groups divide $d \times (X + u_1)$. If this vector is nonzero, its size $O(NB)$ gives a contradiction. If it is zero, two positive value vectors are proportional with distinct coordinates in every block, again giving a forbidden pair relation. Thus the centers lie on a genuine plane

$$((u_2 - u_1) \times (u_3 - u_1)) \cdot X = -((u_2 - u_1) \times (u_3 - u_1)) \cdot u_1.$$

There are only $B^{O(1)}$ such planes, determined by the offsets alone.

For fixed k group first by the squarefree product $Q = \mathcal{G}_k$. The ordered factorizations cost $d_k(Q) \ll_{k,\varepsilon'} Q^{\varepsilon'}$, and the supports cost B^{3k}. CRT fixes one residue class modulo Q for each center. Summing $O((N/Q + 1)^3)$ gives, for $1 \le T \le N \le R \le N^{O(1)}$,

$$\#\{T < \mathcal{G}_k \le R\} \ll_{k,\varepsilon} N^{\varepsilon}(N^3/T^2 + R). \tag{T1.9}$$

For $\mathcal{G}_3 > R \gg NB^2$, use its forced plane. The two largest groups have product $P > R^{2/3}$ and $P \ll NB$. Two free coordinates are fixed by CRT modulo P; the plane determines the third. Summing over the two positive group sizes and their offsets gives

$$\#\{\mathcal{G}_3 > R\} \ll_{\varepsilon} N^{\varepsilon}(N^2 R^{-2/3} + N). \tag{T1.10}$$

Here one uses $\sum_{ab>A}(ab)^{-2} \ll \log(2A)/A$, $\sum_{ab\le U}(ab)^{-1} \ll \log^2(2U)$, and $\#\{ab \le U\} \ll U\log(2U)$. The third numerical group is not enumerated again: it certifies the plane and the lower bound for P.

For $k = 3$, take $R = N^{6/5}$ in (T1.9)–(T1.10). For induction on fixed k, the region $T < \mathcal{G}_k \le N^{6/5}$ is already covered by (T1.9). Beyond it, $\mathcal{G}_{k-1} > N^{6(k-1)/(5k)}$. At $k = 4$ the preceding N^3/T^2 term is $N^{6/5}$; at $k = 5$ it is smaller; for $k \ge 6$ the threshold is at least N and the preceding tail at N suffices. Only a fixed number of induction steps is taken. This proves (T1.8). □

T1.7 The probability of two active good edges

Lemma T1.17 (A positive second moment for good edges)**.** *Let* $\Xi_N = \sum_{e\in\mathcal{E}_N} \prod_{x\in e} J_x$. *Then*

$$\mathbb{E}(\Xi_N)_2 = \beta_N^2 + o_C(\beta_N^2) \ll_C L^2 N^{-4/3}.$$

In particular $0 \le \beta_N - \mathbb{P}(\Xi_N \ge 1) \le \frac{1}{2}\mathbb{E}(\Xi_N)_2$.

Proof. For compatible good edges put $q_e = q(e) = W(e)$ and $W_e = \sum_{x\in e} S_x$. On pairs with $W_e \cap W_f = \{0\}$, their original affine events factor exactly. On an interacting separated pair,

$$|q(e \cup f) - q_e q_f| \le W(e \cup f) :$$

the joint is zero or $W(e \cup f)$, while $q_e q_f \le W(e \cup f)$. The interaction estimate proved in Proposition T1.12 therefore bounds the total separated covariance by $O(N^{-41/30+\varepsilon})$.

The diagonal and strict-contact baseline products have only $NB^{O(1)}$ hosts, using the geometric degree $O(B)$. Each product is at most $N^{-10/3+o(1)}$, giving a total $O(N^{-7/3+o(1)})$; strict contacts have zero original joint probability. Thus

$$\mathbb{E}(\Xi_N)_2 = \beta_N^2 + O_{C,\varepsilon}(N^{-41/30+\varepsilon}).$$

Choose $\varepsilon < 1/30$ and use $\beta_N \asymp_C LN^{-2/3}$. This gives the relative asymptotic and its fixed-constant upper bound. Finally $(m-1)_+ \le m(m-1)/2$ for integers $m \ge 0$ gives the stated Bonferroni inequality after expectation. □

T1.8 Local rank at critical length

We apply the binary Griesmer bound [11, Section 2, equation (2.10)] to the code of local even square relations. The arithmetic input is its minimum nonzero weight; the residual-code proof below keeps the dependence on the parameters explicit.

Lemma T1.18 (Local rigidity)**.** *Let $V = \{U, \dots, U + B - 1\}$, where $U \asymp N$ and $B = O(\log N)$. Put*

$$d_{N,B} = \left\lceil \frac{\log U}{\log(64B)} \right\rceil .$$

If the relative square-relation code has dimension $r \geq t$, then

$$\pi(B) + 1 \geq \sum_{i=0}^{t-1} \left\lceil d_{N,B}/2^i \right\rceil - t. \tag{T1.11}$$

Consequently $r \leq 1$ when $B = \log N/\log 2 + o(\log N)$. More generally, if $B \sim c \log N$ and $c < 2 - 2^{1-t}$, then $r < t$ for all sufficiently large N. For every fixed $\varepsilon > 0$, the relative signs are exactly independent in all windows with $B \leq (1 - \varepsilon) \log N$.

Proof. Every square relation is supported on the D integers whose odd prime valuations are all at primes at most B: a larger prime is private in the window. The even square-relation code on these coordinates has redundancy at most $\pi(B)+1$. Its minimum nonzero weight is at least $d_{N,B}$ by Lemma T1.1.

For completeness, a binary linear code of length D, dimension r and minimum distance d has redundancy at least $\sum_{i<t} \lceil d/2^i \rceil - t$ whenever $r \geq t$. Choose a word of minimum weight $d_0 \geq d$ and project outside its support. The kernel of this projection is its one-dimensional span: another nonzero word supported there would have smaller weight. The projected code has dimension $r-1$ and minimum distance at least $\lceil d_0/2 \rceil$, since one of two complementary lifts uses at most $\lfloor d_0/2 \rfloor$ coordinates on that support. The redundancy decreases by $d_0 - 1$. Repeating t times proves the claim, and hence (T1.11).

The prime number theorem gives $\pi(B) \sim c \log N/\log\log N$, whereas $d_{N,B} \sim \log N/\log\log N$ when $B \sim c \log N$. Thus (T1.11) excludes dimension t under the stated strict inequality. At critical length, $1/\log 2 < 3/2$, so $t = 2$ applies. The assertion for every smaller subcritical window follows by including it in a window of length $\lfloor (1 - \varepsilon/2) \log N \rfloor$ and projecting. □

T2 Regular windows, pair profiles and compression

The first correction to the Poisson law comes from two windows whose integer arguments follow the same rational progression. The relevant quantity is the number of square relations, with their affine signs retained. Counting only the pairs that support a relation would lose both the size of the correction and its dependence on the integer length.

Throughout this section, $I_N = [N, 2N) \cap \mathbb{Z}$, $B = L+1$, $p = 2^{-L}$ and $\lambda = Np$, with $|L - \log_2 N| \leq C$. All constants are uniform in this band. The square-class vector of a positive integer n is denoted by $v(n)$; its prime coordinates are $v_\ell(n) \bmod 2$. For $V_x = \{x - 1, \dots, x + L - 1\}$, let E_x be the space of even subsets of V_x, and let S_x be its image under the square-class map. Thus E_x has dimension L. A selection $A \in E_x$ has affine sign $(-1)^{\mathbf{1}_{\{x-1 \in A\}}}$ for the run-start event $J_x = 1$.

For the arithmetic profiles in this section, write

$$\rho(I) = |I|L - \dim \sum_{x \in I} S_x, \qquad q(I) = \mathbb{P}(J_x = 1 \text{ for every } x \in I).$$

Here ρ includes the local kernels. On the regular set it equals ρ_{rel}; on arbitrary windows the two quantities must be distinguished. The exact-marginal probability argument uses ρ_{rel}. The affine

character on the kernel is either trivial or nontrivial. Accordingly, with $\eta(I) \in \{0,1\}$ denoting compatibility, finite Fourier inversion gives

$$q(I) = p^{|I|}\eta(I)2^{\rho(I)}. \tag{T2.1}$$

This formula includes incompatible events. In particular, $|\eta(I)2^{\rho(I)} - 1| \le 2^{\rho(I)} - 1$ whenever the individual systems have full rank. For a pair, the same inequality holds even before this restriction, since a zero-dimensional kernel has no compatibility obstruction.

T2.1 The regular set and the cost of its complement

For $H \ge 2$ put

$$K_H(n) = \prod_{\substack{\ell > H \text{ prime} \\ v_\ell(n) \text{ odd}}} \ell.$$

An integer with $K_H(n) = 1$ will be called H-defective. When the diameter of an interval is less than H, every other integer in it has a prime coordinate which occurs nowhere else in the interval. Such a coordinate excludes that integer from every square relation.

Proposition T2.1 (Regular set)**.** *Let*

$$G_N = \{x \in I_N : \dim S_x = L\}, \qquad Z_N^\circ = \sum_{x \in G_N} J_x.$$

Then $|I_N \setminus G_N| = N^{o(1)}$*, and*

$$d_{\mathrm{TV}}(\mathcal{L}(Z_N), \mathcal{L}(Z_N^\circ)) + |\mathbb{E}Z_N - \lambda| + |\mathbb{E}Z_N^\circ - \lambda| \ll_C N^{-1+o(1)}. \tag{T2.2}$$

More generally, for every deterministic $A \subset I_N$ *and every choice of affine targets on its windows,*

$$\sum_{x \in A} |\mathbb{P}(A_x X = b_x) - p| \le pN^{o(1)}. \tag{T2.3}$$

If the fixed moments of Z_N *are uniformly bounded, then, for every fixed integer* $j \ge 1$ *and every* $\varepsilon > 0$*,*

$$\left|\mathbb{E}Z_N^j - \mathbb{E}(Z_N^\circ)^j\right| \ll_{C,j,\varepsilon} N^{-1+\varepsilon}. \tag{T2.4}$$

In particular the second factorial cumulants have the same transfer bound.

The proof is given in Section T2.4.1.

T2.2 Sparse hosts and weighted nonaligned pairs

For separated windows, $|x - y| > L$, a prime larger than B occurs with odd valuation at most once in either window. Its equation either pins one vertex to zero or joins two vertices. After the pinned components are removed, each remaining component has one binary degree of freedom. If $D(x,y)$ counts the defective vertices and $c(x,y)$ the other unpinned components, then

$$\rho(x,y) \le D(x,y) + c(x,y), \qquad D(x,y) = o(B). \tag{T2.5}$$

The small prime equations and the two block parities can only reduce this dimension.

A pair is called aligned if, in one orientation, it satisfies $bx - ay = h$ for coprime positive $a, b \le B^3$ and $|h| \le (a+b)B$. Using this slightly enlarged definition includes every alignment produced by two exact units. Two different reduced slopes cannot occur for the same dyadic pair once N is large: distinct slopes are separated by at least B^{-6}, whereas the allowed displacement changes the ratio of the starts by $O(B^{O(1)}/N)$.

Proposition T2.2 (Support and residual mass)**.** *For every $\varepsilon > 0$,*

$$\#\{(x,y) \in I_N^2 : |x-y| > L,\ \rho(x,y) > 0\} \ll_{C,\varepsilon} N^{1+\varepsilon}, \tag{T2.6}$$

$$\sum_{\substack{x,y\in I_N,\ |x-y|>L \\ (x,y)\ \text{nonaligned}}} (2^{\rho(x,y)} - 1) \ll_{C,\varepsilon} N^{1+\varepsilon}. \tag{T2.7}$$

The elementary proof in Section T2.4.2 first gives the sufficient bound $N^{6/5+\varepsilon}$ in both displays. The stronger bounds stated here are completed in Section T7.1, using the conditional half-count and dense-pair confinement. That later strengthening is not an input to either of those arguments; the elementary bounds suffice for the earlier pair-profile calculation.

T2.3 Good edges, compression, and one additional window

Let $\mathcal{E}_N$ consist of the unordered compatible dependent pairs in the channels $2:3$ and $3:4$ for which the distinct normalized values have private prime coordinates above the fixed threshold H. Write $q_e = \mathbb{P}(J_x = J_y = 1)$ for $e = \{x,y\}$ and $\beta_N = \sum_{e\in\mathcal{E}_N} q_e$. The family is deterministic.

Corollary T2.3 (The pair input for the expansion)**.** *One has $|\mathcal{E}_N| = O_C(NB)$ and, for an edge in channel $a:b$,*

$$\rho(e) \le L/b + O(1), \qquad q_e \le p^2 2^{L/b+O(1)}.$$

Moreover,

$$2\beta_N = \alpha_N + O_{C,\varepsilon}(N^{-4/5+\varepsilon}), \tag{T2.8}$$

$$\sum_{\substack{x,y\in G_N,\ x\neq y \\ \{x,y\}\notin\mathcal{E}_N}} |\mathbb{E}[(J_x - p)(J_y - p)]| \ll_{C,\varepsilon} N^{-4/5+\varepsilon} + B/N. \tag{T2.9}$$

In particular $\beta_N \asymp_C LN^{-2/3}$.

The proof is given in Section T2.4.4.

Lemma T2.4 (Exact compression of a good edge)**.** *For a good edge in channel $a:b$, the joint homogeneous event space is the injective square-class image of the even-selection space of*

$$U_e = bV_x \cup aV_y,$$

where repeated values are included only once. In particular its rank is

$$\dim(S_x + S_y) = |U_e| - 1 \ge (2 - 1/b)L - O(1). \tag{T2.10}$$

After the endpoint offsets are fixed, U_e is a subset of an interval of length $O(B)$ depending on one integer translation parameter. There are only $B^{O(1)}$ offset patterns.

Proof. Multiplying every integer in a window by the same constant does not change the square class of an even selection. The two even-selection spaces on bV_x and aV_y therefore map to the original event spaces. They intersect on at least one normalized value: dependence in the good family requires at least two exact units. The sum of the even-selection spaces of two intersecting finite sets is the even-selection space of their union. Its square-class map is injective, because every distinct value in U_e has a private prime. There are at most $L/b + O(1)$ repeated values, which proves (T2.10). □

T2.4 Pair profiles and the regular-set estimates

This section supplies the estimates used in Section T2. The regular-set transfer is used for the explicit channel formulas and for the independent fixed-moment argument. The scalar reconstruction with exact marginals does not discard these exceptional windows.

T2.4.1 Regular-set transfer

Proof of Proposition T2.1. The one-window estimates proved in Lemma T1.3 give $m_B(x) = O_C(B/\log B)$ uniformly, and only $N^{o(1)}$ windows contain two B-defective integers. A nonzero relative relation is an even nonempty selection. Private prime coordinates therefore force it to use at least two defects. Its dimension is at most $m_B(x)$. Consequently

$$\sum_{x\in I_N}(2^{L-\dim S_x}-1)=N^{o(1)}.$$

Fourier inversion proves (T2.3). The sum of the probabilities of active discarded starts is at most $pN^{o(1)}$; the coupling using the original prime signs proves the total-variation assertion. On G_N the marginal is exactly p, giving both assertions about the mean.

For the last statement, let $D=\{Z_N\neq Z_N^\circ\}$. Since $0\le Z_N^\circ\le Z_N$, Hölder's inequality gives, for any fixed $s>1$,

$$\mathbb{E}[Z_N^j\mathbf{1}_D]\le(\mathbb{E}Z_N^{js})^{1/s}\mathbb{P}(D)^{1-1/s}.$$

Choose a sufficiently large integer s, and then use the fixed-moment bound. This last step is used only after that bound has been established; a small coupling error alone would not justify a variance transfer. □

T2.4.2 Sparse hosts and weighted residual mass

The elementary bounds preceding Proposition T2.2. A relative relation that is internal to one window contributes at most $N^{1+o(1)}$ hosts, by Proposition T2.1. Otherwise it selects at least two vertices in each window. Put $T=N^{3/5}$. If all selected kernels are at most T, each window contains two small kernels. The estimate

$$\#\{x\in I_N:\#\{n\in V_x:K_B(n)\le N^\beta\}\ge 2\}\ll_{C,\beta,\varepsilon}N^{\max(\beta,7/12)+\varepsilon}\tag{T2.11}$$

proved in Lemma T1.5 gives $N^{6/5+\varepsilon}$ choices. If one selected kernel exceeds T, Lemma T1.6 gives $N^{3/2+o(1)}T^{-1/2+\tau}$, for every fixed $\tau>0$. Choose τ sufficiently small.

For the weighted assertion, first restrict to $c(x,y)<\eta B$. Equation (T2.5) bounds the weight by $2^{\eta B+o(B)}$; choose η after the desired error exponent. The complementary sector has total weight $O_{C,\eta,\varepsilon}(N^{1+\varepsilon})$ by Lemma T1.8. The latter estimate uses the component equations themselves, rather than just their support: many small components put the pair on a multiquadratic surface, and its low-degree curves are either aligned or absent from the dyadic region. This proves both bounds with exponent $6/5+\varepsilon$, which is all that the pair-profile proof below uses. The exponent $1+\varepsilon$ in the revised statement is proved later in Section T7.1; no late confinement estimate is being inferred from this elementary calculation. □

T2.4.3 The complete pair profile

Proof of Theorem 3.1. For the channel $2:3$, either orientation permits $N/2\le t<2N/3$, up to $O(B)$ endpoint changes. For $3:4$ the corresponding interval is $N/3\le t<N/2$. Both intervals have length $N/6+O(B)$. The two orientations therefore give $N/3$ times the finite sums (3.3)–(3.4). Translations near a boundary contribute $O(B^3 2^{L/b})$.

For either fixed channel choose H a sufficiently large fixed multiple of B and normalize the values to $bV_x\cup aV_y$. Its diameter is $O(B)$. If its distinct values are all nondefective at level H, each has a private prime and the only relations before normalization are the even selections of exact units. Lemma T1.3 shows that the exceptional channel pairs number $N^{1/2+o(1)}$. Lemma T1.9 bounds their nullity by $L/b+o(L)$. Replacing the rational kernel by the full arithmetic kernel consequently costs at most $N^{1/2+1/b+o(1)}$, for either the positive or the signed sum.

There are $NB^{O(1)}$ aligned hosts altogether. The slope of height two forces both starts within $O(B)$ of the dyadic endpoints; its mass is $N^{1/2+o(1)}$. Height one has no exact unit in separated windows. The only reduced slopes of heights three and four that occur in the interior are $2:3$ and $3:4$. Every remaining channel has height at least five and nullity at most $L/5+o(L)$, again by Lemma T1.9; its total contribution is $N^{6/5+o(1)}$. The elementary version just proved controls the nonaligned sector at the same sufficient scale. The inequality following (T2.1) permits exactly the same error estimates for the signed sum.

At distances $1,\dots,L-1$, two run starts are incompatible. At distance L, their edge systems form a tree on an interval of diameter $2L$; Lemma T1.3, at threshold $2L+1$, bounds its nullity by $o(L)$. Hence the entire contact band changes the second factorial moment by $O(N^{-1+o(1)})$. Multiplying (3.6) by p^2, and using $\mu_N-\lambda=O(N^{-1+o(1)})$, proves (3.7). Finally $2^{(L-r)/b}=2^{-r/b}(N/\lambda)^{1/b}$ in (3.3)–(3.4) gives the two equivalents. □

T2.4.4 The good-edge mass

Proof of Corollary T2.3. Each start has $O(B)$ possible partners in a fixed channel. The rational kernel has dimension one less than the number of exact units when that number is positive, giving the rank bound. All pairs excluded from the good family have precisely the error mass estimated in the proof of Theorem 3.1, together with the contact band. For good compatible pairs the centered contribution is q_e-p^2. Since $p^2|\mathcal{E}_N|=O(B/N)$, replacing it by q_e has the stated cost. The asymptotic follows from (3.8). □

T3 Coefficient budgets, coordinate restrictions, and fixed channels

The first two tools below do not depend on the critical-window scaling. They control, respectively, the small-prime coefficients of selected values and the repetition of coordinate forms in a binary relation space. We then give the extension estimates for any fixed rational channel. In that part, the channel and every finite family of channels are fixed before N tends to infinity.

T3.1 A deterministic budget for the small-prime part

For a positive integer n, write

$$u_B(n)=\frac{\mathrm{sf}(n)}{K_B(n)}=\prod_{\substack{\ell\le B\ \text{prime}\\ v_\ell(n)\ \text{odd}}}\ell.$$

Lemma T3.1 (A small-prime coefficient budget)**.** *Let V be any interval of $B\ge 2$ consecutive positive integers. Then*

$$\sum_{n\in V}\log u_B(n)\le B\log B+O(B),$$

with an absolute constant independent of the position of V. For every fixed $A>0$, at most

$$(A^{-1}+O_A(1/\log B))B$$

of its values satisfy $u_B(n)>B^A$.

Proof. Dropping the parity condition gives

$$\sum_{n\in V}\log u_B(n)\le\sum_{\ell\le B}\left(\left\lfloor\frac{B}{\ell}\right\rfloor+1\right)\log\ell\le\log(B!)+\sum_{\ell\le B}\log\ell.$$

Here the sums are over primes. Legendre's formula gives the second inequality, and the elementary Chebyshev bound makes the last sum $O(B)$. Finally $\log(B!) \le B \log B$. Each value counted in the last assertion contributes more than $A \log B$ to the nonnegative sum. □

T3.2 Regularising coordinate forms by independent restrictions

Lemma T3.2 (Coordinate regularisation)**.** *Let $R \subset \mathbb{F}_2^V$ have dimension r, put $n = |V|$, and let $\ell_v \in R^*$ be the coordinate forms. Suppose that s distinct nonzero forms occur among them. For integers $K, d \ge 1$, there is $D \subset V$ such that*

$$R' = \{x \in R : x_v = 0 \text{ for every } v \in D\}, \qquad \operatorname{codim}_R R' = |D| \le \frac{n}{K+1} + \frac{s-r}{d+1},$$

and the restricted coordinate forms have the following properties:

1. *every nonzero form has multiplicity at most K;*
2. *every nonzero form ℓ has at most d unordered decompositions $\ell = a+b$, where a, b are distinct nonzero coordinate forms present on R'.*

Coordinates whose restricted forms vanish may then be removed without any additional loss of dimension.

Proof. The coordinate forms span R^*, and their restrictions continue to span the dual of every subspace encountered below. At an intermediate stage let S be the set of distinct nonzero forms and let r_0 be their rank. Setting a present nonzero form ℓ equal to zero reduces the dimension by one and replaces the forms by their images modulo $\langle \ell \rangle$. The form ℓ disappears. All other collisions are exactly the disjoint unordered pairs $\{a, a+\ell\} \subset S$. If there are $t(\ell)$ such pairs, then

$$|S_{\mathrm{new}}| = |S| - 1 - t(\ell), \qquad (|S| - r_0)_{\mathrm{new}} = |S| - r_0 - t(\ell).$$

Thus the nonnegative potential $|S| - r_0$ never increases.

If a form has multiplicity greater than K, set it equal to zero. This permanently kills at least $K+1$ previously nonzero coordinates. There are at most $n/(K+1)$ such steps, since these sets of killed coordinates are disjoint. When the multiplicity condition holds but a form has more than d decompositions, set that form equal to zero instead. Such a step decreases the potential by at least $d+1$, so there are at most $(s-r)/(d+1)$ of these steps. Each step decreases the dimension, and the procedure therefore ends with both required properties.

At every step select one coordinate carrying the form just killed. Its restriction was nonzero before that step, so its original coordinate form is independent of the previously selected forms. The selected coordinates constitute D, and consequently $\operatorname{codim}_R R' = |D|$. □

Corollary T3.3 (Constraints meeting three blocks)**.** *Suppose V is partitioned into three blocks and each row of a family of binary constraints annihilating R has at most one occurrence in each block. Apply Lemma T3.2 and delete the coordinates whose forms on R' are zero. Every remaining row has weight zero, two, or three. A row of weight two joins equal nonzero forms. A row of weight three has three distinct nonzero forms a, b, c with $a + b + c = 0$.*

The connected components formed by the rows of weight two have at most K vertices. After repeated rows with the same support are identified, the number of ternary rows incident to any vertex is at most $2dK^2$.

Proof. The sum of the forms on a row is zero, since that row annihilates R'. A row of weight one is therefore impossible. At weight two its forms are equal. At weight three a repeated form would force the third form to be zero, which has already been excluded. Binary components stay inside a class of equal forms, of size at most K. At a vertex of form ℓ, a ternary row chooses one

of at most d decompositions $\ell = a + b$, then the other two coordinates from classes of size at most K. Allowing their two assignments to the remaining blocks gives the asserted $2dK^2$ bound. □

Remark T3.4 (Order of restrictions). If some coordinates must first be killed for an arithmetic reason, do this before applying the regularisation lemma. A further restriction after regularisation can identify previously distinct forms and destroy both bounds. The lemma is an existence statement about a subspace; its proof does not enumerate possible relation spaces or histories of restrictions.

T3.3 A finite rank bound for every fixed channel

Fix coprime integers a, b with

$$b \geq 3, \qquad b/2 < a < b.$$

Let $\mathcal{G}_{a:b}$ be the symmetric family of ordered separated pairs in this channel with at least two exact units and with a private prime coordinate on every distinct normalized value in $bV_x \cup aV_y$, above a fixed threshold $H_{a,b} = C_{a,b}B$. Choose the constant large enough for the fixed normalized widths. For $\alpha = (x, y)$ put

$$q_\alpha = \mathbb{P}(J_x = J_y = 1), \qquad M_{a:b} = \sum_{\alpha \in \mathcal{G}_{a:b}} q_\alpha.$$

Pairs of probability zero remain in the family and are omitted from conditional expressions. The private-prime condition makes the relative kernel exactly the canonical even-selection kernel.

Lemma T3.5 (Mass of a fixed channel). *With $S_{a,b}(L)$ as defined before (3.2), and $w_{a,b} = 2(2a - b)/(ab)$, one has*

$$M_{a:b} = p^2 N w_{a,b} S_{a,b}(L) + O_{a,b,C,\varepsilon}(N^{-1+\varepsilon}).$$

In particular $M_{a:b} \asymp_{a,b,C} LN^{-1+1/b}$.

Proof. In either orientation the canonical parameter satisfies $N/a \leq t < 2N/b$, up to fixed multiples of B at the endpoints. The two orientations therefore supply $Nw_{a,b}$ translations per choice of the within-window placements. The diameter and sign calculation giving (3.2) applies for this fixed a, b.

Removing the normalized defective pairs costs $N^{-3/2+1/b+o(1)}$: there are $N^{1/2+o(1)}$ such hosts, their nullity is at most $L/b + o(L)$, and their weights include p^2. Endpoint errors are smaller. On the retained dependent pairs the signed kernel contribution is $q_\alpha - p^2$. Adding back p^2 costs $O_{a,b}(p^2 NB)$, because each start has $O_{a,b}(B)$ possible partners. This proves the displayed formula. Its leading term is positive and of order $LN^{-1+1/b}$, by the finite identity, uniformly in the critical band. □

Define the finite, explicitly computable number

$$\gamma_{a:b} = \max\left(\left\{\frac{1}{b}\right\} \cup \left\{\frac{(c,a)+(c,b)-1}{c} : \begin{array}{l} 1 \leq c \leq b(a+b-1),\ d \geq 1,\ (c,d) = 1, \\ b/2 < c/d < 2a, \quad c \nmid a, \quad c \nmid b \end{array}\right\}\right). \tag{T3.1}$$

All variables in the second set are integers.

Lemma T3.6 (Additional rank in a fixed channel). *One has $\gamma_{a:b} \leq 2/3$. Apart from $B^{O_{a,b}(1)}$ source pairs at the dyadic endpoints, every aligned target $z \neq x, y$ compatible with a source $(x, y) \in \mathcal{G}_{a:b}$ satisfies*

$$\delta_{x,y}(z) := L - \dim \frac{S_x + S_y + S_z}{S_x + S_y} \leq \gamma_{a:b} L + o_{a,b,C}(L).$$

This definition includes any internal rank deficiency of the new window. The bound remains true when the prime signs of a, b are included in the old information.

Proof. The matching calculation before the two channel specializations in the proof of Lemma T1.11 applies to any fixed coprime $b/2 < a < b$. Write $x = at - i_0$, $y = bt - j_0$, and an aligned target as $z = (ct + h)/d$, with $(c, d) = 1$ and polynomially bounded data. Away from the stated $B^{O(1)}$ endpoint sources, dyadic positions and rational separation give $b/2 < c/d < 2a$. Its residual estimate and (T1.7) bound the loss by the displayed matching count plus $o(L)$, also after exposing the prime bits of a, b. Only matches to true old values are counted.

Suppose $c \nmid a$ and $c \nmid b$, and put $g = (c, a)$, $h_1 = (c, b)$. These numbers are coprime and $gh_1 \mid c$. If one equals one, the other is a proper divisor of c, giving density at most $1/2$. Otherwise $g, h_1 \geq 2$ and

$$\frac{g + h_1 - 1}{c} \leq \frac{1}{g} + \frac{1}{h_1} - \frac{1}{gh_1} \leq \frac{2}{3}.$$

The last maximum occurs at the smallest coprime pair $\{2, 3\}$. When $c > b(a + b - 1)$, the density is less than $1/b$. All other nonphysical fractions are included in (T3.1).

If $c \mid a$, a proper divisor has $c \leq a/2 < b/2$, which is impossible here; thus $c = a, d = 1$. Similarly $c \mid b$ leaves only $c = b, d = 1$. These are the physical fractions reproducing one old family. Compatibility gives distance at least L from that source, so the two true windows share at most one value. Matches to the other old family have steps a, b in intervals of width B, and number at most $B/b + O(1)$. This proves the rank bound in these cases as well. □

Proposition T3.7 (Palm mean for a fixed channel)**.** *Put*

$$\eta_{a:b} = \min\{17/24, 1 - \gamma_{a:b}\} \geq 1/3.$$

Then, with the sum over $z \in I_N \setminus \{x, y\}$,

$$R_{a:b} := \frac{1}{M_{a:b}} \sum_{\alpha=(x,y)\in\mathcal{G}_{a:b}} q_\alpha \sum_z \big(\mathbb{P}(J_z = 1 \mid J_x = J_y = 1) - p\big)_+ \ll_{a,b,C,\varepsilon} N^{-\eta_{a:b}+\varepsilon}.$$

Moreover

$$\frac{1}{M_{a:b}} \sum_{\alpha=(x,y)\in\mathcal{G}_{a:b}} q_\alpha \mathbb{E}[Z_N - 2 \mid J_x = J_y = 1] = \lambda + O_{a,b,C,\varepsilon}(N^{-\min(\eta_{a:b},1/2)+\varepsilon}).$$

These bounds are not asserted to be optimal, or uniform in growing channel height.

Proof. The source probability is $O_{a,b,C}(p^2 2^{L/b})$. For the compressed old interval choose a fixed enclosing radius DB. Use this same radius in the alignment test and the complementary aligned-target sector; the large-prime cutoff is chosen accordingly. The nonaligned asymmetric estimate in Lemma T1.10, with the fixed larger widths and $B^{O_{a,b}(1)}$ patterns, bounds the total positive excess by

$$p^3 2^{L/b} N^{31/24+\varepsilon'}.$$

Division by $M_{a:b} \asymp NLp^2 2^{L/b}$ gives $N^{-17/24+\varepsilon}$.

Each source has $B^{O_{a,b}(1)}$ aligned targets. Lemma T3.6 bounds their positive conditional excess by $pN^{\gamma_{a:b}+o(1)}$ per target. The exceptional endpoint sources have total weight $B^{O(1)} p^2 2^{L/b}$. Even bounding every conditional probability by one, their aligned contribution after division by $M_{a:b}$ is $N^{-1+o(1)}$. This proves the first assertion, including all new windows.

For the mean, Proposition T1.15 with two exposed source windows leaves at most $N^{1/2+o(1)}$ exceptional targets for each source. Outside them the new relative vector is uniform and independent of the old values, so its conditional probability is exactly p. The total negative deviation is therefore at most $pN^{1/2+o(1)}$. Combining it with the positive estimate and the removal of the two sources, which costs $2p$, proves the assertion. □

T3.4 The ternary mass generated by a finite family of pairs

Corollary T3.8 (Triples containing a good pair)**.** *Let $\mathcal{F}$ be a nonempty finite fixed family of admissible channels, and set*

$$M_{\mathcal{F}} = \sum_{a:b\in\mathcal{F}} M_{a:b}, \qquad \nu_{\mathcal{F}} = \min\{1/2, \min_{a:b\in\mathcal{F}} \eta_{a:b}\}.$$

The sum of $q(\{x,y,z\})$ over all ordered triples of distinct starts containing at least one good pair from this family is

$$3\lambda M_{\mathcal{F}} + O_{\mathcal{F},C,\varepsilon}(M_{\mathcal{F}} N^{-\nu_{\mathcal{F}}+\varepsilon}).$$

In particular, for fixed $Q \geq 5$, the triples containing a good pair of height $5 \leq b \leq Q$ have total mass $O_{Q,C,\varepsilon}(N^{-4/5+\varepsilon})$.

Proof. Choose which two positions of an ordered triple form the designated pair. For each of the three choices, Proposition T3.7 gives $\lambda M_{\mathcal{F}} + O(M_{\mathcal{F}} N^{-\nu_{\mathcal{F}}+\varepsilon})$. The pair orientations are already included in $M_{\mathcal{F}}$.

This designated-pair sum counts a triple once for each good pair that it contains. If the third window is independent of all old values of the designated pair, it cannot form a dependent good pair with either source. Every overcount therefore has its third start in the exceptional set of Proposition T1.15. For each designated source, the conditional mass of this set is at most

$$pN^{1/2+o(1)} + \sum_{z} \big(\mathbb{P}(J_z = 1 \mid J_x = J_y = 1) - p\big)_+.$$

Its weighted average is $O_{\mathcal{F},C,\varepsilon}(N^{-\nu_{\mathcal{F}}+\varepsilon})$. This bounds the entire overcount, since a triple has at most three pairs. The finite channel families are disjoint for large N, by the separation of distinct reduced slopes. Finally, Lemma T3.5 gives $M_{\mathcal{F}} \ll_{Q,C} LN^{-4/5}$ for heights between five and Q. □

The last corollary concerns the part of the third moment containing a good pair. It makes no assertion about the primitive remainder or about a family whose heights grow with N.

T4 Fixed sources, primitive families, and fixed moments

The boundary estimate will use large families, but its finitely many small sizes require stronger arithmetic input. We establish that input first. All numbers of sources and all moment orders in this section are fixed before N tends to infinity.

For a family I of windows, let $R(I)$ be the kernel of the sum map on the effective spaces and put $\rho_{\mathrm{rel}}(I) = \dim R(I)$. For s fixed sources $\boldsymbol{x} = (x_1, \dots, x_s)$ and a new centre z, write

$$\delta_z = \dim R(\boldsymbol{x}, z) - \dim R(\boldsymbol{x}) = \dim \mathrm{pr}_z R(\boldsymbol{x}, z). \tag{T4.1}$$

If the prescribed affine events at the sources and target are jointly compatible, their conditional probability is $p_z 2^{\delta_z}$; it is zero otherwise. The same formula applies to homogeneous prescriptions. The proofs on regular windows and their effective-space extension in Section 2 establish the following inputs without using a distributional approximation.

T4.1 Regularizing the coordinate labels

The following linear-algebra lemma gives control of every short circuit of coordinate labels, including repeated labels.

Lemma T4.1 (Disjoint recovery). *Let $R \subseteq \mathbb{F}_2^V$, $n = |V|$, and let $\ell_v \in R^*$ be its coordinate forms. For fixed integers $d \geq 1$ and $t \geq 2$, there is a subspace $R' \subseteq R$ with*

$$\operatorname{codim}_R R' \leq 2nt^{-1/d} \tag{T4.2}$$

such that no nonzero $a \in (R')^*$ has t representations

$$a = \sum_{v \in A_i} \ell_v|_{R'}, \qquad 1 \leq |A_i| \leq d,$$

by pairwise disjoint subsets $A_1, \ldots, A_t$ of V. Each nonzero coordinate label occurs at most $K = t - 1$ times. For $1 \leq j \leq d$, a nonzero label has at most $j!t^{j-1}$ representations as a sum of j distinct linearly independent labels.

Proof. Start with $W = 0 \subset R^*$. Whenever R^*/W has a nonzero class with t such disjoint recovery sets, record those sets and adjoin one representative α_i of the class to W. Each step increases $\dim W$ by one. Fix the entire resulting sequence of m steps.

Select each coordinate independently with probability $u \in (0, 1)$, obtaining $A \subseteq V$. At step i, declare success if one of its recorded recovery sets is contained in A. On success, $\alpha_i \in \langle \ell_v : v \in A \rangle + W_{i-1}$. The failure probability is at most $(1 - u^d)^t$, since the recorded sets at that step are disjoint. For each realization,

$$m \leq |A| + \#\{\text{failed steps}\}.$$

Taking expectations and choosing $u = t^{-1/d}$ gives

$$m \leq \frac{nu}{1 - (1 - u^d)^t} \leq 2nt^{-1/d}.$$

The annihilator $R' = W^\perp$ has the stated codimension and satisfies the stopping condition.

The multiplicity bound follows by using singleton recovery sets. To count independent j-term representations of a nonzero a, take a maximal disjoint family of their label supports. It has fewer than t members. Its union has at most $j(t-1)$ labels and meets every representation. Removing one of these labels reduces to a $(j-1)$-term independent representation of a nonzero vector: the removed label cannot equal a, because the remaining independent labels would then sum to zero. Induction from $j = 1$ proves the last bound. □

T4.2 Uniformly few targets of large rank loss

Proposition T4.2 (Large rank loss after fixed sources). *For every fixed $s \geq 1$ and $\eta, \varepsilon > 0$,*

$$\sup_{x_1, \ldots, x_s \text{ distinct}} \#\{z \in I_N \setminus \{x_1, \ldots, x_s\} : \delta_z \geq \eta L\} \ll_{s,C,\eta,\varepsilon} N^\varepsilon. \tag{T4.3}$$

The sources may be dependent or overlapping.

The complete proof is given in Section T4.4.2.

Proposition T4.3 (Conditional sums and propagation). *For every fixed $s \geq 1$ and $\varepsilon > 0$, uniformly over compatible prescribed affine events at s distinct sources,*

$$\sum_{z \notin \{x_1, \ldots, x_s\}} \mathbb{P}(J_z = 1 \mid J_{x_1} = \cdots = J_{x_s} = 1) \ll_{s,C,\varepsilon} N^\varepsilon. \tag{T4.4}$$

If a family of ordered s-tuples has total positive joint mass $W_{s,N}$, the mass of its extensions to ordered distinct r-tuples, for fixed $r \geq s$, is

$$W_{s \to r,N} \ll_{r,C,\varepsilon} N^\varepsilon W_{s,N}. \tag{T4.5}$$

Both (T4.4) and (T4.5) also hold for homogeneous prescriptions. For the original run-start events, the mass of fixed r-tuples containing a local defect or a compatible contact is $O_{r,C,\varepsilon}(N^{-1+\varepsilon})$.

Proof. For $\delta_z < \eta L$, the conditional sum is at most $2Np2^{\eta L} \ll_C N^{\eta}$. For $\delta_z \geq \eta L$, Proposition T4.2 counts $O(N^{\varepsilon/2})$ targets, each of conditional probability at most one. Take $\eta = \varepsilon/2$. The proof uses only the rank loss and compatibility, so it applies unchanged to homogeneous prescriptions. For $r = s$, (T4.5) is immediate. For $r > s$, iterating with margin $\varepsilon/(r-s)$ proves the bound. Allowing a seed at any positions in the ordered tuple multiplies the estimate by a constant depending on r.

For the original run-start events, the local and contact seeds have positive masses $N^{-1+o(1)}$, by the regular-core estimates in Section T2. Contacts at distance strictly less than L have zero joint probability. Propagate the remaining seeds to obtain the last claim. □

T4.3 The correction to every fixed moment

Theorem T4.4 (Fixed factorial moments)**.** *For every fixed integer $r \geq 2$ and every $\varepsilon > 0$,*

$$\mathbb{E}(Z_N)_r = \mu_N^r + \binom{r}{2}\mu_N^{r-2}\alpha_N + O_{r,C,\varepsilon}(N^{-4/5+\varepsilon}). \tag{T4.6}$$

Equivalently, one may use λ_N and $a_N = \mathbb{E}(Z_N)_2 - \lambda_N^2$. All fixed ordinary moments are bounded. More precisely, with $T_r(u) = \mathbb{E}[\mathrm{Pois}(u)^r]$,

$$\mathbb{E}Z_N^r = T_r(\mu_N) + \frac{\alpha_N}{2}T_r''(\mu_N) + O_{r,C,\varepsilon}(N^{-4/5+\varepsilon}). \tag{T4.7}$$

For every fixed $r \geq 3$, the r-th factorial cumulant is $O_{r,C,\varepsilon}(N^{-4/5+\varepsilon})$.

This first estimate is proved independently of the law approximation. Its bounded-moment conclusion is used in the sharp lower bound of Theorem 6.2 in the article. After that reconstruction, Corollary 7.2 improves the scalar moment remainders to $O_{r,C,\varepsilon}(N^{-4/3+\varepsilon})$ when the exact parameters are retained.

The complete proof is given in Section T4.4.4.

For $r \geq 2$, the first discrepancy in (T4.6) has the exact scale $LN^{-2/3}$, by (1.5). The theorem fixes r first. Its proof does not allow r to grow like $\log N/\log\log N$. The uniform certificate tools in Section T11 and the large-order reconstruction in Section 5 supply that separate input.

T4.4 Fixed-source geometry and the independent moment proof

The statements of the fixed-source estimates and the fixed-moment theorem are in Section T4. This section supplies the geometric count and the positive classification used in their proofs. The original window estimates below use the regular-set convention $p = 2^{-L}$; Section 2 gives their extension to effective spaces. The moment proof is independent of the reconstruction of the complete law.

T4.4.1 Uniformly few targets of large rank loss

We first record the one-parameter geometric estimate in the precise form needed here.

Lemma T4.5 (Several disjoint square equations)**.** *Fix $K, A, D_0 > 0$ and $\varepsilon > 0$. There is a fixed integer j such that the following holds. Let $P_1, \dots, P_j$ be nonconstant products of distinct linear factors $Z+u$, of degrees at most K, with $|u| \leq D_0 B$. Suppose their root lists are pairwise disjoint, and let $0 < |D_\nu| \leq N^A$ be integers. Then*

$$\#\{z \asymp N : w_\nu^2 = D_\nu P_\nu(z)\ (1 \leq \nu \leq j) \text{ for some integers } w_\nu\} \ll_{K,A,D_0,\varepsilon} N^{\varepsilon}.$$

The constants are uniform in the displayed coefficients and offsets.

Proof. The square classes of the functions $D_\nu P_\nu(Z)$ are geometrically independent: a root of P_ν has odd valuation in that class and even valuation in every other class. Thus the resulting curve is geometrically integral and has degree 2^j over the Z-line.

Choose an integer $D \geq (K+A)/2$. Give T, S degree one and each W_ν degree D, and use the equations

$$W_\nu^2 = D_\nu S^{2D-\deg P_\nu} \prod_{u \in P_\nu} (T + uS).$$

The graded ring is free of rank 2^j over $\mathbb{Q}[T,S]$; its generic localization is a field. Its D-th Veronese ring is generated in degree one by the monomials in T, S of degree D and the W_ν. The corresponding projective curve has degree $D2^j$, since its hyperplane bundle is the pullback of $\mathcal{O}_{\mathbb{P}^1}(D)$. On $S = 1$, the coordinates are integral and have height $O_{K,D}(N^D)$.

Pila's uniform curve estimate [7, Theorem A] bounds their number by

$$O_{K,A,j,\varepsilon'}((N^D)^{1/(D2^j)+\varepsilon'}) = O(N^{2^{-j}+D\varepsilon'}).$$

Choose j and then ε' to put this exponent below ε. Both choices precede N. □

Lemma T4.6 (A product certificate with fixed sources)**.** *Fix $s, k \geq 1$. Uniformly over sources $x_1, \dots, x_s \in I_N$, the number of centres $z \in I_N$ for which there are pairwise coprime $G_1, \dots, G_k > 1$ satisfying*

$$G_i \mid x_{a_i} + u_i, \qquad G_i \mid z + w_i, \qquad \prod_i G_i > N, \quad u_i, w_i \in \{-1, \dots, L-1\},$$

is $O_{s,k,C,\varepsilon}(N^\varepsilon)$ for every $\varepsilon > 0$.

Proof. There are $(sB)^k B^k = N^{o(1)}$ choices of positions. For fixed positions, the product $Q = \prod_i G_i$ divides $\prod_i (x_{a_i} + u_i) \leq (3N)^k$. Its divisors and their ordered factorizations into k factors have $N^{o(1)}$ possibilities. Coprimality fixes one residue class for z modulo $Q > N$, hence at most one centre in I_N. This counts the certificate without requiring the groups to be fixed in advance. □

T4.4.2 Proof of Proposition T4.2

Proof. Take $H = D_0 B$, with fixed $D_0 > 1$. A prime exceeding H has at most one occurrence of odd valuation in each window. Apply Lemma T4.1 to $R(\boldsymbol{x}, z)$, with $d = s$, choosing t so that the loss of dimension is at most $\eta L/4$. The surviving target projection has dimension at least $3\eta L/4$. Nonzero labels have multiplicity at most $K = t - 1$.

For each prime $q > H$ incident to an active target coordinate whose label has odd multiplicity in the q-row, choose a minimal circuit of distinct labels containing this label. Such a circuit exists: a minimal zero-sum subset containing the label cannot contain a proper zero-sum subset, because it could be removed. Choose incident coordinates for its other labels. They all belong to sources, because q has only one target occurrence. Group primes according to the selected physical support. These groups are pairwise coprime. The number of supports through a fixed target coordinate is at most

$$D = \max\left(1, \sum_{j=2}^{s} j! t^{j-1} K^j\right).$$

Choose the integer j in Lemma T4.5 for degree at most K and coefficient height N^{A_0+2K+2}, where $\prod_{q \leq H} q \leq N^{A_0}$. Choose k large enough that $\theta = KD/k \leq 1$ and $j\theta < \varepsilon/4$. If the product of the k largest nontrivial groups exceeds N, Lemma T4.6 applies. It remains to count the complementary case.

Discard from selection the label classes meeting the target positions of these k groups. At most k classes are discarded; no further restriction of the relation space is made. Every group incident to a remaining target class has size at most $N^{1/k}$. For such a class A, write

$$M_A = \prod_{v\in A \text{ at sources}} n_v, \qquad P_A(Z) = \prod_{v\in A \text{ at target}} (Z + w_v).$$

Its squarefree product $d_A = \mathrm{sf}(M_A P_A(z))$ can be factored as

$$d_A = d_{0,A} d_{1,A} e_A, \qquad d_{0,A} \mid \prod_{q\le H} q, \quad d_{1,A} \mid M_A, \quad 1 \le e_A \le N^{\theta}.$$

Indeed, every remaining prime with odd multiplicity and not dividing M_A lies in one of at most KD incident groups. The choices of $d_{0,A}$ and $d_{1,A}$ are $N^{o(1)}$, uniformly for fixed source positions.

There are at least $3\eta L/4 - k$ available classes meeting the target. Select j of them. Their target polynomials have simple disjoint roots. After choosing their positions and coefficients, the equations

$$w_\nu^2 = d_{A_\nu} M_{A_\nu} P_{A_\nu}(z)$$

are counted by Lemma T4.5. The positions cost $B^{O(1)}$, and the coefficients cost $N^{j\theta+o(1)}$. The stated choices, with the error margin split between these factors, give (T4.3). Only the selected positions and coefficients have been enumerated, not the possible relation spaces or their regularizations. □

T4.4.3 Primitive families at a fixed order

A family is *primitive* if it has a nonzero relative relation and every proper subfamily has none. Its relation space projects injectively to each block. For original windows its nullity is consequently at most L. For the homogeneous prescription define

$$W(I) = p^{|I|} 2^{\rho(I)}.$$

This is an actual joint probability, regardless of the compatibility of the run-start prescriptions. Those have joint probability either $W(I)$ or zero.

Corollary T4.7 (Primitive families at a fixed order)**.** *For every fixed $r \ge 3$ and $\varepsilon > 0$, the sum of the homogeneous weights $p^r 2^{\rho(I)}$ over ordered primitive r-tuples of separated regular windows is $O_{r,C,\varepsilon}(N^{-r/2+\varepsilon})$. The same bound holds for their actual joint mass and absolute centered mass.*

Proof. Apply Proposition T12.4 on G_N, where $q_0(I) = p^r 2^{\rho_{\mathrm{rel}}(I)}$, and multiply by the fixed ordering factor $r!$. The proposition uses the geometric half-count and dense-host construction, independently of the moment proof below. □

The proof is homogeneous throughout. It therefore bounds incompatible configurations as well as compatible ones. This feature will be needed for centered Fourier coefficients.

Proposition T4.8 (Fixed primitive mass after pair compression)**.** *Fix a channel $a : b \in \{2:3, 3:4\}$, and put $\kappa_b = 2 - 1/b$. All original starts lie in G_N and are distinct and pairwise separated. For every fixed $r \ge 3$ and $\varepsilon > 0$, the total homogeneous mass of primitive families consisting of one exactly compressed pair in this channel and $r-1$ further original windows is*

$$O_{r,C,\varepsilon}(N^{-\kappa_b+\varepsilon}). \tag{T4.8}$$

For a compressed geometric good pair and just one further original window with positive relative rank loss, the mass is $O_{C,\varepsilon}(N^{-41/30+\varepsilon})$ in either channel. The affine positive masses satisfy the same bounds.

Proof. Put $t = r - 1 \geq 2$ and fix a compression motif. The compressed block is the even-selection image of a genuine set of $O(B)$ values of width $O(B)$; write r_e for its rank. Use its normalized centre bx; it lies above the original target range and determines the source pair once the motif is fixed. There are $B^{O(1)}$ motifs and $NB^{O(1)}$ source pairs. For its homogeneous mass, Lemma T1.9 gives $q_0(e) = 2^{-r_e} \ll N^{-\kappa_b+\varepsilon'}$, including affine-incompatible sources. Primitivity gives $\rho \leq L$ and the exact homogeneous weight $q_0(e)p^t 2^\rho$.

Every primitive family has a nonzero effective projection in each of its $t+1$ blocks. Theorem T6.1, with no old sources, counts at most $N^{(t+1)/2+\varepsilon'}$ hosts, including the compressed centre. For $t \geq 3$, their total mass is

$$N^{(t+1)/2+\varepsilon'} N^{-\kappa_b - t + 1 + \varepsilon'} = N^{-\kappa_b + (3-t)/2 + 2\varepsilon'} \leq N^{-\kappa_b + 2\varepsilon'}.$$

For $t = 2$, the same count on $\rho \leq L/4$ gives $N^{-\kappa_b - 1/4 + 2\varepsilon'}$. On $\rho \geq L/4$, fix the compressed source image W. Every proper subfamily is direct, so each target is disjoint from W and the two targets are direct. Lemma T12.1 therefore gives conditional homogeneous target mass $O(N^{-1+\varepsilon'})$; the original targets are regular, so their homogeneous marginals are exactly p. The source sum is $\sum_e q_0(e) \ll N^{1-\kappa_b+2\varepsilon'}$, giving total $O(N^{-\kappa_b+3\varepsilon'})$. Reduce the margins and absorb the motif choices to obtain (T4.8).

The one-target clause is part (2) of Proposition T1.12. Original affine joint probabilities are bounded by the homogeneous ones. □

T4.4.4 Proof of Theorem T4.4

Proof. Expand $\mathbb{E}(Z_N)_r$ over ordered distinct centres. We describe all configurations removed before evaluating the two principal classes.

Local defects and contacts have mass $N^{-1+\varepsilon}$ after propagation. A dependent separated pair outside the two retained channels has mass $N^{-4/5+\varepsilon}$, also after propagation. Tuples containing a primitive original subfamily of size at least three have that same bound, by Corollary T4.7. Two distinct good pairs, including pairs sharing a vertex, have positive seed mass $O_C(L^2 N^{-4/3})$ by Proposition T1.12(3): their union belongs to the separated raw-edge population counted there, and its affine mass is bounded by its homogeneous mass. Fixed-order propagation preserves the required precision.

There remains the case of one good pair with an additional relation. Compress this pair exactly. A minimal dependent subfamily of the compressed system has at least two blocks and contains the compressed block, unless an original primitive family has already been removed. Apply Proposition T4.8 to this seed and then propagate. Each of its exponents exceeds $4/5$. This exhausts the dependent residual configurations: after their removal the spaces are either all direct, or one good pair is direct from all remaining singletons.

The support estimate of Proposition T1.14, together with the $O_r(N^{r-1}B)$ contact tuples, shows that the number of full-rank tuples is

$$N^r + O_{r,C,\varepsilon}(N^{r-4/5+\varepsilon}).$$

Their joint probability is exactly p^r, and their contribution is $\lambda_N^r + O(N^{-4/5+\varepsilon})$.

Let A_N be the ordered positive mass of good pairs. The pair calculation gives

$$A_N = a_N + O_{C,\varepsilon}(N^{-4/5+\varepsilon}).$$

Fix one such pair. After any fixed number of already chosen windows, at most $N^{1/2+o(1)}$ target centres are dependent on their relative space, by Proposition T1.15. Thus the number of ordered direct extensions by $r-2$ windows is

$$N^{r-2}(1 + O_{r,C}(N^{-1/2+o(1)})),$$

uniformly in the pair. Their joint weight is $q(e)p^{r-2}$. Placing the ordered pair in two of the r positions gives contribution

$$\binom{r}{2}\lambda_N^{r-2}A_N + O_{r,C}(A_N N^{-1/2+o(1)}).$$

Since $A_N = O_C(LN^{-2/3})$, the last error is smaller than the stated remainder. The two principal classes give the formula with (λ_N, a_N). The mean transfer $\mu_N - \lambda_N = O(N^{-1+\varepsilon})$ and $\alpha_N - a_N = O(N^{-1+\varepsilon})$ give (T4.6).

The Stirling-number identity $Z^r = \sum_{j=0}^{r} \left\{ {r \atop j} \right\} (Z)_j$ proves (T4.7) and boundedness. Finally, to any fixed formal order, the factorial generating function equals

$$e^{\mu_N t}\left(1 + \frac{\alpha_N t^2}{2}\right) + O_{\mathrm{coeff}}(N^{-4/5+\varepsilon}).$$

Taking its formal logarithm is legitimate at that fixed order. Terms quadratic in α_N are $O(N^{-4/3+o(1)})$, which proves the cumulant claim. □

T5 Square products with short selections

The conditional count in the next section must handle a relation even when none of its selected integers has a large prime kernel. For selections of at most a small multiple of $\log\log N$, we instead count square products through moments. The constants in this argument are explicit in the selection size; a fixed-degree result alone would not give the required range.

For a fixed number of unrestricted coordinates, global counts of tuples whose product is a square were obtained by de la Bretèche, Kurlberg, and Shparlinski [16, Theorem 2.2 and Corollary 2.4]. Here the selections lie in short additive windows, and the ensuing conditional estimates are uniform in arbitrary fixed source windows even when the number of selected factors grows with N.

The quadratic kernel argument is related to the decomposition in de la Bretèche, Wang, and Xu [3, Lemmas 6.2–6.3]. We give the estimates needed here, including their dependence on the number and location of the roots.

Write

$$P_A(X) = \prod_{a\in A}(X+a), \qquad A \subset [-DB, DB]\cap\mathbb{Z}, \qquad 2 \le d = |A| \le c\log\log N,$$

where d is even and $D, c > 0$ are fixed. The interval of summation has height and length comparable to N.

Proposition T5.1 (Uniform moments for short split products)**.** *For every fixed even integer $v \ge 2$, with the preceding notation,*

$$\mathbb{E}\left|\sum_{n\asymp N} f(P_A(n))\right|^v \le N^{v/2+c[v(v-1)+2]\log 2+o(1)}. \tag{T5.1}$$

In particular, for each fixed v and $\varepsilon > 0$, one may choose $c > 0$ so that the L^v norm is $O(N^{1/2+\varepsilon})$, uniformly in all the indicated motifs. The summation interval may have any fixed comparable length and positive height.

The proof, in Section T5.1.1, pairs the linear factors into quadratics and controls the common kernel by their resultant. In the moment expansion, primes are assigned to the even subsets of indices on which they occur. Averaging the exposure order then gives the exponent $v/2$. The explicit term $c[v(v-1)+2]\log 2$ is what permits a growing selection size.

Corollary T5.2 (Short projections in an old space)**.** *Fix r, s, D and $\varepsilon > 0$. Let V_{old} be generated by the square classes of all values in s fixed blocks of diameter and cardinality $O_D(B)$. For a sufficiently small fixed $c > 0$, the number of lists of r new centres admitting nonempty even selections A_i, $|A_i| \le c \log\log N$, whose square classes sum to an element of V_{old}, is $O_{r,s,D,C,\varepsilon}(N^{r/2+\varepsilon})$, uniformly in the old blocks.*

Proof. For a square class u, let $\chi_u(f)$ denote its character, and put $\Phi_V = \sum_{u \in V} \chi_u$, summing over distinct classes. Character orthogonality gives

$$\Phi_V \ge 0, \qquad \mathbb{E}\Phi_V = 1, \qquad \|\Phi_V\|_a = 2^{\dim V(1-1/a)} \quad (a \ge 1).$$

For fixed motifs, with $X_i = \sum_z f(P_{A_i}(z))$, the expectation $\mathbb{E}(\Phi_{V_{\mathrm{old}}} \prod_i X_i)$ counts exactly the required lists. For each list, a unique old character cancels the new sum if it belongs to the old space. All terms after expectation are zero or one.

Since $\dim V_{\mathrm{old}} = O_{s,D}(B)$, choose a fixed $a > 1$ sufficiently close to one that $\|\Phi_{V_{\mathrm{old}}}\|_a \le N^{\varepsilon'}$. Next choose a fixed even $v \ge ra/(a-1)$, and then choose c by Proposition T5.1. Hölder gives

$$\left| \mathbb{E}\left(\Phi_{V_{\mathrm{old}}} \prod_i X_i \right) \right| \le N^{\varepsilon'} \prod_i \|X_i\|_{ra/(a-1)} \ll N^{r/2+(r+1)\varepsilon'}.$$

The motifs have $\exp(O_{r,D}(c(\log\log N)^2)) = N^{o(1)}$ choices. Choose the internal margins in the stated ε. There is no enumeration of the $2^{\dim V_{\mathrm{old}}}$ old classes. □

T5.1 Short split products: kernels and moment enumeration

This section proves the short-support moment input from Section T5. The degree is allowed to grow as a sufficiently small multiple of $\log\log N$; all estimates below retain that dependence.

Lemma T5.3 (A quadratic kernel with a prescribed divisor)**.** *Let $1 \le h \le 2DB$, where D is fixed, and $N \le n \le 2N$. Suppose $N^\beta \le q = \mathrm{sf}(n(n+h)) < 2N^\beta$ and a fixed squarefree integer e divides q. Put $\delta = \log e / \log N$. Uniformly in these data, the number of n is at most*

$$N^{c_2(\beta,\delta)+o(1)}, \qquad c_2(\beta,\delta) = \min\{1, \beta-\delta, \max(\beta/2-\delta, 1-\beta/2)\} \le \tfrac{1}{2} + \tfrac{\beta}{4} - \tfrac{\delta}{2}.$$

Every $o(1)$ in this lemma is uniform and may be replaced by an arbitrary fixed positive exponent. The uncapped minimum without the entry 1 also gives a valid counting exponent. Its exact comparison with the affine expression requires $\beta - 2\delta \le 2$; the displayed capped comparison is valid without this additional range restriction.

Proof. For a fixed q, the equation $n(n+h) = qv^2$ becomes $(2n+h)^2 - 4qv^2 = h^2$. Lemma T1.2 gives $N^{o(1)}$ solutions, uniformly for the polynomial heights in question. When $q = 1$, factor the difference of squares instead. Summing over the $O(1 + N^\beta/e)$ possible kernels gives the first exponent.

For the other estimate, $v \asymp N^{1-\beta/2}$ and $ev^2 \mid n(n+h)$. If $\rho_h(M) = \#\{a \bmod M : a(a+h) \equiv 0 \pmod M\}$, then

$$\rho_h(\ell^a) \le 2\ell^{\min(v_\ell(h), \lfloor a/2 \rfloor)}, \qquad \rho_h(M) \le h\, 2^{\omega(M)}.$$

For $a \le 2v_\ell(h)$, a root is divisible by $\ell^{\lceil a/2 \rceil}$; for larger a, the two separated roots give the displayed bound. The Chinese remainder theorem proves the second inequality. Here $M = ev^2 \le N^{2+o(1)}$, so summing residue classes gives

$$N^{o(1)} \sum_{v \asymp N^{1-\beta/2}} \left(\frac{N}{ev^2} + 1 \right) \le N^{o(1)} \{ N^{\beta/2-\delta} + N^{1-\beta/2} \}.$$

The trivial $O(N)$ count supplies the entry 1 in the minimum. If $\beta - \delta \geq 1$, the maximum is $\beta/2 - \delta$, and the capped exponent is at most

$$\min(1, \beta/2 - \delta) \leq \frac{1 + \beta/2 - \delta}{2} = \frac{1}{2} + \frac{\beta}{4} - \frac{\delta}{2}.$$

Otherwise the maximum is $1 - \beta/2$, and the minimum of this number and $\beta - \delta$ is at most their average, with the same affine value. This proves the exact comparison, including admissible cases with $\beta > 2$. □

Necessity of the exponent cap. In Lemma T5.3, the counting exponent is capped:

$$c_2(\beta, \delta) = \min\{1, \beta - \delta, \max(\beta/2 - \delta, 1 - \beta/2)\}.$$

Its affine upper bound is then valid for every real β, δ. The uncapped expression still gives a valid counting estimate, but its exact affine comparison requires $\beta - 2\delta \leq 2$ when $\delta \geq 0$. This restriction is necessary on actual positive-divisor data: squarefree consecutive $N, N+1$ with $h = e = 1$ give $\beta = \log(N(N+1))/\log N > 2$ and $\delta = 0$. The extra entry 1 is supplied by the trivial $O(N)$ count.

Lemma T5.4 (Kernels of a split product)**.** *Let*

$$P_A(X) = \prod_{a \in A} (X + a), \qquad A \subset [-DB, DB] \cap \mathbb{Z}, \qquad 2 \leq d = |A| \leq c \log\log N,$$

where d is even and $D, c > 0$ are fixed. For a fixed squarefree e, the number of $n \asymp N$ with $q = \mathrm{sf}(P_A(n)) \asymp N^\beta$ and $e \mid q$ is at most

$$N^{1/2 + \beta/(2d) - \delta/d + c\log 2 + o(1)}, \qquad \delta = \frac{\log e}{\log N}. \tag{T5.2}$$

For a fixed kernel q, the fibre has at most $N^{2c\log 2 + o(1)}$ elements. All estimates are uniform in A.

Proof. Pair the d linear factors into $m = d/2$ quadratics Q_j, put $q_j = \mathrm{sf}(Q_j(n))$, and write

$$D_A = \prod_{\substack{a,b \in A \\ a<b}} |a - b|.$$

Its logarithm is $O_D(d^2 \log B) = o(\log N)$. Outside D_A, the q_j are pairwise coprime, and

$$\frac{\prod_j q_j}{q} \mid D_A^2. \tag{T5.3}$$

Indeed, a prime occurring in s of the q_j divides at least s distinct linear factors. Its exponent in the quotient is $2\lfloor s/2 \rfloor$, whereas its exponent in D_A^2 is at least $2\binom{s}{2}$.

Remove the primes of D_A from e, obtaining $e' = e/\gcd(e, \mathrm{rad}\, D_A)$. There is a factorization $e' = \prod_j e_j$, with $e_j \mid q_j$, such that the factors are pairwise coprime. Set $\beta_j = \log q_j/\log N$ and $\delta_j = \log e_j/\log N$. Then

$$\sum_j \beta_j = \beta + o(1), \qquad \sum_j \delta_j = \delta + o(1).$$

For at least one j, $\beta_j - 2\delta_j \leq (\beta - 2\delta)/m + o(1)$. Lemma T5.3 bounds its quadratic count by $N^{1/2 + \beta/(2d) - \delta/d + o(1)}$. We enumerate only that j, the divisor $e_j \mid e'$, and its dyadic size. This costs at most $m\tau(e')O(\log N)$, not the number of all factorizations of e'. Since $e' \leq (C_D N)^d$, the product of its distinct prime factors is at least $\omega(e')!$. Stirling's estimate therefore gives

$$\omega(e') \leq (c + o(1)) \log N, \qquad \tau(e') \leq N^{c\log 2 + o(1)}.$$

This proves (T5.2).

For a fixed q, choose two distinct offsets a, b. Both $\mathrm{sf}(n+a)$ and $\mathrm{sf}(n+b)$ divide $\mathrm{rad}(qD_A)$, so there are at most $4^{\omega(qD_A)} \le N^{2c\log 2+o(1)}$ choices. The effective coefficients u, v satisfy $u, v \ll_D N$. For each such choice, $us^2 - vt^2 = a - b \neq 0$ has $N^{o(1)}$ solutions by Lemma T1.2, including its factorizable case. Thus Pell is applied to coefficients of height $O(N)$, even though q itself may be as large as $N^{O(d)}$. □

T5.1.1 Proof of Proposition T5.1

Proof. The moment counts the tuples for which $\prod_i P_A(n_i)$ is a square. Write $q_i = \mathrm{sf}(P_A(n_i))$. There is a unique representation

$$q_i = \prod_{\substack{S\subseteq\{1,\dots,v\}\\ |S|\ge 2 \text{ even},\ i\in S}} d_S$$

with pairwise coprime squarefree d_S. Fix dyadic size classes for these numbers, and put $\alpha_S = \log d_S/\log N$ and $\beta_i = \sum_{S\ni i}\alpha_S$. There are $N^{o(1)}$ size profiles, since v is fixed.

Expose $v-1$ of the integers in a chosen order. Before exposing i, the product of the d_S whose sets meet an earlier index is a prescribed divisor $e_i \mid q_i$. If its exponent is δ_i, Lemma T5.4 gives the cost

$$N^{1/2+\beta_i/(2d)-\delta_i/d+c\log 2+o(1)}.$$

After choosing n_i, assign the new prime factors of q_i/e_i to the at most 2^{v-1} possible sets S. Their number is at most $N^{c(v-1)\log 2+o(1)}$; inconsistent assignments are discarded. Every S has at least two members, so the last kernel is now fixed and its fibre costs $N^{2c\log 2+o(1)}$. The enumeration costs in total at most $N^{c[v(v-1)+2]\log 2+o(1)}$.

For the remaining exponent, average over orders. A set S of size a has coefficient

$$\frac{1}{d}\left(1-\frac{a}{2}+\frac{1}{2}\mathbf{1}_{\{\text{last index}\in S\}}\right)$$

in $\sum_{i\neq\text{last}}(\beta_i/(2d)-\delta_i/d)$. Its average is $d^{-1}(1-a/2+a/(2v))$, which is $1/(dv)$ for $a=2$ and nonpositive for $a\ge 4$. As $2\sum_{|S|=2}\alpha_S \le \sum_i\beta_i \le vd+o(1)$, some order has total exponent at most $v/2+o(1)$. The order can depend on the size profile; its $v!$ choices do not affect the estimate. This proves (T5.1). □

T6 A half power for each new window

Relations with long selections require a different count. The useful alternative is concrete: either several large odd prime kernels determine another centre by congruences, or many selected values have large smooth parts. Those smooth parts then determine a rare centre by a modulus larger than N. Iterating this alternative reaches the square-root scale.

In this section the number of blocks and their widths are fixed before N. A block consists of distinct integers in a translate of an interval of diameter $O_D(B)$; its values are comparable to N, and its centre varies in an interval of length $O_D(N)$. Repeated values in a normalized block are merged. These conventions include original windows and the fixed-width blocks obtained by compressing a rational pair. Let S_z denote the image of its even selections.

Theorem T6.1 (Conditional full-support count)**.** *Fix $r\ge 1$, $s\ge 0$, $D>0$, and $\varepsilon>0$. Let V_{old} be the space generated by all values in s fixed source blocks. Uniformly in those sources, the number of lists of distinct new centres $z_1,\dots,z_r$ such that*

$$u_i \in S_{z_i}\setminus\{0\}\quad (1\le i\le r),\qquad \sum_i u_i \in V_{\mathrm{old}} \tag{T6.1}$$

for some u_i, is $O_{r,s,D,C,\varepsilon}(N^{r/2+\varepsilon})$. The u_i need not remain nonzero modulo V_{old}. The source blocks may overlap or be dependent.

The condition asks for a nonzero projection in each new effective space. It imposes neither affine compatibility nor a separation condition on physical values.

For $H = C_D B$ larger than every block diameter, write

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

This is the odd prime kernel after the small primes have been removed.

Why the half power appears. For a fixed witness, short selections are counted by Corollary T5.2. A long selection either has many large odd kernels, which determine another centre by congruences, or has many large smooth parts, whose least common multiple exceeds the length of the centre interval. In the second case the centre lies in a sparse deterministic set.

The complete proof in Section T6.1.3 iterates these two alternatives. At a stage indexed by $x_j \le 1/2$, a surviving source has at most $N^{1-x_{j-1}+o(1)}$ possibilities and its receiver degree is at most $N^{x_j+o(1)}$. Removing this source–receiver pair and applying induction to the other $r-2$ blocks costs $N^{r/2+x_j-x_{j-1}+o(1)}$. Small fixed increments and a final $N^{1/2+o(1)}$ exceptional source set give the theorem. The induction is simultaneous in every fixed number of old sources, so adding exposed blocks to that space introduces no dependence assumption.

Corollary T6.2 (Connected and primitive families at fixed order)**.** *For every fixed $r \ge 2$, connected families of r original windows have $O(N^{r/2+\varepsilon})$ hosts. Their independent mass, and in particular the independent mass of primitive families, is $O(N^{-r/2+\varepsilon})$.*

Proof. A primitive family has a relation with nonzero projection in each block, so Theorem T6.1 applies with no sources. A connected family contains a minimal dependent subfamily; expose it first. If A is the set already exposed, connectedness implies

$$\left(\sum_{i\in A} S_i\right) \cap \left(\sum_{j\notin A} S_j\right) \neq \{0\}.$$

Choose a minimal nonempty subset U of the complement whose sum meets the old sum nontrivially. A nonzero vector in that intersection has a representation with all projections on U nonzero, by minimality. The theorem counts these new blocks by $N^{|U|/2+\varepsilon'}$. Continue until the whole family is exposed. The group sizes sum to r, and the number of possible ordered partitions of the index set depends only on r. Their product count is $N^{r/2+O_r(\varepsilon')}$. Multiplying by $\prod p_x \ll_C N^{-r}$ gives the mass. □

Corollary T6.3 (Dependent pairs)**.** *The number of pairs of blocks with nonzero intersection of their effective even-selection spaces is $O(N^{1+\varepsilon})$, uniformly for fixed logarithmic widths and fixed normalizations.*

Proof. Use Theorem T6.1 with $r = 2, s = 0$. The two equal nonzero classes give the required relation. □

Sharpness and range. For $r = 2k$, take k pairs of adjacent windows, with the pairs far apart. In each pair select the same two consecutive values in the overlap. Their square class is nonzero, since $n(n+1)$ lies strictly between n^2 and $(n+1)^2$, and the equal projections cancel. There are $\gg N^k$ such lists. Thus $r/2$ is optimal for the general full-support count at even orders. For $k > 1$ these examples can be decomposable; they make no optimality claim for connected or primitive subclasses. Every order and width in this section remains fixed.

T6.1 Conditional half-counts and the certificate iteration

We prove Theorem T6.1 with its stated uniformity in arbitrary fixed source blocks. Blocks and effective spaces have the meaning specified in Section T6; in particular, the nonzero projections are required in the individual new spaces, not in their quotients by the old space.

T6.1.1 Small kernels and smooth moduli

Choose $H = C_D B$ larger than every block diameter, and set

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

For $q > 0$, let $\mathcal{X}_q$ be the set of centres containing distinct values n_1, n_2 with $K_H(n_1)K_H(n_2) \le N^q$.

Lemma T6.4 (Two small kernels)**.** *For every fixed $q > 0$ and $\varepsilon > 0$, $\#\mathcal{X}_q \ll_{q,D,C,\varepsilon} N^{q+\varepsilon}$.*

Proof. Write $n_i = c_i k_i a_i^2$, where $k_i = K_H(n_i)$ and c_i is squarefree and supported on primes at most H. The c_i have $2^{2\pi(H)} = N^{o(1)}$ choices, while the pairs $k_1 k_2 \le N^q$ have $N^{q+o(1)}$ choices. Fix the two distinct offsets. Their nonzero difference h gives

$$c_1 k_1 a_1^2 - c_2 k_2 a_2^2 = h.$$

The effective coefficients are $O_D(N)$. Lemma T1.2 gives $N^{o(1)}$ solutions, including the factorizable case, and each solution determines the centre. The $O_D(B^2)$ offsets are absorbed. □

For $Y = (\log N)^A$, with fixed $A \ge 2$, write

$$n = s_Y(n) a_Y(n)^2 K_Y(n).$$

Here $s_Y(n)$ is the entire part supported on primes at most Y; $a_Y(n)$ contains the halves of the valuations at larger primes; and $K_Y(n)$ is their odd squarefree part.

Lemma T6.5 (A large smooth modulus)**.** *Fix t, $c > 0$, and $A \ge 2$. There is a deterministic set $\mathcal{M}_Y$ of centres, of cardinality $O(N^{1-1/A+\varepsilon})$, with the following property. If a selection has more than $c \log\log N$ values, its centre lies outside $\mathcal{X}_{1/2}$, and it contains fewer than t values with $K_Y(n) > N^{1/5}$, then its centre belongs to $\mathcal{M}_Y$.*

Proof. Outside $\mathcal{X}_{1/2}$, at most one value satisfies $a_Y(n)^2 > N^{3/4+1/100}$: two such values would give

$$K_H(n_1)K_H(n_2) \le \frac{O_D(N^2)}{a_Y(n_1)^2 a_Y(n_2)^2} < N^{1/2}.$$

Remove this possible value and the at most $t-1$ large kernels. For sufficiently large N, select 26 remaining values. Their smooth parts satisfy

$$s_i \gg_D N^{1-3/4-1/100-1/5} = N^{1/25}.$$

Since $\gcd(s_i, s_j) \le |n_i - n_j| = O_D(B)$,

$$\operatorname{lcm}(s_1, \dots, s_{26}) \ge \frac{\prod_i s_i}{\prod_{i<j} \gcd(s_i, s_j)} \ge N^{26/25 - o(1)} > C_1 N,$$

where $C_1 N$ exceeds the length of the centre interval. Multiply its prime factors, one at a time and with multiplicity, until the first passage above $C_1 N$. The resulting smooth modulus satisfies

$$C_1 N < D_0 \le C_1 N Y, \qquad D_0 \mid \operatorname{lcm}(s_1, \dots, s_{26}).$$

For fixed $1 - 1/A < \sigma < 1$, Rankin's elementary bound counts such moduli by

$$(C_1 NY)^\sigma \prod_{\ell \le Y} (1 - \ell^{-\sigma})^{-1} = N^{\sigma + o(1)}.$$

Indeed the logarithm of the product is $O_\sigma(Y^{1-\sigma}) = o(\log N)$. Assign each maximal prime power of D_0 to one of the 26 offsets on which it occurs. This costs $26^{\omega(D_0)} = N^{o(1)}$. The Chinese remainder theorem determines at most one centre in its interval, since $D_0 > C_1 N$. The offsets themselves have $B^{26} = N^{o(1)}$ choices. Choose σ sufficiently close to $1 - 1/A$. □

T6.1.2 Several anchors force a congruence certificate

Lemma T6.6 (Incoming and outgoing certificate counts)**.** *Suppose a relation in* (T6.1) *selects t values in one new block whose kernels exceed N^b, where $b = 1/5$ and $tb > r + s - 1$. Some other old or new block receives a congruence certificate whose modulus exceeds a fixed multiple of NB. For fixed r, s, t, A, the directed graph of these source–receiver certificates has both incoming and outgoing degrees at most $N^{1/A + o(1)}$.*

Proof. Represent the old vector by a selection of actual old values; parity in the old blocks is not required. Together with the new selections this is a square product. Primes above Y cannot occur in two distinct values of one block. The t anchor kernels are therefore pairwise coprime, and their product Q satisfies $Q > N^{tb}$, $Q \ll_D N^t$. Assign each prime of Q to one of its other odd occurrences. The $r + s - 1$ receiver moduli have product Q. For sufficiently large N, at least one exceeds $C_2 NB$.

In that modulus keep the primes only up to its first passage above $C_2 NB$. The retained list has at most

$$k = \left\lfloor \frac{\log(C_2 NB)}{\log Y} \right\rfloor + 1$$

members. For a fixed source, the t offsets and a subset of the $O_t(\log N / \log Y)$ available primes cost $N^{o(1)}$. Assigning their destinations among $O_D(B)$ receiver positions costs at most $B^k = N^{1/A + o(1)}$. The congruences determine the receiver.

For a fixed receiver, its values contain $R = O_D(B \log N / \log Y)$ primes above Y. Choose at most k of them and assign each to an anchor: the number of codes is at most

$$B^t \sum_{j \le k} \binom{R}{j} t^j = N^{1/A + o(1)}.$$

When $R \ge 2k$, use $(k+1)(etR/k)^k$; when $R < 2k$, use $(1+t)^R = N^{o(1)}$. The retained modulus again exceeds the source interval. Forgetting all other primes enlarges the count and does not remove the original full-support relation from the remaining blocks. □

T6.1.3 The iteration

Proof of Theorem T6.1. We induct strongly on r, simultaneously for every fixed number s of sources. Put $T_0 = 1$. The case $r = 1, s = 0$ is empty, since its sole projection is required to be nonzero. For all other cases, choose

$$t = 5(r + s - 1) + 1.$$

First choose a small fixed step $\delta > 0$ within the final error margin. Let

$$x_0 = 0, \qquad x_j = \min\{\tfrac{1}{2}, x_{j-1} + \delta\}, \qquad A_j = x_j^{-1}, \qquad Y_j = (\log N)^{A_j}.$$

There are $J = \lceil 1/(2\delta) \rceil$ stages. All these numbers are fixed before N; choose the internal margins after J.

Corollary T5.2 counts lists possessing a witness with all new selections short. For every remaining list choose one witness, without enumerating it. It has at least one long new selection. At the start of stage j, every surviving long new source outside $\mathcal{X}_{1/2}$ belongs to a set of size $N^{1-x_{j-1}+\varepsilon'}$. This is trivial for $j = 1$.

If such a source selects t large K_{Y_j} kernels, Lemma T6.6 supplies a receiver. When the receiver is new, count the source–receiver pair by $N^{1-x_{j-1}+x_j+O(\varepsilon')}$. The other $r-2$ projections are still nonzero in their own spaces; their sum belongs to the old space enlarged by these two fixed blocks. The induction hypothesis counts them by $N^{(r-2)/2+\varepsilon'}$. The total is

$$N^{r/2+x_j-x_{j-1}+O(\varepsilon')} \le N^{r/2+\delta+O(\varepsilon')}.$$

If the receiver is old, it is already fixed. The incoming degree counts the new source by $N^{x_j+O(\varepsilon')}$. Adjoin that source to the old space and apply induction to the $r-1$ remaining targets. The cost is at most

$$N^{x_j+(r-1)/2+O(\varepsilon')} \le N^{r/2+O(\varepsilon')}.$$

Only a fixed number of choices of roles is involved.

Remove all these saturated cases. Every surviving long selection outside $\mathcal{X}_{1/2}$ has fewer than t large kernels, and Lemma T6.5 places its centre in a set of size $N^{1-x_j+\varepsilon'}$. This proves the stage invariant. At the final stage every long source lies in $\mathcal{X}_{1/2} \cup \mathcal{M}_{Y_J}$, whose size is $N^{1/2+\varepsilon'}$. Fix one such centre and use the induction hypothesis for the other $r-1$ targets. This costs $N^{1/2+\varepsilon'} N^{(r-1)/2+\varepsilon'}$.

Every application of induction uses fewer targets, although the fixed number of old sources can increase. Thus the simultaneous induction is legitimate. Choosing δ and then all internal margins within ε proves the theorem. □

T7 Confining large common divisors

At small relative rank, the host count already controls triple probabilities. At macroscopic rank, many large prime factors must be shared between the windows. An interpolation argument confines the offsets of these common divisors to a fixed-degree curve. Unless the pair has an exact rational alignment, that curve contains too few offsets to sustain many large groups.

Fix $H = C_1 B$ larger than the diameters and use the odd kernel K_H from Section T6. For this section, a pair is called aligned if two distinct offset pairs give the same ratio of true values. Equivalently, two points (i, j) lie on a line through $(-x, -y)$. Reducing the ratio bounds its height by $O(B)$; the enlarged height-B^3 convention used elsewhere in the paper includes these alignments.

The confinement lemma, proved in Section T7.1, starts from pairwise coprime groups $g_a > N^\tau$ dividing $x+i_a$ and $y+j_a$. Determinants of bounded-degree monomials force their offset points onto an algebraic curve. Nonlinear components have $o(B)$ grid points, while a line containing many groups must be an exact alignment. After removing these few groups, the surviving coordinates have small odd kernels and the half-count machinery makes their centres rare. The argument never enumerates the auxiliary curve or a chosen relation subspace.

For each fixed arity, we also use an anchored version: after coordinate regularization, each active anchor value meets only a fixed number of minimal prime supports. Confinement removes the large supports at a cost of $o(B)$ anchor coordinates. Two remaining small kernels make that centre rare. On a primitive family this works at every window, giving N^ε dense hosts and the fixed-size weights in Proposition T12.4. The arity and density proportion are fixed before N. The

same section completes the stronger pair profile of Proposition T2.2 without using that profile as an input.

The two forms needed to partition compatible triples are as follows. The argument also proves the primitive-triple and dense-pair versions, Theorem T7.4 and Corollary T7.5.

Proposition T7.1 (Dense triples with weak proper intersections)**.** *Fix $\eta > 0$ and $0 < \theta < \eta/100$. Admissible triples satisfying*

$$\rho_{\mathrm{rel}}(T) \geq \eta B, \qquad d_{xy} \leq \theta B \quad \text{for every pair}$$

have $O_{\eta,\theta,C,\varepsilon}(N^{\varepsilon})$ hosts and total homogeneous mass $O(N^{-2+\theta+\varepsilon})$.

Its proof is given in Section T7.1.1.

Proposition T7.2 (A strong nonaligned pair in a triple)**.** *Fix $\theta, \sigma, \varepsilon > 0$. Compatible connected triples with a nonaligned pair of rank at least θL, and $\rho_{\mathrm{rel}}(T) \leq \sigma L$, have total joint mass*

$$O_{\theta,\sigma,C,\varepsilon}(N^{-\min\{3/2,\,3-\sigma\}+\varepsilon}).$$

Its proof is given in Section T7.1.2.

The last proposition preserves the actual conditional probabilities. Replacing them by independent weights would lose the large-rank case. What remains is a pointwise bound on the possible pair ranks and a separate calculation for rationally aligned triples.

T7.1 Polynomial confinement and dense arithmetic profiles

We use the block and alignment conventions of Section T7. The group argument treats nonaligned pairs and primitive families at every fixed arity, then proves the two triple interfaces stated there.

Lemma T7.3 (Polynomial confinement of common divisors)**.** *Fix $\tau > 0$, $0 < \delta < 1/2$, and $C_0 > 0$. Let $x, y \asymp N$, and let pairwise coprime integers $g_a > N^{\tau}$ satisfy*

$$g_a \mid x + i_a, \quad g_a \mid y + j_a, \qquad |i_a|, |j_a| \leq C_0 B.$$

Suppose no line through $(-x, -y)$ contains two distinct points among the (i_a, j_a). Then their number, with multiplicity, is $O_{\tau,\delta,C_0}(B^{1/2+\delta}) = o(B)$. The same estimate holds after points on exact alignments have been removed.

Proof. Choose a fixed degree d such that, for $M = (d+1)(d+2)/2$, $(M-1)\tau > d$. Let $v_d(s,t)$ list the M monomials of total degree at most d, including 1. If M of the vectors $v_d(i_a, j_a)$ form a basis, replace each row of that basis in turn by $v_d(-x, -y)$. Subtracting the new row from the other rows shows that the resulting determinant is divisible by the product of their $M-1$ pairwise coprime divisors. Its absolute value is at most

$$C_d N^d B^{d(M-1)} < N^{(M-1)\tau}.$$

All the determinants vanish. The vector $v_d(-x, -y)$ would then belong to the intersection of the M coordinate hyperplanes of this basis, which is $\{0\}$. Its constant coordinate is one, a contradiction.

Thus a nonzero polynomial $F(S,T)$ of degree at most d vanishes on all the offset points. For a geometrically integral component of degree at least two, Pila's uniform estimate [7, Theorem A] gives $O_d(B^{1/2+\delta})$ integer points of height $O(B)$. A rationally irreducible component that is not geometrically integral has only $O_d(1)$ rational points, by intersecting two conjugate components.

A line through at least two distinct grid points has a primitive equation $aS + bT = c$ with $|a|, |b| \ll B$, $|c| \ll B^2$. Every group on it divides $ax + by + c$. This integer is nonzero by the

hypothesis, and has size $O(NB)$. Coprimality bounds the number of these groups by $O_\tau(1)$. A line through only one grid point needs no separate point count. Finally, all groups at one point divide its true value $x + i_a = O(N)$; their multiplicity is $O_\tau(1)$. There are at most d components. These observations prove the bound. □

The groups in this lemma can be composite. Only divisibility and coprimality are used, and the auxiliary curve is not enumerated. Its degree is chosen after τ and before N.

An anchored small-kernel witness at fixed arity. Fix integers $r \geq 2$, $t \geq 2$ and real numbers $\eta, \tau > 0$. Let M be a binary space of physical square relations on r true windows with centres in I_N: its words select an even set in each window and their total product has even prime valuations. Write $\ell_v \in M^*$ for the coordinate forms and call a position active when $\ell_v \neq 0$. Designate one window as the anchor. Assume that:

- $\dim M \geq \eta L$ and the physical projection into the anchor is injective;
- no nonzero form on M has t representations by pairwise disjoint coordinate sets of size between 1 and $r - 1$;
- between the anchor and each other block, no two distinct pairs of active offsets have the same ratio of true values.

For every cutoff $H \geq B = L + 1$, the anchor then contains two distinct values n_1, n_2 with $K_H(n_i) \leq N^{D\tau}$, where

$$D = \sum_{j=1}^{r-1} j!(t-1)^{j-1} 2^{j(t-1)}.$$

One threshold, chosen after r, t, C, η, τ, precedes the centres, M and H. No injectivity or non-alignment between the other blocks is assumed.

Proof. The recovery-exclusion hypothesis gives multiplicity at most $t - 1$ for every nonzero coordinate label, by considering singleton recovery sets. The final counting argument of Lemma T4.1 also bounds the independent j-label representations of a nonzero label by $j!(t-1)^{j-1}$: a maximal disjoint family of their supports has fewer than t members; its union, of size at most $j(t-1)$, meets every representation, and induction removes a label. A zero remaining sum is excluded by independence. Permitting every subset of the j label fibres gives at most $2^{j(t-1)}$ physical lifts. Thus at most D physical supports of size at most $r - 1$ can represent a fixed active anchor label.

A prime $\ell > H$ has at most one odd occurrence in each true window. If it occurs at an active anchor position v, its parity row on M expresses ℓ_v as a sum of at most $r - 1$ other labels. Remove zero labels and choose a minimal subrepresentation. Its labels are distinct, nonzero and independent: otherwise a nonempty zero-sum subset could be removed without changing the represented nonzero form. Choose their actual positions in the prime row. The resulting support A consists of v and positions in other blocks and uses only active coordinates.

Assign every prime with an active anchor occurrence to one such support, and let g_A be the product assigned to A. Distinct nontrivial groups are coprime and divide every true value selected by their support. Every anchor position meets at most D groups, and its entire odd kernel above H divides their product. Assign each group to one other block it meets. For each of the $r - 1$ choices, Lemma T7.3 gives $o(B)$ groups larger than N^τ, with their full multiplicities. Its no-line hypothesis is exactly the stated active-offset condition. Equal projected points are covered by the multiplicity bound in that lemma.

Impose zero only on the anchor positions touched by these large groups. Each group has a unique anchor position, so the dimension cost is at most the number of such groups, namely $o(B)$. A nonzero space remains for large N and its anchor projection is still injective. A nonzero even

projected word uses at least two positions. Each of those positions was active on the original space M and met only groups at most N^{τ}, so its odd kernel is at most $N^{D\tau}$. The incidence bound is used on M before erasure; no recovery condition or prime-row classification on the erased space is assumed. □

Dense primitive hosts and their exact weights. For each fixed $r \geq 3$, $\eta > 0$ and $\varepsilon > 0$, primitive families of r true windows with $\rho_{\mathrm{rel}} \geq \eta L$ have $O(N^{\varepsilon})$ hosts. Indeed, every effective relation-space projection is injective. Choose linear sections of the local even-selection maps to lift that space without increasing its dimension or adding local square kernels. Lemma T4.1, with $d = r - 1$ and t chosen after r, η, loses at most $2rB\, t^{-1/(r-1)} \leq \eta L/4$ dimensions. The surviving physical space has dimension at least $3\eta L/4$, retains all injective projections, and has the required recovery exclusion. Two exact units between any pair would give equal nonzero even square classes in those two blocks: the selected two-value products are nonsquares by Lemma T1.1. This contradicts primitivity, since the pair is a proper subfamily. Thus the active-offset condition holds for each choice of anchor. Apply the witness with $H = B$ and density parameter $\eta/2$ at each of the r windows, with the same fixed constant D. Every centre lies in $\mathcal{X}_{2D\tau}$. Lemma T6.4 therefore counts at most $N^{2rD\tau + r\varepsilon'}$ unordered families. Choose τ and ε' after r, η, ε so that this exponent is at most ε. No auxiliary support or chosen subspace is enumerated in this last count.

The conversion of this host count to homogeneous probabilities and the three original weights is given in Proposition T12.4. That proposition uses the host count proved here, together with the conditional half-count, and treats the weight comparison at fixed arity.

Theorem T7.4 (Dense primitive triples)**.** *For fixed $\eta, \varepsilon > 0$, primitive triples of true windows with $\rho_{\mathrm{rel}}(T) \geq \eta B$ have $O_{\eta,C,\varepsilon}(N^{\varepsilon})$ hosts. Their homogeneous mass is $O(N^{-2+\varepsilon})$. The total absolute centered activity of all primitive triples is $O(N^{-3/2+\varepsilon})$.*

Proof. Apply Proposition T12.4 with $r = 3$. The dense condition $\rho_{\mathrm{rel}}(T) \geq \eta B$ implies $\rho_{\mathrm{rel}}(T) \geq \eta L$, giving the host count and the homogeneous mass $N^{-2+\varepsilon}$. The global absolute-centered bound is $N^{-3/2+\varepsilon}$. □

Here homogeneous mass is the sum of the probabilities of the homogeneous events, with zero right-hand side in place of the affine start prescription. It must be distinguished from a weight using the original marginals p_x, which can vanish. For instance, a start at $x = 3$ of length $L = 7$ is impossible: the square 4 forces the common bit on $3, \ldots, 9$ to be zero, and $6 = 2 \cdot 3$ then forces the bit at 2 to be zero, contradicting the left-boundary equation. The corresponding homogeneous event is compatible and has positive probability.

Corollary T7.5 (Dense nonaligned pairs)**.** *For fixed $\eta, \varepsilon > 0$, nonaligned pairs with $d_{xy} \geq \eta B$ have $O(N^{\varepsilon})$ hosts.*

Proof. Realize their effective relation space with injective projections. Regularize with Lemma T4.1, now with $d = 1$, losing at most $\eta B/4$. Every active prime row has two equal labels. A coordinate belongs to at most $K = t - 1$ active supports. Confinement bounds the groups above N^{τ} by $o(B)$. Kill both endpoints of every such group. The remaining positive-dimensional space has two active positions in each block, with kernels at most $N^{K\tau}$. The centres lie in $\mathcal{X}_{2K\tau}$. Their count is at most $N^{4K\tau + 2\varepsilon'}$; choose τ, ε' sufficiently small. □

Completion of the quasi-linear pair profile. We now prove the stronger bounds in Proposition T2.2, retaining its full arithmetic nullity $\rho(x, y) = 2L - \dim(S_x + S_y)$. If both windows are regular, positive ρ is exactly a nonzero effective intersection; Corollary T6.3 counts these pairs by $N^{1+\varepsilon'}$. A locally deficient coordinate gives at most $2N\,|I_N \setminus G_N| = N^{1+o(1)}$ further ordered pairs. This proves the support bound, even before imposing separation.

For the weighted nonaligned sum, split at $d_{xy} = \eta L$ with fixed $\eta > 0$. Local rigidity gives, uniformly,

$$\rho(x,y) = d_{xy} + (L - d_x) + (L - d_y) \le d_{xy} + 2, \qquad \rho(x,y) \le L + 2.$$

In the low sector, the just proved support bound and $2^{\eta L+2} \ll_C N^{\eta}$ give $N^{1+\eta+\varepsilon'}$. In the high sector, Corollary T7.5, with a smaller fixed density parameter if necessary to replace L by B, gives $N^{\varepsilon'}$ hosts, each of weight $O_C(N)$. Choose η, ε' after the final ε. This proves both assertions with $N^{1+\varepsilon}$ and with a common threshold before all pairs. Only conditional half-counts, local rigidity and confinement were used; the pair-profile theorem and its elementary 6/5 proof are not inputs to this completion.

T7.1.1 Proof of Proposition T7.1

Proof. The kernel of each projection of the relation space has dimension at most θB. Intersect three complements to these kernels. The resulting subspace loses at most $3\theta B$ dimensions and has all projections injective. Use linear sections of the local even-selection maps to realize this subspace in physical coordinates, preserving its dimension and the injectivity of its block projections.

In a given pair, two exact units impose a reduced rational alignment of height $O(B)$. Two distinct such alignments would bound the centres by $O(B^3)$, which is impossible at the present scale. The alignment is therefore unique. If it has u units, their even-selection space has dimension $u - 1$; its image loses at most one more dimension by Lemma T1.18. Thus $u \le d_{xy} + 2 \le \theta B + 2$. Kill all the positions of these units in the three pairs, at a cost of at most $6\theta B + 12$ dimensions. The remaining space still has macroscopic dimension and injective projections.

Apply Lemma T4.1 with $d = 2$, choosing a fixed integer $t \ge 2$ such that $6t^{-1/2} \le \eta/8$. There are at most $3B$ physical coordinates, so this regularization loses at most $\eta B/8$ further dimensions. Denote the resulting physical relation space by M. For all sufficiently large N,

$$\dim M \ge \left(\frac{7\eta}{8} - 9\theta\right) B - 12 \ge \frac{\eta L}{2},$$

since $\theta < \eta/100$.

Apply the anchored small-kernel witness of Section T7.1 separately to each of the three windows, using the same space M, the fixed cutoff H, and density parameter $\eta/2$. Its block projections are injective, its coordinate labels satisfy the recovery exclusion, and its active offsets have no remaining exact alignments. Let $D_* = D(3,t)$ be the support-incidence constant in that witness. Each centre therefore belongs to $\mathcal{X}_{2D_*\tau}$, and Lemma T6.4 bounds the number of triples by

$$O\left(N^{6D_*\tau + 3\varepsilon'}\right).$$

Choose $\tau, \varepsilon' > 0$ after the requested ε so that $6D_*\tau + 3\varepsilon' \le \varepsilon/2$. In each application, the incidence bound is used before the final anchor-coordinate erasure. No auxiliary subspace or deletion history is enumerated.

Finally, $\rho_{\mathrm{rel}}(T) \le d_z + d_{xy} \le L + \theta B$. The homogeneous weight is at most $N^{-2+\theta+o(1)}$. Absorbing the $o(1)$ in the remaining exponent margin proves the mass estimate. □

T7.1.2 Proof of Proposition T7.2

Proof. Choose the strong pair $e = \{x, y\}$. Put $\delta_z = \dim(S_z \cap (S_x + S_y)) = \rho_{\mathrm{rel}}(T) - d_{xy}$. Connectedness gives $\delta_z > 0$, and compatibility gives $q(T) = q(e) p_z 2^{\delta_z}$.

For $0 < \delta_z < \eta' L$, the conditional count in Theorem T6.1 gives $N^{1/2+\varepsilon'}$ targets of weight at most $N^{-1+\eta'+o(1)}$. Lemma T1.8, together with Corollary T6.3 for the independent part, bounds the total joint mass of strong nonaligned pairs by $N^{-1+\varepsilon'}$. This branch costs $N^{-3/2+\eta'+O(\varepsilon')}$.

For $\delta_z \geq \eta' L$, the effective-space form of Proposition T4.2, explained after (2.7), gives $N^{\varepsilon'}$ targets after the fixed pair. Corollary T7.5 also gives $N^{\varepsilon'}$ source pairs. Each triple has $q(T) \ll N^{-3+\sigma+o(1)}$. Choose η' and the internal margins after ε. □

T8 Short square relations and the rank of a pair

A large intersection of the square-class spaces of two windows contains many relations, but those relations need not initially have short support. We first extract a large subspace with a short basis. Two independent basis vectors then produce an algebraic curve through the two window centres. A resultant argument puts the whole subspace on one curve, while the relations force many integer points in the small grid of offsets. A point count uniform in the degree rules out this configuration unless the centres lie on a rational alignment. The distinction between a fixed degree and a growing degree is essential: the basis supports grow with L throughout the argument.

The combination of resultant elimination and degree-uniform curve bounds also appears in Shparlinski and Sleiman's work on consecutive multiplicatively dependent triples [14, Theorem 1.1 and Section 3]. Here the common curve must accommodate an entire space of square-class relations and the integer fibres forced by that space in the offset grid.

All estimates in this section are uniform in $x, y \in I_N$, $|L - \log_2 N| \leq C$, for fixed C, and we put $B = L + 1$. A logarithm without a subscript is natural. The conclusion used in the probability argument is Corollary T8.8.

T8.1 Square classes on all windows

We use the true window $V_x = \{x-1, x, \dots, x+L-1\}$, including its left endpoint. Let E_x be the space of even selections from V_x, and let

$$\phi_x : E_x \longrightarrow \mathbb{Q}_{>0}^{\times}/\mathbb{Q}_{>0}^{\times 2}, \qquad I \longmapsto \left[\prod_{n \in I} n\right]$$

be the square-class map. Its image is the space S_x already used above. Put $d_{xy} = \dim(S_x \cap S_y)$ and, for any finite set of centres T, define the *relative nullity*

$$\rho_{\mathrm{rel}}(T) = \sum_{x \in T} \dim S_x - \dim\left(\sum_{x \in T} S_x\right). \tag{T8.1}$$

On the regular set, where every S_x has dimension L, this is the earlier nullity $\rho(T)$. Formula (T8.1) also applies to the remaining windows and separates their local rank defects from dependence between distinct windows.

A *complete relation* of a pair is a pair of even selections with the same square class. We represent its selections by offsets $I, J \subset \{-1, 0, \dots, L-1\}$ and attach the polynomial

$$F_w(X, Y) = \prod_{u \in I}(X + u) \prod_{v \in J}(Y + v), \qquad w = (I, J). \tag{T8.2}$$

Thus $F_w(x, y)$ is an integer square. The word w is a vector over $\mathbb{F}_2$; addition means symmetric difference in each block. The two blocks remain disjoint formal copies even when the actual windows overlap. Consequently, independent words give independent classes of their polynomials in $\mathbb{Q}(X, Y)^{\times}/\mathbb{Q}(X, Y)^{\times 2}$: the factors $X + u$ and $Y + v$ are distinct irreducible polynomials.

We call x, y *aligned* if there are coprime positive integers $a, b \leq B^3$ such that

$$|bx - ay| \leq (a + b)B. \tag{T8.3}$$

This definition includes nearby centres. Otherwise they are nonaligned. Alignment is a property of the centres; compatibility of their run events is a separate condition.

T8.2 Extracting short relations

Primes larger than $10B$ meet each physical block at most once. Their parity equations therefore join positions into components; a component meeting a prime only once is forced to zero. Local square-product separation shows that there are only $O_C(B/\log B)$ defective positions. Thus the number of free components meeting both blocks is $d_{xy} + O_C(B/\log B)$; see Lemmas T8.10 and T8.11.

For medium primes, columns of weight two can be kept as edges of a graph. Its short cycles cancel their equations. Discarding only columns of weight at least three costs one third of their total weight, explaining the factor $2/3$ in the following extraction.

Theorem T8.1 (A short basis with a controlled loss of dimension)**.** *For every fixed $a \in (0,1)$ there is a subspace $\mathcal{K}_a$ of complete relations of the pair, with injective projections onto each physical block and onto the common square class, such that*

$$\dim \mathcal{K}_a \geq d_{xy} - \frac{2}{3} B \log(1/a) - O_{a,C}\left(\frac{B}{\sqrt{\log B}}\right). \tag{T8.4}$$

It has a basis whose vectors have total physical support at most

$$O_{a,C}\left(\frac{B^a \log B}{\log\log B}\right). \tag{T8.5}$$

Every nonzero vector selects a nonempty even set in each block and has nonzero common square class. Neither compatibility nor nonalignment is assumed.

The full extraction proof, including the final small-prime equations and projection injectivity, is in Section T8.9.3.

T8.3 One curve for a whole relation space

Two independent short words first yield a nonzero polynomial through the window centres. Proposition T8.13 constructs it as the norm of a biquadratic auxiliary function, with degree and logarithmic height linear in the support size. Polynomial gcd removal is performed before specialization, so vanishing at the integer point does not destroy the norm argument.

The simultaneous step below is stronger than constructing a separate curve for each word. Pairwise resultants have quadratic height cost in the support, independent of the number of basis vectors.

Theorem T8.2 (A common curve)**.** *There is an absolute constant $c > 0$ with the following property. Let V be a space of even formal words of dimension $t \geq 2$, with a basis $w_1, \dots, w_t$ satisfying $|I_{w_i}|, |J_{w_i}| \leq 2q_0$, where $1 \leq q_0 \leq B$. Suppose all $F_{w_i}(x,y)$ are integer squares and $x, y \in [N, 2N)$. If $N > (B+1)^{cq_0^2}$, there is a primitive irreducible polynomial $C_0 \in \mathbb{Z}[X,Y]$, depending on both variables, such that*

$$C_0(x,y) = 0, \qquad \deg C_0 \leq 48q_0, \qquad H(C_0) \leq (B+1)^{cq_0}. \tag{T8.6}$$

In the function field $K_0 = \mathbb{Q}(C_0 = 0)$, the image of

$$V \longrightarrow K_0^\times / K_0^{\times 2}, \qquad w \longmapsto [F_w]$$

has dimension at most one. The constant is independent of t.

The proof is given in Section T8.9.5. Its key point is that every norm polynomial must share the same irreducible factor through (x,y): otherwise a resultant would give a nonzero polynomial of height $(B+1)^{O(q_0^2)}$ with integer root $x \geq N$. Primitivity of the auxiliary coefficients then forces the functional square-class image to have dimension at most one.

T8.4 Functional squares force integer fibres

The centres have size N, but the next step counts points in a box of side $O(B)$. The bridge is that a square function has even valuation at every point of the normalized curve. A simple zero belonging to a selected first-block offset must therefore be matched by a selected second-block offset.

Proposition T8.3 (Integer fibres in the offset grid)**.** *Under the hypotheses of Theorem T8.2, assume both physical projections of V are injective. Put $D = \deg C_0$, $d_X = \deg_X C_0$, $d_Y = \deg_Y C_0$. For an absolute constant c,*

$$\#\{(u,v) \in [-B,B]^2 \cap \mathbb{Z}^2 : C_0(-u,-v) = 0\} \geq \max(d_X, d_Y)(t - 1 - cD^2). \tag{T8.7}$$

A negative right side may be replaced by zero.

Proof. The kernel V_0 of the map to functional square classes has dimension at least $t-1$. The leading coefficients and discriminants of C_0 in each variable are nonzero polynomials, of total degree $O(D^2)$ in the remaining variable. Exclude the offsets at which one of these polynomials vanishes. Requiring the corresponding coordinates of a word to be zero leaves a subspace V_1 of dimension at least $t - 1 - cD^2$.

Let $\mathcal{U}_X$ be the union of first-block offsets used by V_1. Projection injectivity gives $|\mathcal{U}_X| \geq \dim V_1$. For $u \in \mathcal{U}_X$, select a word containing $X+u$. The noncritical fibre $X = -u$ has exactly d_Y distinct affine roots on the geometric normalization. At each, $X+u$ has a simple zero and no other first-block factor vanishes. The product for the selected word is a square function, so a selected factor $Y+v$ must vanish there. Every root is thus an integer $-v$ in the offset interval. Distinct u give disjoint fibres, proving at least $d_Y \dim V_1$ grid points. Interchanging X, Y proves the bound with d_X as well. The argument may be performed after extension to $\mathbb{C}$ and normalization of the components; geometric integrality is not required. □

T8.5 The point count at growing degree

We use the integer-point part of Binyamini–Cluckers–Novikov [2, Theorem 2, equation (27), in the specified author version].

Theorem T8.4 (Binyamini–Cluckers–Novikov)**.** *For an integral algebraic curve $\mathcal{C} \subset \mathbb{A}^n_K$ of degree D, where $n > 1$ and $K \subset \mathbb{C}$, there are a constant $c(n)$ and an absolute constant κ such that*

$$\#\{z \in \mathcal{C} \cap \mathbb{Z}^n : \max_i |z_i| \leq H\} \leq c(n) D^2 H^{1/D} (\log H)^\kappa, \qquad H > 2.$$

The constants are independent of the curve's coefficients and degree. Integrality over K is sufficient.

For curves over $\mathbb{Q}$, a corresponding affine estimate is also given by Binyamini, Cluckers, and Kato [15, Corollary 8, equation (23)]. The external theorem above asserts the existence of an absolute logarithmic exponent. The admissible value $\kappa = 17$ is supplied by the pinned Lean declaration `article_13_4` in the namespace `PaperP1.IntegralAffineCurvePointCount`. Only the existence of a fixed exponent is used below.

Corollary T8.5 (Sublinear counts in the offset box)**.** *Fix $0 \leq a < 1/2$ and $c_0 > 0$. Let $D_{\max} = D_{\max}(B)$ be any fixed degree envelope such that, for every $\delta > 0$, $D_{\max}(B) \leq B^{a+\delta}$ for all sufficiently large B. Uniformly for integral plane curves of degree $2 \leq D \leq D_{\max}(B)$,*

$$\#(\mathcal{C} \cap [-c_0 B, c_0 B]^2 \cap \mathbb{Z}^2) \ll_{c_0} (B^{1/2} + D_{\max}^2)(\log B)^\kappa = o(B).$$

Proof. Set $H = \max(3, c_0 B)$. The logarithmic derivative of $u^2 H^{1/u}$ is $(2u - \log H)/u^2$, so its maximum on $[2, M]$, $M \geq 2$, is attained at an endpoint. Put $M' = \max(M, \log H)$. Applying this observation on $[2, M']$, and using $H^{1/M'} \leq e$, gives, for every $2 \leq D \leq M$,

$$D^2 H^{1/D} \leq 4H^{1/2} + e(M')^2 \leq 4H^{1/2} + e\{M^2 + (\log H)^2\}.$$

The last logarithmic square is absorbed into $B^{1/2}$. Apply Theorem T8.4 with $n = 2$ and $M = D_{\max}(B)$. Choose a fixed $\delta > 0$ with $2(a + \delta) < 1$. Both $B^{1/2}(\log B)^\kappa$ and $B^{2(a+\delta)}(\log B)^\kappa$ are $o(B)$. The threshold is chosen after c_0, a, the envelope and the requested error, and before the curve and its coefficients. No lower growth of $D_{\max}$ is needed; bounded and polylogarithmic envelopes are included. □

This application uses the theorem's degree-uniform constant. A bound with an unspecified constant depending on D would not give the conclusion of Corollary T8.5.

T8.6 Nonaligned pairs have rank less than half a window

Theorem T8.6 (A uniform nonaligned rank bound)**.** *For every fixed critical band,*

$$\limsup_{N\to\infty} \sup_{\substack{x,y\in I_N \\ x,y \text{ nonaligned}}} \frac{d_{xy}}{B} \leq \frac{2}{3}\log 2 < \frac{1}{2}. \tag{T8.8}$$

Compatibility of the two events is not needed.

Proof. Suppose along a sequence that $d_{xy} \geq ((2/3)\log 2 + \delta)B$ for some fixed $\delta > 0$. Choose a fixed $a < 1/2$, sufficiently close to $1/2$, that $(2/3)\log(1/a) < (2/3)\log 2 + \delta/2$. Theorem T8.1 gives a space V of dimension at least $\delta B/2 - o(B)$ with injective physical projections and a basis whose half-cardinalities are at most $q_0 = B^{a+o(1)}$. The quantitative bridge to the common-curve theorem is

$$q_0^2 \log(B + 1) = B^{2a+o(1)} = o(B) = o(\log N). \tag{T8.9}$$

Thus Theorem T8.2 supplies an irreducible C_0 through (x, y) of degree $D \leq B^{a+o(1)}$ and logarithmic height $o(B)$, with functional image of dimension at most one.

If $D \geq 2$, Proposition T8.3 gives at least $\delta B/2 - o(B)$ integer points in the offset box, because $D^2 = o(B)$. Corollary T8.5 gives $o(B)$ points, a contradiction.

If $D = 1$, write $C_0 = \alpha X + \beta Y + \gamma$. Both α, β are nonzero. Since $H(C_0) = \exp(o(B)) = N^{o(1)}$ and $x, y \geq N$, they have opposite signs; otherwise γ could not cancel the two terms. The line therefore has the form $X = \lambda Y + \nu$ with $\lambda > 0$. The kernel of the functional-class map has positive dimension. Choose a nonzero word in it. Each selection is nonempty and even by projection injectivity. On the line, the distinct simple roots from the two blocks must pair, each pairing giving $\nu + u = \lambda v$. There are at least two pairings, hence

$$\lambda = \frac{u_1 - u_2}{v_1 - v_2} = \frac{p}{q}, \qquad (p, q) = 1, \qquad 1 \leq p, q \leq 2B.$$

Evaluating one pairing at the centres gives $qx - py = pv - qu$, of absolute value at most $(p + q)B$. This is an alignment as in (T8.3), contrary to the assumption. Both degree cases are impossible, and δ was arbitrary. □

The number $(2/3)\log 2$ is a sufficient upper threshold arising from the extraction and degree bounds. No matching rank construction is asserted for this constant.

T8.7 Aligned pairs and the exact contact distance

For an aligned pair the line is already known. Substitution into it costs a height linear in the support, so the extraction exponent may approach one without the quadratic resultant threshold.

Proposition T8.7 (Rank on a known alignment)**.** *Suppose $bx - ay = h$, where $(a, b) = 1$, $1 \le a, b \le B^3$, and $|h| \le (a + b)B$. For every fixed $\tau \in (0, 1)$,*

$$d_{xy} \le \frac{B}{\max(a, b)} + \frac{2}{3} B \log(1/\tau) + O_{\tau,C}(B/\sqrt{\log B}). \tag{T8.10}$$

For distinct compatible centres this implies, uniformly over aligned pairs, $d_{xy} \le B/2 + o_C(B)$. If such a pair has an alignment with $a = b = 1$, then $d_{xy} = o_C(B)$.

The proof in Section T8.9.6 counts the matched linear factors on that line. At height one, compatibility imposes $|x - y| \ge L$; distance exactly $L = B - 1$ leaves only one shared true position. A nonzero even relation needs at least two matches, so this contact contributes no macroscopic intersection rank. Thus the bound includes the minimum compatible separation exactly.

T8.8 Uniform triple nullity and its probabilistic cost

Corollary T8.8 (Triple rank and maximum joint probability)**.** *Every compatible triple T of distinct centres in the fixed critical band satisfies*

$$\rho_{\mathrm{rel}}(T) \le \frac{3}{2} L + o_C(L). \tag{T8.11}$$

If any of its pairs is nonaligned, the stronger bound $\rho_{\mathrm{rel}}(T) \le (1 + (2/3)\log 2)L + o_C(L)$ holds. In particular, with $q(T) = \mathbb{E} \prod_{x \in T} J_x$,

$$\max_{\substack{T \subset I_N \\ |T| = 3}} q(T) \le N^{-3/2 + o_C(1)}. \tag{T8.12}$$

Proof. For three spaces of dimensions at most L, the dimension identity and the intersection bound give

$$\begin{aligned} \rho_{\mathrm{rel}}(T) &= \dim(S_1 \cap S_2) + \dim((S_1 + S_2) \cap S_3) \\ &\le L + \dim(S_1 \cap S_2). \end{aligned}$$

Permuting the indices gives $\rho_{\mathrm{rel}}(T) \le L + \min_{i<j} d_{ij}$. Every pair in a compatible triple is compatible. Theorem T8.6 and Proposition T8.7, with $B = L + 1$, prove the stated nullity bounds.

For clarity, the passage to probabilities uses exact affine constraints. The prime signs induce a uniform character on each finite square-class space. The constraints for $J_x = 1$ have classes $\{x - 1, x\}$ and $\{x, x + j\}$, $1 \le j < L$, with prescribed values -1 and $+1$ respectively. These selections span E_x. Thus an admissible window has $p_x := \mathbb{E} J_x = 2^{-\dim S_x}$, and a compatible family has

$$q(T) = 2^{-\dim \sum_{x \in T} S_x} = 2^{\rho_{\mathrm{rel}}(T)} \prod_{x \in T} p_x. \tag{T8.13}$$

Every local relation is supported on defective positions, so Lemma T8.10 implies $L - O_C(B/\log B) \le \dim S_x \le L$. Consequently $p_x = 2^{-L + o_C(L)}$ for admissible windows; inadmissible windows contribute zero. Substitution in (T8.13) gives $q(T) \le 2^{-3L/2 + o_C(L)} = N^{-3/2 + o_C(1)}$ uniformly, which is (T8.12). □

The estimate (T8.12) is the arithmetic input needed by the reconstruction from three-window activities. Its proof uses the true windows and exact marginals, so no removal of locally exceptional windows is required at this stage.

T8.9 The short-basis and common-curve constructions

The notation is that of Section T8. This section supplies the complete local estimates, basis extraction, and algebraic construction used by the main rank argument. The final degree-uniform point count and the nonaligned contradiction are proved in this section.

T8.9.1 Local square relations and defective positions

For primes larger than the window length, a local square relation can use only positions having no odd valuation at any such prime. The next elementary argument bounds those positions without an admissibility assumption.

Lemma T8.9 (A local square threshold)**.** *Let $B \geq 2$ be an integer, let $I \subset [-B, B] \cap \mathbb{Z}$ have positive even cardinality m, and let x be a positive integer. If $x > (B+1)^{5m}$, then $\prod_{a \in I}(x+a)$ is not an integer square.*

Proof. The assumed threshold implies $x \geq 2B$. If the product were a square, Lemma T1.1, with $R = B$ and $d = m$, would give $x \leq (64B)^m$. For $B \geq 2$, we have $64B \leq (B+1)^5$. Thus $x \leq (B+1)^{5m}$, contrary to the hypothesis. □

Lemma T8.10 (Few defective positions in every critical window)**.** *The number of $n \in V_x$ for which*

$$\prod_{\substack{p > 10B \\ v_p(n) \text{ odd}}} p = 1$$

is $O_C(B/\log B)$, uniformly over all critical windows.

Proof. Apply the local assertion of Lemma T1.3 with $H = 10B$ to V_x. In the critical band its diameter is $O(B)$ and its values lie in $[N/2, 3N]$ for all sufficiently large N. The comparison constants are uniform in x. □

We use the elementary bound $\pi(t) \ll t/\log t$. For completeness, the product of the primes in $(u, 2u]$ divides $\binom{2u}{u} \leq 4^u$ for integer u; dyadic summation gives $\vartheta(t) = \sum_{p \leq t} \log p \ll t$. Splitting at $\sqrt{t}$ gives the stated bound for $\pi(t)$.

We will also need the elementary prime harmonic estimate

$$\sum_{B^a < p \leq 10B} \frac{1}{p} = \log(1/a) + O_a(1/\log B), \qquad 0 < a < 1 \text{ fixed.} \tag{T8.14}$$

Here is a derivation recording its error term. Chebyshev's bound above implies $\psi(t) = \sum_{p^j \leq t} \log p \ll t$. The identity $\log(n!) = \sum_{d \leq n} \Lambda(d)\lfloor n/d \rfloor$ then gives $\sum_{d \leq t} \Lambda(d)/d = \log t + O(1)$. The contribution of prime powers with exponent at least two is bounded, so $A(t) = \sum_{p \leq t} (\log p)/p = \log t + O(1)$. Partial summation yields

$$\sum_{p \leq t} \frac{1}{p} = \frac{A(t)}{\log t} + \int_2^t \frac{A(u)}{u(\log u)^2}\, du = \log\log t + c_0 + O(1/\log t).$$

The difference at $10B$ and B^a proves (T8.14).

T8.9.2 Large primes turn intersections into components

Construct a graph on the disjoint formal copies of V_x, V_y. A prime $p > 10B$ has at most one occurrence of odd valuation in either block. If it has two occurrences, join their vertices; if it has only one, mark that vertex as forced to zero. The corresponding binary equation says that the endpoints of an edge have equal coordinates, or that a marked coordinate is zero. A component is free when it contains no marked vertex.

Lemma T8.11 (Free components)**.** *Let c be the number of free components containing an edge. Then*

$$c = d_{xy} + O_C(B/\log B).$$

Each such component meets both blocks and is selected in its entirety or omitted in any solution to the large-prime equations.

Proof. The isolated unmarked vertices are precisely the defective positions. Their total number D_0 is $O_C(B/\log B)$ by Lemma T8.10. The kernel of the large-prime equations has dimension $c + D_0$. Imposing the remaining prime equations and the two block parities costs at most $\pi(10B) + 2$ dimensions. If $\mathcal{R}$ is the space of complete relations and $k_x = \dim \ker \phi_x$, then the map taking a relation to its common class gives an exact sequence with kernel $\ker \phi_x \oplus \ker \phi_y$. Hence

$$c + D_0 - \pi(10B) - 2 \le \dim \mathcal{R} = k_x + k_y + d_{xy} \le c + D_0.$$

A local relation can use only defective positions, so $k_x + k_y \le D_0$. These inequalities prove the assertion. □

T8.9.3 Proof of Theorem T8.1

Proof. A graph estimate. A multigraph with n vertices and $e/n \to \infty$ has a cycle of length $O(1+\log(n+1)/\log(e/n))$. Loops and parallel edges give cycles of length one and two. Otherwise delete vertices of degree less than $t = \lfloor e/(2n) \rfloor$. Not all vertices can disappear, since fewer than $nt \le e/2$ edges would then be deleted. The remaining graph has minimum degree at least t. If its girth exceeds $2r + 1$, a breadth first search contains at least $1 + t\sum_{j<r}(t-1)^j$ vertices. This bounds r and proves the estimate.

Consequently, when $n \ll B/\log B$ and $e \le B$, the binary cycle space contains at least $e - O(B/\sqrt{\log B})$ independent cycles of length $O(\log B/\log\log B)$. Indeed, while at least $B/\sqrt{\log B}$ edges remain, record a short cycle and delete just one of its edges. The recorded cycles are independent: the first deleted edge belongs to no later cycle, and the same argument applies successively to all deleted edges.

Medium-prime columns. Start with the free nontrivial components of Lemma T8.11, and discard those with more than $M = \lfloor \log B \rfloor$ vertices. This loses at most $2B/(M+1)$ components. The remaining number is $d_{xy} + O_C(B/\log B)$. Put $z = B^a$. To each retained component attach its parity column on the primes $z < p \le 10B$, and let w_j be that column's Hamming weight. Each nonzero entry consumes a divisible position, whence

$$\sum_j w_j \le \sum_{z<p\le 10B} \#\{n \in V_x \sqcup V_y : p \mid n\} \le 2B \sum_{z<p\le 10B} \frac{1}{p} + 2\pi(10B) \le 2B\log(1/a) + O_a(B/\log B).$$

Discard only columns of weight at least three. Their number is at most one third of this sum, so at least

$$e \ge d_{xy} - \frac{2}{3}B\log(1/a) - O_{a,C}(B/\log B)$$

columns remain. There is no factor M in this count: it bounds the number of discarded columns, not their physical size.

Use the primes $z < p \le 10B$ and one extra vertex ∞ as vertices of a multigraph. A column of weight two joins its two primes; a column of weight one joins its prime to ∞; a zero column is a loop at ∞. Its binary incidence matrix has the same kernel as the parity matrix, because the row at ∞ is the sum of the other rows. The graph has $n \le \pi(10B) + 1 \ll B/\log B$ vertices and $e \le B$ edges. The preceding graph argument therefore supplies a space $\mathcal{U}$ of dimension at least $e - O(B/\sqrt{\log B})$ with a basis of short cycles. Each basis vector uses $O(\log^2 B/\log\log B)$ physical positions.

Small primes and parity. All large and medium primes already cancel on $\mathcal{U}$. The primes at most z and the two block parities impose at most $s = \pi(z)+2 \ll_a B^a/\log B+2$ further equations on its cycle coordinates. Their kernel is $\mathcal{K}_a$; since $s = o(B/\sqrt{\log B})$, it has the dimension in (T8.4). A fundamental kernel basis for a matrix with at most s rows uses at most $s+1$ columns per vector: express each nonpivot column in terms of pivot columns. Thus each resulting vector uses at most $s+1$ short cycles, giving (T8.5).

Finally, the original free nontrivial components are disjoint and meet both blocks. Every nonzero selection of them contains a component carrying a large prime which occurs once in each block and in no other component. Its common square class is therefore nonzero. This also proves injectivity of both physical projections and of the common-class map. The imposed parity equations complete the proof. □

T8.9.4 Two relations produce a polynomial of controlled height

The short basis will have support $B^{a+o(1)}$. We therefore need an auxiliary polynomial whose degree and logarithmic height grow linearly with the support, and a simultaneous construction whose threshold is quadratic in that support. The following elementary factor bound will keep the height under control when common polynomial factors are removed.

Lemma T8.12 (Height of a bivariate factor)**.** *Let $P \in \mathbb{Z}[X,Y] \setminus \{0\}$, with partial degrees d_X, d_Y. Every primitive integer factor Q of P, and the corresponding integer quotient, satisfy*

$$H(Q) \le 2^{d_X+d_Y}\|P\|_2 \le 2^{d_X+d_Y}\sqrt{(d_X+1)(d_Y+1)}\,H(P).$$

Proof. The multivariate Mahler measure $M(P)$ is multiplicative, at most $\|P\|_2$, and at least one for a nonzero integer polynomial. Jensen's formula in one variable, followed by the leading coefficient in that variable, proves the last assertion by iteration. The one-variable coefficient bound also iterates: integrating first in X and then in Y gives

$$|[X^iY^j]P| \le \binom{d_X}{i}\binom{d_Y}{j}M(P).$$

Thus $H(Q) \le 2^{\deg_X Q+\deg_Y Q}M(Q)$; multiplicativity and $M(P/Q) \ge 1$ prove the result. The quotient has the same bound, with the integer content handled separately. □

Proposition T8.13 (Two complete witnesses)**.** *There is an absolute constant $c > 0$ with the following property. Let u, v be independent formal words, even in each block, with offsets in $[-B, B]$. For $w = u, v, u+v$, write $|I_w| = 2r_w$, $|J_w| = 2s_w$, and assume $\max(r_w, s_w) \le q$, where $1 \le q \le B$. Suppose that $F_u(x,y)$ and $F_v(x,y)$ are integer squares, with $x, y \in [N, 2N)$. If $N > (B+1)^{cq}$, there is a nonzero $R \in \mathbb{Z}[X,Y]$ such that*

$$R(x,y) = 0, \qquad \deg_X R, \deg_Y R \le 12q, \qquad \deg R \le 24q, \qquad H(R) \le (B+1)^{cq}. \tag{T8.15}$$

Moreover, R is the norm of a biquadratic auxiliary function whose four polynomial coefficients have no common nonconstant factor.

Proof. Set $F_1 = F_u$, $F_2 = F_v$, $G = \gcd(F_1, F_2)$, with monic gcd, and $F_3 = F_{u+v} = F_1F_2/G^2$. Formal independence makes

$$K = \mathbb{Q}(X, Y)(W_1, W_2), \qquad W_i^2 = F_i,$$

a degree-four extension. With $W_0 = 1$ and $W_3 = W_1W_2/G$, the four W_χ form a basis. At the given integer point use positive square roots. Then all $W_\chi(x, y)$ are integers: $F_3(x, y)$ is an integer which is a rational square, hence an integer square, and $W_3(x, y) = \sqrt{F_1(x, y)}\sqrt{F_2(x, y)}/G(x, y)$.

The linear system at infinity. Put $r_0 = s_0 = 0$, $k = 3q$, and consider

$$\mathcal{A} = \sum_{\chi=0}^{3} A_\chi(X, Y)W_\chi, \qquad \deg_X A_\chi \le k - r_\chi, \quad \deg_Y A_\chi \le k - s_\chi.$$

There are

$$n = \sum_{\chi=0}^{3}(k - r_\chi + 1)(k - s_\chi + 1) \ge 21q^2 + 18q + 4$$

unknown coefficients, and $n \le 4(3q + 1)^2$. Writing $\xi = 1/X$, $\eta = 1/Y$, the positive formal branches are

$$W_\chi = X^{r_\chi}Y^{s_\chi} \prod_{u\in I_\chi}(1 + u\xi)^{1/2} \prod_{v\in J_\chi}(1 + v\eta)^{1/2}.$$

Thus $f(\xi, \eta) = \xi^k\eta^k\mathcal{A}$ is a formal power series. Require its coefficients of total degree at most $2k$ to vanish. This gives at most

$$t = \frac{(2k + 1)(2k + 2)}{2} = 18q^2 + 9q + 1, \qquad n - t \ge 3q^2 + 9q + 3, \qquad \frac{t}{n - t} < 6$$

homogeneous linear equations.

All coefficients required from the binomial series have denominator dividing 4^{2k}. Cauchy's estimate on circles of radius $(2B)^{-1}$ bounds their numerators, after this common denominator is cleared, by $(B+1)^{c_0q}$ for an absolute c_0. An elementary pigeonhole form of Siegel's lemma yields a nonzero integer solution of height $(B + 1)^{c_1q}$. Indeed, if the integer matrix has height A, the images of $\{0, \dots, M\}^n$ take at most $(2nAM + 1)^t$ values. Choosing $M = \lceil(3nA)^7\rceil$ makes two images equal, because $t/(n - t) < 6$. Their difference is a nonzero kernel vector; $\log n = O(\log B)$ is absorbed in the stated height.

Primitive coefficients. Let D be the primitive polynomial gcd of the four A_χ, omitting zero coefficients, and write $A_\chi = DA'_\chi$. The factor bound of Lemma T8.12 keeps $H(A'_\chi) \le (B+1)^{O(q)}$. Write $a = \deg_X D$, $b = \deg_Y D$, $e = \deg D$. Then

$$f = \widetilde{D}f', \qquad \widetilde{D} = \xi^a\eta^b D(\xi^{-1}, \eta^{-1}), \qquad f' = \xi^{k-a}\eta^{k-b}\sum_\chi A'_\chi W_\chi.$$

These are power series, and $\operatorname{ord}\widetilde{D} = a + b - e$ for total order. Since order is additive in the power-series domain,

$$\operatorname{ord} f' \ge 2k + 1 - a - b + e = (k - a) + (k - b) + e + 1. \tag{T8.16}$$

Thus removing D preserves vanishing beyond the remaining pole weight. This is a formal division, never a division by the possibly zero value $D(x, y)$. Replace the auxiliary by its primitive version from now on.

Integer separation and the norm. On the bidisc $|\xi|, |\eta| \le (4B)^{-1}$, the analytic numerator is bounded by $(B + 1)^{O(q)}$. Cauchy's estimate and the order of vanishing give, in the comparable range $x, y \in [N, 2N)$,

$$|\mathcal{A}(x, y)| \le \frac{(B + 1)^{c_2q}}{N}. \tag{T8.17}$$

To see the power of N explicitly before removing D, the tail of the numerator is bounded by its coefficient bound times

$$\sum_{\ell \geq 2k+1} (\ell+1)(4B/N)^{\ell} \ll (k+1)(4B/N)^{2k+1}.$$

Multiplication by $x^k y^k \leq (2N)^{2k}$ gives (T8.17). After removing D, both pole weights decrease, and (T8.16) gives the same estimate with the new height bound.

The specialization $\mathcal{A}(x, y)$ is an integer, so it vanishes for the stated threshold. The auxiliary is nonzero in K by independence of the four characters. Its norm is therefore nonzero. An explicit polynomial expression is

$$\begin{aligned} U &= A_0^2 + A_1^2 F_1 - A_2^2 F_2 - A_3^2 F_3, \\ V &= A_0 A_1 - A_2 A_3 (F_2/G), \\ R &= U^2 - 4F_1 V^2. \end{aligned}$$

Here F_2/G is a polynomial. The partial degrees of U are at most $2k$, and those of V are at most $2k - r_1$, $2k - s_1$. For the second term in V, use $r_3 = r_1 + r_2 - \deg_X G$ and its Y analogue. It follows that the partial degrees of R are at most $4k = 12q$. The coefficient ℓ^1 norms of all factors in this formula are $(B+1)^{O(q)}$, which proves the height bound and completes (T8.15). □

T8.9.5 Proof of Theorem T8.2

Proof. For $i < j$, the half-cardinalities of $w_i, w_j, w_i + w_j$ are at most $q = \min(2q_0, B)$, since an even subset of $[-B, B] \cap \mathbb{Z}$ has size at most $2B$. Proposition T8.13 gives a nonzero polynomial R_{ij} of degree at most $48q_0$ and height $(B+1)^{O(q_0)}$, vanishing at (x, y) and arising from an auxiliary with primitive coefficients. Choose an irreducible primitive factor C_0 of R_{12} which vanishes at this point. The factor-height lemma preserves its height bound. If C_0 depended on only one variable, its integer root of size at least N would contradict Cauchy's root bound. Thus it depends on both.

If C_0 failed to divide some R_{ij}, the resultant $\mathrm{Res}_Y(C_0, R_{ij})$ would be a nonzero integer polynomial vanishing at x. Its height is $(B+1)^{O(q_0^2)}$: its Sylvester matrix has size $O(q_0)$, entries of coefficient norm $(B+1)^{O(q_0)}$, and $O(q_0)!$ determinant terms. This again contradicts the root bound under the stated threshold. Hence C_0 divides every R_{ij}.

No factor $X + u$ or $Y + v$ vanishes identically on this curve, since $C_0(x, y) = 0$ with both coordinates larger than B. The displayed square classes are therefore defined in K_0. If the classes of F_{w_i}, F_{w_j} were independent there, they would still generate a degree-four field extension. Specializing the polynomial norm formula gives zero, since $C_0 \mid R_{ij}$; in a field this forces the auxiliary to vanish. Independence of its four characters would force all four coefficient polynomials to be divisible by C_0, contrary to their primitivity. Thus the images of every pair of basis vectors are dependent. Over $\mathbb{F}_2$ this means that all nonzero images are equal, so the image has dimension at most one. □

T8.9.6 Proof of Proposition T8.7

Proof. Fix $\tau < 1$. Theorem T8.1 gives a space $V = \mathcal{K}_\tau$ of dimension

$$t \geq d_{xy} - \frac{2}{3} B \log(1/\tau) - O_{\tau, C}(B/\sqrt{\log B}),$$

with a basis of half-cardinalities at most $q_0 = B^{\tau + o(1)}$. The case $t < 2$ is covered by the displayed error term. For $t \geq 2$, apply Proposition T8.13 to each pair of basis vectors. Its threshold holds because $q_0 \log(B+1) = o(B) = o(\log N)$.

Put $\ell(X,Y) = bX - aY - h$. Every resulting norm polynomial R_{ij} is divisible by ℓ. Otherwise, with $d_Y = \deg_Y R_{ij}$, the nonzero integer polynomial

$$a^{d_Y} R_{ij}\left(X, \frac{bX-h}{a}\right)$$

would vanish at x. Its height is $(B+1)^{O(q_0)}$ since $\max(a,b,|h|) \le 2B^4$, contradicting the integer root bound for $x \ge N$. The primitive-norm argument in Theorem T8.2 now applies to this line and shows that the image of V in $\mathbb{Q}(\ell=0)^\times/\mathbb{Q}(\ell=0)^{\times 2}$ has dimension at most one.

Let V_0 be its kernel, of dimension at least $t-1$. On the line, $X+u$ and $Y+v$ have the same root precisely when

$$h + bu = av. \tag{T8.18}$$

Every word of V_0 must pair all its first-block roots with second-block roots. The integer solutions of (T8.18) advance by (a,b) in two intervals of B consecutive offsets, so their number r satisfies

$$r \le 1 + \frac{B-1}{\max(a,b)}.$$

Injectivity of the physical projection gives $t-1 \le \dim V_0 \le r$, proving (T8.10) after absorbing the bounded terms.

For compatible distinct centres with $\max(a,b) \ge 2$, this gives $d_{xy} \le B/2 + (2/3)B\log(1/\tau) + o_{\tau,C}(B)$. Choose τ sufficiently close to one after fixing a desired error margin, then let N tend to infinity. This order of choices proves the uniform $B/2 + o_C(B)$ bound.

For $a = b = 1$, compatibility implies $|x-y| \ge L$: if $x < y < x+L$, the first run makes $f(y-1) = f(y)$, whereas the second requires their signs to be opposite. At distance exactly $L = B-1$, the true windows share one position; at larger distance they are disjoint. Thus (T8.18) has at most one solution. A nonzero word in V_0 would need at least two pairings, because its selections are nonempty and even. Consequently $V_0 = 0$, $t \le 1$, and the dimension estimate followed by $\tau \uparrow 1$ gives $d_{xy} = o_C(B)$. □

T9 Separating compatible triples from cancellations

A count of hosts does not by itself control their probabilities: compatibility multiplies an independent weight by $2^{\rho_{\mathrm{rel}}}$. For incompatible triples, however, an exact centered identity removes this multiplier. The remaining cost is a pair coefficient followed by a dependent target. We estimate that cost before using signs.

Proposition T9.1 (The independent cost of one target after a pair)**.** *Define*

$$D_{2,1} = \sum_{\substack{e=\{x,y\}\subseteq\mathcal{I}_N \\ d_{xy}>0}} |c(e)| \sum_{\substack{z\in\mathcal{I}_N\setminus e \\ S_z\cap(S_x+S_y)\neq\{0\}}} p_z.$$

Then $D_{2,1} \ll_{C,\varepsilon} N^{-3/2+\varepsilon}$.

The proof in Section T9.1.1 partitions pairs into finitely many fixed rational channels, weak pairs, and an exceptional class. A fixed-source half-count handles the last two classes. For a good channel, exact compression turns the pair and its target into two logarithmic blocks; their total host count is $N^{1+\varepsilon}$ and their individual weights give an even smaller $N^{-5/3+\varepsilon}$ contribution. This bound uses independent weight only for the new target, while preserving the pair activity.

Theorem T9.2 (Absolute activity of incompatible triples)**.** *For every* $\varepsilon > 0$,

$$\sum_{\substack{T\subseteq\mathcal{I}_N,\ |T|=3 \\ q(T)=0}} |c(T)| \ll_{C,\varepsilon} N^{-3/2+\varepsilon}.$$

Both signs of the centered coefficient are included.

Proof. A decomposable triple has a singleton component, hence centered coefficient zero. For a connected incompatible triple, expansion of its product gives exactly

$$c(T) = -\prod_{i\in T} p_i - \sum_{\substack{e\subset T\\ |e|=2}} c(e) \prod_{i\in T\setminus e} p_i.$$

The first terms have total magnitude $O(N^{-3/2+\varepsilon'})$ by Corollary T6.2. A direct pair has $c(e) = 0$; for any other pair in a connected triple the remaining space meets its sum nontrivially. The pair terms are therefore bounded by three copies of $D_{2,1}$. Proposition T9.1 concludes. □

Proposition T9.3 (The third cumulant and the positive activity)**.** *Every compatible triple has* $c(T) \geq 0$*, and*

$$\kappa_{3,N} = 6K_3^+ + O_{C,\varepsilon}(N^{-3/2+\varepsilon}), \qquad A_3 = K_3^+ + O_{C,\varepsilon}(N^{-3/2+\varepsilon}). \tag{T9.1}$$

In the second equality the difference $A_3 - K_3^+$ *is nonnegative.*

Proof. Order the pair ranks of a compatible triple as $a \geq b \geq c$. Their relation spaces show $\rho_{\mathrm{rel}}(T) \geq a + b$, so

$$\frac{c(T)}{\prod_{x\in T} p_x} = 2^{\rho_{\mathrm{rel}}(T)} - 2^a - 2^b - 2^c + 2 \geq 0.$$

If $b = 0$, this is immediate. Otherwise use $(2^a - 1)(2^b - 1) \geq 2^c - 1$. The same formula includes the primitive case $a = b = c = 0$. Decomposable triples again have zero coefficient.

In (2.7), the repeated-index terms are bounded by $O(p_* A_2 + p_*^2 \mu_N) = O(N^{-5/3+\varepsilon'})$. Apply Theorem T9.2 to the other triples. This proves both formulas. □

For example, three mutually exclusive events have $c(T) = 2p_x p_y p_z > 0$ despite being incompatible. Thus it is the absolute estimate for incompatible triples, rather than an assumption on their sign, that permits the reduction to K_3^+.

T9.1 Pair partitions and the cost of one dependent target

The signed triple identities in Section T9 use one quantitative input: the independent target deficit following a pair. We first prove the finite partition needed for that estimate.

Lemma T9.4 (A finite partition of the pair activity)**.** *For each fixed* $\eta > 0$*, dependent pairs can be partitioned into good pairs from a finite fixed set of rational channels, a weak class, and an exceptional class, with*

$$A_{\mathrm{w}} \ll_{C,\eta,\varepsilon} N^{-1+\eta+\varepsilon}, \qquad A_{\mathrm{s}} \ll_{C,\eta,\varepsilon} N^{-1+\varepsilon}.$$

A good pair in a channel $a : b$*,* $b \geq 3$*, has*

$$|c(e)| \ll_{a,b,C} N^{-(2-1/b)+o(1)}$$

and an exact compression into a block of fixed logarithmic width. Moreover $A_2 \ll_{C,\varepsilon} N^{-2/3+\varepsilon}$*.*

Proof. Choose Q with $1/(Q+1) < \eta/2$, and retain the channels $3 \leq b \leq Q$, $b/2 < a < b$, $(a, b) = 1$. Lemma T1.9 bounds the rank in all larger aligned channels by $L/(Q+1) + o(L)$. Put these pairs, nonaligned pairs of rank at most a sufficiently small fixed fraction of ηL, and affinely incompatible pairs in the weak class. Their coefficients have magnitude at most $N^{-2+\eta+o(1)}$. Corollary T6.3 counts their dependent hosts by $N^{1+\varepsilon'}$, proving the weak bound.

Nonaligned pairs of macroscopic rank have total activity $O(N^{-1+\varepsilon'})$ by Lemma T1.8. The local defects contribute only bounded factors under (2.1). Pairs meeting the irregular set have the same bound: its marginal mass is $N^{-1+o(1)}$, by Proposition T2.1, and fixed-source propagation

(2.8) bounds their joint mass. The independent products have the same cost. At each retained fixed height, normalized defective pairs cost $N^{-3/2+1/b+o(1)}$, as in the proof of Lemma T3.5. Endpoint height-two pairs and non-strict height-one contacts cost $N^{-1+o(1)}$, by the same rank and endpoint estimates. These are the exceptional class. The remaining pairs have the private coordinates defining the good fixed-channel compression. Its rank is at most $L/b + O(1)$, giving the stated individual weight.

The finite channel sums are $O(N^{-2/3+\varepsilon'})$, by Lemma T3.5. Together with the other two classes, for a sufficiently small fixed η, this also proves the assertion about A_2. □

T9.1.1 Proof of Proposition T9.1

Proof. After any fixed pair, Theorem T6.1 with one target counts at most $N^{1/2+\varepsilon'}$ centres. Their total independent mass is at most $N^{-1/2+\varepsilon'}$. The weak and exceptional pair classes thus cost $N^{-3/2+\eta+O(\varepsilon')}$.

For a good pair in a fixed channel, multiply its two windows by the channel coefficients and merge identical values. An even selection in each original window is multiplied by a square, so its square class is unchanged. Merging cancels repeated occurrences and preserves total even parity. Consequently every nonzero class in $(S_x + S_y) \cap S_z$ is represented by a nonempty even selection in the normalized block and in the target. The normalized block has fixed logarithmic width. Corollary T6.3 counts these two centres by $N^{1+\varepsilon'}$. The offsets and their inverse normalization have $B^{O_{a,b}(1)}$ multiplicity. Each host carries weight at most $N^{-(2-1/b)-1+o(1)}$, so its contribution is

$$N^{1+\varepsilon'} N^{-(2-1/b)-1+o(1)} \ll N^{-5/3+2\varepsilon'}.$$

Sum over the finite channels, then choose η and ε' within the final margin. □

T10 The third factorial cumulant and rational triples

The quadratic reference removes the first two factorial cumulants. To identify the remaining third-order contribution, two tasks remain distinct: we must bound every compatible triple, and we must show that the positive rational examples survive in the signed cumulant. The incompatible activity estimate in Section T9 supplies the second step. The first step combines rational overlap geometry with the nonaligned pair bound proved above.

Recall the exact marginals $p_x = \mathbb{E}J_x$ and the relative nullity

$$\rho_{\mathrm{rel}}(T) = \sum_{x \in T} \dim S_x - \dim\Big(\sum_{x \in T} S_x\Big).$$

Write $d_{xy} = \dim(S_x \cap S_y)$ and $c(T) = \mathbb{E}\prod_{x \in T}(J_x - p_x)$. A triple is connected when its spaces have no nontrivial direct-sum partition. For compatible triples, $q(T) = 2^{\rho_{\mathrm{rel}}(T)} \prod_{x \in T} p_x$ and $0 \le c(T) \le q(T)$, by the identities established in Section T9.

The third factorial cumulant and its optimal exponent are stated in Theorem 3.3 of the article. We give the geometric and probabilistic proof here.

T10.1 The largest canonical overlap of three rational windows

Consider integer slopes $u < v < w$ with $(u, v, w) = 1$ and $w < 2u$. In the common parameter, the three windows are intervals of lengths $L/u, L/v, L/w$ on the lattices $u^{-1}\mathbb{Z}$, $v^{-1}\mathbb{Z}$, and $w^{-1}\mathbb{Z}$. Translation of a window moves its interval and leaves the lattice origin fixed. The densities of pairwise coincidence lattices are $(u, v), (u, w), (v, w)$; the density of the triple coincidence lattice is one.

Let I_u, I_v, I_w denote the three parameter intervals. Up to a bounded endpoint error, the number of dimensions lost when equal values are merged is at most

$$\begin{aligned}&(u,v)|I_u \cap I_v| + (u,w)|I_u \cap I_w| + (v,w)|I_v \cap I_w| - |I_u \cap I_v \cap I_w| \\ &\quad = (u,v)|I_u \cap I_v| + ((u,w)-1)|I_u \cap I_w| \\ &\qquad + ((v,w)-1)|I_v \cap I_w| + |I_w \cap (I_u \cup I_v)|.\end{aligned}$$

All coefficients in the last expression are nonnegative. Bounding its four lengths by $L/v, L/w, L/w, L/w$ gives

$$D(u,v,w)L + O(1), \qquad D(u,v,w) = \frac{(u,v)}{v} + \frac{(u,w)+(v,w)-1}{w}. \tag{T10.1}$$

Nested intervals attain the coefficient D. The endpoint error is absolute: each of the four lattice counts has a bounded rounding error, even when the integer slopes vary.

Lemma T10.1 (A sharp rational overlap inequality)**.** *For all integers $u < v < w < 2u$ with $(u,v,w) = 1$,*

$$D(u,v,w) \le \frac{19}{30}. \tag{T10.2}$$

Equality holds for $(u,v,w) = (10,12,15)$. This is an inequality for the canonical coincidence geometry; arithmetic relations among the distinct normalized values require a separate bound.

Proof. If the reduced height of every pair is at least five, each pairwise intersection density is at most $1/5$. Dropping the subtracted triple intersection in (T10.1) gives $D \le 3/5$. Otherwise a pair has reduced ratio $2:3$ or $3:4$. Reduced heights one and two are impossible for distinct slopes whose extreme ratio is strictly less than two.

For a pair $2:3$, write the three slopes before reordering as $(2d, 3d, c)$, with $(c,d) = 1$ and $3/2 < c/d < 4$. The pair itself contributes at most $L/3 + O(1)$ lost dimensions. In the third window, positions that can coincide with either old family have density at most

$$\frac{(c,2)+(c,3)-1}{c}.$$

This is at most $1/4$ except for the possible numerators $c = 9, 12$, after removing physical repetitions and endpoint ratios. For completeness, if $5 \le c \le 15$, direct evaluation of $(c,2)+(c,3)-1$ leaves just $9, 12$ above $c/4$; the numerator is at most four for $c \ge 16$. The smaller numerators $2, 3$ give physical repetitions, and $4, 6$ have no coprime denominator in the strict interior range apart from endpoint cases. The exceptional fractions and their values of (T10.1) are

c/d	(u,v,w)	$D(u,v,w)$
$9/4$	$(8,9,12)$	$11/18$
$9/5$	$(9,10,15)$	$17/30$
$12/5$	$(10,12,15)$	$19/30$
$12/7$	$(12,14,21)$	$4/7$.

Every other case has $D \le 1/3 + 1/4 = 7/12$.

For a pair $3:4$, write $(3d, 4d, c)$ with $(c,d) = 1$ and $2 < c/d < 6$. The new density is at most $((c,3)+(c,4)-1)/c$. It is at most $1/3$ except for $c = 8, 12$, after the same exclusions: evaluating the integers $c \le 17$ suffices, and the numerator is at most six for $c \ge 18$. The exceptional reduced fractions are $8/3$ and $12/5$, giving $(8,9,12)$ and $(12,15,20)$, of densities $11/18$ and $3/5$. All other cases have $D \le 1/4 + 1/3 = 7/12$. The largest value in these cases is $19/30$. □

T10.2 All triples above a strong aligned pair

A fixed rational pair can be compressed into its true normalized values even when some of these values lack private prime coordinates. Injectivity is not needed for the inclusion of event spaces. This is what allows the next estimate to cover all pairs in the chosen channels, including the arithmetically exceptional ones.

Proposition T10.2 (Triples meeting finitely many aligned channels)**.** *Fix a finite set $\mathcal{F}$ of reduced interior channels $a : b$, with $b \geq 3$ and $b/2 < a < b$, before N. Then*

$$\sum_{\substack{|T|=3,\ T \text{ connected},\ q(T)>0 \\ \text{some pair of } T \text{ is aligned in } \mathcal{F}}} q(T) \ll_{\mathcal{F},C,\varepsilon} N^{-41/30+\varepsilon}. \tag{T10.3}$$

The sum allows irregular admissible windows, bad normalization of the source pair, and contacts at distance exactly L.

The complete normalization argument is in Section T10.5.1. After exact merging, the residual kernel has dimension $o(L)$ uniformly even when the last rational coefficients have polynomial height in B. Distinct physical slopes then lose at most $19L/30 + O(1)$ dimensions through canonical coincidences, so each triple has probability at most $N^{-71/30+o(1)}$. Their $NB^{O(1)}$ hosts give $N^{-41/30+\varepsilon}$. The proof separately controls nonaligned targets, repeated physical slopes, and dyadic endpoint populations; none of these cases requires good normalization of the source pair.

Corollary T10.3 (Strong aligned pairs)**.** *For every fixed $\theta > 0$ and every $\varepsilon > 0$,*

$$\sum_{\substack{|T|=3,\ T \text{ connected},\ q(T)>0 \\ \exists \{x,y\} \subset T \text{ aligned: } d_{xy} \geq \theta B}} q(T) \ll_{C,\theta,\varepsilon} N^{-41/30+\varepsilon}. \tag{T10.4}$$

Proof. The uniform bound $d_{xy} \leq L/b + o(L)$ forces $b \leq 2/\theta$ for sufficiently large N. Thus only a finite collection of heights is relevant. Heights at least three satisfy Proposition T10.2. Height two forces the source pair to the dyadic endpoints: there are $B^{O(1)}$ pairs of weight at most $N^{-3/2+o(1)}$. Proposition T4.3 adds a third window with a loss $N^{\varepsilon'}$, which is harmless. At height one, compatible distinct windows have disjoint interiors and at most one exact unit, so their relative intersection is $o(L)$. They cannot reach the threshold θB. □

T10.3 A positive family that survives in the total cumulant

The equality case $(10, 12, 15)$ of the overlap inequality also supplies a population of actual compatible run starts. Its role is stronger than merely maximizing a geometric formula: its centered activity is positive, and there are linearly many admissible translations.

Proposition T10.4 (The rational lower bound)**.** *For all sufficiently large N in the critical band,*

$$K_3^+ := \sum_{\substack{|T|=3,\ T \text{ connected} \\ q(T)>0}} c(T) \gg_C N^{-41/30}. \tag{T10.5}$$

In particular $\kappa_{3,N} \geq c_C N^{-41/30}$ for a suitable $c_C > 0$.

Proof. Take integers t with $N \leq 10t < 12t < 15t < 2N$. There are $N/30 + O(1)$ choices. Multiplication of the three windows by $6, 5, 4$ respectively gives the true normalized sets

$$U_1 = 60t + 6\{-1, 0, \dots, L-1\}, \quad U_2 = 60t + 5\{-1, 0, \dots, L-1\}, \quad U_3 = 60t + 4\{-1, 0, \dots, L-1\}.$$

Choose $H = 10B$, larger than the diameter of their union. By (T1.1), only $N^{1/2+o(1)}$ integers in the relevant range have $K_H(n) = 1$. For each of the $O(B)$ possible normalized offsets there is at

most one parameter t per such integer. Hence, outside $N^{1/2+o(1)}$ choices of t, every distinct value in $U_1 \cup U_2 \cup U_3$ has a private prime above H. Its square-class map is then injective.

The three sets meet at $60t$. The sum of their even-selection spaces is consequently the even-selection space of their union. The pairwise intersection sizes are

$$1+\left\lfloor\frac{L-1}{6}\right\rfloor, \qquad 1+\left\lfloor\frac{L-1}{3}\right\rfloor, \qquad 1+\left\lfloor\frac{L-1}{5}\right\rfloor,$$

and the triple intersection has size $1+\lfloor(L-1)/15\rfloor$. Inclusion–exclusion therefore gives the exact relative nullity

$$\rho_L = \left\lfloor\frac{L-1}{6}\right\rfloor + \left\lfloor\frac{L-1}{3}\right\rfloor + \left\lfloor\frac{L-1}{5}\right\rfloor - \left\lfloor\frac{L-1}{15}\right\rfloor = \frac{19}{30}L + O(1). \tag{T10.6}$$

The normalized predecessors have offsets $-6, -5, -4$, each belonging to just one set. Assign sign -1 to a private prime of each of these three values and sign $+1$ to every other prime occurring in the windows and multipliers. The chosen primes are distinct and occur in no other normalized value. Each original predecessor then has sign -1 and every internal value has sign $+1$. This verifies affine compatibility explicitly. For all these parameters,

$$q(\{10t, 12t, 15t\}) = p^3 2^{\rho_L}. \tag{T10.7}$$

The three pair nullities are the first three floor terms in (T10.6). Since all marginals are p on this private-prime population, the centered identity is exact:

$$\begin{aligned} c(\{10t, 12t, 15t\}) = p^3\Big\{&2^{\rho_L} - 2^{\lfloor(L-1)/6\rfloor} - 2^{\lfloor(L-1)/3\rfloor} \\ &- 2^{\lfloor(L-1)/5\rfloor} + 2\Big\}. \end{aligned} \tag{T10.8}$$

The largest subtracted exponent is $L/3+O(1)$, whereas the first exponent is $19L/30+O(1)$. Thus $c(T) \geq q(T)/2 > 0$ for large N. This also ensures connectedness: a direct singleton split would make the centered coefficient vanish. All compatible connected triple coefficients are nonnegative, and the triples obtained from distinct t are different. Consequently

$$K_3^+ \gg N p^3 2^{19L/30} \gg_C N^{-41/30}.$$

Finally (T9.1) gives $\kappa_{3,N} = 6K_3^+ + O(N^{-3/2+\varepsilon'})$. Choose $\varepsilon' < 3/2 - 41/30 = 2/15$ to make the error negligible. □

When $N = 2^L$ and $L = 30m$, the contribution of the private-prime population in this proof to K_3^+ is $(1/120 + o(1))N^{-41/30}$, since $\rho_L = 19m - 2$. This coefficient belongs to the displayed subfamily; it is not a coefficient for the full cumulant.

T10.4 Partitioning every compatible triple

Proof of Theorem 3.3. Proposition T10.4 supplies the lower bound. For the upper bound, fix $\eta = 1/30$ and then a positive $\theta < \eta/200$ for which Proposition T7.1 applies. All these parameters are fixed before N. We bound K_3^+ by partitioning its compatible connected triples.

If $\rho_{\mathrm{rel}}(T) < \eta B$, the host bound of Corollary T6.2 and $p_x \leq 2p$ give

$$\sum q(T) \ll N^{3/2+\varepsilon'} N^{-3+\eta+o(1)} = N^{-22/15+\varepsilon'+o(1)}.$$

If $\rho_{\mathrm{rel}}(T) \geq \eta B$ and every proper pair has $d_{xy} < \theta B$, Proposition T7.1 gives total homogeneous mass $O(N^{-2+\theta+\varepsilon'})$, which bounds the compatible mass as well.

It remains to consider a pair with $d_{xy} \geq \theta B$. If any such pair is aligned, Corollary T10.3 gives $O(N^{-41/30+\varepsilon'})$. Otherwise choose a strong nonaligned pair. Theorem T8.6 gives

$$d_{xy} \leq \left(\frac{2}{3}\log 2 + o(1)\right) L.$$

For the third space,

$$\rho_{\mathrm{rel}}(T) = d_{xy} + \dim(S_z \cap (S_x + S_y)) \leq d_{xy} + L \leq \left(1 + \frac{2}{3}\log 2 + o(1)\right) L.$$

Since $(2/3)\log 2 < 1/2$, this is eventually at most $3L/2$. Proposition T7.2, with $\sigma = 3/2$, therefore bounds this sector by $O(N^{-3/2+\varepsilon'})$. This is where the unconditional nonaligned-pair rigidity enters the third-cumulant theorem.

Each exponent other than $41/30$ is strictly larger than $41/30$. Shrink the internal error margins to obtain $K_3^+ \ll N^{-41/30+\varepsilon}$. The identity (T9.1) transfers the estimate to $\kappa_{3,N}$, completing the upper bound on the full critical band. □

The gap $41/30 - 4/3 = 1/30$ explains why third-order activity is below the optimal quadratic remainder scale. This conclusion does not identify the optimal error for a reference that also retains cubic or quartic terms; such a reference requires its own reconstruction and lower-bound argument.

The rank threshold $\eta = 1/30$ and the weak-pair threshold $\theta < \eta/200$ are distinct from the accuracy parameters. Using the convention in Section 1.3, prescribe $\varepsilon_{\mathrm{out}} > 0$ and choose $0 < \tau < \min(\varepsilon_{\mathrm{out}}/4, 1/960)$. Allocate smaller internal losses so that each completed sector loses at most τ. The nonaligned target above a rational pair retains $1/120 - \tau > 0$, and the incompatible-triple error retains $2/15 - \tau > 0$. These budgets include offset choices and residual-rank losses before N is taken large. The latter error is therefore negligible in the positive rational lower bound as well.

T10.5 All triples above finitely many rational pair channels

The canonical overlap inequality and the positive $10 : 12 : 15$ witness are proved in Section T10. The remaining technical input is uniformity of the upper bound across every triple above a fixed finite set of aligned channels, including arithmetically exceptional source pairs. We prove that input here.

T10.5.1 Proof of Proposition T10.2

Proof. Fix one channel $a : b$ and orient the source pair $e = \{x, y\}$. Choosing $B^{O_{a,b}(1)}$ offsets, write $x = at - i_0$, $y = bt - j_0$ and $X = abt$. Merge repeated values in $bV_x \cup aV_y$ to obtain $U_X \subset X + [-DB, DB]$, where D depends only on the fixed channel. Use radius DB for alignment with this compressed source, and choose the large-prime cutoff above $(2D+1)B$. The old centre lies in a fixed expanded interval starting at N, while the new centre remains in I_N, as required by the expanded version of Lemma T1.10. Multiplying an even selection by a constant does not change its square class, so

$$S_x + S_y \subseteq \mathrm{span}\{v(n) : n \in U_X\}. \tag{T10.9}$$

This inclusion holds regardless of injectivity on U_X. The uniform aligned rank bound of Lemma T1.9 gives

$$q(e) \leq N^{-\kappa_b + o(1)}, \qquad \kappa_b = 2 - \frac{1}{b} \geq \frac{5}{3}. \tag{T10.10}$$

For fixed height the residual defect is uniform, so the same bound includes admissible local rank defects.

Targets not aligned with the compressed source. For $z \notin e$, put $\delta = \dim(S_z \cap (S_x + S_y))$. Connectedness of $T = e \cup \{z\}$ implies $\delta > 0$. The inclusion (T10.9) bounds δ by the asymmetric rank loss $\delta^*(X, z)$ of Lemma T1.10. That lemma permits an arbitrary true set U_X of this fixed logarithmic width; it does not require injectivity. It gives

$$\sum_{X,z \text{ nonaligned}} (2^{\delta^*(X,z)} - 1) \ll N^{31/24+\varepsilon'}.$$

Since $q(T) = q(e)p_z 2^\delta$, $p_z \le 2p$, and $2^\delta \le 2(2^{\delta^*} - 1)$, these targets have total mass at most

$$N^{-\kappa_b - 1 + 31/24 + \varepsilon' + o(1)} \ll N^{-11/8+\varepsilon}. \tag{T10.11}$$

The number of source-offset patterns is polynomial in B and is absorbed into the exponent margin.

The residual kernel for an aligned target. This sector includes every target aligned at the same enlarged radius DB. Enlarging that fixed radius changes only the constants in the polynomial offset and height bounds below. Write an aligned target as $z = (ct + h)/d$, with $(c, d) = 1$ and $c, d, |h| \le B^{O_{a,b}(1)}$, as in the proof of Lemma T3.6. Before this final normalization, U_X and V_z each have fixed logarithmic width. Choose $H > C_D B$ larger than both diameters and remove the prime rows at most H and those dividing the normalization multipliers. Each remaining prime row meets at most one distinct old value and at most one new value. This uses the fact that repeated old values were merged before the target was added.

After the common normalization, merge equal values once more. There are at most $3B$ distinct integers, all at height at least a constant multiple of N and in an interval of diameter $B^{O_{a,b}(1)}$. The remaining prime rows have weight at most two, and leave one variable per unpinned component. We give the uniform residual estimate, since the coefficients c, d need not be fixed.

For a fixed integer K, restrict to components of at most K vertices. The removed prime rows and a total-parity row form a code on these components of redundancy

$$s \le \pi(H) + O_{a,b}(\log B) + 1 = O_{a,b}(B/\log B).$$

A nonzero word in its kernel gives an even square product of distinct normalized integers. Lemma T1.1, applied with the polynomial-in-B diameter just specified, forces it to contain $\gg_{a,b} \log N/\log B$ physical vertices and hence $\gg_{a,b,K} \log N/\log B$ components. If c_K is the number of these components, the Hamming inequality with $t \asymp_{a,b,K} \log N/\log B$ gives

$$\sum_{j=0}^{t} \binom{c_K}{j} \le 2^s, \qquad c_K = O_{a,b,K}(B/\log B).$$

Here t is chosen below half the component minimum distance. Components larger than K number at most $3B/K$. Choosing K arbitrarily large before N therefore bounds the entire value kernel on the distinct normalized set by $o_{a,b}(L)$, uniformly over c, d, h.

The relevant map factors as

$$\bigoplus_{v \in T} E_v \longrightarrow \{\text{merged normalized coordinates}\} \longrightarrow \{\text{square classes}\}.$$

Its nullity is at most the nullity of merging plus the residual $o(L)$ just proved. Internal defects of the old windows are included in this composition and do not require any assumption of good normalization.

Counting the exact coincidences. In the common parameter the physical coefficients are ad, bd, c. Suppose first that they are distinct. Divide by their common gcd and order them as $u < v < w$. Unless the parameter belongs to a $B^{O_{a,b}(1)}$ endpoint population, dyadic geometry gives $w < 2u$. Indeed a rational ratio of polynomial height cannot cross two by $B^{O(1)}/N$, and equality forces the extreme windows to the two dyadic endpoints. The lattice calculation (T10.1) and Lemma T10.1 show that merging loses at most $19L/30 + O(1)$ dimensions. The residual estimate therefore gives

$$\dim(S_x + S_y + S_z) \geq \frac{71}{30}L - o(L), \qquad q(T) \leq N^{-71/30+o(1)}.$$

There are $NB^{O_{a,b}(1)}$ aligned triples of this form, including the choices of offsets and rational coefficients. Their total mass is $O(N^{-41/30+\varepsilon})$.

If two physical coefficients are equal, compatibility makes the interiors of the corresponding parameter intervals disjoint, since distinct compatible starts have separation at least L. Their combined coincidences with the third lattice are bounded by the length of that third interval: at most $L/a + O(1)$ or $L/b + O(1)$ units. As $a \geq 2$, the total merging nullity is at most $L/2 + O(1)$. Including the residual, this gives mass $O(N^{-3/2+\varepsilon})$. Endpoint populations have only $B^{O_{a,b}(1)}$ source and target choices; the bound $q(T) \leq q(e)$ from (T10.10) gives $O(N^{-5/3+\varepsilon})$. These estimates include exact contacts. Strict contacts are incompatible and do not enter the sum.

The least of the exponents obtained is $41/30$. Sum over the fixed finite family $\mathcal{F}$ to conclude. □

Homogeneous variant for separated triples. The same $O_{C,\mathcal{F},\varepsilon}(N^{-41/30+\varepsilon})$ bound holds for the sum of $q_0(T) = 2^{-\dim \sum_{x\in T} S_x}$ over connected triples $T \subset G_N$ with distinct, pairwise separated starts and containing a source pair in $\mathcal{A}_{\mathrm{raw}}$ in one of the fixed channels. Here q_0 is the homogeneous probability, even when the original affine prescriptions are incompatible.

To see this, repeat the proof above with q_0 in place of q. The source estimate (T10.10) is a rank estimate and holds for $q_0(e)$; the target marginal is exactly p on G_N. The nonaligned identity is $q_0(T) = q_0(e)p2^{\delta}$, giving the same $N^{-11/8+\varepsilon}$ bound. In the aligned sector all residual-kernel and merging estimates are phase-free. When two physical coefficients coincide, the imposed separation makes the corresponding interval interiors disjoint, so the $N^{-3/2+\varepsilon}$ estimate still holds without compatibility. The distinct-coefficient estimate is $N^{-41/30+\varepsilon}$, and the endpoint estimate follows from $q_0(T) \leq q_0(e)$ as before. Exact contacts at distance L remain allowed. This proves the stated homogeneous variant independently of the moment and reconstruction theorems.

T10.6 Positive Palm costs on all target windows

A good edge changes the law of a third start by introducing an additional intersection of square-class spaces. Its positive effect can be bounded by the same positive triple mass, even when the target is locally irregular.

Proposition T10.5 (Positive Palm cost of a fixed channel)**.** *Fix a reduced interior channel $a : b$, with $0 < a < b < 2a$, and an integer $D \geq 6b$. Choose an integer private-prime threshold with $6bB \leq H \leq DB$. Let $\mathcal{E}_{a,b}$ be its compatible dependent good edges, write $J_e = \prod_{x\in e} J_x$ and $q_e = \mathbb{P}(J_e = 1)$, and put $\beta_{a,b} = \sum_{e\in\mathcal{E}_{a,b}} q_e$. For a deterministic set of targets $G \subseteq I_N$, define*

$$M_{\mathrm{ex}}(G) = \sum_{e\in\mathcal{E}_{a,b}} \sum_{z\in G\setminus e} \big(q(e\cup\{z\}) - q_e p_z\big)_+,$$

$$M_{\mathrm{nom}}(G) = \sum_{e\in\mathcal{E}_{a,b}} q_e \sum_{z\in G\setminus e} \big(\mathbb{P}(J_z = 1 \mid J_e = 1) - p\big)_+.$$

For every fixed critical band and $\varepsilon > 0$, uniformly in G and in the stated range of H, for all sufficiently large N,

$$\begin{aligned} &\beta_{a,b} > 0, \qquad M_{\mathrm{ex}}(G) + M_{\mathrm{nom}}(G) \ll_{a,b,C,D,\varepsilon} N^{-41/30+\varepsilon}, \\ &\frac{M_{\mathrm{ex}}(G) + M_{\mathrm{nom}}(G)}{\beta_{a,b}} \ll_{a,b,C,D,\varepsilon} N^{-11/30-1/b+\varepsilon}. \end{aligned} \tag{T10.12}$$

No regularity condition is imposed on the target windows.

Proof. For $z \notin e$, put $d_e(z) = \dim(S_z \cap \sum_{x\in e} S_x)$. The finite affine rank bound gives

$$q(e \cup \{z\}) \le q_e p_z \, 2^{d_e(z)}.$$

The inequality includes incompatible joint prescriptions and zero marginals. Thus $p_z = 0$ forces $q(e \cup \{z\}) = 0$, and $d_e(z) = 0$ makes the exact positive excess vanish. A positive summand of M_{ex} therefore has $d_e(z) > 0$ and $q(e \cup \{z\}) > 0$. Compatibility forces all three pairwise distances to be at least L, including possible equality. The source is a dependent pair, so neither source singleton can split off as a direct component. Positive crossing prevents the target singleton from splitting off. The triple is connected. Each summand is at most $q(e \cup \{z\})$, and a triple has at most three choices of a marked source pair. Proposition T10.2, which includes irregular windows and exact contacts, gives

$$M_{\mathrm{ex}}(G) \le 3 \sum_{\substack{|T|=3,\ T \text{ connected},\ q(T)>0 \\ T \text{ contains a pair in the fixed channel}}} q(T) \ll N^{-41/30+\eta}$$

for every fixed $\eta > 0$.

All conditioned sources have $q_e > 0$. The positive-part inequality gives

$$\big(q(e \cup \{z\}) - q_e p\big)_+ \le \big(q(e \cup \{z\}) - q_e p_z\big)_+ + q_e (p_z - p)_+.$$

Using the marginal estimate (T2.3),

$$M_{\mathrm{nom}}(G) \le M_{\mathrm{ex}}(G) + \beta_{a,b} \sum_{z \in I_N} (p_z - p)_+ \ll N^{-41/30+\eta} + \beta_{a,b} N^{-1+\eta}.$$

Lemma T3.5 gives, in the stated cutoff range, $\beta_{a,b} \asymp_{a,b,C,D} LN^{-1+1/b} > 0$. Each unordered compatible edge has two ordered orientations, so its mass is half the ordered mass in that lemma. Monotonicity of the private-prime condition sandwiches the mass between the phases at $H = 6bB$ and $H = DB$, proving uniformity throughout the stated range. Since $b \ge 3$, the second raw term is $O(LN^{-2+1/b+\eta})$, smaller than the $41/30$ scale. Division by the positive phase bounds the normalized sum by a constant times

$$L^{-1} N^{-11/30-1/b+\eta} + N^{-1+\eta}.$$

The second exponent is smaller because $11/30 + 1/b \le 7/10 < 1$. Choose the internal margin within ε and enlarge the scale threshold to absorb fixed factors. All parameters precede N, L, H, G. □

For the two principal channels, one common choice $24B \le H \le DB$ with fixed $D \ge 24$ gives

$$\frac{M_{\mathrm{nom},2:3}(G)}{\beta_{2,3}} \ll_{C,D,\varepsilon} N^{-7/10+\varepsilon}, \tag{T10.13}$$

$$\frac{M_{\mathrm{nom},3:4}(G)}{\beta_{3,4}} \ll_{C,D,\varepsilon} N^{-37/60+\varepsilon}. \tag{T10.14}$$

These are bounds for positive increases in the target probabilities; no optimality or absolute-discrepancy statement is asserted.

T11 Uniform certificate tools

Fixed-order estimates leave a separate problem: summing orders that grow with N. The arithmetic certificates below solve it by counting congruences and graph components with explicit cardinality weights. Their definitions concern the actual values in the windows and apply before any affine prescription is imposed. In particular, the geometric counts also permit irregular windows; replacing p by $p_x \le 2p$ only changes a fixed weight. The final parameter choices for the optimal reconstruction are made in Section 5. The complete elementary proofs are collected in Section T11.3.

T11.1 Small kernels and moderate congruence certificates

For $Y \ge B$, write

$$K_Y(v) = \prod_{\substack{\ell > Y \text{ prime} \\ v_\ell(v) \text{ odd}}} \ell.$$

The following elementary form of the Rankin argument will be used with its exponent fixed.

Lemma T11.1 (Counting a small truncated squarefree kernel)**.** *For fixed $1/2 < \sigma < 1$, fixed $C_1 > 0$, and $T \ge 1$,*

$$\#\{v \le C_1 N : K_Y(v) \le T\} \ll_{\sigma, C_1} N^\sigma T^{1-\sigma} \exp(C_\sigma Y^{1-\sigma}). \tag{T11.1}$$

The complete proof is given in Section T11.3.1.

Fix

$$A_0 > 1, \quad 0 < \beta < 1, \quad \tfrac{1}{2} < \sigma < 1, \quad A_0(1-\sigma) < 1, \tag{T11.2}$$

and put $H = \log N$, $Y = H^{A_0}$, $T = N^\beta$ and $\delta = (1-\sigma)(1-\beta) > 0$. Mark a window if one of its values has $K_Y(v) \le T$. Each value lies in at most B windows, so the number M of marked centres satisfies

$$M \le N^{1-\delta+o(1)}. \tag{T11.3}$$

Here and below the $o(1)$ depends on the fixed parameters, but not on the size of a family subsequently summed over.

Choose two distinct values in an unmarked anchor window. Their kernels are coprime, because a prime exceeding $Y \ge B$ cannot divide their nonzero difference. Thus their product Q is squarefree and $T^2 < Q \le 9N^2$. If a relation selects these values, each prime of Q occurs to odd order in at least one other window. Assign it to one such occurrence. For each receiving block j, let D_j be the product of the assigned primes. The D_j are coprime and $\prod_j D_j = Q$. Call the resulting certificate *moderate* when every $D_j \le 6NB$. After specifying a certificate we may discard the relation itself: the congruences alone bound the possible centres.

Lemma T11.2 (Counting moderate certificates)**.** *Let $\mathcal{U}_r$ be the ordered lists of r distinct original windows having such a certificate, and set $m = \lceil 2\log(3N)/\log Y \rceil$. Then*

$$|\mathcal{U}_r| \le 36 r B^4 2^{r-1} N^r T^{-2} (rB)^m. \tag{T11.4}$$

For every fixed $u > 0$,

$$\sum_{r \ge 1} \frac{(up)^r}{r!} |\mathcal{U}_r| \le N^{-2\beta + 4/A_0 + o(1)}. \tag{T11.5}$$

The complete proof is given in Section T11.3.2.

T11.2 Saturation has small degree in both directions

Put a directed edge $x \to y$ when two values in window x have a kernel product possessing a divisor $D > 6NB$ all of whose primes divide values in window y. We impose no direct-sum condition on S_x, S_y. Let Γ be the underlying simple undirected graph.

Lemma T11.3 (Incoming and outgoing saturation degrees)**.** *Uniformly in the fixed vertex,*

$$\deg^+(x), \deg^-(x) \le N^{2/A_0+o(1)}, \qquad \Delta(\Gamma) \le N^{2/A_0+o(1)}. \tag{T11.6}$$

The complete proof is given in Section T11.3.3.

If a primitive family has no moderate certificate, every unmarked vertex has a neighbour in its induced saturation graph. To see this, choose a nonzero relation, select two of its values at that vertex and assign the primes to exterior occurrences. Some receiver must have $D_j > 6NB$. The choices made at different vertices need not agree. Their existence in one deterministic graph is sufficient.

Lemma T11.4 (A generating function with marked isolated vertices)**.** *Let a simple graph have n vertices, maximum degree Δ and M marked vertices. If $\mathcal{F}$ consists of the vertex sets whose induced isolated vertices are all marked, then, for $z \ge 0$,*

$$\sum_{D \in \mathcal{F}} z^{|D|} \le \exp\left\{ Mz + nz\big((1+z)^{\Delta} - 1\big) \right\}. \tag{T11.7}$$

More generally, for nonnegative vertex weights w_v and marked set $\mathcal{M}$,

$$\sum_{D \in \mathcal{F}} \prod_{v \in D} w_v \le \exp\left\{ \sum_{v \in \mathcal{M}} w_v + \sum_{v} w_v \left(\prod_{u \sim v} (1 + w_u) - 1 \right) \right\}.$$

The complete proof is given in Section T11.3.4.

T11.3 Proofs of the uniform certificate tools

All definitions and parameter conditions are those of Section T11. The counts are geometric, and all constants are uniform in the size of the family being summed.

T11.3.1 Proof of Lemma T11.1

Proof. Write uniquely $v = ksa^2$, where $k = K_Y(v)$ and s is the squarefree product of the primes at most Y occurring to odd order. The inequality $1 \le (C_1 N/(ksa^2))^{\sigma}$ on the range being counted gives an upper bound

$$(C_1 N)^{\sigma} \left(\sum_{k \le T} k^{-\sigma} \right) \left(\sum_{a \ge 1} a^{-2\sigma} \right) \prod_{\ell \le Y} (1 + \ell^{-\sigma}).$$

Enlarging the sum over k to all positive integers only increases it. The three factors are respectively $O_\sigma(T^{1-\sigma})$, $\zeta(2\sigma)$, and at most $\exp(\sum_{j \le Y} j^{-\sigma})$. This proves the bound. □

T11.3.2 Proof of Lemma T11.2

Proof. There are at most rNB^2 choices for the anchor and its two offsets. The product Q has at most m prime factors, and each has at most rB destinations. Once these assignments are fixed, the Chinese remainder theorem leaves at most $1 + N/D_j$ possibilities for centre j. At most two

of the D_j exceed N, since $Q \le 9N^2$. For those two modules, replacing $1+N/D_j$ by $2N/D_j$ costs at most $(6B)^2$. Therefore

$$\prod_{j\neq\text{anchor}} (1+N/D_j) \le 2^{r-1}(6B)^2 N^{r-1}T^{-2}.$$

This proves (T11.4); allowing collisions and incompatible assignments only enlarges the count.

After multiplication by $(up)^r/r!$, the remaining series is, up to a fixed factor,

$$N^{-2\beta}B^{m+4}\sum_{r\ge 0}\frac{z^r r^{m+1}}{r!}, \qquad z = 2uNp = O_u(1).$$

Since $r^{m+1} \le (m+1)!e^r$, this series is at most $N^{-2\beta}B^{m+4}(m+1)!\exp(ez)$. Finally

$$m = \left(\frac{2}{A_0}+o(1)\right)\frac{H}{\log H}, \qquad \log(B^m(m+1)!) = \left(\frac{4}{A_0}+o(1)\right)H.$$

This yields (T11.5). The factorial here divides the orderings of a set of distinct windows; it is not an assumption on the growth of a fixed-order arithmetic constant. □

T11.3.3 Proof of Lemma T11.3

Proof. For an outgoing edge, choose two anchor offsets in B^2 ways. Each of the at most m primes is either unused or assigned to one of B offsets in the receiver. A qualifying code has modulus $D > N$ and therefore determines at most one receiving centre in I_N. This gives $B^2(B+1)^m = N^{2/A_0+o(1)}$ possibilities.

For the incoming degree, fix the receiving window y and let Ω_y be the primes greater than Y dividing any of its values. Writing $R = \lceil B\log(3N)/\log Y\rceil$, we have $|\Omega_y| \le R$. A code chooses two anchor offsets, at most m primes from Ω_y, and one of the two offsets for each chosen prime. For a qualifying code their product exceeds N, so its congruences determine at most one anchor centre. Thus the number of possible incoming neighbours is at most

$$B^2\sum_{j=0}^{m}\binom{R}{j}2^j \le B^2(m+1)(2eR/m)^m = N^{2/A_0+o(1)}.$$

The displayed binomial bound holds for large N, since $m \le R/2$. Nonqualifying codes may be included in the number of codes; the one-centre conclusion is applied only to qualifying ones. No receiving offset needs to be chosen in this incoming count: the receiver has already been fixed. Symmetrisation adds the two degree bounds. □

T11.3.4 Proof of Lemma T11.4

Proof. Fix an admissible set D. Its isolated vertices are marked singletons. On the other vertices choose an inclusion-minimal edge cover of the induced graph. Every edge of this cover has an endpoint of degree one in the cover, since otherwise that edge could be removed. Its components are therefore nontrivial stars. Thus D has a partition into marked singletons and stars, each with one centre and a nonempty subset of its neighbours as leaves.

Write $w(A) = \prod_{v\in A} w_v$. Choose one such partition for each D. The product over all possible blocks of $1+w(A)$ contains a distinct term of weight $w(D)$ for this chosen partition: its union recovers D. Dropping disjointness only adds nonnegative terms. Since $1+u \le e^u$ for $u \ge 0$, the weighted sum is at most the exponential of the total block weight. The singleton contribution is $\sum_{v\in\mathcal{M}} w_v$. For a fixed star centre v, its leaf sum is exactly

$$w_v \sum_{\varnothing\neq A\subseteq N(v)}\prod_{u\in A} w_u = w_v\left(\prod_{u\sim v}(1+w_u)-1\right).$$

Counting the same two-vertex star from both centres is harmless. This proves the heterogeneous assertion. Set $w_v = z$ and use $\deg(v) \le \Delta$ to obtain (T11.7). There is no small-activity hypothesis. When $\Delta z < 2$, the elementary bound $e^u - 1 \le 2u/(2-u)$ also gives the convenient majorant $\exp\{Mz + 2n\Delta z^2/(2-\Delta z)\}$. □

T12 Arithmetic estimates for the exact reconstruction

We use the notation of Section 5 of the article. All original windows with $p_x > 0$ remain present. In this section, ρ means the effective relative nullity ρ_{rel}. The estimates culminate in separate bounds for each fixed connected order and for the entire remaining range of cardinalities. The article then assembles them using the positive stopped expansion.

T12.1 Two dense targets above a prescribed space

The independent counts do not control a large rank gain: a family of nullity d carries an additional factor 2^d. We now isolate the conditional geometric estimate that pays for this factor. Only two centers remain free. Every prime circuit meeting at least three physical blocks must then meet one of the fixed sources, which makes its coefficient enumerable by a divisor bound.

Lemma T12.1 (Conditional dense pair estimate)**.** *Fix a number s of source windows and an effective subspace W of their value space. Suppose $S_y \cap W = S_z \cap W = 0$, and write*

$$d = \dim W + \dim S_y + \dim S_z - \dim(W + S_y + S_z).$$

For every fixed $\eta, \varepsilon > 0$ there exists a fixed $\theta > 0$ such that, uniformly in the sources and W,

$$\sum_{\substack{y \neq z,\ y,z \text{ outside the sources} \\ d \ge \eta L \\ (y,z) \text{ nonaligned or } \dim(S_y \cap S_z) \le \theta L}} p_y p_z 2^d \ll_{s,C,\eta,\varepsilon} N^{-1+\varepsilon}. \tag{T12.1}$$

In particular the hypothesis on the pair is satisfied when S_y, S_z are direct. For a primitive family of fixed size $r \ge 3$, with a prescribed collection of $r-2$ sources having at most $N^{a+\varepsilon}$ possibilities, its dense homogeneous mass is at most

$$\sum_{\rho \ge \eta L} \Big(\prod_{i \in I} p_i\Big) 2^{\rho(I)} \ll N^{-(r-1)+a+\varepsilon}. \tag{T12.2}$$

The kernel defining d projects injectively onto each target, so its coordinate labels force many binary components in each target separately. After regularization and removal of the weak aligned components, let m_t count the remaining components meeting target t, and let $n_{1,t}$ count those using exactly one of its $B = L+1$ coordinates. Then $n_{1,t} \ge 2m_t - B$. Above $d = (1/2+\delta)L$, this forces a positive proportion of singleton occurrences in both targets, allowing the stronger Pell or degree-one curve count. The resulting bounds are $N^{1/2+\varepsilon'}$ target pairs in the lower stratum and $N^{\varepsilon'}$ in the upper stratum; multiplying by $p_y p_z 2^d$ gives $N^{-1+\varepsilon}$ after choosing the internal margins. Section T12.7.1 supplies the separate component counts, the coefficient enumeration from fixed sources, and the removal of small aligned curves, including effective local sections.

Lemma T12.2 (A dense compressed target)**.** *For fixed sources and a fixed channel of good pairs, the number of compressed centers X for which $\dim(S(U_X) \cap W) \ge \eta L$ is $O_{s,C,\eta,\varepsilon}(N^{\varepsilon})$.*

Proof. The compression has fixed width $O(B)$ and an injective even-selection map. In Proposition T4.2, take H larger than this width. Its regularization loses a fixed small fraction of ηL. A prime above H has at most one target occurrence, so all other occurrences in its circuit are source

occurrences. The circuit products are divisors of fixed source values, and the fixed-source CRT and split-product steps in that proposition are unchanged. The proof uses the width bound and the target projection rank, not a target space of dimension exactly L. The finitely many channel dilations and $B^{O(1)}$ offset patterns are absorbed in the margins. □

T12.2 The exact weights of primitive families

The primitive estimate below combines the dense-host construction of Section T7.1, the homogeneous envelope of Lemma 2.3, and the conditional half-count through Corollary T6.2. We first record a separate small-kernel estimate for related counting arguments; it is not used in the proof of Proposition T12.4. For $Y \ge H$, put

$$K_Y(n) = \prod_{\substack{\ell > Y \\ v_\ell(n) \text{ odd}}} \ell, \qquad a_Y(n) = \prod_{\ell > Y} \ell^{\lfloor v_\ell(n)/2 \rfloor}.$$

Then $n = s_Y(n) a_Y(n)^2 K_Y(n)$, where $s_Y(n)$ is Y-smooth.

Lemma T12.3 (Several small kernels in one window)**.** *Fix $A > 2$, $0 < b, u < 1$, and an integer m with $m(1-b) > 1$. For $Y = (\log N)^A$, the number of windows containing m distinct values with $K_Y(n) \le N^b$ and $a_Y(n)^2 \le N^u$ is*

$$\ll N^{\max(1-1/A,(1+u)/2)+\varepsilon}. \tag{T12.3}$$

Also, with the previously fixed cutoff $H = C_D B$, the set $\mathcal{X}_q$ of windows containing distinct n_1, n_2 with $K_H(n_1)K_H(n_2) \le N^q$ has cardinality $O(N^{q+\varepsilon})$, for fixed $q \ge 0$.

The complete proof is given in Section T12.7.2.

Proposition T12.4 (Homogeneous mass at fixed primitive arity)**.** *Fix an integer $r \ge 3$, a critical band, and $\varepsilon > 0$. Let $\mathcal{P}_{r,N}$ be the primitive r-element families of original windows in I_N. With the four weights of Lemma 2.3,*

$$\sum_{I \in \mathcal{P}_{r,N}} q_0(I) \ll_{r,C,\varepsilon} N^{-r/2+\varepsilon}, \qquad \sum_{I \in \mathcal{P}_{r,N}} w(I), \qquad \sum_{I \in \mathcal{P}_{r,N}} q(I), \qquad \sum_{I \in \mathcal{P}_{r,N}} |c(I)| \ll_{r,C,\varepsilon} N^{-r/2+\varepsilon}. \tag{T12.4}$$

For each fixed $\eta > 0$, the subfamily

$$\mathcal{P}_{r,N}(\eta) = \{I \in \mathcal{P}_{r,N} : \rho_{\mathrm{rel}}(I) \ge \eta L\}$$

has $O_{r,C,\eta,\varepsilon}(N^\varepsilon)$ members, and each of its four total weights is $O_{r,C,\eta,\varepsilon}(N^{1-r+\varepsilon})$. One scale threshold controls these conclusions and every deterministic subfamily. No regularity, separation, affine compatibility, or positivity of the original marginals is required. The arity is fixed before that threshold is chosen.

Proof. The dense-host construction in Section T7.1 proves the required count. This part of the construction is geometric: it counts centres before assigning probability weights, and applies to all true windows. Primitivity makes each relation-space projection injective, whence

$$\rho_{\mathrm{rel}}(I) \le \min_{x \in I} d_x \le L.$$

Lemma 2.3 therefore gives

$$q_0(I) \le (2p)^r 2^L \ll_{r,C} N^{1-r}.$$

Multiplying by the dense host count proves the dense homogeneous bound. The other three bounds follow pointwise from the same lemma.

For the full homogeneous sum, choose fixed small $\xi, \delta > 0$ and split at $\rho_{\mathrm{rel}} = \xi L$. Corollary T6.2 counts at most $O_{r,C,\delta}(N^{r/2+\delta})$ primitive hosts. On the lower part, $q_0(I) \ll_{r,C} N^{-r+\xi}$, so their contribution is $O(N^{-r/2+\xi+\delta})$. On the upper part, the dense bound just proved with parameter ξ gives $O(N^{1-r+\delta})$. Because $r \geq 3$, $1-r < -r/2$. Choose ξ, δ within the requested margin and take the maximum of the finitely many scale thresholds involved. Pointwise domination again gives the other three sums, and restriction to any deterministic subfamily preserves every estimate. □

On the regular set, $q_0(I) = p^r 2^{\rho_{\mathrm{rel}}(I)}$. For ordered lists of distinct starts, the same bounds acquire only the fixed factor $r!$. Estimates at growing cardinalities will instead use the independent certificate tail below.

The exact centered expansion is (2.9). The retained measure B_2 uses its terms of cardinalities zero and two; the stopping identity controls the remainder through a positive suffix.

T12.3 An independent extension estimate at all cardinalities

For a fixed number s of source windows and an effective subspace W of their value space, let

$$E_W(u) = \sum_{\substack{U \subset \mathcal{I}_N \setminus \{\text{sources}\} \\ W, (S_x)_{x \in U} \text{ not direct}}} u^{|U|} \prod_{x \in U} p_x, \qquad u > 0.$$

The weights here are products of marginal probabilities. They are not conditional probabilities of dependent extensions.

Lemma T12.5 (Uniform independent deficit)**.** *For fixed s, u, C and every $\varepsilon > 0$,*

$$E_W(u) \ll_{s,u,C,\varepsilon} N^{-1/2+\varepsilon}, \tag{T12.5}$$

uniformly in the sources and W. The sum includes all cardinalities.

Proof. Set $Y = B^2$ and

$$a(n) = \prod_{\substack{\ell > Y \\ \ell \mid n}} (1 + \ell^{-1/2}), \qquad A_x = \prod_{\substack{\ell > Y \\ \ell \mid n \text{ for some } n \in V_x}} (1 + \ell^{-1/2}).$$

Hölder applied to the B positions in a window gives

$$\sum_{N \leq x < 2N} A_x \leq \prod_{i=-1}^{L-1} \left(\sum_{N \leq x < 2N} a(x+i)^B \right)^{1/B}.$$

Write $a(n)^B = \sum_{d \mid n} h_B(d)$, where h_B is multiplicative, supported on squarefree integers with primes exceeding Y, and $h_B(\ell) = (1 + \ell^{-1/2})^B - 1 \ll B\ell^{-1/2}$. The coefficients are nonnegative, so

$$\sum_{n \leq 3N} a(n)^B \leq 3N \prod_{Y < \ell \leq 3N} \left(1 + \frac{h_B(\ell)}{\ell}\right) \leq 3N \exp\left(O\left(B \sum_{m > Y} m^{-3/2}\right)\right) \ll N.$$

Thus $\sum_x A_x \ll N$, with a constant independent of B. The small-prime contribution satisfies

$$A_{\leq Y} = \prod_{\ell \leq Y} (1 + \ell^{-1/2}) \leq \exp(O(B/\log B)) = N^{o(1)}.$$

Chebyshev's bound and partial summation suffice for this estimate.

Let P_0 contain the small primes and the prime factors of the fixed sources. Their total logarithmic size is $O_s(B\log N)$. Using the same $Y=B^2$, we have

$$\sum_{\substack{\ell\in P_0\\ \ell>Y}}\ell^{-1/2}\le\frac{\sum_{\ell\in P_0}\log\ell}{\sqrt{Y}\log Y}=O_s(B/\log B).$$

The primes up to Y have the same bound by partial summation. Consequently $A_0=\prod_{\ell\in P_0}(1+\ell^{-1/2})=\exp(O_s(B/\log B))=N^{o(1)}$. If P_k also includes the primes of an arbitrary list $x_1,\dots,x_k$, then

$$\prod_{\ell\in P_k}(1+\ell^{-1/2})\le A_0A_{\le Y}\prod_{j=1}^{k}A_{x_j}.$$

The elementary proof of the exceptional-target estimate in Proposition T1.15, before restricting the number of sources, gives

$$D(x_1,\dots,x_k)\ll B(s+k)+B\sqrt{N}\prod_{\ell\in P_k}(1+\ell^{-1/2}). \tag{T12.6}$$

Contacts account for the first term. Outside them, a selected target value whose odd prime factors all belong to P_k has the form av^2, with a squarefree and supported on P_k. Summing $\sqrt{N/a}$ over these a gives the second term, with B possible placements. This upper bound also includes local defects.

Put $w_x=up_x$, $w_*\ll N^{-1}$, $M_0=\sum_x w_x=O_{u,C}(1)$, and $M_1=\sum_x w_xA_x=O_{u,C}(1)$. Every non-direct extension of size h has an exceptional exposure in each of its $h!$ orders. Sum over the possible first exceptional position, allowing repeated centers for an upper bound. The first term of (T12.6) contributes at most

$$O(Bw_*)\sum_{h\ge1}\frac{1}{h!}\sum_{i=1}^{h}(s+i-1)M_0^{h-1}=O(B/N).$$

The second contributes at most

$$O(B\sqrt{N}w_*A_0A_{\le Y})\sum_{h\ge1}\frac{1}{h!}\sum_{i=1}^{h}M_1^{i-1}M_0^{h-i}=N^{-1/2+o(1)}.$$

The double sum is bounded by $e^{\max(1,M_0,M_1)}$. The varying number of sources has therefore been averaged explicitly; no constant from a fixed-order theorem has been used at a growing order. □

Lemma T12.6 (A fixed number of extensions of a pair)**.** *For every fixed $k\ge1$,*

$$\sum_{|e|=2}|c(e)|\sum_{\substack{U\cap e=\varnothing,\ |U|=k\\ S_e,(S_z)_{z\in U}\ \text{not direct}}}\prod_{z\in U}p_z\ll_{k,C,\varepsilon}N^{-3/2+\varepsilon},\qquad S_e=\sum_{i\in e}S_i. \tag{T12.7}$$

Proof. If the new windows are not direct among themselves, choose a primitive subfamily. A primitive of fixed size $h\ge2$ has independent mass $O(N^{-h/2+\eta})$ by Theorem T6.1 with no prescribed sources. The largest such cost is $N^{-1+\eta}$. After multiplication by A_2 and summation of the remaining independent windows this is $N^{-5/3+O(\eta)}$.

Otherwise choose a minimal nonempty $A\subseteq U$ with $S_e\cap\sum_{a\in A}S_a\ne0$. If $|A|=1$, the total cost is the one-target deficit of Proposition T9.1. If $h=|A|\ge2$, a relation meeting S_e has a nonzero projection in each of the h target spaces, by minimality. Conditional on e, Theorem T6.1 gives $N^{h/2+\eta}$ possibilities. The product of the h target marginals costs $O_h(N^{-h})$; summing the pair weights therefore gives $N^{-2/3-h/2+O(\eta)}$. The remaining $k-h$ independent weights have bounded sum. There are only finitely many choices of h and roles because k is fixed. □

T12.4 Uniform certificate tails and the deficit of all pairs

We record the part of the prime-certificate argument that is uniform in the number of blocks. A block may be an original window or one of a fixed finite collection of geometric compressions of width $O(B)$. Its center ranges over $O(N)$ values and its independent weight is bounded by a fixed multiple of p. A relation of full support selects a nonzero even vector in every block.

Lemma T12.7 (Independent full-support tails)**.** *Fix a critical band, fixed bounds on the block widths and centre ranges, a weight $v > 0$, and an exponent $a < 2$. There exist a fixed integer $s \geq 3$ and a scale threshold such that, for every N above that threshold,*

$$\sum_{\substack{|F|\geq s \\ F \text{ has a full-support relation}}} \prod_{x\in F} (vp_x) \leq N^{-a}. \tag{T12.8}$$

The same conclusion holds with one block of any of the fixed compressed types, weighted by vp, and with the remaining factors vp_x. The cutoff and threshold depend only on the fixed parameters, not on N, L or the chosen blocks. In particular the cutoff is fixed before any later exponent margin.

Proof. Here are the parameter calculations in the certificate argument of Lemmas T11.2–T11.4. Choose $Y = (\log N)^{A_0}$ and a kernel threshold $T = N^b$, with $A_0 > 1$ and $0 < b < 1$. A block containing a value whose odd kernel above Y is at most T is marked. If $1/2 < \sigma < 1$ and $A_0(1-\sigma) < 1$, the Rankin count gives $N^{1-\delta+o(1)}$ marked centers, where $\delta = (1-b)(1-\sigma) > 0$. Every unmarked block in a full-support relation has two selected values of kernel exceeding T. Their kernels are coprime, since Y exceeds the fixed block diameter.

For a certificate in which every receiving modulus is at most $O(NB)$, the CRT and the exponential sum over cardinalities give independent mass

$$N^{-2b+4/A_0+o(1)}.$$

Here $m = \lceil 2\log(O(N))/\log Y \rceil$ and

$$\log(B^m(m+1)!) = (4/A_0 + o(1))\log N.$$

If no such certificate exists, each unmarked block has a saturated neighbor. Its directed degree is $N^{2/A_0+o(1)}$ in both directions, by Lemma T11.3. When $2/A_0 + \delta < 1$, the component generator of Lemma T11.4, tilted by $N^{\delta/2}$, stays bounded. The mass at cardinality at least s is consequently $N^{-s\delta/2+o(1)}$. Choose A_0 large and then b close enough to one that $2b - 4/A_0 > a$; choose an allowed σ and then s with $s\delta/2 > a$. All these choices are made before N.

The proof uses the actual prime divisors of selected values and a fixed bound on the block width. Changing one block to a fixed geometric type changes only constant factors and $B^{O(1)}$ offset choices. Its two anchors, receiving congruences, and degree bounds obey the same estimates. Thus the same parameter choices prove the compressed assertion.

The size cutoff can be chosen before any final ε. Indeed, fix $a < 2$, put $a' = (a+2)/2$ and $\varepsilon' = (2-a)/4$, and apply the established estimate with these fixed parameters. Since $-a' + \varepsilon' < -a$, it yields the pure bound N^{-a} above a suitable threshold, for both versions with the same fixed cutoff. This argument retains $a < 2$ and a fixed bound on the compressed block width. □

For $e = \{x, y\}$ put $E_e(u) = E_{S_e}(u)$ and $D_2(u) = \sum_{|e|=2} |c(e)| E_e(u)$.

Proposition T12.8 (The pair deficit over all cardinalities)**.** *For fixed $u > 0$ and every $\varepsilon > 0$,*

$$D_2(u) \ll_{u,C,\varepsilon} N^{-3/2+\varepsilon}. \tag{T12.9}$$

In particular it is smaller than the $N^{-4/3}$ scale required below.

Proof. First consider extensions with an internal primitive family of new windows. At each fixed size their independent mass is $O(N^{-1+\eta})$ by the half-exponent count; at all larger sizes apply Lemma T12.7 with any $a > 1$. Summing free complements costs at most $e^{u\mu}$. Their contribution to D_2 is $O(A_2 N^{-1+\eta}) = O(N^{-5/3+2\eta})$.

For a direct extension, choose its minimal subfamily A crossing S_e. Use Lemma T9.4. Weak and exceptional pairs have total activity $O(N^{-1+\eta})$ and $O(N^{-1+\eta})$, respectively; Lemma T12.5 bounds their complete contribution by $N^{-3/2+O(\eta)}$. There remain good pairs in a finite set of fixed channels. The cases $|A| < s$ are covered by Lemma T12.6. For $|A| \ge s$, compress the good pair geometrically. Minimality and directness of A give a full-support relation on the compressed block and all blocks of A. The independent weight assigned to the compressed center in Lemma T12.7 is p, whereas the required pair weight is at most $N^{-\kappa_b+o(1)}$, $\kappa_b = 2 - 1/b \ge 5/3$. The ratio is at most $N^{1-\kappa_b+o(1)}$. Choosing $a = 1$ in that lemma bounds this tail by $N^{-\kappa_b+o(1)} \le N^{-5/3+o(1)}$. Each good pair is represented by a fixed-width pattern and one center; the $B^{O(1)}$ patterns are absorbed in the margins. Free complements again cost a constant. Choose the finitely many internal margins after ε to conclude. □

T12.5 Every fixed connected order reduces to triples

Proposition T12.9 (Reduction of connected activities)**.** *For every fixed $r \ge 4$ and every $\varepsilon > 0$,*

$$\sum_{\substack{|I|=r \\ I \text{ connected}}} |c(I)| \ll_{r,C,\varepsilon} N^{\varepsilon} A_3 + N^{-4/3+\varepsilon}. \tag{T12.10}$$

The incompatible part, for every fixed $r \ge 4$, is $O_{r,C,\varepsilon}(N^{-3/2+\varepsilon})$.

Proof. We first bound compatible families by their positive probabilities. Let T_3^+ denote the sum of $q(T)$ over connected compatible triples. For a primitive triple, $q(T) \le 2c(T)$. For a nonprimitive compatible triple,

$$q(T) = c(T) + \sum_{e \subset T,\ |e|=2} c(e) p_{T \setminus e} + \prod_{i \in T} p_i.$$

The pair terms sum to the one-target deficit, and the independent term has mass $O(N^{-3/2+\eta})$ by the half-exponent count. Therefore

$$T_3^+ \le 2A_3 + O(N^{-3/2+\eta}). \tag{T12.11}$$

Families containing such a triple have mass $N^{\eta} T_3^+$ after positive propagation (2.8). Families containing a primitive of size at least four have mass $N^{-4/3+\eta}$, by Proposition T12.4 and the same propagation. All seed sizes here are bounded by the fixed r.

Consider a remaining compatible connected family. It contains a dependent pair e; otherwise a minimal dependent subfamily would be one of the seeds just treated. Put $U = S_e$ and choose a nonempty $A \subset I \setminus e$ minimal with $U \cap \sum_{a \in A} S_a \ne 0$. A single target gives a connected triple, so $h = |A| \ge 2$. If A contains a proper primitive P, then $U \cap \sum_{i \in P} S_i = 0$ by minimality. Thus $q(e \cup P) = q(e)q(P)$. Only $|P| = 2$ remains, since larger primitives were already treated. Compatible dependent pairs satisfy $q(e) \le 2c(e)$, so the total seed mass is at most $4A_2^2$. Propagate it to I.

Suppose next that A is direct. The nullity d of the compressed family $U, (S_a)_{a \in A}$ is at most L, and

$$q(e \cup A) = q(e)\Big(\prod_{a \in A} p_a\Big) 2^d.$$

For $d < \eta L$, the conditional half-exponent count of the h targets gives a cost at most $A_2 N^{-h/2+O(\eta)}$, hence $N^{-5/3+O(\eta)}$. For $d \ge \eta L$, fix $h - 2$ members of A. The last two targets

are each direct from the sum of U and these fixed targets, and are direct from one another. Otherwise there would be an internal dependence in A or an earlier crossing of U. Lemma T12.1 gives conditional weighted cost $N^{-1+\eta}$. The independent weights of the $h-2$ additional sources sum to a constant; multiplication by the pair mass gives $N^{-5/3+O(\eta)}$ again. If A is itself primitive of size at least three, it was already one of the excluded seeds.

It remains that $A = f = \{y, z\}$ is a dependent pair. By minimality, $U \cap S_y = U \cap S_z = 0$. Put

$$d_e = \dim(S_x \cap S_{x'}), \quad d_f = \dim(S_y \cap S_z), \quad \delta = \dim(U \cap S_f), \qquad e = \{x, x'\}.$$

If $\delta < \eta L$, then $q(e \cup f) = q(e)q(f)2^{\delta}$, and the total is $O(N^{\eta}A_2^2)$. For $\delta \geq \eta L$, choose θ from the dense-target lemma. If f is nonaligned or $d_f \leq \theta L$, that lemma gives $N^{-5/3+O(\eta)}$ after summing e. If f is a good pair of a fixed channel, fix e and use Lemma T12.2; there are only N^{η} choices for the compressed center of f. Linear algebra gives

$$d_f + \delta \leq L, \qquad q(e \cup f) \leq p_* q(e). \tag{T12.12}$$

Indeed the projection of $U \cap (S_y + S_z)$ modulo S_y is injective, and its image has dimension at most $\dim S_z - d_f$. This case also costs $N^{-5/3+O(\eta)}$.

Every aligned pair with rank greater than θL has bounded channel height, by Lemma T1.9. Outside the good pairs at those finitely many heights, the normalization estimates in Theorem 3.1 and Lemma T3.5 give at most $N^{1/2+\eta}$ exceptional aligned pairs. The endpoint pairs of height two and the pairs meeting an irregular window satisfy the same bound; height one has rank $o(L)$ when compatible. Call this exceptional collection $\mathcal{B}$. If only f belongs to $\mathcal{B}$, exchange e and f and use one of the preceding cases. This exchange is valid because a target in e intersecting S_f would form a connected triple, already removed. If both belong to $\mathcal{B}$, there are at most $N^{1+2\eta}$ hosts. The symmetric forms of (T12.12) give

$$d_e + d_f + \delta \leq L + \min(d_e, d_f) \leq \tfrac{3}{2}L + o(L),$$

where the final inequality uses Theorems T8.6 and T8.7. The joint probability is at most $N^{-5/2+o(1)}$, so this contribution is $N^{-3/2+O(\eta)}$. These cases exhaust the possible minimal sets A. Positive propagation adds the remaining windows, proving the compatible part of (T12.10).

For incompatible families we use induction on r, starting with the triple estimate. The exact identity is

$$c(I) = -\prod_{i \in I} p_i - \sum_{\substack{J \subsetneq I \\ |J| \geq 2}} c(J) \prod_{i \in I \setminus J} p_i. \tag{T12.13}$$

The independent term over connected r-tuples is $O(N^{-r/2+\eta})$. The pair terms cost $O(N^{-3/2+\eta})$ by Lemma T12.6. For a fixed proper J of size at least three, connectedness of I requires a non-direct extension of its total space; its independent cost is $N^{-1/2+\eta}$ by Lemma T12.5. Canonical factorization of $c(J)$ has no singleton factors. By induction a factor of size at least three costs $N^{\eta}A_3 + N^{-4/3+\eta}$; a decomposition consisting only of pairs has at least two of them and costs $O(A_2^2)$. There are finitely many partitions at each fixed size. Multiplying by the independent extension cost proves (T12.10) at order r. Inserting the arithmetic bound $A_3 \ll N^{-41/30+\eta}$ from Corollary 3.4 into the same induction gives the stated $N^{-3/2+\varepsilon}$ incompatible activity. No law approximation or lower bound has been used in this induction. □

T12.6 A uniform tail after 256 blocks

The preceding proposition has constants depending on its fixed order. We now use the certificate generator to handle all larger orders in the stopped expansion. The number 256 is a sufficient fixed cutoff, not an optimal threshold.

Lemma T12.10 (Certificate budgets with several anchors)**.** *Choose an even integer $t \geq 2$, $A_0 > 1$, $0 < b < 1$, and $1/2 < \sigma < 1$ with $A_0(1-\sigma) < 1$. Put*

$$\delta = (1-b)(1-\sigma), \qquad \theta = t/A_0.$$

Suppose each unmarked block in a family has a relation selecting at least t values in that block; the relation may depend on the block. Mark a block when some value has odd kernel above $Y = (\log N)^{A_0}$ at most N^b. Under any fixed independent weight per block, the moderate certificates have total mass

$$N^{-tb+2t/A_0+o(1)}. \tag{T12.14}$$

If $\theta + \delta < 1$, the saturated families of size at least s have total independent mass

$$N^{-s\delta/2+o(1)}. \tag{T12.15}$$

Proof. This is the calculation in Lemma T12.7 with t anchors. Their pairwise coprime kernels have product exceeding N^{tb} and at most $m_t = \lceil t \log(O(N))/\log Y \rceil$ prime factors. The CRT bound and exponential cardinality sum have combinatorial cost $B^{m_t}(m_t+1)! = N^{2t/A_0+o(1)}$, proving (T12.14). The saturated graph has degree $N^{\theta+o(1)}$ in both directions; the number of marked vertices is $N^{1-\delta+o(1)}$. After tilting by $N^{\delta/2}$, the three terms of the component generator of Lemma T11.4 are

$$N^{-\delta/2+o(1)}, \quad N^{-1+\theta+\delta/2+o(1)}, \quad N^{-1+\theta+\delta+o(1)}.$$

They remain bounded, and in fact tend to zero, under the stated condition. Removing the tilt proves (T12.15). The inequality $p_x \leq 2p$ is absorbed by doubling the fixed weight. The arguments concern true prime factors and permit contacts and irregular source windows. □

For a family D, let $\mathcal{K}(D)$ be the kernel of the sum map $\bigoplus_{x\in D} S_x \to \sum_{x\in D} S_x$, and set

$$d_x(D) = \dim \operatorname{proj}_{S_x} \mathcal{K}(D) = \rho(D) - \rho(D \setminus \{x\}).$$

Call D a 10-core when each $d_x(D) \geq 9$. The effective projection admits a linear section in the formal even coordinates. Setting nine pivot coordinates to one gives a word of weight at least nine; even parity makes its weight at least ten. Thus the preceding budgets apply with $t = 10$. For a compatible family Fourier expansion gives exactly

$$\frac{c(D)}{q(D)} = \mathbb{P}_{v \text{ uniform in } \mathcal{K}(D)}(\operatorname{proj}_{S_x} v \neq 0 \text{ for every } x \in D) \geq 1 - |D|2^{-9}. \tag{T12.16}$$

In particular $q(D) \leq 2c(D)$ for $|D| < 256$.

Proposition T12.11 (Uniform connected and primitive tails)**.** *For fixed $w > 0$, every $\varepsilon > 0$, and every $K_N = o(L)$,*

$$\sum_{\substack{I \text{ connected, } 256<|I|\leq K_N \\ \rho_*(I)\leq 2L}} w^{|I|}|c(I)| \ll N^{-4/3+\varepsilon}, \tag{T12.17}$$

$$\sum_{\substack{P \text{ primitive} \\ |P|>256}} w^{|P|}|c(P)| \ll N^{-4/3+\varepsilon}. \tag{T12.18}$$

The second sum includes all cardinalities. Here $\rho_*(I) = \max\{\rho(A) : A \subseteq I,\ q(A) > 0\}$.

Proof. For compatible 10-cores of nullity at most $2L$, choose

$$t = 10, \quad A_0 = 12, \quad b = 31/60, \quad \sigma = 23/25, \quad \delta = 29/750, \quad \theta = 5/6.$$

The condition $\theta + \delta < 1$ is strict. Multiplying the independent budgets by $2^{2L} = O_C(N^2)$ gives, at sizes at least 256,

$$\sum_D w^{|D|} q(D) \ll N^{-3/2+o(1)} + N^{-1106/375+o(1)}. \tag{T12.19}$$

Indeed $2 - 10(31/60) + 20/12 = -3/2$ and $2 - 256(29/1500) = -1106/375$.

If $\rho_*(I) \le L/60$, the centered expansion gives

$$|c(I)| \le 2^{|I|}\Big(\prod_{x\in I} p_x\Big)2^{L/60}.$$

A nonzero centered coefficient has a full-support Fourier relation, so two anchors are available in every block. Use

$$t = 2, \quad A_0 = 12, \quad b = 17/20, \quad \sigma = 23/25, \quad \delta = 3/250, \quad \theta = 1/6.$$

Starting at cardinality 256, the two costs after the rank factor are

$$N^{-27/20+o(1)} + N^{-2279/1500+o(1)}. \tag{T12.20}$$

The four costs in (T12.19) and (T12.20) leave the exact margins

$$\frac{3}{2} - \frac{4}{3} = \frac{1}{6}, \quad \frac{1106}{375} - \frac{4}{3} = \frac{202}{125},$$
$$\frac{27}{20} - \frac{4}{3} = \frac{1}{60}, \quad \frac{2279}{1500} - \frac{4}{3} = \frac{93}{500}.$$

Thus the smallest margin is $1/60$. For the fixed certificate parameters and fixed weight w, every displayed $o(1)$ loss is less than $1/120$ for sufficiently large N; no certificate parameter is allowed to vary with N.

For the remaining connected I, choose deterministically a compatible $A \subseteq I$ attaining $\rho_*(I) > L/60$. Delete blocks of projection dimension at most eight until a 10-core D remains. Since $8K_N = o(L)$, its nullity exceeds $L/120$ for large N, so $|D| \ge 2$. With $r = |I|$ and $h = |D|$, the centered expansion and maximality of $\rho_*(I)$ imply

$$|c(I)| \le 2^r 2^{8(r-h)}\Big(\prod_{x\in I\setminus D} p_x\Big)q(D). \tag{T12.21}$$

The elements outside A are already in this product; there is no enumeration of deletion orders. All fixed powers of two can be absorbed in the independent weight of the complements.

If $h \ge 256$, sum the complements freely and apply (T12.19). If $3 \le h < 256$, then $q(D) \le 2c(D)$ by (T12.16). Proposition T12.9, the arithmetic triple bound, and canonical factorization show that the sum of $|c(D)|$ at each such fixed size is $O(N^{-1+\eta})$. Because D is proper in the connected I, its complement is a non-direct independent extension. Lemma T12.5 adds $N^{-1/2+\eta}$. If $h = 2$, then $D = e$ is a compatible dependent pair and $q(e) \le 2c(e)$; the complement is exactly the pair deficit of Proposition T12.8. These bounds prove (T12.17).

A primitive family has nullity at most L. If it is incompatible, $|c(P)| = \prod_P p_x$ and the independent two-anchor budgets suffice. If it is compatible with nullity at most $L/60$, use (T12.20). Otherwise each projection has dimension $\rho(P) > L/60$, so it is itself a 10-core for large N. Apply (T12.19) directly, without an upper cardinality cutoff. This proves (T12.18). □

T12.7 Conditional dense targets and small-kernel counts

We prove the conditional dense-pair estimate of Lemma T12.1 and then the separate auxiliary small-kernel estimate of Lemma T12.3. The first supplies a fixed-source input to the reconstruction; the second completes the auxiliary statement recorded above. All source counts, widths, and rank thresholds are fixed before N tends to infinity. Neither proof uses the finite-transfer theorem or the law approximation.

T12.7.1 Proof of Lemma T12.1

Proof. We first count target pairs, keeping the number of binary components in *each* target explicit. All component counts below include isolated active coordinates.

The two injective projections. Put

$$R = \ker\bigl(W \oplus S_y \oplus S_z \longrightarrow W + S_y + S_z\bigr), \qquad \dim R = d.$$

If an element of R has zero y-projection, its other two entries belong to $W \cap S_z = 0$. Hence the projection $R \to S_y$ is injective; the other projection is injective by $W \cap S_y = 0$. In particular $d \le \min(\dim S_y, \dim S_z) \le L$. Represent W by a linear section in the physical source coordinates, and each target space by a section in its formal even-selection space. These sections preserve both injective target projections and all actual prime-valuation equations, including on an irregular window.

Fix $0 < \delta < 1/2$, to be chosen inside the final exponent margin, and choose fixed positive ξ, θ with

$$\xi \le \tfrac{1}{4}\min(\eta, \delta), \qquad \theta \le \tfrac{1}{8}\min(\eta, \delta).$$

Apply Lemma T4.1 to obtain $R' \subseteq R$ with $d' := \dim R' \ge d - \xi L$. Choose its recovery parameter large enough to give this loss and to control circuits in the $s+2$ physical blocks. Every nonzero coordinate label on R' then has a fixed bounded multiplicity K.

For each prime $\ell > H \asymp B$, pair equal active labels in its row. All edges of the resulting binary graph join equal nonzero labels, so every component lies within one label class and has at most K vertices. The coordinate labels in either target span $(R')^*$, since its projection is injective. A component contributes only one label to this span, even if it meets that target several times. Consequently each target meets at least d' binary components. This is a separate lower bound for each target, not just for their union.

Enumerating coefficients from the fixed sources. After the equal labels have been paired, the remaining distinct labels in a prime row split into minimal circuits. Each such circuit uses at least three physical blocks: $\ell > H$ has at most one occurrence in any block. Since only two blocks are free, every circuit contains a source occurrence. The product of primes assigned to a circuit support therefore divides a fixed source value. The recovery bound gives a fixed bound on the number of circuit supports incident to any selected coordinate. For a fixed number of selected binary components, enumerate their positions, these circuit supports, one source occurrence on each support, and divisors of the corresponding source values. The cost is $B^{O(1)}N^{o(1)} = N^{o(1)}$. Small-prime parities cost $2^{O(\pi(H))} = N^{o(1)}$ as well. Thus the squarefree coefficients of their component equations have $N^{o(1)}$ choices and height N^{A_0} for a fixed A_0. The height bound per component is fixed once s, the recovery parameters and the width multiplier are fixed; it does not increase when more disjoint component equations are selected. The number of equations required by Lemma T1.7 is chosen afterwards. No subspace or full deletion pattern is enumerated.

Removing the small aligned curves. Two exact units between the targets determine a reduced alignment of height $O(B)$. Two distinct alignments of this height would force $y, z = O(B^3)$, impossible in the critical range. If there are fewer than two exact units, no removal is needed. Otherwise let $u \geq 2$ be the number of units on the unique alignment. Their even selections have dimension $u-1$ before passage to square classes, and a target has at most one local rank defect. Their common image therefore gives

$$u - 2 \leq \dim(S_y \cap S_z) \leq \theta L$$

in the weak aligned case. Keep at most one of these units and exclude from selection every binary component meeting an endpoint of another unit. At most $2(u-1) \leq 2\theta L + 2$ components are excluded. This deletes components from the selectable collection; it does not restrict R' again. The surviving positions contain no two exact units on an alignment, which excludes the small-curve alternatives in the proof of Lemma T1.7. No removal is needed for a nonaligned pair.

For $t = y, z$, write m_t for the number of selectable components meeting target t. The separate projection bounds now give

$$m_t \geq d' - 2\theta L - 2 \geq d - (\xi + 2\theta)L - 2.$$

Thus $m_t \geq \eta L/3$ for large N whenever $d \geq \eta L$; if $d > (1/2 + \delta)L$, then $m_t > (1/2 + \delta/3)L$ for both targets.

The first stratum: a positive rank proportion. A component type records its number of coordinates in each physical block. There are finitely many types, because the component size is at most K. If a type meeting both targets occurs sufficiently often, select the fixed number of its components required by the two-variable geometric count. With the small curves excluded above and coefficients of height N^{A_0}, the proof of Lemma T1.7 gives $N^{1/2+\varepsilon'}$ target pairs, including the $N^{o(1)}$ pattern and coefficient choices.

Otherwise only a bounded number of components meet both targets. Each target still has a positive proportion of L components avoiding the other target. In each target choose a univariate type occurring with positive frequency, and then the fixed number of its components needed by the one-variable part of Lemma T1.7, taking at least two. Their component equations form a one-variable system with fixed coefficients and at least two distinct selected offsets; different components have disjoint vertices. In the degree-one case, these are precisely the two linear factors for Lemma T1.2. The one-variable count gives $N^{\varepsilon'}$ target pairs after allocating the margin between the two centers and the coefficient choices. Every selected set has fixed size, so its even-selection square-class map is injective for large N by Lemma T1.1; irregularity of the entire window is irrelevant to this local use. We have proved

$$\#\{(y, z) : d \geq \eta L\} \ll N^{1/2+\varepsilon'}.$$

Both counts here and below impose the pair restriction in the lemma.

The second stratum: rank greater than half a window. Suppose $d > (1/2 + \delta)L$. Let $n_{1,t}$ be the number of selectable components using exactly one coordinate in target t. Every other component meeting that target uses at least two of its $B = L + 1$ coordinates. Hence, separately for $t = y, z$,

$$B \geq n_{1,t} + 2(m_t - n_{1,t}), \qquad n_{1,t} \geq 2m_t - B > \frac{2\delta}{3}L - 1.$$

If a mixed type with one coordinate in either target occurs sufficiently often, the degree-one case of the geometric count gives $N^{\varepsilon'}$ pairs directly. Otherwise only boundedly many of the $n_{1,t}$ components are mixed. The displayed inequality then provides two components for each target,

each using one coordinate of that target and no coordinate of the other. The resulting two distinct linear factors give the uniform two-value Pell count for each center. Thus

$$\#\{(y,z) : d > (1/2+\delta)L\} \ll N^{\varepsilon'}.$$

This explains the threshold $L/2$: above it, the coordinate budget forces singleton occurrences in each target.

Finally $p_y, p_z \ll N^{-1}$ and $2^L \asymp_C N$. The stratum $\eta L \le d \le (1/2+\delta)L$ has weighted cost $N^{1/2+\varepsilon'}N^{-2+1/2+\delta}$, and the stratum above it has cost $N^{\varepsilon'}N^{-2+1}$. Choose $\delta > 0$ and then $\varepsilon' > 0$ so that $\delta + \varepsilon' < \varepsilon$; all preceding choices are fixed before N. This proves (T12.1). In a primitive family, each target is direct from the prescribed $r-2$ sources separately, as well as from the other target; the full nullity is at most L. Multiplication by the $r-2$ source marginals and summation over the prescribed source collection gives (T12.2), after dividing the final exponent margin between the source count and the conditional estimate. □

T12.7.2 Proof of Lemma T12.3

Proof. For the first assertion fix the m offsets, at a polylogarithmic cost, and write $d_i = n_i/K_Y(n_i) = s_i a_i^2$. Since $d_i \gg N^{1-b}$ and $\gcd(d_i, d_j) \ll B$,

$$\operatorname{lcm}(d_1,\dots,d_m) \ge \frac{\prod_i d_i}{\prod_{i<j}\gcd(d_i,d_j)} \ge N^{m(1-b)-o(1)} > N.$$

The a_i are pairwise coprime. If $\prod_i a_i^2 > N$, multiply the square factors in a fixed order until their product first exceeds N. The resulting square D lies between N and N^{1+u}, so it has at most $N^{(1+u)/2}$ possible roots. Assign its complete prime powers to the m offsets; the number of assignments is $m^{\omega(D)} = N^{o(1)}$. The CRT leaves at most one center modulo $D > N$ in the interval.

If $\prod_i a_i^2 \le N$, start with this product and add the small prime factors from the least common multiple until the first crossing of N. Now $N < D \le NY$ and $D = sa^2$ with s Y-smooth. For fixed $\sigma > \max(1/2, 1-1/A)$, the number of such modules is at most

$$(NY)^\sigma \sum_{a\ge 1} a^{-2\sigma} \prod_{\ell \le Y} (1-\ell^{-\sigma})^{-1} = N^{\sigma+o(1)}.$$

Choose σ just above $1-1/A$. Assigning the complete prime powers again costs $N^{o(1)}$ and the CRT gives at most one center. This proves (T12.3).

For the second assertion, use $H = C_D B$ as in Lemma T6.4. That lemma gives the result for $q > 0$. For $q = 0$, use $\mathcal{X}_0 \subseteq \mathcal{X}_{\varepsilon/2}$ and apply the lemma with error margin $\varepsilon/2$. □

T13 Uniform independent support counts

T13.1 Full-support relations at all cardinalities

For $G \subseteq G_N$ and $r \ge 2$, let $\mathcal{H}_{r,N}(G)$ consist of the ordered lists of r distinct starts in G, with pairwise distances greater than L, for which there exist

$$s_i \in S_{x_i} \setminus \{0\}, \qquad s_1 + \cdots + s_r = 0.$$

The projections are nonzero in their effective square-class spaces. Minimality and affine compatibility are not required.

Proposition T13.1 (A uniform weighted count of full supports)**.** *For a fixed critical band, fixed $u > 0$ and $\varepsilon > 0$, uniformly in deterministic $G \subseteq G_N$ and integers $s \geq 2$,*

$$\sum_{r=s}^{|G|} \frac{u^r}{r!} \sum_{\boldsymbol{x} \in \mathcal{H}_{r,N}(G)} p^r \ll_{u,C,\varepsilon} N^{-\min\{s/2,2\}+\varepsilon}. \tag{T13.1}$$

The scale threshold and implicit constant are independent of s and G. Empty sums are zero.

Proof. Fix $0 < \eta < \min\{1, \varepsilon\}$. Lemma T12.7, applied with $v = u$ and $a = 2 - \eta$, gives a fixed cutoff. Increase it to $R \geq 4$, still independent of s, G. On G_N the exact marginals equal p, and each unordered family of r distinct starts has exactly $r!$ orderings. Thus all terms with $r \geq R$, and any positive subfamily of them, have total mass at most $N^{-2+\eta}$.

For each of the finitely many $2 \leq r < R$, Proposition T1.14, with all r effective projections nonzero, gives

$$|\mathcal{H}_{r,N}(G)| \ll_{r,C,\eta} N^{r/2+\eta}.$$

Since $p \ll_C N^{-1}$, the corresponding weighted layer is $O_{r,u,C,\eta}(N^{-r/2+\eta})$. Take the maximum of these finitely many thresholds and sum their constants before varying s or G. For $s = 2$, $s = 3$, and $4 \leq s < R$, the small layers are respectively $O(N^{-1+\eta})$, $O(N^{-3/2+\eta})$, and $O(N^{-2+\eta})$. For $s \geq R$ only a subfamily of the same tail remains. These bounds prove the assertion with a constant independent of s, G. □

The fixed cutoff may depend on the requested exponent margin. Accordingly the proposition gives the exponent $2 - \varepsilon$ in its saturated range, without asserting the endpoint N^{-2}.

T14 Poisson differences and reference coefficients

T14.1 Computing the correction coefficient

The coefficient in the total variation equivalent can be evaluated from four Poisson masses. We record the elementary Poisson–Charlier calculation, together with a bound for every order. These are standard identities for signed Poisson approximations. The orthogonality bound and the exact second-difference norm appear in Zacharovas and Hwang [10, (2.13), Remark 4.4]; see also [1] for the Poisson–Charlier expansion framework. We include the calculation to fix the normalization of Δ, the endpoint conventions, and the factor relating the ℓ^1 norm to total variation.

Proposition T14.1 (Norms of Poisson differences)**.** *For $\mu > 0$, put $\pi_\mu(k) = e^{-\mu}\mu^k/k!$ when $k \geq 0$ and $\pi_\mu(k) = 0$ when $k < 0$. Define*

$$u_\pm = \mu + \frac{1}{2} \pm \frac{\sqrt{4\mu+1}}{2}, \qquad a = \lfloor u_- \rfloor + 1, \qquad b = \lceil u_+ \rceil - 1.$$

Then

$$\|\Pi_\mu * \Delta^{*2}\|_1 = 2\{\pi_\mu(a-1) - \pi_\mu(a-2) + \pi_\mu(b-1) - \pi_\mu(b)\}. \tag{T14.1}$$

More generally, for every integer $r \geq 0$,

$$\|\Pi_\mu * \Delta^{*r}\|_1 \leq \min\left\{2^r, \frac{\sqrt{r!}}{\mu^{r/2}}\right\}. \tag{T14.2}$$

In particular,

$$\|\Pi_2 * \Delta^{*2}\|_1 = \frac{10}{3e^2}, \qquad \mathcal{C}(2) = \frac{5}{6e^2}. \tag{T14.3}$$

The proof, including the normalization of the Charlier polynomials and the telescoping at integer roots, is given in Section T14.2.

The function $\mathcal{C}$ is continuous and strictly positive on $(0,\infty)$. Indeed, coupling two Poisson variables by an independent increment gives $\|\Pi_\mu - \Pi_\nu\|_1 \le 2(1-e^{-|\mu-\nu|})$. Convolving with Δ^{*2} proves continuity; the mass at zero gives positivity, as in (7.6). Thus the finite formula remains continuous when one of its two roots crosses an integer.

For completeness, let $H_{\mu,\theta} = \mathcal{L}(\mathrm{Pois}(\mu-\theta) + 2\,\mathrm{Pois}(\theta/2))$, with independent summands and $0 \le \theta < \mu$. The expansion of its convolution exponential gives the parameter-uniform estimate

$$|d_{\mathrm{TV}}(H_{\mu,\theta}, \Pi_\mu) - \theta\mathcal{C}(\mu)| \le \theta^2 e^{2\theta}. \tag{T14.4}$$

To verify the constant, write $H_{\mu,\theta} = \Pi_\mu * \exp_*(\theta\Delta^{*2}/2)$. The ℓ^1 norm of the exponent is at most 2θ, so the remainder after its linear term has norm at most $2\theta^2 e^{2\theta}$. The reverse triangle inequality and $d_{\mathrm{TV}} = \|\cdot\|_1/2$ prove (T14.4). In particular, the factor $1/4$ in $\mathcal{C}$ consists of the factor $1/2$ in the quadratic correction and the factor $1/2$ in the total variation convention. Neither (T14.2) nor the remainder constant in (T14.4) is asserted to be optimal.

T14.2 Norms of Poisson differences

The following calculation proves Proposition T14.1. It fixes both the signed-difference convention and the factor of two between ℓ^1 distance and total variation.

Proof of Proposition T14.1. Write $v_k = (\Pi_\mu * \Delta^{*2})\{k\}$. Direct calculation gives

$$v_k = \pi_\mu(k) - 2\pi_\mu(k-1) + \pi_\mu(k-2) = \frac{\pi_\mu(k)}{\mu^2}\big((k-\mu)^2 - k\big).$$

The quadratic has roots u_-, u_+ and is strictly negative exactly at the integers $a,\dots,b$. These form a nonempty interval; integer roots contribute zero and are excluded by the definitions of a, b. Since $\sum_k v_k = 0$,

$$\begin{aligned}\sum_k |v_k| &= -2\sum_{k=a}^{b} v_k \\ &= 2\{\pi_\mu(a-1) - \pi_\mu(a-2) + \pi_\mu(b-1) - \pi_\mu(b)\},\end{aligned}$$

by telescoping. This proves (T14.1). For $\mu = 2$, the quadratic is $(k-1)(k-4)$: the negative masses are $v_2 = -e^{-2}$ and $v_3 = -2e^{-2}/3$. Their sum, multiplied by -2, gives (T14.3). The same calculation at $\mu = 1$ gives the value $3/e$ used in (7.13).

For the general bound, define the normalized Charlier polynomials

$$C_r(k;\mu) = \sum_{j=0}^{r} \binom{r}{j} (-1)^{r-j} \frac{(k)_j}{\mu^j}.$$

Then $(\Pi_\mu * \Delta^{*r})\{k\} = \pi_\mu(k) C_r(k;\mu)$. Their exponential generating function is

$$\sum_{r\ge 0} C_r(k;\mu)\frac{t^r}{r!} = e^{-t}(1+t/\mu)^k.$$

For $P \sim \mathrm{Pois}(\mu)$, the expectation of the product of the generating functions with variables s, t is

$$\mathbb{E}\Big[e^{-s-t}(1+s/\mu)^P (1+t/\mu)^P\Big] = e^{st/\mu}.$$

Coefficient comparison, justified in a neighborhood of $(s,t) = (0,0)$ by the Poisson exponential moments, yields

$$\mathbb{E}[C_r(P;\mu)C_\ell(P;\mu)] = \mathbf{1}_{\{r=\ell\}}\,\frac{r!}{\mu^r}.$$

Cauchy–Schwarz now gives the second bound in (T14.2). The first follows from $\|\Pi_\mu\|_1 = 1$, $\|\Delta\|_1 = 2$, and submultiplicativity of the convolution norm. □

T15 Fixed moments do not give a polynomial rate

This example isolates the logical role of the growing-order estimate. It is a family of probability laws, with no claimed multiplicative realization.

Proposition T15.1 (A perturbation invisible at each fixed order)**.** *For $0 \le a \le 1$ put*

$$Q_a = \Pi_1 + \frac{a}{2}\Pi_1 * \Delta^{*2}.$$

For every integer $m \ge 3$, define

$$\epsilon_{m,a} = \frac{(1-a/2)e^{-1}}{2m!}, \qquad P_{m,a} = Q_a + \epsilon_{m,a}\Delta^{*m}.$$

Then $P_{m,a}$ is a probability measure satisfying $\frac{1}{2}Q_a \le P_{m,a} \le \frac{3}{2}Q_a$ pointwise. Its factorial moments agree exactly with those of Q_a at every order other than m, and

$$d_{\mathrm{TV}}(P_{m,a}, Q_a) = \frac{(1-a/2)2^{m-2}}{e\,m!}.$$

For every fixed $t \ge 0$, its exponential moment $\mathbb{E}_{P_{m,a}} e^{tX}$ is bounded uniformly in m, a.

Proof. The second Charlier polynomial at mean one is $k^2 - 3k + 1$, whose minimum on $\mathbb{N}$ is -1. Thus

$$Q_a\{k\} \ge (1-a/2)e^{-1}/k!.$$

For $0 \le k \le m$,

$$\frac{\epsilon_{m,a}\binom{m}{k}}{Q_a\{k\}} \le \frac{1}{2(m-k)!} \le \frac{1}{2}.$$

Outside this range the perturbation is zero. Positivity and the two pointwise bounds follow. The perturbation has total mass zero, so $P_{m,a}$ has mass one.

The probability generating function of Δ^{*m} is $(z-1)^m$. Its derivatives at $z = 1$ vanish at every order other than m, where the derivative is $m!$. This proves the factorial-moment statement. Its ℓ^1 norm is 2^m, yielding the exact total variation distance. Finally,

$$\mathbb{E}_{P_{m,a}} e^{tX} \le \frac{3}{2}\exp(e^t - 1)\left(1 + \frac{1}{2}(e^t-1)^2\right),$$

which is independent of m, a. □

Take $m_N = \lfloor\sqrt{\log N}\rfloor$ and $a_N^\circ = (\log N)N^{-2/3}$, for N sufficiently large. At every fixed order r, the factorial moments of P_{m_N,a_N° eventually equal $1 + \binom{r}{2}a_N^\circ$ exactly. Nevertheless, Stirling's formula gives

$$\log d_{\mathrm{TV}}(P_{m_N,a_N^\circ}, Q_{a_N^\circ}) = -\tfrac{1}{2}\sqrt{\log N}\log\log N + O(\sqrt{\log N}).$$

This distance exceeds $N^{-\gamma}$ for every fixed $\gamma > 0$, eventually. Fixed-order information, even with the displayed exponential-moment control, therefore cannot replace the uniform estimate on the boundary of the stopped expansion.

Formal proofs and reproducibility

The accompanying Lean development provides proofs for the 118 numbered results in the article and companion (26 and 92, respectively). Its statement index records the correspondence of clauses, definitions and inherited hypotheses, including the conditional premise of Theorem 5.2. The documentation of the recorded replay is fixed at commit `a1c8092f7e25`. The linked source

manifest and validation record identify the proof source used for the recorded full replay as `aa12e797ce89`. That record specifies Lean 4.36.0-rc2 and Mathlib commit `9d581dfff2e4`, and includes compilation results, extracted statement types and the kernel-axiom audit. The documentation snapshot and the replayed proof source have distinct roles.

The correspondence between the written and formal statements was examined with AI assistance. It is not a machine translation of the written proofs or a substitute for independent human review. The conditional-sum and propagation clauses of Proposition T4.3 allow homogeneous prescriptions; its small-contact clause uses the original run-start events. The formal interfaces distinguish these scopes as well.

The final five-dossier allocation selects 42 formal statements covering every clause of the article's 26 numbered results. It preserves the four issued registrations and merges the two previously unregistered remainder scopes. The allocation is as follows, with numbers referring to the article:

Dossier	Main results	Statements
Quadratic approximation and Poisson scales	1.1–1.2, 7.2–7.4	10
Cumulants and optimality	3.3–3.4, 6.1–6.2	8
Pair profiles and channel hierarchy	3.1–3.2, 8.1–8.3	13
Stopping, conversion and reconstruction	4.1–4.2, 5.1–5.3	8
Uniform support and positive reference limits	2.1–2.5, 7.1, 8.4	3

The last dossier uses three compound statements, retaining all thirteen remaining main-result clauses and directly selecting companion Propositions T13.1 and T15.1, with the latter's displayed sequence corollary. The former controls the weighted full-support sum at all cardinalities with one threshold before every subset and minimum size. The latter shows that fixed factorial moments and uniform exponential-moment control cannot replace growing-order estimates for a polynomial total-variation rate; no multiplicative realization of that counterexample is claimed. The other companion results supply proof dependencies, without being claimed as all directly registered. The exact clause allocation fixes these scopes.

As of 13 November 2026, all five entries are published, in table order:

`PALOMAR-2026-11-12-000013`; `PALOMAR-2026-11-13-000001`;
`PALOMAR-2026-11-13-000002`; `PALOMAR-2026-11-13-000008`;
`PALOMAR-2026-11-13-000020`.

The first entry fixes source `f9a2d213265a`; entries 2–4 fix `d6a0f6c51098`; the final dossier fixes `eeda0e055bc1`. The final interface has a separate audit of 111 declarations and fresh statement comparison under the same Lean and Mathlib pins. Individual public entries provide the actual mechanical reports, rendered statements, preserved sources and published warnings.

Statement on AI assistance

ChatGPT (GPT-6 Astra) was used to organize the research material, develop and rewrite mathematical arguments, inspect references, and prepare the article and its technical companion. Parallel AI-assisted reviews examined selected identities, constants, proof dependencies and the exposition. The associated Lean development also used AI assistance for proof construction and statement comparison. AI-assisted review and machine checking are distinct from independent human verification. The author remains responsible for the mathematical claims, their correspondence with the formal statements, and any submission.

References

[1] A. D. Barbour, E. Kowalski, and A. Nikeghbali, *Mod-discrete expansions*, Probability Theory and Related Fields **158** (2014), 859–893. doi:10.1007/s00440-013-0498-8. Section numbers refer to the 26-page author text, Sections 5 and 8.1: `https://people.math.ethz.ch/~kowalski/mod-discrete.pdf`.

[2] G. Binyamini, R. Cluckers, and D. Novikov, *Bounds for rational points on algebraic curves, optimal in the degree, and dimension growth*, International Mathematics Research Notices **2024** (2024), no. 11, 9256–9265. doi:10.1093/imrn/rnae034. The degree-uniform integer-point estimate is cited from arXiv:2302.04209v2 (20 September 2023), Theorem 2, equation (27).

[3] R. de la Bretèche, V. Y. Wang, and M. W. Xu, *Random multiplicative functions and making squares from polynomial values*, arXiv:2607.06398v1 (2026), Theorem 1.2 and Lemmas 6.2–6.3.

[4] P. Erdős and J. Turk, *Products of integers in short intervals*, Acta Arithmetica **44** (1984), no. 2, 147–174. doi:10.4064/aa-44-2-147-174. Article text.

[5] D. R. Heath-Brown, *Square-free values of* n^2+1, Acta Arithmetica **155** (2012), no. 1, 1–13. doi:10.4064/aa155-1-1. Preprint version: arXiv:1010.6217v3, Lemma 2.

[6] M. Munsch, I. E. Shparlinski, Y.-C. Sun, and Y. Xiao, *A large sieve inequality for sums of Legendre symbols over short intervals*, arXiv:2604.23661v3 (5 August 2026), Lemma 3.3 and its proof.

[7] J. Pila, *Density of integral and rational points on varieties*, in *Columbia University Number Theory Seminar (New York, 1992)*, Astérisque **228** (1995), 183–187. Numdam: AST_1995___228___183_0.

[8] T. Reuss, *Pairs of k-free numbers, consecutive square-full numbers*, arXiv:1212.3150v2 (19 March 2014), Sections 2.1–2.3.

[9] P. Salberger, *Counting integral points of affine hypersurfaces*, arXiv:2311.05561v1 (9 November 2023), Lemma 2.

[10] V. Zacharovas and H.-K. Hwang, *A Charlier–Parseval approach to Poisson approximation and its applications*, Lithuanian Mathematical Journal **50** (2010), no. 1, 88–119. doi:10.1007/s10986-010-9073-5. Preprint version used for equation and remark numbers: arXiv:0810.4756v1, (2.13) and Remark 4.4.

[11] J. H. Griesmer, *A bound for error-correcting codes*, IBM Journal of Research and Development **4** (1960), no. 5, 532–542. doi:10.1147/rd.45.0532. Original article scan, Section 2, equation (2.10).

[12] J. Cilleruelo and M. Z. Garaev, *Concentration of points on two and three dimensional modular hyperbolas and applications*, Geometric and Functional Analysis **21** (2011), no. 4, 892–904. doi:10.1007/s00039-011-0127-6. Numbering refers to arXiv:1007.1526v2, Section 3, Lemma 4 and Proposition 1.

[13] T. Tao, *Products of consecutive integers with unusual anatomy*, arXiv:2603.27990v3 (26 September 2026), Lemma 2.10.

[14] I. E. Shparlinski and N. Sleiman, *Counting consecutive multiplicatively dependent triples*, arXiv:2609.29408v1 (24 September 2026), Theorem 1.1, Lemmas 2.1–2.2, and Section 3.

[15] G. Binyamini, R. Cluckers, and F. Kato, *Sharp bounds for the number of rational points on algebraic curves and dimension growth, over all global fields*, Proceedings of the London Mathematical Society **130** (2025), no. 1, e70016. Published online 16 December 2024. doi:10.1112/plms.70016. Corollary 8, equation (23).

[16] R. de la Bretèche, P. Kurlberg, and I. E. Shparlinski, *On the number of products which form perfect powers and discriminants of multiquadratic extensions*, International Mathematics Research Notices **2021** (2021), no. 22, 17140–17169. Published online 30 December 2019. doi:10.1093/imrn/rnz316. Author text, Theorem 2.2 and Corollary 2.4.