Reconstruction and sign prediction from sampled prime factors

Brice Pouly

Prepared with AI assistance

hexagon:2611.00894v1

Abstract

How much of an integer can be recovered from a few of its prime factors? We study exact prime labels sampled with logarithmic weights from an integer of harmonic law, distinguishing reconstruction, certification without error, and prediction of a prescribed sign. With one observed integer, sampling with replacement has its transition at order $\log M$, whereas sequential removal has a uniform termination profile and a Gaussian transition at order $\log\log M$. For the two related integers n and $an+1$, $a \in \{1,3\}$, one conditionally independent label from each gives optimal reconstruction probability $1/2 + o(1)$; the optimal zero-error certificate has mass $3 - 4\log 2 + o(1)$. One diverging budget and one auxiliary label suffice for reconstruction with probability tending to one under every coupling with the specified conditional marginals. Under conditional independence of the two sampling chains, we determine fixed-budget reconstruction and entropy profiles for the with-replacement, without-replacement and mixed experiments, with matching reconstruction upper bounds over all decoders. Two labels also predict every prescribed sign with success at least $3/4 - o(1)$. For either Liouville sign we give a strict asymptotic upper bound below $21293/26208$, using large-sieve cancellation, smooth-supported dispersion, and an explicit extension to rational residue classes. The remaining signed posterior is identified precisely. Finally, when one observer knows a rough cofactor and another knows three complete small-prime parts, reconstruction and one-message Liouville-sign prediction are governed by a four-variable Dickman probability, for which we give a reproducible numerical enclosure.

Contents

1 Introduction and main results **2**

I Reconstruction from sampled labels **8**

2 A small-prime factorization shared by the two experiments **8**

3 One anchor with replacement **9**

4 One anchor without replacement: the termination profile **14**

5 Two anchors: reconstruction and dependence **18**

6 One growing observation and one auxiliary prime **21**

7 Information carried by one auxiliary prime **23**

8 Exact two-anchor profiles with fixed replacement budgets **25**

9 Fixed budgets without replacement and mixed protocols **28**

10 Comparison of protocols and the positive-budget criterion **31**

11 First-order entropies for fixed budgets **32**

II Predicting a sign from two labels **36**

12 Predicting a sign on an observable confidence region **36**

13 The joint law of the logarithmic prime sizes **39**

14 The signed remainder beyond the universal guarantee **42**

15 Large-sieve cancellation on a region of mass $3/8$ **44**

16 A further gain from smooth cofactors **49**

17 The smooth gain at the second anchor **55**

18 An explicit region without a smoothness restriction **59**

19 Combining the rational and smooth regions **65**

20 A fixed sign attaining the universal guarantee **68**

21 Zero-error certificates and dependence of the labels **69**

22 Reconstruction, stability, and the quartic mean **71**

III Complete small-prime information and communication **74**

23 Reconstruction from a cofactor and three small-prime parts **74**

24 Cancellation when the congruence information is too small **77**

25 The exact cost of one message for a Liouville sign **84**

26 The Dickman curve and a reproducible numerical value **88**

AI assistance and verification **89**

References **90**

1 Introduction and main results

Knowing a prime divisor of an integer gives both a congruence and a piece of its factorisation. These two uses of the same observation lead to different reconstruction mechanisms. Repeated observations of a single integer must eventually determine its missing small valuations. Observations of two neighbouring integers can instead combine through the Chinese remainder theorem, long before either factorisation has been fully revealed. This paper quantifies the best decisions available from the exact observed labels.

A small example explains why the decision itself must be specified. Take an origin in $[10, 100)$, and suppose that the observations are $5 \mid n$ and $7 \mid n+1$. The compatible origins are $20, 55, 90$. Returning the least one is correct when the origin is 20, even though the same transcript cannot certify that answer without error. Predicting a sign is a third problem: a transcript may be informative about a sign even when it does not determine the integer. All three questions concern the posterior distribution, but optimise different functionals of it.

1.1 The sampled-label experiments

For an integer $M \geq 3$, set

$$I_M = [M, M^2) \cap \mathbb{N}, \qquad T_M = \sum_{n \in I_M} \frac{1}{n}, \qquad \mu_M(n) = \frac{\mathbf{1}_{I_M}(n)}{nT_M}, \qquad L = \log M, \quad V = \log\log M. \tag{1}$$

The hidden origin is $n \sim \mu_M$. Write $v_p(n)$ for its prime valuations, $\Omega(n) = \sum_p v_p(n)$, $\omega(n) = \#\{p : p \mid n\}$ and $\lambda(n) = (-1)^{\Omega(n)}$. All logarithms are natural unless a base is written explicitly. For laws on the same space, $d_{\mathrm{TV}}(P, Q) = \frac{1}{2}\|P - Q\|_1$.

At a nontrivial integer u, one prime label has distribution

$$\kappa_u(p) = \frac{v_p(u)\log p}{\log u}. \tag{2}$$

With replacement, each observation uses this kernel at the original u. Without replacement, one occurrence of the selected prime is removed and the kernel is applied to the remaining cofactor; after termination the transcript is padded by a distinguished symbol. The exact prime labels, their order and their repetitions are observed. The unrevealed cofactor, $\log u$, and the total number of prime factors are not supplied. When a final termination bit is included, its role will be stated explicitly; we also prove the corresponding limits without it.

For a transcript T, define the optimal reconstruction probability and sign advantage by

$$\mathcal{R}_M(\mathsf{T}) = \sup_{\widehat{n}} \mathbb{P}(\widehat{n}(\mathsf{T}) = n), \qquad \mathcal{A}_M(W; \mathsf{T}) = \|\mathbb{E}[W(n) \mid \mathsf{T}]\|_1, \tag{3}$$

where $W : I_M \to \{-1, 1\}$ is a prescribed sign. The best probability of a correct sign decision is $(1 + \mathcal{A}_M(W; \mathsf{T}))/2$. An error-free certificate may output $\perp$, but every integer it outputs must be correct on every positive-probability transcript. There is no computational restriction on any decoder.

1.2 One anchor: two observation clocks

One anchor already displays the effect of the sampling rule. With replacement, the transition occurs at $r \asymp \log M$. Section 3.4 identifies the complete critical profiles through a product of geometric–Poisson channels, including the averaging over the unobserved logarithmic size. Their small-time asymptotics are

$$I(\tau) \sim \frac{\pi^2}{12}\tau^2, \qquad J(\tau) \sim (\log 2)\tau \qquad (\tau \downarrow 0),$$

where I is the sign-advantage profile and J the reconstruction profile. The limit in M is taken first. Below this scale both quantities vanish; above it both tend to one.

Without replacement the correct clock is the total number of prime occurrences. In the notation of Section 4,

$$\sup_{r \geq 0} \left| \mathcal{R}_M^-(r) - \mathbb{P}(\Omega(n) \leq r) \right| \longrightarrow 0, \tag{4}$$

and the same comparison holds for the Liouville advantage. Consequently, for fixed $t \in \mathbb{R}$ and $r = V + t\sqrt{V} + o(\sqrt{V})$, reconstruction tends to $\Phi(t)$ and sign-prediction success to $(1 + \Phi(t))/2$. The nontrivial part of (4) is its upper bound: every decoder has negligible extra success before termination. The proof separates a finite set of small valuations, shows that they are unlikely to be observed sufficiently far before termination, and then lets this finite set grow. Erdős–Kac supplies the Gaussian shape after this decision comparison has been established.

1.3 Two anchors: the exact labels form a useful joining

Take anchors $u_0 = n$ and $u_1 = an + 1$, with $a \in \{1, 3\}$. Each chain follows one of the two rules above. For positive budgets, one budget tending to infinity already suffices for reconstruction with probability tending to one, for every coupling having the specified conditional marginals. An explicit asymmetric decoder combines the almost-complete factor information at one anchor with one sufficiently large auxiliary prime at the other.

Exact fixed-budget profiles require conditional independence of the two chains given n. Write $+$ for with replacement and $-$ for without replacement. For positive fixed r, s,

$$\mathcal{R}_M^{+,+}(r,s) \longrightarrow 1 - \binom{r+s}{r}^{-1}, \tag{5}$$

$$\mathcal{R}_M^{-,-}(r,s) \longrightarrow 1 - \int_{x+y\le 1} f_r^-(x) f_s^-(y)\, dx\, dy, \tag{6}$$

$$\mathcal{R}_M^{+,-}(r,s) \longrightarrow 1 - (r+1)^{-s}, \tag{7}$$

where $x, y > 0$ in the integral and

$$f_r^-(x) = \frac{[-\log(1-x)]^{r-1}}{(r-1)!}, \qquad 0 < x < 1.$$

The density for the with-replacement observed logarithmic product is $f_r^+(x) = rx^{r-1}$. These are classical size-biased partition densities. The arithmetic step is to count the two-anchor observation jointly on the region where its congruence modulus is smaller than the origin, and then show that no decoder extracts a positive limiting reconstruction mass from this region. Full proofs appear below, together with the leading entropy and mutual-information coefficients. No independence of the two complete affine factorisations is assumed.

In particular, one label at each anchor gives optimal reconstruction $1/2 + o(1)$. The optimal error-free certificate instead has mass

$$3 - 4\log 2 + o(1). \tag{8}$$

The gap comes from the distinction illustrated by the opening example: a decoder may correctly choose one of several possible origins, while an error-free certificate must abstain there.

1.4 A universal sign guarantee and the Liouville remainder

With one conditionally independent label at each anchor, there is an observable acceptance region of mass $1/2 + o(1)$ on which the decoder's conditional error tends uniformly to zero. Returning any prescribed sign of the decoded origin, and guessing fairly elsewhere, gives success at least $3/4 - o(1)$, uniformly over the prescribed sign. A fixed periodic sign, for example $(-1)^n$, attains this limit.

For the arithmetic signs $W(n) = \lambda(n)$ or $\lambda(an + 1)$, the optimal success probability has the more informative form

$$\mathrm{Succ}_M^*(W) = \frac{3}{4} + \frac{1}{2}\Lambda_M(W) + o(1), \tag{9}$$

where $\Lambda_M(W)$ is the absolute mass of the signed posterior on the unreconstructed sector, defined in Section 14. This isolates the quantity that would have to vanish to make the universal guarantee exact for Liouville. Large-sieve cancellation, smooth-supported dispersion and a rational-residue extension give, for both signs,

$$\frac{3}{4} \le \liminf_M \mathrm{Succ}^*_M(W) \le \limsup_M \mathrm{Succ}^*_M(W) \le \frac{13}{16} - \frac{1}{26208} - \frac{3}{8}\gamma_\eta < \frac{21293}{26208}. \tag{10}$$

The positive constant γ_η is a specified Dickman integral. The supplementary saving involving it is not assigned a numerical value, because the smooth-dispersion input supplies its exponent by existence. The rational-sector saving $1/26208$ is explicit. The equality $\mathrm{Succ}^*_M(\lambda) = 3/4 + o(1)$ is not proved here.

1.5 A different observation: complete small-prime parts

Part III studies an experiment with substantially more structured information. Alice receives $r_y(n)$; Bob receives the complete z-smooth parts of $n+1, 3n+1, 3n+2$. The harmonic window is now $(U, V]$, with $1 \le U < V$, $y, z \to \infty$, $\ell = \log(V/U) \to \infty$ and $\log(2\max(y,z)) = o(\ell)$. If $b/\log_2 y \to t$ and $\log z/\log y \to \tau$, the optimal reconstruction probability is

$$P(t,\tau) = \mathbb{P}\{D_0 < t,\ D_0 < \tau(D_1 + D_2 + D_3)\},$$

for independent Dickman probability variables. The reconstruction upper bound permits interaction, while one message attains it. For one message from Bob to Alice, optimal Liouville-sign success tends to $(1 + P(t,\tau))/2$. At equal thresholds and unlimited normalised budget, $P(\infty, 1)$ lies in $(0.9302864610, 0.9302864868)$. This different numerical value belongs to the complete-table experiment, rather than to the two-prime-label experiment.

Table 1 collects the observation, budget and decision assumptions. In particular, the three constants for one label at each anchor concern three different loss functions. The numerical Dickman constant concerns the richer complete-table experiment.

1.6 Previous work and the role of the arithmetic inputs

The probabilistic structure of prime factors has a substantial classical literature. Billingsley's work [2] concerns the distribution of large prime divisors; Donnelly and Grimmett [4] establish the stick-breaking limit for logarithmically size-biased factors removed successively. Arratia [5] develops couplings with Poisson–Dirichlet factor sizes, and Haddad and Koukoulopoulos [7] prove an optimal-order refinement of the logarithmic-factor coupling. The GEM densities and Ewens partition weights used below are part of the partition theory described by Pitman [6]. Bach's algorithm [3] addresses the related generative problem of producing a uniform random integer together with its factorisation. In the present experiments, the integer is hidden and only specified samples of its factors are observed. The factor-size limits are inputs to the decision analysis; they do not determine the posteriors of the exact prime labels. Likewise, the Gaussian theorem for the number of prime factors is due to Erdős and Kac [1]; its transfer to the harmonic window here follows by partial summation.

The harmonic Euler model in Part III also has a direct antecedent. Pinsky [8, Proposition 1 and Theorem 1] gives the exact law proportional to $1/a$ on integers supported on a finite initial set of primes, its independent geometric valuations, and the Dickman limit of the normalised logarithm, with density $e^{-\gamma}\rho$. Jaramillo and Yang [9, Proposition 4.1 and Section 4.3] discuss the representation of a harmonic integer in $[1, N]$ by independent geometric valuations conditioned on their product being at most N. Further developments concern conditioned Dickman limits for additive functionals [10] and attenuated Poisson–Dirichlet approximations for the prime-factor configuration of a harmonic sample [11]. Thus neither the Euler model nor its Dickman marginal

Table 1: Observations, budgets and decision criteria. All displayed performances are asymptotic; local computation is unrestricted. For the sampled-label experiments, $n \sim \mu_M$ and $M \to \infty$, the budgets are deterministic, and the two anchors are n and $an + 1$, $a \in \{1, 3\}$. The complete-table experiment has the different window and threshold assumptions of Theorem 23.1.

Observation and budget	Limit or guarantee	Decision and hypotheses
One anchor, with replacement; $r/\log M \to \tau > 0$	Reconstruction $J(\tau)$; Liouville advantage $I(\tau)$.	Independent draws given n; $\log n$ is hidden. The uniform-in-budget statements are absolute-error limits (Section 3).
One anchor, without replacement; deterministic r	Reconstruction and Liouville advantage equal $\mathbb{P}(\Omega(n) \le r) + o(1)$, uniformly in r.	Sequential removal. If $r = V + o(\sqrt{V})$, reconstruction tends to $1/2$ and sign success to $3/4$. Termination conventions are treated explicitly (Section 4).
Two anchors; fixed positive budgets r, s	Exact reconstruction profiles (5)–(7); corresponding entropy coefficients.	Conditional independence of the two chains given n. Budgets stay fixed as $M \to \infty$ (Sections 8–11).
Two anchors; positive budgets with $\max(r, s) \to \infty$	Reconstruction probability tends to one.	Sufficiency holds for every coupling with the specified conditional marginals. Necessity uses conditional independence (Corollaries 6.3 and 10.1).
One label at each anchor	Optimal reconstruction $1/2$; optimal zero-error certificate $3 - 4\log 2$; success at least $3/4$ for every prescribed sign.	Conditional independence given n. The certificate must never return a false integer; the other criteria allow error. Liouville has the bounds (10), with equality at $3/4$ left open.
Complete small-prime data; $b/\log_2 y \to t$, $\log z/\log y \to \tau$	Optimal reconstruction $P(t, \tau)$. At $t = \infty$, $\tau = 1$, it lies in $(0.9302864610, 0.9302864868)$.	Alice knows $r_y(n)$; Bob knows the three neighbour parts. The budget counts Bob's bits to Alice; Alice may send unlimited bits in return. One message attains the bound (Theorem 23.1).
The same complete data; one message Bob $\to$ Alice	Optimal $\lambda(n)$-prediction success $(1 + P(t, \tau))/2$.	Bob's encoder uses his table and independent public randomness; there is no preliminary message from Alice. The sign upper bound is proved for this protocol (Theorem 25.1).

is a contribution of this paper. The additional work in Part III is the joint separation of the small-prime parts from the remaining logarithmic location on a moving harmonic window, followed by the decision and communication analysis. General total-variation comparisons with geometric models for sequences of positive level of distribution are developed by Gorodetsky [12]; the simultaneous weighted cell estimates needed here are stated and proved in the form used by our protocols.

Information-theoretic arguments involving prime factorisation appear, for example, in Kontoyiannis [25]. Delsol et al. [30, Sections 6.3–6.4] study reconstruction from full prime signatures and congruence data, with observations and priors different from the sampled-label experiments. The correlated-source coding theorem of Slepian and Wolf [29] provides a broader context for communication with side information; it concerns repeated samples from a fixed joint source, whereas our profiles concern one arithmetic sample from a growing family. The public hashing step uses the classical universal-hashing principle of Carter and Wegman [26]. The binary Bayes-error/entropy inequalities are classical as well; see Hellman and Raviv [27]. In each case the paper's task is to identify the relevant arithmetic posterior or collision law. The inversion formula used in the certified Dickman calculation is the classical formula of Gil-Pelaez [28].

For the two-anchor logarithmic limit, the analytic input is Teräväinen's binary correlation theorem [14]. We give the transfer from fixed thresholds to thresholds depending on the origin, and the restriction needed for the anchor $3n+1$. Independence of the limiting normalised logarithms does not imply independence of the exact labels: their joint entropy and total variation retain substantial dependence. Recent preprints [32, 33] address ordinary joint Dickman laws for the largest prime factors of consecutive integers and ordinary two-point correlations of multiplicative functions. They are included for context; the proofs here use the explicitly stated Teräväinen and dispersion inputs. Results for fixed affine forms do not by themselves supply the uniform estimates over growing moduli and residues required by the signed posteriors.

The sign upper bounds use the classical Liouville exponential-sum estimate recalled in [15], the smooth-support theorem of Drappeau–Granville–Shao [16], Green's bilinear reduction [18], and Bettin–Chandee's Kloosterman-fraction bound [20]. Later results improve several distribution ranges. Fouvry and Radziwiłł [19, Corollary 1.3] reach level $17/33-\varepsilon$ for multiplicative functions under their Siegel–Walfisz hypothesis, with a common integer residue and error $O_\varepsilon(x/(\log x)^{1-\varepsilon})$; their prime-modulus convolution theorem has a stronger error in its own range. Pascadi's published Theorem 1.7 [17] reaches level $66/107-\varepsilon$ for one-bounded completely multiplicative functions with smooth support and a Siegel–Walfisz condition, allowing specified rational classes; its smoothness and residue ranges remain part of the hypothesis. His Theorem 4.2 makes the removal of small-conductor characters explicit. Wright's Theorem 2.1 [21] treats Kloosterman fractions with a partially fixed denominator and is structurally relevant to the denominator factors arising here. Applying these results to our posteriors would require verifying their coefficient norms, parameter ranges, character terms, and frequency sums. We retain the proved denominator cost and the range $39\vartheta+22\kappa<20$ obtained from the stated Green–Bettin–Chandee argument; we do not identify that range with the best distribution level available in other settings.

The smooth-number approximation is credited to de Bruijn [22]. Friable averages of oscillating functions form another established part of the analytic background; see de la Bretèche and Tenenbaum [23]. The communication proofs use the classical additive and multiplicative large sieves and the fixed-dimensional fundamental lemma of the sieve [24], with their uniformity requirements recorded below.

The results are therefore formulated at the level of the specified observation experiments: their optimal decision profiles, the effect of combining two anchors, and the signed dispersion transfers that control the remaining ambiguity. The reconstruction target is a hidden integer. In the quadratic-probe model [36], a public panel of randomly chosen primes probes the Legendre symbols of $n, n+1, 3n+1, 3n+2$. Here the conditional sampling law selects prime divisors of the forms themselves, with weights given by their valuations and logarithms. In [31], the

target is instead a multiplicative function observed on prescribed quadratic traces. Finally, a high probability of reading a Liouville sign does not estimate its unconditional mean; Section 22 records the covariance that remains for a quartic correlation.

1.7 Organisation and limits

Part I builds the one-anchor profiles, the two-anchor decoders, and the fixed-budget reconstruction and entropy formulas. Part II passes from reconstruction to sign decisions, proves the cancellation estimates, and distinguishes observable confidence from a zero-error certificate. Part III gives a self-contained treatment of the complete-table communication experiment and its numerical constant. In all fixed-budget asymptotics, the budgets are held fixed while $M \to \infty$; no growing-budget uniformity is inferred from them. The uniform one-anchor profiles and the coupling-independent reconstruction estimates have their own proofs. This separation of contracts is also used whenever a parameter must be sent to zero after the arithmetic limit.

Part I
Reconstruction from sampled labels

2 A small-prime factorization shared by the two experiments

The obstruction to reconstructing an integer from a few sampled factors is already present in any fixed finite set of small primes. We first separate those valuations from the remaining cofactor. Throughout, $M \geq 3$ is an integer, $I_M = [M, M^2) \cap \mathbb{N}$, $L = \log M$ and $\mu_M(n) = \mathbf{1}_{I_M}(n)/(nT_M)$, where

$$T_M = \sum_{n \in I_M} \frac{1}{n} = L + O(M^{-1}), \qquad T_M \geq L, \qquad \mathbb{P}_{\mu_M}(a \mid n) \leq \frac{2}{a} \quad (a \geq 1). \tag{11}$$

For the divisibility bound, write $n = ak$ and use the upper bound $1 + L$ for the harmonic sum over $k \in [M/a, M^2/a)$.

Fix a finite set S of primes before taking $M \to \infty$, and put

$$P_S = \prod_{p \in S} p, \qquad c_S = \prod_{p \in S} (1 - p^{-1}), \qquad d = D_S(n) = \prod_{p \in S} p^{v_p(n)}, \qquad m = n/d.$$

An integer is called S-smooth if all its prime factors belong to S. The probability measure

$$\pi_S(d) = \frac{c_S}{d} \mathbf{1}_{\{d \text{ is } S\text{-smooth}\}}$$

has independent geometric valuations: $\mathbb{P}(v_p(d) = e) = (1 - p^{-1})p^{-e}$ for $e \geq 0$. This classical small-prime model is discussed in [5, §1.1]. For a harmonic integer on $[1, N]$, Jaramillo and Yang [9, Proposition 4.1] give an exact representation by independent geometric valuations conditioned on their product being at most N. The comparison below also separates the remaining cofactor and pays for moving the harmonic window; its product reference law is not conditioned on that product bound. Define

$$a_{M,S} = \mu_M((n, P_S) = 1), \qquad \rho_{M,S} = \mu_M(\cdot \mid (n, P_S) = 1).$$

For all sufficiently large M, depending on S, the conditional measure is defined and $a_{M,S} = c_S + O_S((ML)^{-1})$.

Lemma 2.1 (An independent small part)**.** *The law $\eta_M = \mathcal{L}(D_S(n), n/D_S(n))$ satisfies*

$$\beta_{M,S} := d_{\mathrm{TV}}(\eta_M, \pi_S \otimes \rho_{M,S}) = O_S(L^{-1} + M^{-1/2}). \tag{12}$$

The product dm under the reference law need not belong to I_M; the displayed error includes this change of window.

Proof. The unique decomposition $n = dm$ gives

$$\eta_M(d,m) = \frac{\mathbf{1}_{\{d \text{ is } S\text{-smooth}\}}\mathbf{1}_{(m,P_S)=1}\mathbf{1}_{M\le dm<M^2}}{dmT_M}.$$

Replace the last indicator by $\mathbf{1}_{M\le m<M^2}$. If $d \le M$, the harmonic sum over the symmetric difference of the two windows is at most $2\log d + O(d/M)$; if $d > M$, it is $O(L+1)$. The convergent series

$$\sum_{d\ S\text{-smooth}} \frac{\log d}{d} < \infty, \qquad \sum_{d\ S\text{-smooth}} d^{-1/2} = \prod_{p\in S}(1-p^{-1/2})^{-1} < \infty$$

show that the change has ℓ^1 cost $O_S(L^{-1} + M^{-1/2})$. Indeed, the number of S-smooth $d \le M$ is $O_S(\sqrt{M})$, and the reciprocal sum over $d > M$ is $O_S(M^{-1/2})$. Inclusion–exclusion at the fixed modulus P_S gives

$$H_{M,S} := \sum_{\substack{m\in I_M\\(m,P_S)=1}} \frac{1}{m} = c_S L + O_S(M^{-1}).$$

The modified measure has mass $H_{M,S}/(c_S T_M) = 1 + O_S((ML)^{-1})$, and its normalization is exactly $\pi_S \otimes \rho_{M,S}$. □

We shall use the finite constants

$$h_S := \mathbb{E}_{\pi_S}\Omega(d) = \sum_{p\in S}\frac{1}{p-1}, \qquad b_S := \mathbb{E}_{\pi_S}\log d = \sum_{p\in S}\frac{\log p}{p-1}. \tag{13}$$

Two different properties of the small part control the two decision problems:

$$\max_d \pi_S(d) = c_S \longrightarrow 0, \qquad \ell_S := \mathbb{E}_{\pi_S}\lambda(d) = \prod_{p\in S}\frac{p-1}{p+1} \longrightarrow 0$$

as S increases to the set of all primes. The first bound controls exact reconstruction; the second controls parity. Both limits follow from $\sum_p 1/p = \infty$.

Appending the same observation kernel contracts total variation. Thus the error in a success probability is at most $\beta_{M,S}$, whereas the error in the expectation of a function bounded in absolute value by one is at most $2\beta_{M,S}$.

3 One anchor with replacement

Given $n \sim \mu_M$, draw independent prime labels with common law

$$\kappa_n(q) = \frac{v_q(n)\log q}{\log n}, \qquad \mathcal{P}_r = (P_1,\dots,P_r) \sim \kappa_n^{\otimes r}.$$

The observation contains the exact labels, their order and their repetitions. It does not contain $\log n$. Write

$$\mathcal{R}_M^+(r) = \sup_{\widehat{n}} \mathbb{P}\{\widehat{n}(\mathcal{P}_r) = n\}, \qquad \mathcal{A}_M^+(r) = \left\|\mathbb{E}[\lambda(n) \mid \mathcal{P}_r]\right\|_1.$$

The optimal sign-prediction probability is $(1 + \mathcal{A}_M^+(r))/2$. Decision rules may depend on M, r and have unrestricted computational complexity. Independent randomization does not improve either optimum.

3.1 Small valuations below the logarithmic scale

Append the exact kernel $\kappa_{dm}^{\otimes r}$ to both laws in Lemma 2.1. Under the product law, $m \geq M$ and $(d, m) = 1$, so, for $d > 1$,

$$\kappa_{dm} = \frac{\log m}{\log(dm)}\kappa_m + \frac{\log d}{\log(dm)}\kappa_d.$$

A coordinatewise coupling replaces the kernel by $\kappa_m^{\otimes r}$ at cost at most $r \log d/L$; the cost is zero when $d = 1$. Since $\mathbb{E}_{\pi_S} \log d = b_S < \infty$, the true joint law of $(d, m, \mathcal{P}_r)$ is within total variation

$$O_S\left(\frac{r+1}{L} + M^{-1/2}\right) \tag{14}$$

of the experiment in which $d \sim \pi_S$ and $m \sim \rho_{M,S}$ are independent and then $\mathcal{P}_r \sim \kappa_m^{\otimes r}$. This second kernel replacement will only be used when $r = o(L)$.

In the reference experiment, d is independent of $(m, \mathcal{P}_r)$. For any decoder and fixed $(m, \mathcal{P}_r)$, the equality $\widehat{n}(\mathcal{P}_r) = dm$ selects at most one value of d. Consequently,

$$\mathcal{R}_M^+(r) \leq c_S + O_S\left(\frac{r+1}{L} + M^{-1/2}\right). \tag{15}$$

For any test $|\varphi| \leq 1$, complete multiplicativity gives

$$|\mathbb{E}_{\mathrm{ref}}\lambda(dm)\varphi(\mathcal{P}_r)| \leq |\mathbb{E}_{\pi_S}\lambda(d)| = \prod_{q\in S}\frac{1-q^{-1}}{1+q^{-1}}.$$

Conditional duality therefore yields

$$\mathcal{A}_M^+(r) \leq \prod_{q\in S}\frac{1-q^{-1}}{1+q^{-1}} + O_S\left(\frac{r+1}{L} + M^{-1/2}\right). \tag{16}$$

For $r_M = o(L)$, first let $M \to \infty$ with S fixed and then let $S = \{q : q \leq Q\}$ increase through the primes. We obtain

$$\mathcal{R}_M^+(r_M) \longrightarrow 0, \qquad \mathcal{A}_M^+(r_M) \longrightarrow 0. \tag{17}$$

The small-prime set remains fixed during the arithmetic limit.

3.2 An explicit decoder above the logarithmic scale

Let $N_q = \#\{j : P_j = q\}$ and $R_r = \prod_{q:N_q>0} q$. For $r \geq 1$ the transcript alone defines

$$\widehat{e}_q = \left\lfloor \frac{N_q \log R_r}{r \log q} + \frac{1}{2} \right\rfloor, \qquad \widehat{n}_r = \prod_{q:N_q>0} q^{\widehat{e}_q}. \tag{18}$$

There are finite absolute constants C_1, C_2 such that

$$\mathbb{P}(\widehat{n}_r \neq n) \leq C_1\frac{L}{r} + \frac{C_2}{L} \qquad (M \geq 3, r \geq 1). \tag{19}$$

Here is a proof retaining the dependence on the sample size. Define the radical defect

$$D(n) = \log\frac{n}{\mathrm{rad}(n)} = \sum_q \sum_{k\geq 2} (\log q)\mathbf{1}_{q^k|n}$$

and the convergent constants

$$U_1 = \sum_q \frac{\log q}{q(q-1)}, \qquad U_2 = \sum_q \frac{(q+1)(\log q)^2}{q(q-1)^2}, \qquad S_0 = \sum_q \frac{1}{(q-1)\log q}.$$

The first two series converge by comparison with sums over integers; the third follows from Chebyshev's bound $\pi(x) \ll x/\log x$, summed over dyadic intervals. The divisibility estimate (11) implies

$$\mathbb{E}D \le 2U_1, \qquad \mathbb{E}D^2 \le 2(U_1^2 + U_2), \qquad \max_q v_q(n) \le 1 + D/\log 2. \tag{20}$$

For the diagonal contribution to the second moment, use $\sum_{k,j\ge 2} q^{-\max(k,j)} = (q+1)/(q(q-1)^2)$; for distinct primes use the product of their powers.

In the proof only, introduce $\widetilde{e}_q = N_q \log n/(r \log q)$ for $q \mid n$. The conditional binomial law gives

$$\mathbb{E}[\widetilde{e}_q \mid n] = e_q := v_q(n), \qquad \mathrm{Var}(\widetilde{e}_q \mid n) \le \frac{2e_q L}{r \log q}.$$

Chebyshev's inequality at threshold $1/3$, followed by a union bound and $\mathbb{E}v_q(n) \le 2/(q-1)$, yields

$$\mathbb{P}(G^c) \le 36 S_0 L/r, \qquad G = \{\text{for every } q \mid n : |\widetilde{e}_q - e_q| < 1/3\}.$$

On G, every prime divisor has been observed and $R_r = \mathrm{rad}(n)$. The difference between $\widetilde{e}_q$ and the quantity rounded in (18) is at most

$$(e_q + 1/3)\frac{D}{\log n} \le \frac{1}{L}\left(\frac{4}{3}D + \frac{D^2}{\log 2}\right).$$

If the right side is less than $1/6$, all the rounded valuations are correct. Markov's inequality and (20) prove (19), for example with

$$C_1 = 36 S_0, \qquad C_2 = 16 U_1 + \frac{12(U_1^2 + U_2)}{\log 2}.$$

Thus $r_M/L \to \infty$ implies $\mathcal{R}_M^+(r_M) \to 1$ and $\mathcal{A}_M^+(r_M) \to 1$. More generally, any deterministic sign $W_M(n)$ can then be predicted with vanishing error by evaluating $W_M(\widehat{n}_r)$, extending it arbitrarily outside I_M if necessary. These are averaged statements under μ_M. Pointwise reconstruction from the sampling law is impossible, since $\kappa_{n^k} = \kappa_n$.

3.3 Recovering the logarithmic size before the integer

The product of the distinct observed labels satisfies

$$\log \frac{n}{R_r} = D(n) + \sum_{q \mid n} (\log q) \mathbf{1}_{N_q = 0}.$$

Put $a = r/(2L) > 0$. A prime of valuation e is missed with conditional probability at most q^{-ae}. Since $\mathbb{P}(v_q(n) = e) \le 2q^{-e}$,

$$\mathbb{E} \log \frac{n}{R_r} \le 2U_1 + 2\sum_q \frac{\log q}{q^{1+a} - 1} = 2U_1 - 2\frac{\zeta'(1+a)}{\zeta(1+a)} \ll 1 + L/r. \tag{21}$$

The last estimate follows by partial summation from $\sum_{k \le x} \Lambda(k) \ll x$. Hence $\log R_{r_M}/\log n \to 1$ in L^1 whenever $r_M \to \infty$. In the range $r_M \to \infty$ and $r_M = o(L)$, almost all of the logarithmic size is recovered, although exact reconstruction has vanishing success probability and the optimal Liouville prediction has success tending to $1/2$.

3.4 The geometric–Poisson channel at the critical scale

For $a \geq 0$, let $\mathrm{Poi}_a(k) = e^{-a}a^k/k!$, with $\mathrm{Poi}_0(0) = 1$. For a prime q and $u \geq 0$, define

$$w_q(e,k;u) = (1-q^{-1})q^{-e}\,\mathrm{Poi}_{ue\log q}(k), \qquad e,k \geq 0, \tag{22}$$

$$I_q(u) = \sum_{k\geq 0}\left|\sum_{e\geq 0}(-1)^e w_q(e,k;u)\right|, \qquad J_q(u) = \sum_{k\geq 0}\max_{e\geq 0} w_q(e,k;u). \tag{23}$$

These are respectively the optimal parity advantage and exact-value success for a geometric valuation observed through a Poisson count. For $\tau > 0$, put

$$I(\tau) = \int_1^2 \prod_q I_q(\tau/s)\,ds, \qquad J(\tau) = \int_1^2 \prod_q J_q(\tau/s)\,ds. \tag{24}$$

Theorem 3.1 (Critical profiles with replacement)**.** *For every fixed $\tau > 0$,*

$$\mathcal{A}_M^+(\lfloor \tau L\rfloor) \longrightarrow I(\tau), \qquad \mathcal{R}_M^+(\lfloor \tau L\rfloor) \longrightarrow J(\tau). \tag{25}$$

Both profiles are continuous and nondecreasing, take values strictly between zero and one on $(0,\infty)$, and tend to one at infinity. After the limit in M, they satisfy

$$I(\tau) \sim \frac{\pi^2}{12}\tau^2, \qquad J(\tau) \sim (\log 2)\tau \qquad (\tau \downarrow 0). \tag{26}$$

An auxiliary cofactor that can be estimated. Fix $Q \geq 2$ and $S = \{q : q \leq Q\}$. Temporarily reveal $m = n/D_S(n)$, and denote the enriched optima by $\mathcal{A}_{M,Q}^{\mathrm{aux}}$ and $\mathcal{R}_{M,Q}^{\mathrm{aux}}$. The original transcript already provides $\widehat{m}_Q = \prod_{q>Q,N_q>0} q$. If $r = \lfloor \tau L\rfloor$, then, for large M, $r \geq \tau L/2$, and a prime $q > Q$ of valuation e is missed with probability at most $q^{-\tau e/4}$. Thus

$$\mathbb{P}(\widehat{m}_Q \neq m) \leq 2\sum_{q>Q}\left(q^{-2} + \frac{1}{q^{1+\tau/4}-1}\right) =: \delta_Q(\tau) \longrightarrow 0. \tag{27}$$

The two summands pay respectively for repeated large prime factors and for unobserved large primes. Replacing the auxiliary cofactor by its estimate in an optimal rule gives

$$\begin{aligned} 0 \leq \mathcal{A}_{M,Q}^{\mathrm{aux}} - \mathcal{A}_M^+(r) \leq 2\delta_Q(\tau),\\ 0 \leq \mathcal{R}_{M,Q}^{\mathrm{aux}} - \mathcal{R}_M^+(r) \leq \delta_Q(\tau). \end{aligned} \tag{28}$$

These comparisons are uniform over all decision rules. Lemma 2.1 next replaces the law of (d,m) by $\pi_S \otimes \rho_{M,S}$ while retaining the exact kernel $\kappa_{dm}^{\otimes r}$. Its error tends to zero by contraction, even when $r \asymp L$. The subcritical replacement of that kernel by $\kappa_m^{\otimes r}$ is not made here.

The small counts are sufficient. Write $e_q = v_q(d)$ and $K_q = \#\{j : P_j = q\}$ for $q \in S$. Conditional on (d,m), these counts and the count outside S are multinomial with probabilities

$$\alpha_q = \frac{e_q \log q}{\log m + \log d}, \qquad \alpha_0 = \frac{\log m}{\log m + \log d}.$$

Given m and the small counts, the category positions are uniform and the outside labels have law κ_m, independently of d. The rest of the transcript therefore provides no further information about d. This sufficiency statement holds for the finite experiment.

Set $s = \log m/L \in [1,2]$, and first restrict to $d \leq D$. Uniformly in m,

$$\sum_{q\in S}\alpha_q = O_D(L^{-1}), \qquad r\alpha_q = (\tau/s)e_q\log q + O_{D,\tau}(L^{-1}).$$

The total count of the small categories can be coupled to a Poisson variable at cost at most $2r(\sum \alpha_q)^2 = O_{D,\tau}(L^{-1})$. Given this total, the allocation among categories is unchanged, so the vector couples to independent Poisson variables of means $r\alpha_q$. Changing these means to $(\tau/s)e_q \log q$ costs at most the sum of their differences. The tail $d > D$ has mass $\pi_S(d > D) \to 0$, uniformly in m. Taking $M \to \infty$ and then $D \to \infty$ proves total-variation convergence of the joint law of the small valuations and counts to the finite product of the channels (22), uniformly in m.

The known sign $\lambda(m)$ can be removed. Absolute convergence and product factorization give optimal values $\prod_{q\in S} I_q(u)$ and $\prod_{q\in S} J_q(u)$, where $u = \tau/s$. Each local value is continuous: truncate the valuation, use total-variation continuity of Poisson laws, and then remove the geometric tail. Under $\rho_{M,S}$, the variable $s = \log m/L$ converges to the uniform distribution on $[1, 2]$, by harmonic inclusion–exclusion at the fixed modulus P_S. Consequently,

$$\lim_M \mathcal{A}^{\mathrm{aux}}_{M,Q} = \int_1^2 \prod_{q\le Q} I_q(\tau/s)\, ds, \qquad \lim_M \mathcal{R}^{\mathrm{aux}}_{M,Q} = \int_1^2 \prod_{q\le Q} J_q(\tau/s)\, ds.$$

The total number of samples has remained deterministic. Poisson limits have only been used for a fixed finite collection of small prime counts.

Infinite products and removal of the cofactor. Put $x = q^{-1}$ and $y = q^{-1-u} \le x$. The parity rule that answers $+1$ at count zero and -1 otherwise has correlation

$$A_q(u) = (1 - x)\left(\frac{2}{1 + y} - \frac{1}{1 + x}\right) = 1 - \frac{2(y + x^2)}{(1 + x)(1 + y)} > 0.$$

The valuation rule that answers zero at count zero and one otherwise has success

$$B_q(u) = (1 - x)(1 + x - y) = 1 - x^2 - (1 - x)y > 0.$$

It follows that

$$1 \ge I_q \ge A_q, \quad 1 - I_q \le 2y + 2x^2, \qquad 1 \ge J_q \ge B_q, \quad 1 - J_q \le y + x^2.$$

The defects are uniformly summable for $u \ge u_0 > 0$, so the infinite products are positive and continuous on $(0, \infty)$. They are strictly less than one: when the count at prime 2 is zero, several valuations of different parity still have positive posterior mass. Letting $Q \to \infty$ in (28) establishes (25) for the original observation.

Small intensity. At count zero the signed mass is $(1 - x)/(1 + y) > 0$. At a positive count the valuation-one contribution has negative sign. The improvement over the preceding parity rule is therefore bounded by twice the mass of valuations $e \ge 2$. Similarly, zero is the most likely valuation at count zero, and the possible improvement over the choice one at positive counts is bounded by the mass of $e \ge 2$. Hence

$$0 \le I_q(u) - A_q(u) \le 2x^2, \qquad 0 \le J_q(u) - B_q(u) \le x^2.$$

Uniformly for $u \ge 0$,

$$\frac{I_q(u)}{(1 - q^{-1-u})^2} = 1 + O(q^{-2}), \qquad \frac{J_q(u)}{1 - q^{-1-u}} = 1 + O(q^{-2}).$$

The corresponding regularized products H_I, H_J are continuous down to zero. Since $I_q(0) = (1 - x)/(1 + x)$ and $J_q(0) = 1 - x$,

$$H_I(0) = \prod_q (1 - q^{-2})^{-1} = \zeta(2), \qquad H_J(0) = 1.$$

Euler's product and the simple pole of ζ at one give

$$\prod_q I_q(u) = \zeta(1+u)^{-2} H_I(u) \sim \zeta(2) u^2, \qquad \prod_q J_q(u) = \zeta(1+u)^{-1} H_J(u) \sim u.$$

These equivalents hold uniformly for $u \in [\tau/2, \tau]$ as $\tau \downarrow 0$. Integrating s^{-2} and s^{-1} over $[1,2]$ proves (26).

Monotonicity and large intensity. When $u \le v$, thinning a Poisson count of mean $ve \log q$ produces a count of mean $ue \log q$. Extra information cannot decrease an optimal decision value; the local values and both profiles are therefore nondecreasing. Rounding $K/(u \log q)$ estimates the valuation with error at most $4/[u(q-1)\log q]$, by Chebyshev's inequality. A parity error costs at most twice the reconstruction error, so

$$1 - J(\tau) \le \frac{6 S_0}{\tau}, \qquad 1 - I(\tau) \le \frac{12 S_0}{\tau}.$$

This proves both limits at infinity and completes the theorem.

3.5 Uniformity over deterministic budgets

Extend I, J by $I(0) = J(0) = 0$. Pointwise convergence, continuity and monotonicity imply the absolute uniform limits

$$\begin{aligned} \sup_{r \in \mathbb{N}_0} |\mathcal{A}_M^+(r) - I(r/L)| &\longrightarrow 0, \\ \sup_{r \in \mathbb{N}_0} |\mathcal{R}_M^+(r) - J(r/L)| &\longrightarrow 0. \end{aligned} \tag{29}$$

Indeed, longer transcripts contain the shorter ones, so the finite optimal values are nondecreasing in r. On a compact interval $r/L \in [a,b] \subset (0,\infty)$, bracket them by their values on a finite grid and use uniform continuity of the profiles. For $r/L < a$, bound above by the value near aL and then take $a \downarrow 0$. For $r/L > b$, bound below by the value near bL and then let $b \to \infty$. All values lie in $[0,1]$.

The convergence in (29) is absolute. It does not turn (26), whose limit in τ follows the limit in M, into a relative asymptotic for a moving $\tau_M \to 0$. Moreover, I is an L^1 posterior advantage; it does not by itself identify the optimal quadratic risk. Its product representation uses $\lambda(dm) = \lambda(d)\lambda(m)$ and concerns this one-anchor observation.

4 One anchor without replacement: the termination profile

Removing an observed occurrence changes the sampling weights after every step. In this experiment the number of prime occurrences, rather than the logarithmic size, determines the transition. The proof again separates finitely many small valuations, and uses exponential clocks to show that they are unlikely to have been observed well before termination.

4.1 The observation and the optimal profiles

Starting from $n \sim \mu_M$, remove one prime occurrence p from the current cofactor $m > 1$, assigning each occurrence weight $\log p / \log m$. The transcript T_r records the first r ordered labels, with an absorbing symbol after exhaustion, and a final bit indicating whether exhaustion has occurred by step r. Budgets $r \in \mathbb{N}_0$ are deterministic. Neither $\Omega(n)$ nor a nontrivial cofactor is revealed. Put

$$F_r = \{\Omega(n) \le r\}, \qquad \mathcal{R}_M^-(r) = \sup_{\widehat{n}} \mathbb{P}\{\widehat{n}(T_r) = n\}, \qquad q_{M,r} = \mathbb{E}[\lambda(n) \mid T_r].$$

The decoder may abstain, and may use independent randomization; neither option changes the optimal probability of a correct answer. Let $V = \log\log M$, and let Φ be the standard normal distribution function.

Theorem 4.1 (Uniform exact reconstruction)**.** *As $M \to \infty$,*

$$\sup_{r\geq 0}\left|\mathcal{R}_M^-(r) - \mathbb{P}_{\mu_M}(\Omega(n)\leq r)\right| \longrightarrow 0. \tag{30}$$

Consequently,

$$\sup_{r\geq 0}\left|\mathcal{R}_M^-(r) - \Phi\left(\frac{r-V}{\sqrt{V}}\right)\right| \longrightarrow 0. \tag{31}$$

Theorem 4.2 (The Liouville posterior)**.** *Uniformly over deterministic budgets,*

$$\sup_{r\geq 0}\mathbb{E}\,|q_{M,r} - \lambda(n)\mathbf{1}_{F_r}| \longrightarrow 0. \tag{32}$$

In particular, uniformly in r,

$$\|q_{M,r}\|_1 = \mathbb{P}(F_r) + o(1), \qquad \mathbb{E}q_{M,r}^2 = \mathbb{P}(F_r) + o(1).$$

If $r = V + t\sqrt{V} + o(\sqrt{V})$ for fixed t, the optimal reconstruction probability tends to $\Phi(t)$, while the optimal sign-prediction probability tends to $(1+\Phi(t))/2$. The minimal absolute and quadratic prediction risks for $\lambda(n)$ both tend to $1-\Phi(t)$.

Theorem 4.1 is an upper bound for every decoder, including those that try to infer the unobserved factors before termination. It is attained asymptotically by the rule that multiplies the observed labels when the final bit certifies exhaustion.

4.2 Small prime factors remain hidden before termination

Fix the set S from Section 2, and use the quantities $c_S, a_{M,S}, \beta_{M,S}, h_S, b_S$ defined there. Let $K_r(u)$ be the law of the without-replacement transcript, including its final bit, when the initial integer is $u \geq 1$.

Attach independent exponential clocks of rate $\log p$ to all prime occurrences of u. Listing them in increasing order realizes the required sampling law: after each removal the total rate of the remaining clocks is the logarithm of the current cofactor. This is the usual exponential-clock representation of size-biased sampling [5, §1.3].

Let $Z_\tau(m)$ count the clocks belonging to m that have not rung by time τ. Keep those clocks and add the clocks of d to couple $K_r(m)$ and $K_r(dm)$. The added total rate is $\log d$. On

$$\{\Omega(m) > r + K\} \cap \{Z_{1/K}(m) \leq K\}$$

and in the absence of an added arrival before $1/K$, the first r labels coincide and neither transcript has terminated. Therefore, for every $K \geq 1$,

$$\mathbf{1}_{\{\Omega(m)>r+K\}} d_{\mathrm{TV}}(K_r(dm), K_r(m)) \leq \frac{\mathbb{E}[Z_{1/K}(m)\mid m]}{K} + \frac{\log d}{K}.$$

This remains a valid upper bound when its right side exceeds one. The harmonic divisibility estimate gives

$$\mathbb{E}_{\rho_{M,S}} v_p(m) \leq \frac{2}{a_{M,S}(p-1)}.$$

Hence

$$\mathbb{E}_{\rho_{M,S}} Z_{1/K}(m) \leq \frac{2}{a_{M,S}}\sum_p \frac{p^{-1/K}}{p-1} \leq \frac{2}{a_{M,S}}\{\log(K+1)+1\}.$$

For the last inequality, use $\sum_p p^{-1-\tau} \le \log\zeta(1+\tau) \le \log(1+1/\tau)$ and $\sum_p 1/[p(p-1)] \le 1$. Averaging over $d \sim \pi_S$ yields, uniformly in r,

$$\mathbb{E}_{\pi_S\otimes\rho_{M,S}}\left[\mathbf{1}_{\{\Omega(m)>r+K\}} d_{\mathrm{TV}}(K_r(dm), K_r(m))\right] \le C_{M,S,K}, \tag{33}$$

where

$$C_{M,S,K} = \frac{2\{\log(K+1)+1\}/a_{M,S} + b_S}{K}.$$

The variables (d, m) are only a device for comparing experiments; they are not included in the actual observation.

4.3 Moving the nonterminated event

Define the uniform Erdős–Kac discrepancy

$$\Delta_M = \sup_{x\in\mathbb{R}} \left| \mathbb{P}_{\mu_M}\left(\frac{\Omega(n)-V}{\sqrt{V}} \le x\right) - \Phi(x)\right|.$$

We verify below that $\Delta_M \to 0$. An interval of length $2K$ then satisfies

$$\sup_r \rho_{M,S}\{r-K < \Omega(m) \le r+K\} \le \frac{2\Delta_M + 2K/\sqrt{2\pi V}}{a_{M,S}}.$$

Under the product law set $G_r = \{\Omega(dm) > r\}$ and $H_r = \{\Omega(m) > r+K\}$. Since $H_r \subseteq G_r$ and

$$G_r \setminus H_r \subseteq \{\Omega(d) > K\} \cup \{r-K < \Omega(m) \le r+K\},$$

Markov's inequality gives

$$\mathbb{P}_{\pi_S\otimes\rho_{M,S}}(G_r \setminus H_r) \le B_{M,S,K} := \frac{h_S}{K} + \frac{2\Delta_M + 2K/\sqrt{2\pi V}}{a_{M,S}}. \tag{34}$$

This estimate is uniform in r.

4.4 Proof of the reconstruction theorem

Fix a decoder $\widehat{n}$. Its success on F_r is at most $\mathbb{P}(F_r)$. For its success on F_r^c, first use Lemma 2.1, then replace G_r by H_r using (34), and finally use the kernel comparison (33). Each test is an indicator, so each total-variation cost is paid once:

$$\begin{aligned}\mathbb{P}\{\widehat{n}(T_r) = n, F_r^c\} \le{}& \beta_{M,S} + B_{M,S,K} + C_{M,S,K} \\ &+ \mathbb{E}_{\pi_S\otimes\rho_{M,S}K_r(m)}\left[\mathbf{1}_{H_r}\mathbf{1}_{\{\widehat{n}(T_r)=dm\}}\right].\end{aligned}$$

In the last experiment d is independent of (m, T_r) and of the decoder's independent randomization. For fixed values of those variables, success selects at most one d, with mass at most c_S. Since multiplying the labels reconstructs n on the observable event F_r, we have for every r

$$0 \le \mathcal{R}_M^-(r) - \mathbb{P}(F_r) \le c_S + \beta_{M,S} + B_{M,S,K} + C_{M,S,K}. \tag{35}$$

Take $K = \max(1, \lfloor V^{1/4}\rfloor)$. For fixed S, $a_{M,S} \to c_S > 0$ and all the displayed error terms tend to zero. It follows that

$$\limsup_{M\to\infty} \sup_{r\ge 0}\{\mathcal{R}_M^-(r) - \mathbb{P}(F_r)\} \le c_S.$$

Now let $S = \{p : p \le y\}$ and then $y \to \infty$. Since $c_S \to 0$, this proves (30). The proof uses the order $M \to \infty$ followed by S increasing through the primes.

4.5 Proof of the signed posterior theorem

For any test $|\varphi| \le 1$ of the transcript, the same transport applied to signed functions gives

$$\begin{aligned}\left|\mathbb{E}[\lambda(n)\mathbf{1}_{F_r^c}\varphi(T_r)]\right| \le{}& 2\beta_{M,S} + B_{M,S,K} + 2C_{M,S,K} \\ &+ \left|\mathbb{E}_{\pi_S\otimes\rho_{M,S}K_r(m)}[\lambda(d)\lambda(m)\mathbf{1}_{H_r}\varphi(T_r)]\right|.\end{aligned}$$

The last term factors through the average of $\lambda(d)$ and is at most $\ell_S = \prod_{p\in S}(p-1)/(p+1)$. Thus

$$\left|\mathbb{E}[\lambda(n)\mathbf{1}_{F_r^c}\varphi(T_r)]\right| \le \ell_S + 2\beta_{M,S} + B_{M,S,K} + 2C_{M,S,K}. \tag{36}$$

With the same order of limits this tends to zero, uniformly in r, φ. Since F_r is observable and $q_{M,r} = \lambda(n)$ on F_r, conditional duality identifies the supremum over φ with $\mathbb{E}[\mathbf{1}_{F_r^c}|q_{M,r}|]$. This proves (32). The second-moment conclusion follows from $|q_{M,r}| \le 1$. For a sign target, the minimal absolute and quadratic risks are respectively $1 - \mathbb{E}|q_{M,r}|$ and $1 - \mathbb{E}q_{M,r}^2$, giving the asserted limits.

4.6 Erdős–Kac on the harmonic window

The classical theorem [1] first applies to $\omega(n)$ under uniform sampling from $[1, x]$. In that experiment,

$$\mathbb{E}(\Omega(n) - \omega(n)) \le \sum_p \frac{1}{p(p-1)} < \infty,$$

so Markov's inequality replaces ω by Ω on the scale $\sqrt{\log\log x}$. Continuity of the limiting normal distribution then gives uniform convergence of the distribution functions. For $x \in [M, M^2]$, the centering $\log\log x$ differs from V by at most $\log 2 + o(1)$, and the standard deviations are uniformly equivalent. The discrepancy therefore tends to zero simultaneously for all such x and all thresholds.

Apply partial summation to $A_r(x) = \sum_{n\le x}\mathbf{1}_{\{\Omega(n)\le r\}}$. The boundary terms in the normalized harmonic sum are $O(1/T_M)$, and the integral has weight $dt/(tT_M)$ on $[M, M^2]$ and total mass $1 + o(1)$. The preceding uniform convergence implies $\Delta_M \to 0$, uniformly in the integer threshold r. This also proves (31).

4.7 Removing the final bit and translating the anchor

Without the final bit, $\{\Omega(n) < r\}$ remains observable through the absorbing symbol or the transcript length. Multiplying the labels on this event gives success at least $\mathbb{P}(\Omega(n) < r)$, and the optimum is bounded above by the optimum with the bit. Moreover,

$$\sup_r \mathbb{P}(\Omega(n) = r) \le 2\Delta_M + \frac{1}{\sqrt{2\pi V}} \longrightarrow 0.$$

Thus the same reconstruction profile holds without the bit. For the posterior $\widetilde{q}$ without the bit, the tower property and convexity give

$$\mathbb{E}|\widetilde{q} - \lambda(n)\mathbf{1}_{F_r}| \le \mathbb{E}[\mathbf{1}_{F_r^c}|q_{M,r}|] + 2\mathbb{P}(\Omega(n) = r).$$

The stronger signed conclusion (32) therefore also survives.

The law of $n+1$ for $n \sim \mu_M$ has total variation exactly $1/(MT_M)$ from μ_M on the integer being observed. This follows by telescoping the differences between $1/(u-1)$ and $1/u$ on the common support and including the two endpoints. Appending a common sampling kernel does not increase this distance. The reconstruction and signed profiles consequently transfer to the anchor $n+1$, uniformly in r.

The vanishing posterior on the nonterminated event is a statement about its unnormalized mass. At a fixed Gaussian threshold that event has limiting mass $1-\Phi(t)>0$, so one may condition on it. The proof does not assert uniformity after division by a vanishing event probability, nor does it cover stopping rules that reveal hidden occurrence counts or nontrivial cofactors.

5 Two anchors: reconstruction and dependence

We now observe $u_0 = n$ and $u_1 = an+1$, where $a \in \{1,3\}$ is fixed and $n \sim \mu_M$. Anchor i supplies r_i labels according to either sampling protocol. Within an anchor sampled with replacement, the draws are conditionally independent. The two chains may be coupled arbitrarily, provided that each has its prescribed conditional marginal law.

Without replacement, let D_i be the product of all removed occurrences; with replacement, let D_i be the product of the distinct labels seen. In either case $D_i \mid u_i$. Set

$$m_i = u_i/D_i, \qquad R = \log m_0 + \log m_1, \qquad s_M = \frac{1}{r_0+1} + \frac{1}{r_1+1}.$$

Only D_0, D_1 are observed. The cofactors m_i enter the proof.

5.1 A logarithmic remainder estimate

Lemma 5.1 (Unobserved logarithmic mass)**.** *For $u = n, n+1$ or $3n+1$, either sampling protocol satisfies*

$$\mathbb{E}\log\frac{u}{D_r} \le \frac{120L}{r+1} + 42 \qquad (M \ge 3,\ r \ge 0). \tag{37}$$

In particular, $\mathbb{E}R \le 120Ls_M + 84$.

Proof. For all three forms, $u < M^3$ and $\mathbb{P}(d \mid u) \le 7/d$. The bound $2/d$ already holds for $u = n$. For $u = n+1 = dk$, $1/(dk-1) \le 4/(3dk)$; for $u = 3n+1 = dk$, $1/(dk-1) \le 10/(9dk)$. Dropping the congruence restrictions and summing over intervals of multiplicative length at most M gives respectively $(4/3d)(1+1/L)$ and $(10/3d)(1+1/L)$, both at most $7/d$. Therefore $\mathbb{E}v_p(u) \le 7/(p-1)$.

At a fixed u, mark the prime occurrences (p,j) and draw these objects independently with weights $\log p/\log u$. The order of first appearance of distinct objects has the without-replacement law. Every occurrence seen during the first r independent draws has therefore been removed within the first r removals. It follows that

$$\mathbb{E}\left[\log\frac{u}{D_r}\middle| u\right] \le \sum_p v_p(u)\log p\left(1-\frac{\log p}{\log u}\right)^r.$$

For $r \ge 1$ the last power is at most $p^{-r/(3L)}$. The elementary Chebyshev estimate $\vartheta(x) \le C_\vartheta x$, with $C_\vartheta = 4\log 2$, and partial summation imply

$$\sum_p \frac{\log p}{p^{1+\alpha}} \le C_\vartheta(1+1/\alpha), \qquad \sum_p \frac{\log p}{p(p-1)} < 3.$$

The first estimate follows from $\prod_{m<p\le 2m} p \mid \binom{2m}{m} \le 2^{2m}$ and dyadic summation; the second follows by comparison with $2\sum_{k\ge 2}(\log k)/k^2$. Thus

$$\mathbb{E}\log(u/D_r) \le 7\{C_\vartheta(1+3L/r)+3\} \le \frac{120L}{r+1} + 42.$$

Here $1/r \le 2/(r+1)$, $168\log 2 < 120$, and $28\log 2 + 21 < 42$. The case $r = 0$ follows from $\log u \le 3L$.

With replacement,

$$\log(u/D_r) = \sum_p (v_p(u) - 1)_+ \log p + \sum_{p|u} (\log p)\mathbf{1}_{\{p \text{ is not seen}\}}.$$

The first term has expectation at most 21; conditional independence bounds the probability in the second by $p^{-r/(3L)}$. The same constants follow. Finally, the two remainder expectations add without any assumption on the coupling between anchors. □

5.2 Reconstruction with no false answer

Because $(D_0, D_1) = (a, D_1) = 1$, the observations specify one residue class modulo $Q = D_0 D_1$:

$$n \equiv 0 \pmod{D_0}, \qquad an \equiv -1 \pmod{D_1}.$$

If $D_1 = 1$, the second congruence imposes no condition. Let $\mathcal{D}_M$ return the unique representative in I_M when there is one, and otherwise abstain. The true origin is always a candidate, so any returned value is correct. The sufficient condition $Q \geq M^2$ makes the candidate unique.

Theorem 5.2 (Certified reconstruction from two anchors)**.** *For either protocol at each anchor and every admissible coupling,*

$$e_M := \mathbb{P}(\mathcal{D}_M = \perp) \leq \min\left\{1,\ 2\sqrt{60 s_M + 42/L} + \frac{1}{ML}\right\}. \tag{38}$$

In particular, $r_0, r_1 \to \infty$ is sufficient for reconstruction with probability tending to one.

Proof. The identity $Q = n(an+1)e^{-R} \geq n^2 e^{-R}$ gives uniqueness when $n \geq M^{1+\delta}$ and $R \leq 2\delta L$, where $0 < \delta \leq 1$. The lower part of the harmonic window has mass at most $\delta + 1/(ML)$. Markov's inequality and Lemma 5.1 give

$$e_M \leq \delta + \frac{1}{ML} + \frac{60 s_M + 42/L}{\delta}.$$

Optimize in δ when $60 s_M + 42/L \leq 1$, and otherwise use the trivial bound one. □

The budgets may diverge arbitrarily slowly. On the observable success event, every deterministic sign $W_M(n) \in \{-1, 1\}$ is known, so

$$\|\mathbb{E}[W_M(n) \mid \mathcal{O}]\|_1 \geq 1 - e_M, \tag{39}$$

where $\mathcal{O}$ denotes the joint observation. For $a = 1$, take $r_0 = r_1 = r_M \to \infty$ with $r_M = o(\log\log M)$ without replacement, or $r_M = o(\log M)$ with replacement. Each separate transcript has vanishing optimal reconstruction probability, by the one-anchor results and the translation comparison above. Their joint observation reconstructs the origin with probability tending to one.

The same mechanism applies to two fixed affine forms $u_i = a_i n + b_i$, with $a_i \geq 1$ and $\Delta = a_0 b_1 - a_1 b_0 \neq 0$. The reduced moduli $q_i = D_i/(a_i, D_i)$ determine a compatible class whose modulus is at least $D_0 D_1/(a_0 a_1 |\Delta|)$, since $(D_0, D_1) \mid \Delta$. For large M, $u_i \geq a_i n/2$. With $R = \log(u_0/D_0) + \log(u_1/D_1)$, the abstention probability is therefore at most

$$\delta + \frac{1}{ML} + \frac{\mathbb{E}R + \log(4|\Delta|)}{2\delta L}.$$

Thus $\mathbb{E}R = o(L)$ suffices in this setting as well.

5.3 Separation from the product of the marginals

Let $P_M = \mathcal{L}(D_0, D_1)$ and $Q_M = \mathcal{L}(D_0) \otimes \mathcal{L}(D_1)$. For $H > 1$, let A_H be the observable event that there exist $n \in I_M$ and integers $1 \le k_0, k_1 \le H$ such that

$$D_0 k_0 = n, \qquad D_1 k_1 = an + 1.$$

The true origin witnesses this event whenever $R \le \log H$, hence

$$P_M(A_H) \ge 1 - \frac{120 L s_M + 84}{\log H}.$$

On the other hand, $\mathbb{P}(D_i = d) \le \mathbb{P}(d \mid u_i) \le 7/d$. A union bound over candidate origins and cofactors gives

$$\begin{aligned} Q_M(A_H) &\le 49 \sum_{n=M}^{M^2-1} \frac{1}{n(an+1)} \left(\sum_{k \le H} k \right)^2 \\ &\le 49 H^4 \sum_{n=M}^{\infty} \frac{1}{n(n+1)} = \frac{49 H^4}{M}. \end{aligned}$$

Taking $H = M^{1/8}$ yields

$$d_{\mathrm{TV}}(P_M, Q_M) \ge 1 - 960 s_M - \frac{672}{L} - \frac{49}{\sqrt{M}}. \tag{40}$$

Thus the distance tends to one when both budgets diverge. By projection the same conclusion holds for the full transcripts. The comparison is with the product of the unconditional marginals; conditional independence given n does not remove the dependence created by the common origin.

5.4 Mutual information of two revealed divisors

All entropies in this paper use natural logarithms. For every admissible coupling,

$$\frac{I(D_0; D_1)}{\log M} \longrightarrow \frac{3}{2} \qquad (r_0, r_1 \to \infty). \tag{41}$$

Indeed, sum–integral comparison gives

$$H(n) = \mathbb{E} \log n + \log T_M = \frac{3}{2} L + \log L + O(M^{-1}).$$

At fixed D_i, the cofactor m_i and n determine one another. Gibbs' inequality with the probability $q(m) = 1/[\zeta(2) m^2]$ therefore implies

$$H(n \mid D_i) = H(m_i \mid D_i) \le 2 \mathbb{E} \log m_i + \log \zeta(2).$$

The three-variable information identity gives

$$I(D_0; D_1) \ge H(n) - H(n \mid D_0) - H(n \mid D_1) \ge H(n) - 2 \mathbb{E} R - 2 \log \zeta(2).$$

For the reverse bound, D_0 belongs to the set of divisors of n. The chain rule and Jensen's inequality yield

$$I(D_0; D_1) \le H(n) + H(D_0 \mid n) \le H(n) + \log \mathbb{E} \tau(n) \le H(n) + \log(4L + 2),$$

using $\mathbb{E} \tau(n) \le 2 \sum_{d \le M^2} 1/d \le 4L + 2$. Since $\mathbb{E} R = o(L)$, the two bounds prove (41). The upper bound uses the divisor support; it is not an upper bound for arbitrarily long transcripts with additional shared randomness.

6 One growing observation and one auxiliary prime

The preceding certified decoder uses two growing budgets. A decoder allowed to make an error needs only one. Give an anchor u_j r observations, and let the other anchor reveal only its first prime label P. Let $D \mid u_j$ be the product of removed occurrences without replacement, or of distinct observed labels with replacement. Put $m = u_j/D$. The first-label law is the same in both protocols:

$$\mathbb{P}(P = p \mid n) = \frac{v_p(u_{1-j}) \log p}{\log u_{1-j}}.$$

The coupling between the two observations may be arbitrary with the prescribed conditional marginals. The cofactor m is not revealed.

6.1 A small cofactor and a large auxiliary label

Lemma 5.1 gives, in both orientations,

$$\mathbb{E} \log m \le B_{M,r} := \frac{120L}{r+1} + 42. \tag{42}$$

Lemma 6.1 (The size of the auxiliary label)**.** *For every $t > 0$ and every prime p,*

$$\mathbb{P}(P \le e^t) \le \min\left\{1, \frac{20t + 42}{L}\right\}, \tag{43}$$

$$\mathbb{P}(P = p) \le \frac{7 \log p}{L(p-1)} \le \frac{14 \log p}{Lp}. \tag{44}$$

Proof. The point-mass bound follows from $\log u_i \ge L$ and $\mathbb{E}v_p(u_i) \le 7/(p-1)$. For $H \ge 2$, Chebyshev's bound and partial summation give

$$\sum_{p \le H} \frac{\log p}{p-1} = \sum_{p \le H} \frac{\log p}{p} + \sum_{p \le H} \frac{\log p}{p(p-1)} \le C_\vartheta(1 + \log H) + 3.$$

It follows that

$$\mathbb{P}(P \le H) \le \frac{7C_\vartheta \log H + 7(C_\vartheta + 3)}{L} \le \frac{20 \log H + 42}{L}.$$

For $1 < H < 2$, the probability is zero. Combine this with the trivial upper bound one. □

6.2 The least compatible cofactor

Suppose first that $D \mid n$ and $P \mid an + 1$. Since $(aD, P) = 1$, there is a unique $h \in \{1, \dots, P-1\}$ such that

$$aDh \equiv -1 \pmod{P}. \tag{45}$$

Return $\widehat{n} = Dh$ if it belongs to I_M, and otherwise abstain. In the reverse orientation, $D \mid an + 1$, $P \mid n$, and $(D, P) = 1$. Use the unique $h \in \{1, \dots, P-1\}$ satisfying

$$Dh \equiv 1 \pmod{P}, \tag{46}$$

and return $\widehat{n} = (Dh - 1)/a$ when this is an integer in I_M. Under the true law the inverses exist. In either orientation, the actual cofactor m satisfies the congruence, so $m < P$ implies $m = h$ and exact reconstruction.

Theorem 6.2 (Asymmetric reconstruction)**.** *For both protocols, either orientation and every admissible coupling, the preceding decoder satisfies, for all $t > 0$,*

$$\mathbb{P}(\widehat{n} \neq n) \leq \frac{B_{M,r}}{t} + \frac{20t + 42}{L}. \tag{47}$$

In particular,

$$\mathbb{P}(\widehat{n} \neq n) \leq \min\left\{1,\ 2\sqrt{\frac{2400}{r+1} + \frac{840}{L}} + \frac{42}{L}\right\}. \tag{48}$$

Thus $r = r_M \to \infty$ and a single label at the other anchor suffice for reconstruction with probability tending to one.

Proof. On $\{m \leq e^t < P\}$ the decoder is correct. A union bound, Markov's inequality and Lemma 6.1 give

$$\mathbb{P}(\widehat{n} \neq n) \leq \mathbb{P}(\log m > t) + \mathbb{P}(P \leq e^t) \leq B_{M,r}/t + (20t + 42)/L.$$

No independence of m and P is used. Choose $t = \sqrt{B_{M,r}L/20}$ to obtain (48). □

This decoder does not certify its own success. For example, take $a = 1$, $M = 8$ and $n = 50$. Two removals with labels $(2, 5)$ give $D = 10$, with a nontermination bit, while the anchor 51 can reveal $P = 3$. The congruence gives $h = 2$ and returns $20 \in [8, 64)$. The origin 20 admits the same nonterminated transcript $(2, 5)$ and the auxiliary prime 3 at 21. The theorem controls the average error; it does not provide a zero-error certificate.

Corollary 6.3 (Positive unbalanced budgets)**.** *Suppose both budgets are positive for all sufficiently large M. If $\max(r_0(M), r_1(M)) \to \infty$, the joint observation reconstructs n in probability for every admissible coupling.*

Use the larger-budget anchor and the first label of the other. The constants are common to both orientations, so the choice of orientation may vary with M. For any deterministic sign $W_M(n)$, assign an arbitrary answer on abstention. Then

$$\|\mathbb{E}[W_M(n) \mid D, P]\|_1 \geq 1 - 2\mathbb{P}(\widehat{n} \neq n) \longrightarrow 1. \tag{49}$$

The factor two accounts for possibly incorrect returned values. Section 10 will prove necessity of the diverging-maximum condition under conditional independence.

6.3 Dependence persists with one auxiliary label

Let $P_M = \mathcal{L}(D, P)$ and $Q_M = \mathcal{L}(D) \otimes \mathcal{L}(P)$. For $H > 1$ and $K > H$, let $C_{H,K}$ be the event that $P > K$ and there exists an integer $1 \leq h \leq H$ for which

$$N(D, h) = \begin{cases} Dh, & j = 0, \\ (Dh - 1)/a, & j = 1 \end{cases}$$

is an integer in I_M and P divides its other anchor.

Proposition 6.4 (Asymmetric separation)**.** *For every $t > 0$,*

$$d_{\mathrm{TV}}(P_M, Q_M) \geq 1 - \frac{B_{M,r}}{t} - \frac{40t + 42}{L} - 42e^{-t}. \tag{50}$$

In particular, this distance tends to one if $r_M \to \infty$.

Proof. Under the true law, the actual cofactor witnesses $C_{H,K}$ when $m \le H$ and $P > K$. Thus

$$P_M(C_{H,K}) \ge 1 - \frac{B_{M,r}}{\log H} - \frac{20 \log K + 42}{L}.$$

Under Q_M, condition on D. For each admissible h, the other anchor v of the candidate is less than M^3. By (44),

$$\sum_{\substack{p \mid v \\ p > K}} \mathbb{P}(P = p) \le \frac{14}{LK} \sum_{p \mid v} \log p \le \frac{14 \log v}{LK} \le \frac{42}{K}.$$

A union bound over at most H candidates gives $Q_M(C_{H,K}) \le 42H/K$. Set $H = e^t$ and $K = e^{2t}$. For the limit, choose $t = \sqrt{B_{M,r} L/40}$. Since $B_{M,r} \ge 42$, this tends to infinity, whereas t/L and $B_{M,r}/t$ tend to zero if $r_M \to \infty$. □

The same separation holds for full transcripts by projection. For $a = 1$, the regime $r_M \to \infty$ with $r_M = o(\log \log M)$ without replacement, or $r_M = o(\log M)$ with replacement, is especially instructive. Both individual observations have vanishing optimal reconstruction probability, while their union succeeds. Reconstruction permits evaluation of a prescribed sign at the recovered origin; it does not calculate that sign's arithmetic mean.

7 Information carried by one auxiliary prime

The mutual information of the asymmetric experiment has a different leading coefficient from that of two growing divisor observations. It can be computed from a single additive moment, without a joint limit law for the factorizations of the two anchors.

Theorem 7.1 (Asymmetric mutual information)**.** *For both orientations, either protocol at the growing anchor and every admissible coupling,*

$$\frac{I(D; P)}{\log M} \longrightarrow \frac{3}{4} \qquad (r_M \to \infty). \tag{51}$$

The marginal law of the auxiliary prime satisfies

$$\mathbb{E} \log P = \frac{3}{4} L + O(1), \qquad H(P) = \frac{3}{4} L + O(\log L), \tag{52}$$

with constants common to the anchors $n, n+1, 3n+1$.

7.1 An additive moment

The elementary form of Mertens' theorem is

$$\sum_{p \le x} \frac{\log p}{p} = \log x + O(1). \tag{53}$$

For completeness, combine $\log(N!) = \sum_{d \le N} \Lambda(d) \lfloor N/d \rfloor$, $\sum_{d \le N} \Lambda(d) = O(N)$ and $\log(N!) = N \log N + O(N)$ to obtain $\sum_{d \le N} \Lambda(d)/d = \log N + O(1)$. Prime powers of exponent at least two contribute a bounded amount. This proves (53); see also [13, Remark 1]. Partial summation gives

$$\sum_{p \le x} \frac{(\log p)^2}{p} = \frac{1}{2} (\log x)^2 + O(\log x). \tag{54}$$

Define $F(u) = \sum_p v_p(u)(\log p)^2$. For $(a,b) \in \{(1,0),(1,1),(3,1)\}$,

$$A_{a,b}(x) := \sum_{n \le x} F(an+b) = \frac{1}{2}x(\log x)^2 + O_a(x \log x). \tag{55}$$

Indeed, if $p \nmid a$, the congruence $an + b \equiv 0 \pmod{p^k}$ has one residue class, so its count is $x/p^k + O(1)$. Primes dividing a do not occur. With $Y = a\lfloor x \rfloor + b \ll_a x$,

$$A_{a,b}(x) = x \sum_{\substack{p^k \le Y \\ p \nmid a}} \frac{(\log p)^2}{p^k} + O\left(\sum_{p^k \le Y} (\log p)^2 \right).$$

The $k = 1$ main term is given by (54); the higher-power main term is $O(x)$. The $k = 1$ error is at most $(\log Y)\vartheta(Y) = O_a(x \log x)$, and the higher-power error is $O_a(\sqrt{x}(\log x)^3) = O_a(x \log x)$. Excluding the finitely many primes dividing a and replacing $\log Y$ by $\log x + O_a(1)$ preserves the stated error.

7.2 Harmonic averaging of the prime size

Conditionally on the origin,

$$\mathbb{E}[\log P \mid n] = \frac{F(an+b)}{\log(an+b)}.$$

Since $\log(an+b) = \log n + O_a(1)$ and $F(an+b) \le (\log(an+b))^2$, replacing the denominator by $\log n$ costs $O_a(1)$ in the harmonic expectation. Apply partial summation to (55), with $w(t) = 1/(t \log t)$, on $[M, M^2]$. The error contributes $O_a(L)$, because its boundary terms are $O_a(1)$ and $\int_M^{M^2} t \log t \, |w'(t)| \, dt = O(L)$. The derivative of $\frac{1}{2}t(\log t)^2$ is $\frac{1}{2}(\log t)^2 + \log t$, giving

$$\sum_{M \le n < M^2} \frac{F(an+b)}{n \log n} = \int_M^{M^2} \left(\frac{\log t}{2t} + \frac{1}{t} \right) dt + O_a(L) = \frac{3}{4}L^2 + O_a(L).$$

Division by $T_M = L + O(M^{-1})$ proves the first part of (52).

7.3 From label size to entropy

Write $q_M(p) = \mathbb{P}(P = p)$. Since $p < M^3$, (44) implies $q_M(p) \le 42/p$, hence

$$H(P) \ge \mathbb{E} \log P - \log 42. \tag{56}$$

For the reverse bound, apply Gibbs' inequality to the probability $\nu_M(p) = 1/(Z_M p)$ on primes $p < M^3$, where $Z_M = \sum_{p < M^3} 1/p \le 1 + 3L$. Then

$$H(P) \le \mathbb{E} \log P + \log Z_M \le \mathbb{E} \log P + \log(1 + 3L). \tag{57}$$

This proves the entropy statement in (52).

At a fixed D, the cofactor m and origin n determine one another. Gibbs' inequality with $k \mapsto 1/[\zeta(2)k^2]$ gives $H(n \mid D) \le 2B_{M,r} + \log \zeta(2)$. Given n, the auxiliary prime belongs to at most $3L/\log 2$ possible labels. The entropy chain rule yields

$$\begin{aligned} 0 \le H(P) - I(D;P) = H(P \mid D) &\le H(n \mid D) + H(P \mid n, D) \\ &\le 2B_{M,r} + \log \zeta(2) + \log(3L/\log 2). \end{aligned} \tag{58}$$

The right side is $o(L)$ when $r_M \to \infty$, proving (51) for every admissible coupling. The coefficients $3/4$ here and $3/2$ in (41) refer to different observed variables.

8 Exact two-anchor profiles with fixed replacement budgets

Fix positive integers r_0, r_1, and put $R = r_0 + r_1$. At $u_0 = n, u_1 = an + 1$, $a \in \{1, 3\}$, observe r_i draws with replacement. All draws, including those at different anchors, are independent conditionally on n:

$$\mathbb{P}(P_{i,j} = p \mid n) = \frac{v_p(u_i) \log p}{\log u_i}.$$

Let D_i be the product of the distinct labels at anchor i and $D = D_0 D_1$. Write $\mathcal{R}_M^{+,+}(r_0, r_1)$ for the optimal reconstruction probability from the two ordered transcripts.

Theorem 8.1 (The reciprocal-binomial deficit)**.** *For fixed $r_0, r_1 \geq 1$,*

$$\mathcal{R}_M^{+,+}(r_0, r_1) = 1 - \frac{r_0!\, r_1!}{(r_0 + r_1)!} + O_{a,r_0,r_1}(1/\log M). \tag{59}$$

More precisely,

$$\mathbb{P}(D \leq n) = \binom{r_0 + r_1}{r_0}^{-1} + O_{a,r_0,r_1}(1/L), \tag{60}$$

and uniformly over all decoders,

$$\mathbb{P}\{\widehat{n} = n,\ D \leq n\} = O_{a,r_0,r_1}(1/L). \tag{61}$$

The decoder that returns the least positive representative of

$$z \equiv 0 \pmod{D_0}, \qquad az \equiv -1 \pmod{D_1}, \tag{62}$$

abstaining if it lies outside I_M, succeeds exactly on $D > n$. It attains (59).

One prime per anchor gives success $1/2 + O_a(1/L)$. Budgets $(1, k)$ give $k/(k+1) + O_{a,k}(1/L)$, while $(2, 2)$ and $(3, 3)$ give respectively $5/6 + O_a(1/L)$ and $19/20 + O_a(1/L)$. These errors are not asserted uniformly for growing budgets.

8.1 A positive approximation to the likelihood

For $u = n, n+1, 3n+1$, put $\Delta(u) = \log(u/\operatorname{rad}(u))$. Since $\mathbb{P}(d \mid u) \leq 7/d$,

$$\mathbb{E}\Delta(u) \leq 7 \sum_p \sum_{k \geq 2} \frac{\log p}{p^k} = 7 \sum_p \frac{\log p}{p(p-1)} = O(1). \tag{63}$$

Replace each factor $v_p(u_i)$ in the likelihood by $\mathbf{1}_{p \mid u_i}$. This gives a positive submeasure whose conditional mass at n is

$$\left(\frac{\log \operatorname{rad}(n)}{\log n}\right)^{r_0} \left(\frac{\log \operatorname{rad}(an+1)}{\log(an+1)}\right)^{r_1}.$$

The inequality $1 - x^r \leq r(1-x)$ for $0 \leq x \leq 1$ shows that the removed mass is $O_R(1/L)$. Since $0 \leq \log(an+1) - \log n \leq \log(a+1)$, replacing $(\log n)^{r_0}(\log(an+1))^{r_1}$ by $(\log n)^R$ costs $O_{a,R}(1/L)$ in ℓ^1. The true joint law of the origin and transcripts is therefore within $O_{a,R}(1/L)$ in ℓ^1 of

$$\begin{aligned} \widetilde{w}_M(n, \mathbf{p}, \mathbf{q}) = \frac{\mathbf{1}_{I_M}(n)}{n T_M (\log n)^R} &\prod_{j=1}^{r_0} (\mathbf{1}_{p_j \mid n} \log p_j) \\ &\cdot \prod_{j=1}^{r_1} (\mathbf{1}_{q_j \mid an+1} \log q_j). \end{aligned} \tag{64}$$

This applies to every bounded test, including the success indicator of a decoder depending on M.

8.2 Repetitions and weighted prime convolutions

Each ordered transcript determines a partition π_i of $\{1,\dots,r_i\}$: two positions are in the same block exactly when they carry the same prime. Order blocks by their least position. The partition pair and the assigned block primes then specify the transcript uniquely. If the block sizes are $b_1,\dots,b_k$, then $\sum_j b_j = R$. All block primes are distinct since the anchors are coprime. Set

$$W_{\mathbf{b}}(\mathbf{p}) = \prod_{j=1}^{k} (\log p_j)^{b_j}, \qquad d(\mathbf{p}) = \prod_{j=1}^{k} p_j.$$

Lemma 8.2 (Weighted simplex sums)**.** *For fixed $\mathbf{b} = (b_1,\dots,b_k)$, with $b_j \ge 1$ and $R = \sum_j b_j$,*

$$F_{\mathbf{b}}(x) := \sum_{\substack{p_1\cdots p_k \le x \\ p_j \text{ distinct}}} \prod_{j=1}^{k} \frac{(\log p_j)^{b_j}}{p_j} = \frac{\prod_j (b_j-1)!}{R!}(\log x)^R + O_{\mathbf{b}}((1+\log x)^{R-1}), \tag{65}$$

$$W_{\mathbf{b}}(x) := \sum_{p_1\cdots p_k \le x} \prod_{j=1}^{k} (\log p_j)^{b_j} \ll_{\mathbf{b}} x(1+\log x)^{R-1}. \tag{66}$$

The first estimate remains valid when fixed finite sets of primes are excluded from selected coordinates; the error constant may depend on those sets.

Proof. By (53), $A(t) := \sum_{p \le e^t} (\log p)/p = t + O(1)$ for $t \ge 0$. Without distinctness and with $T = \log x$, the first sum is

$$\int_{t_1+\cdots+t_k \le T} \prod_{j=1}^{k} t_j^{b_j-1}\, dA(t_1)\cdots dA(t_k).$$

Replace $dA(t_j)$ by dt_j, one coordinate at a time. Keeping the others fixed, integration by parts against $d(A(t)-t)$ on $[0,U]$ bounds the change by $O_{b_j}((1+U)^{b_j-1})$. The other weighted masses total $O_{\mathbf{b}}((1+T)^{R-b_j})$. All replacements therefore cost $O_{\mathbf{b}}((1+T)^{R-1})$. The remaining Dirichlet integral is

$$\int_{t_1+\cdots+t_k \le T} \prod_{j} t_j^{b_j-1}\, dt = \frac{\prod_j (b_j-1)!}{R!} T^R.$$

A collision $p_i = p_j$ contributes at most

$$O_{\mathbf{b}}\left((1+T)^{R-b_i-b_j} \sum_{p} \frac{(\log p)^{b_i+b_j}}{p^2}\right) = O_{\mathbf{b}}((1+T)^{R-2}).$$

Excluding finitely many primes in coordinate j costs $O_{\mathbf{b}}((1+T)^{R-b_j})$, with dependence on the excluded set, and is absorbed by the stated error.

For the second estimate, Chebyshev gives $\sum_{p\le y}(\log p)^b \ll_b y(1+\log y)^{b-1}$. Sum first over p_k and then use the weighted-convolution upper bound for the other $k-1$ coordinates:

$$W_{\mathbf{b}}(x) \ll_{\mathbf{b}} x(1+\log x)^{b_k-1}(1+\log x)^{R-b_k}.$$

For $k=1$ this is the weighted Chebyshev bound itself. □

8.3 The mass below the origin

Fix a partition pair $\pi = (\pi_0,\pi_1)$ with block sizes $\mathbf{b}$, assign its primes to the two anchors, and let d_i be their product at anchor i. Put

$$A_{\pi,a}(x) = \sum_{n\le x} \sum_{\substack{\mathbf{p} \text{ distinct} \\ d_0 \mid n,\ d_1 \mid an+1 \\ d_0 d_1 \le n}} W_{\mathbf{b}}(\mathbf{p}).$$

When no prime at the second anchor divides a, the divisibilities specify one class modulo $d = d_0 d_1$. Its number of representatives in $[d, x]$ is $x/d + O(1)$, uniformly in the class and modulus. Lemma 8.2 gives

$$\begin{aligned} A_{\pi,a}(x) &= c_\pi x(\log x)^R + O_{a,\pi}(x(1+\log x)^{R-1}), \\ c_\pi &= \frac{\prod_{B\in\pi_0\cup\pi_1}(|B|-1)!}{R!}. \end{aligned} \tag{67}$$

The unweighted estimate (66) absorbs the CRT counting error. Excluding primes dividing a from the second anchor does not change the leading term. Partial summation with $f(x) = 1/[x(\log x)^R]$ on $[M, M^2]$ gives main term $c_\pi(L + R\log 2)$ and error $O_{a,\pi}(1)$, using

$$-f'(x) = \frac{1}{x^2(\log x)^R} + \frac{R}{x^2(\log x)^{R+1}}.$$

Changing integer endpoints costs $O_{a,R}(1/M)$, since the numerator at one origin is at most $(\log \mathrm{rad}(n))^{r_0}(\log \mathrm{rad}(an+1))^{r_1} \ll_{a,R} (\log n)^R$. After division by T_M, the mass of $D \le n$ with partition π under (64) is $c_\pi + O_{a,\pi}(1/L)$.

The constants sum by

$$\sum_{\pi \text{ a partition of } \{1,\dots,r\}} \prod_{B\in\pi} (|B|-1)! = r!. \tag{68}$$

A permutation decomposes uniquely into cycles; a prescribed block of b elements supports $(b-1)!$ cycles. These are the weights in the Ewens formula with parameter one [6, formula (2.19)]. Hence

$$\sum_{\pi_0,\pi_1} c_{(\pi_0,\pi_1)} = \frac{r_0! r_1!}{R!} = \binom{R}{r_0}^{-1}.$$

The number of partitions is fixed. Restoring the true law proves (60).

8.4 Every decoder has negligible success below the origin

For each transcript with partition π, a deterministic decoder selects at most one origin. On $D = d \le n$, it satisfies $n \ge \max(M, d)$ and $\log n \ge L$. Its success mass under (64) is at most

$$\frac{1}{T_M L^R} \sum_{\substack{\mathbf{p} \text{ distinct} \\ d \le M^2}} \frac{W_{\mathbf{b}}(\mathbf{p})}{\max(M,d)} \le \frac{1}{T_M L^R} \left\{ \frac{W_{\mathbf{b}}(M)}{M} + F_{\mathbf{b}}(M^2) \right\} = O_{a,\pi}(1/L).$$

Sum over the partitions and restore the true law to obtain (61). The uniformity in deterministic decoders also covers their independent randomized mixtures.

8.5 Optimality of the least CRT representative

Under the true law, (62) has a representative $z \in \{1, \dots, D-1\}$; its class is nonzero since $D_1 \ge 2$. If $D > n$, it equals n. If $D \le n$, it is strictly smaller. Equality $D = n$ is impossible, since then D_1 would divide both n and $an+1$. The decoder thus succeeds exactly on $D > n$. Every decoder also satisfies

$$\mathbb{P}(\widehat{n} = n) \le \mathbb{P}(D > n) + \mathbb{P}(\widehat{n} = n, D \le n).$$

Equations (60)–(61) give the matching upper bound, proving Theorem 8.1.

The same profile survives any additional observation taking at most K values, with K fixed independently of M. For each label transcript it allows at most K candidate origins, multiplying

the low-product bound by at most K. For randomized side information, append its conditional kernel to both measures in the ℓ^1 comparison. The CRT decoder can ignore it.

Conditional independence is used in the product likelihood (64). The deficit is not asserted for arbitrary couplings. Also, $D > n$ involves the hidden origin, so the least-representative decoder does not supply an observable zero-error certificate.

9 Fixed budgets without replacement and mixed protocols

The anchors remain $n, an+1$ with $a \in \{1,3\}$ and $n \sim \mu_M$, and the two chains are independent given n. All budgets in this section are positive and fixed before $M \to \infty$. For $0 < x < 1$, define

$$f_r^-(x) = \frac{[-\log(1-x)]^{r-1}}{(r-1)!}, \qquad f_r^+(x) = rx^{r-1}.$$

The signs $-$ and $+$ denote sampling without and with replacement. These marginal densities belong to the classical theory of size-biased partitions and successive fractions [4, 5, 6]. The proof below derives the required measures by counting and identifies the two-anchor decision limit on the small-modulus region; it makes no independence assumption on the full arithmetic factorizations.

Theorem 9.1 (Fixed without-replacement and mixed profiles)**.** *For r removals at the first anchor and s at the second,*

$$\mathcal{R}_M^{-;-}(r,s) = 1 - \gamma_{r,s} + o_{r,s,a}(1), \qquad \gamma_{r,s} = \int_{\substack{x,y>0\\ x+y\le 1}} f_r^-(x) f_s^-(y)\, dx\, dy. \tag{69}$$

In particular,

$$\mathcal{R}_M^{-;-}(r,1) = 1 - 2^{-r} + o_{r,a}(1), \tag{70}$$

$$\mathcal{R}_M^{-;-}(2,2) = \frac{7}{4} - \frac{\pi^2}{12} + o_a(1). \tag{71}$$

For r draws with replacement at one anchor and s removals at the other, in either orientation,

$$\mathcal{R}_M^{+;-}(r,s) = 1 - \frac{1}{(r+1)^s} + o_{r,s,a}(1), \tag{72}$$

where each budget is attached to the listed protocol. These conclusions hold with or without final termination bits.

Thus $(r,s) = (2,1)$ without replacement gives success $3/4 + o(1)$, and mixed budgets $(2,2)$ give $8/9 + o(1)$. The little-oh terms are not asserted uniformly for growing budgets.

9.1 An arithmetic measure on the simplex

Lemma 9.2 (Weighted counting below the origin)**.** *Fix $k \ge 1$ and assign k labels to the two anchors, or to a single form chosen from $n, n+1, 3n+1$. A tuple of distinct primes $(p_1, \dots, p_k)$ is admissible at n if each prime divides its assigned form. Let $d = \prod_j p_j$ and $S_k = \{x \in \mathbb{R}^k_{\ge 0} : \sum_j x_j \le 1\}$. The finite measure*

$$\nu_M(F) = \frac{1}{T_M} \sum_{n \in I_M} \frac{1}{n(\log n)^k} \sum_{\substack{(p_1,\dots,p_k)\ \text{admissible}\\ d \le n}} \left(\prod_j \log p_j\right) F\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_k}{\log n}\right) \tag{73}$$

converges weakly to Lebesgue measure on S_k. The same convergence holds for bounded functions whose discontinuities form a Lebesgue-null set.

Proof. It suffices to check monomials on the compact simplex. Fix nonnegative integers $m_1, \dots, m_k$, put $D = k + \sum_j m_j$ and $c_{\boldsymbol{m}} = \prod_j m_j!/D!$. Lemma 8.2, with $b_j = m_j + 1$, gives

$$\sum_{\substack{\prod_{p_j \le X} \\ p_j \text{ distinct}}} \frac{\prod_j (\log p_j)^{m_j+1}}{\prod_j p_j} = c_{\boldsymbol{m}} (\log X)^D + O_{\boldsymbol{m}}((1+\log X)^{D-1}), \tag{74}$$

$$W_{\boldsymbol{m}}(X) := \sum_{\prod p_j \le X} \prod_j (\log p_j)^{m_j+1} \ll_{\boldsymbol{m}} X(1+\log X)^{D-1}. \tag{75}$$

Excluding primes dividing affine coefficients does not alter the leading term. For each admissible tuple the divisibilities specify one class modulo d, with $X/d + O(1)$ representatives in $[d, X]$. Consequently,

$$A_{\boldsymbol{m}}(X) := \sum_{n \le X} \sum_{\substack{(p_j) \\ \text{admissible} \\ d \le n}} \prod_j (\log p_j)^{m_j+1} = c_{\boldsymbol{m}} X (\log X)^D + O_{\boldsymbol{m},a}(X(1+\log X)^{D-1}).$$

Partial summation with $1/[x(\log x)^D]$ on $[M, M^2]$ gives

$$\nu_M \left(\prod_j x_j^{m_j} \right) = c_{\boldsymbol{m}} + O_{\boldsymbol{m},a}(1/L).$$

The integer endpoints contribute no larger error, since the numerator at one origin is bounded by a constant times $(\log(an+1))^D$. As $c_{\boldsymbol{m}} = \int_{S_k} \prod_j x_j^{m_j}\, dx$, polynomial approximation proves weak convergence. Approximation outside a neighborhood of the discontinuity set proves the extension to bounded almost-everywhere continuous tests. □

The restriction $d \le n$ controls the sum of the $O(1)$ errors in the CRT counts through (75); it is part of the lemma's contract.

9.2 Marginal submeasures and sampling denominators

At an anchor u, distinguish one occurrence of each prime divisor. Retain only paths selecting these distinguished occurrences. Their labels are distinct and they form a submeasure of the true law; the marks are not observed. The path $(p_1, \dots, p_r)$ has weight

$$\frac{\prod_{j=1}^r \log p_j}{\prod_{j=1}^r (\log u - \sum_{h<j} \log p_h)}.$$

Put $K_r(x) = \prod_{j=1}^r (1 - x_1 - \cdots - x_{j-1})^{-1}$. For fixed $0 < \delta < 1$, also impose

$$\prod_{j=1}^r p_j \le n, \qquad \sum_{h<r} \log p_h \le (1-\delta) \log n.$$

Every denominator is then at least $\delta \log n$. Replacing $\log u = \log n + O_a(1)$ in the normalizers changes the weight by a relative factor $1 + O_{r,a,\delta}(1/L)$. Lemma 9.2 gives retained limiting mass

$$I_r(\delta) = \int_{S_r} K_r(x) \mathbf{1}_{x_1 + \cdots + x_{r-1} \le 1-\delta}\, dx. \tag{76}$$

The successive-fraction substitution $z_j = x_j/(1 - x_1 - \cdots - x_{j-1})$, $0 < z_j < 1$, sends $K_r(x)\, dx$ to $dz_1 \cdots dz_r$. Hence

$$\int_{S_r} K_r = 1, \qquad 1 - x_1 - \cdots - x_r = \prod_{j=1}^r (1 - z_j). \tag{77}$$

Thus $I_r(\delta) \uparrow 1$ as $\delta \downarrow 0$. More explicitly, $I_1(\delta) = 1$, and for $r \geq 2$,

$$1 - I_r(\delta) = \delta \sum_{j=0}^{r-2} \frac{(\log(1/\delta))^j}{j!}. \tag{78}$$

The limsup of the mass omitted from the true law is at most $1 - I_r(\delta)$. This controls the small sampling denominators, repetitions, early absorption and, at the affine anchor, observed products between n and $an+1$. Since the product of r independent uniforms on $(0,1)$ has density $(-\log z)^{r-1}/(r-1)!$, the projection $x \mapsto \sum_j x_j$ of $K_r(x)\,dx$ has density f_r^-.

9.3 The two-anchor low-product mass

Let D_i be the product of occurrences actually removed, ignoring absorbing symbols, and $d = D_0 D_1$. Conditional independence permits multiplication of the preceding submeasures at each origin. For their conditional masses $k_0(n), k_1(n) \in [0,1]$,

$$\mathbb{E}[1 - k_0(n)k_1(n)] \leq \mathbb{E}[1 - k_0(n)] + \mathbb{E}[1 - k_1(n)].$$

The limsup of the loss is at most $2 - I_r(\delta) - I_s(\delta)$. On $d \leq n$, all primes in these submeasures are distinct. Apply the simplex lemma to the bounded truncated weight $K_r(x)K_s(y)$. The retained limiting mass is

$$\int_{\sum_i x_i + \sum_j y_j \leq 1} K_r(x)K_s(y)\mathbf{1}_{\sum_{i<r} x_i \leq 1-\delta}\mathbf{1}_{\sum_{j<s} y_j \leq 1-\delta}\,dx\,dy.$$

Let $M \to \infty$ and then $\delta \downarrow 0$. Monotone convergence yields

$$\mathbb{P}(D_0 D_1 \leq n) \longrightarrow \int_{\sum x_i + \sum y_j \leq 1} K_r(x)K_s(y)\,dx\,dy = \gamma_{r,s}. \tag{79}$$

The product integral is restricted to the low-product region. No product law for the full factorizations is used.

9.4 Optimality on the low-product region

Fix $\delta > 0$ and a deterministic decoder. In the truncated submeasures a transcript specifies an ordered tuple of $R = r + s$ distinct primes. If $d \leq n$, both final bits are zero: termination would give $D_i = u_i \geq n$ and the other positive budget gives $D_{1-i} \geq 2$. At most one origin is selected per tuple, and it is at least $\max(M, d)$. Its mass is at most

$$\frac{C_{r,s,\delta}}{T_M L^R} \frac{\prod_{j=1}^R \log p_j}{\max(M,d)}.$$

Equations (74)–(75) give, uniformly over decoders,

$$\begin{aligned}\mathbb{P}(\widehat{n} = n, d \leq n) \leq\ & 2 - I_r(\delta) - I_s(\delta) + o_{r,s,a,\delta}(1) \\ & + \frac{C_{r,s,\delta}}{T_M L^R}\left\{\frac{W_{\mathbf{0}}(M)}{M} + \sum_{\prod p_j \leq M^2} \frac{\prod_j \log p_j}{\prod_j p_j}\right\}.\end{aligned}$$

The explicit term is $O_{r,s,\delta}(1/L)$. Taking $M \to \infty$ and then $\delta \downarrow 0$ proves

$$\sup_{\widehat{n}} \mathbb{P}(\widehat{n} = n, D_0 D_1 \leq n) = o_{r,s,a}(1). \tag{80}$$

Independent randomized decoders are covered by mixtures.

The least positive CRT representative for $z \equiv 0 \pmod{D_0}$, $az \equiv -1 \pmod{D_1}$ lies in $\{1, \ldots, d-1\}$ and equals n exactly when $d > n$. This includes repeated and absorbed paths: D_i always remains a divisor of its anchor. Returning this representative when it belongs to I_M gives success $1 - \mathbb{P}(d \leq n)$. Together with (80), this proves (69).

9.5 Constants and the mixed experiment

Since $f_1^- = 1$,

$$\gamma_{r,1} = \frac{1}{(r-1)!}\int_0^1 (1-x)[-\log(1-x)]^{r-1}\,dx = 2^{-r}.$$

For $|u|, |v| < 1/2$,

$$\begin{aligned}\sum_{r,s\ge 1}\gamma_{r,s}u^{r-1}v^{s-1} &= \int_{x+y\le 1}(1-x)^{-u}(1-y)^{-v}\,dx\,dy \\ &= \frac{1}{(1-u)(1-v)} - \frac{\Gamma(1-u)\Gamma(1-v)}{\Gamma(3-u-v)}.\end{aligned} \tag{81}$$

The coefficient of uv is $\pi^2/12 - 3/4$, proving (71).

In the mixed experiment, take r replacement draws and again distinguish one occurrence of each prime. The same distinguished occurrence can be selected repeatedly. For a repetition partition of block sizes $b_1, \dots, b_k$, the reduced weight is $\prod_j (\log p_j)^{b_j}/(\log u)^r$. The simplex lemma gives density $\prod_j x_j^{b_j-1}$ on S_k, with mass $t^r \prod_j (b_j-1)!/r!$ on $\sum_j x_j \le t$. The cycle identity (68) shows that the sum of these submeasures has limiting mass one and logarithmic-product density $f_r^+(x) = rx^{r-1}$.

Multiply these by the removal submeasures of budget s. Conditional independence and the simplex lemma give

$$\int_{x+y\le 1} f_r^+(x) f_s^-(y)\,dx\,dy = \frac{1}{(s-1)!}\int_0^1 (1-y)^r[-\log(1-y)]^{s-1}\,dy = \frac{1}{(r+1)^s}.$$

For the matching upper bound on success, use (74)–(75) with $m_j = b_j - 1$ at replacement blocks and $m_j = 0$ at removal labels. The total degree is $r+s$. Summing the finitely many partitions preserves the $O_{r,s,\delta}(1/L)$ low-product success bound, while the omitted mass tends to zero after truncation is removed. The CRT decoder uses the distinct-label product at the replacement anchor and the occurrence product at the removal anchor, and succeeds exactly when their product exceeds n. This proves (72) in either orientation.

10 Comparison of protocols and the positive-budget criterion

The fixed-budget reconstruction limits are summarized below. In the mixed column, r is the replacement budget and s the removal budget, irrespective of which anchor uses them.

Budgets (r,s)	With/with	With r/without s	Without/without
$(1,1)$	$1/2$	$1/2$	$1/2$
$(2,1)$	$2/3$	$2/3$	$3/4$
$(1,2)$	$2/3$	$3/4$	$3/4$
$(2,2)$	$5/6$	$8/9$	$7/4 - \pi^2/12$

For a fixed total budget $R = r+s$ and replacement at both anchors, the limiting success is maximized by balancing the budgets: this is exactly the maximization of $\binom{R}{r}$. This is a comparison after taking $M \to \infty$ at fixed R.

Corollary 10.1 (The obstruction at fixed positive budgets)**.** *Under conditional independence, for fixed positive r_0, r_1 and any of the four choices of protocols,*

$$\limsup_{M\to\infty} \mathcal{R}_M^{\text{joint}}(r_0, r_1) \le 1 - \frac{1}{(r_0+r_1)!}. \tag{82}$$

For deterministic budgets that are both positive for all large M,

$$\mathcal{R}_M^{\mathrm{joint}}(r_0(M), r_1(M)) \longrightarrow 1 \quad \Longleftrightarrow \quad \max(r_0(M), r_1(M)) \longrightarrow \infty. \tag{83}$$

Sufficiency holds for every admissible coupling; necessity here is proved under conditional independence.

Proof. With replacement, the exact deficit $r_0!r_1!/(r_0+r_1)!$ is at least $1/(r_0+r_1)!$. Without replacement, $K_{r_0}K_{r_1} \geq 1$ on the simplex of dimension r_0+r_1, whose volume is $1/(r_0+r_1)!$. For a mixed pair, retain the singleton partition at the replacement anchor and use $K_s \geq 1$ at the other anchor. This proves (82).

Sufficiency in (83) is Corollary 6.3, with the larger-budget anchor and one label from the other. If the maximum does not tend to infinity, there is a fixed $B \geq 1$ and a subsequence on which both budgets are at most B. Extend both chains to budget B. This enriches the observation and preserves conditional independence. For a removal chain, the final bit at a smaller budget $r < B$ is recoverable from the symbol at step $r+1$; at $r = B$, use the final bit of the extended transcript. Hence the original transcript is a function of the extended one. Along this subsequence, optimal success has limsup at most $1 - 1/(2B)!$, contradicting convergence to one. □

A zero budget is a one-anchor experiment and is not characterized by (83). A coupling depending on the origin may itself carry information; the necessity assertion is therefore not extended to arbitrary couplings.

11 First-order entropies for fixed budgets

Fix $r, s \geq 1$ and write $\sigma, \tau \in \{+, -\}$ for the protocols at the first and second anchors. Let $\mathcal{T}_i$ denote the ordered transcript and D_i its divisor: the distinct-label product with replacement, or the removed-occurrence product without replacement. Define

$$m_r^+ = \frac{r}{r+1}, \qquad m_r^- = 1 - 2^{-r}, \tag{84}$$

$$K_{r,s}^{\sigma,\tau} := \int_{\substack{x,y>0 \\ x+y\leq 1}} (1-x-y) f_r^\sigma(x) f_s^\tau(y)\, dx\, dy. \tag{85}$$

The factor $1-x-y$ measures the logarithmic size still undetermined when the combined divisor is below the origin. Thus $K_{r,s}^{\sigma,\tau}$ differs from the probability mass of that region.

Theorem 11.1 (Entropy coefficients)**.** *Under conditional independence of the two chains given n,*

$$H(\mathcal{T}_0) = \frac{3}{2} m_r^\sigma L + o(L), \qquad H(\mathcal{T}_1) = \frac{3}{2} m_s^\tau L + o(L), \tag{86}$$

$$H(\mathcal{T}_0, \mathcal{T}_1) = \frac{3}{2}(1 - K_{r,s}^{\sigma,\tau}) L + o(L), \tag{87}$$

$$I(\mathcal{T}_0; \mathcal{T}_1) = \frac{3}{2}(m_r^\sigma + m_s^\tau - 1 + K_{r,s}^{\sigma,\tau}) L + o(L). \tag{88}$$

When both anchors are sampled with replacement, every $o(L)$ remainder can be replaced by $O_{a,r,s}(\log L)$. Adding termination bits, or side-information alphabets of fixed sizes at the respective anchors, does not change the coefficients.

The constants include

$$K^{+,+}_{r,s} = \frac{r!s!}{(r+s+1)!}, \tag{89}$$

$$K^{+,-}_{r,s} = \frac{1}{(r+1)(r+2)^s}, \qquad K^{-,+}_{r,s} = \frac{1}{(s+1)(s+2)^r}, \tag{90}$$

$$K^{-,-}_{r,1} = \frac{1}{2\cdot 3^r}, \qquad K^{-,-}_{2,2} = \frac{31}{108} - \frac{\pi^2}{36}. \tag{91}$$

11.1 Entropy before taking the arithmetic limit

The following comparison holds without conditional independence.

Lemma 11.2 (The entropy of a revealed divisor)**.** *For the fixed budgets above, uniformly over all admissible couplings between the chains,*

$$H(\mathcal{T}_i) = \mathbb{E}\log D_i + O_{a,r,s}(\log L), \tag{92}$$

$$H(\mathcal{T}_0, \mathcal{T}_1) = \mathbb{E}\log \min(n, D_0 D_1) + O_{a,r,s}(\log L), \tag{93}$$

$$I(\mathcal{T}_0; \mathcal{T}_1) = \mathbb{E}\left[\left(\log \frac{D_0 D_1}{n}\right)_+\right] + O_{a,r,s}(\log L). \tag{94}$$

The identities also hold with termination bits.

Proof. We first treat the labels. For a one-anchor transcript t, let $D(t)$ be its divisor. Compatible origins belong to one residue class modulo $D(t)$. They are at least $D(t)$ at the first anchor and at least $D(t)/(2a)$ at the affine anchor. Separating the first term of the progression and comparing the rest with an integral gives a harmonic sum at most $(2a+L)/D(t)$. The observation kernel is bounded by one. Since $T_M \geq L$,

$$\mathbb{P}(\mathcal{T}_i = t) \leq \frac{C_a}{D(t)}. \tag{95}$$

It follows that $H(\mathcal{T}_i) \geq \mathbb{E}\log D_i - \log C_a$.

For the upper bound, all labels are at most $4M^2$. With replacement, group transcripts by their repetition partitions; for each partition, the sum of $1/D(t)$ is bounded by a product of at most r_i harmonic sums. Without replacement, sum over the lengths $0 \leq k \leq r_i$ before absorption and ordered k-tuples of primes, allowing repetitions. Their weight is the product of the reciprocals of the removed occurrences. In either case,

$$Z_i := \sum_{t \text{ possible}} \frac{1}{D(t)} \leq C_{r_i}(1 + \log(4M^2))^{r_i}. \tag{96}$$

Gibbs' inequality with the reference law $1/[Z_i D(t)]$ proves $H(\mathcal{T}_i) \leq \mathbb{E}\log D_i + \log Z_i$ and hence (92).

For a joint transcript $t = (t_0, t_1)$, put $d = D_0(t_0)D_1(t_1)$ and let $z(t) \in \{1, \ldots, d-1\}$ be its CRT representative. Set $B = \{d \leq n\}$, $G = \{d > n\}$ and define subprobabilities

$$\ell(t) = \mathbb{P}(\mathcal{T} = t, B), \qquad h(t) = \mathbb{P}(\mathcal{T} = t, G).$$

With $S(v) = \sum_t v(t)\log(1/v(t))$, the exact mixing identity is

$$0 \leq S(h) + S(\ell) - H(\mathcal{T}) = H(\mathbf{1}_G \mid \mathcal{T}) \leq \log 2. \tag{97}$$

On B, possible origins are $z(t) + kd$ with $k \geq 1$, so

$$\ell(t) \leq \frac{1}{T_M} \sum_{1 \leq k < M^2/d} \frac{1}{kd} \leq \frac{3}{d}.$$

This gives $S(\ell) \geq \mathbb{E}[\log d\, \mathbf{1}_B] - \log 3$. The reference weight $1/d$ on the support of ℓ has normalizer at most $Z_0 Z_1$. If $b = \mathbb{P}(B) > 0$, Gibbs' inequality applied to ℓ/b gives

$$S(\ell) \leq \mathbb{E}[\log d\, \mathbf{1}_B] + b \log(Z_0 Z_1) - b \log b.$$

The case $b = 0$ has zero contribution. Thus

$$S(\ell) = \mathbb{E}[\log d\, \mathbf{1}_B] + O_{r,s}(\log L). \tag{98}$$

On G, the only possible origin is $n = z(t)$. Let $k_n(t)$ denote the conditional joint kernel, for any admissible coupling. Then

$$S(h) = \mathbb{E}[\log n\, \mathbf{1}_G] + \mathbb{P}(G) \log T_M + \sum_{n \in I_M} \mu_M(n) \sum_{t:\, d(t) > n} k_n(t) \log \frac{1}{k_n(t)}.$$

The last term is nonnegative and at most $\mathbb{E}H(k_n)$. At a fixed origin, every label belongs to the prime divisors of its anchor, a set of size at most $\log(4M^2)/\log 2$. Including the absorbing symbol increases that bound by one. The conditional support therefore has size at most $(1 + CL)^{r+s}$, so its entropy is at most $(r + s)\log(1 + CL)$, irrespective of how small its nonzero probabilities are. Hence

$$S(h) = \mathbb{E}[\log n\, \mathbf{1}_G] + O_{a,r,s}(\log L). \tag{99}$$

Combining (97)–(99) proves (93). Subtracting from the two marginal entropies proves (94), since

$$\log(D_0 D_1) - \log \min(n, D_0 D_1) = [\log(D_0 D_1/n)]_+.$$

Finally, adding an alphabet of size A_i at anchor i changes the marginal entropy by at most $\log A_i$ and the joint entropy by at most $\log(A_0 A_1)$. This includes termination bits and proves the last assertion. □

11.2 The harmonic scale factor in the moments

To use the previous lemma we need a joint moment with $\log n$. Convergence of $\log D/\log n$ alone does not justify multiplying its limit by $\mathbb{E}\log n$.

Lemma 11.3 (The simplex weighted by the origin scale)**.** *In the notation of* (73)*, the measure*

$$\widehat{\nu}_M(F) = \frac{1}{LT_M} \sum_{n \in I_M} \frac{1}{n(\log n)^{k-1}} \sum_{\substack{(p_j)\ \text{admissible} \\ \prod p_j \leq n}} \left(\prod_j \log p_j \right) F\left(\frac{\log p_1}{\log n}, \ldots, \frac{\log p_k}{\log n} \right)$$

converges weakly to $\frac{3}{2}$ times Lebesgue measure on S_k. The conclusion also holds for bounded functions with a Lebesgue-null discontinuity set. For a fixed polynomial the error is $O(1/L)$.

Proof. For a monomial of multi-degree $\boldsymbol{m}$, set $D = k + \sum_j m_j$ and $c_{\boldsymbol{m}} = \prod_j m_j!/D!$. The proof of Lemma 9.2 gave

$$A_{\boldsymbol{m}}(X) = c_{\boldsymbol{m}} X (\log X)^D + O_{\boldsymbol{m},a}(X(1 + \log X)^{D-1}).$$

Partial summation with $1/[x(\log x)^{D-1}]$ now gives main term

$$c_{\boldsymbol{m}} \int_M^{M^2} \frac{\log x + D}{x}\, dx = \frac{3}{2} c_{\boldsymbol{m}} L^2 + O_{\boldsymbol{m}}(L)$$

and error $O_{\boldsymbol{m},a}(L)$. Divide by $LT_M = L^2 + O(L/M)$. Endpoint estimates and approximation on the compact simplex then finish the proof as before. □

Apply this lemma to the submeasures used for the fixed-budget profiles. With replacement, the repetition partitions have weights $\prod_j x_j^{b_j-1}$ and project to density f_r^+. Deleting the extra valuations and then restricting the affine anchor to $D_i \le n$ loses $O_{a,r}(1/L)$. For the latter statement, the retained mass is $1 + O_{a,r}(1/L)$ by the polynomial moment calculation, while the mass before restriction is $1 - O_r(1/L)$. It follows that

$$\mathbb{E}\log D_i = \frac{3}{2}\left(\int_0^1 x f_r^+(x)\,dx\right) L + O_{a,r}(1). \tag{100}$$

Without replacement, first impose the prefix truncation in (76). Its weight is bounded at fixed δ, and the lemma applies. The lost mass has limsup at most $1 - I_r(\delta)$, while $0 \le \log D_i/L \le 2 + o(1)$. Taking $M \to \infty$ followed by $\delta \downarrow 0$ gives

$$\frac{\mathbb{E}\log D_i}{L} \longrightarrow \frac{3}{2}\int_0^1 x f_r^-(x)\,dx. \tag{101}$$

These integrals are $r/(r+1)$ and $1 - 2^{-r}$, respectively.

On $B = \{D_0 D_1 \le n\}$, use the same argument with the bounded weight $1 - x - y$:

$$\frac{1}{L}\mathbb{E}\left[\log\frac{n}{D_0 D_1}\,\mathbf{1}_B\right] \longrightarrow \frac{3}{2}K_{r,s}^{\sigma,\tau}. \tag{102}$$

Conditional independence is needed here to multiply the two submeasures. Their lost mass is bounded by the sum of the marginal losses, and the normalized logarithmic weight lies between zero and two. This justifies removal of the truncation without a joint product law beyond $x + y \le 1$.

When both protocols use replacement, there is no prefix singularity. For each partition pair the needed integral is

$$\int_{S_k}\left(1 - \sum_j x_j\right)\prod_j x_j^{b_j-1}\,dx = \frac{\prod_j (b_j-1)!}{(r+s+1)!}.$$

Summing the partitions gives $r!s!/(r+s+1)!$. The polynomial moment errors and the ℓ^1 comparison give an $O_{a,r,s}(1/L)$ error in (102). Finally,

$$\mathbb{E}\log\min(n, D_0 D_1) = \mathbb{E}\log n - \mathbb{E}\left[\log\frac{n}{D_0 D_1}\,\mathbf{1}_B\right], \qquad \mathbb{E}\log n = \frac{3}{2}L + O(1).$$

These identities and Lemma 11.2 prove Theorem 11.1.

11.3 Explicit constants and successive limits

Formula (89) is the Dirichlet integral above. For the mixed case, first integrate the replacement variable:

$$\int_0^{1-y}(1-x-y)rx^{r-1}\,dx = \frac{(1-y)^{r+1}}{r+1}.$$

The substitution $t = -\log(1-y)$ gives $K_{r,s}^{+,-} = 1/[(r+1)(r+2)^s]$. For one auxiliary prime without replacement,

$$K_{r,1}^{-,-} = \frac{1}{2}\int_0^1 (1-x)^2 f_r^-(x)\,dx = \frac{1}{2\cdot 3^r}.$$

For the remaining constants, when $|u|, |v| < 1/2$,

$$\begin{aligned}\sum_{r,s\ge 1} K_{r,s}^{-,-} u^{r-1} v^{s-1} &= \int_{x+y\le 1}(1-x-y)(1-x)^{-u}(1-y)^{-v}\,dx\,dy \\ &= \frac{uv-u-v}{(1-u)(2-u)(1-v)(2-v)} + \frac{\Gamma(1-u)\Gamma(1-v)}{\Gamma(4-u-v)}.\end{aligned} \tag{103}$$

Indeed, integration in y gives $1/(2-v) - x/(1-v) + x^{2-v}/[(1-v)(2-v)]$; integration in x gives the displayed expression. Its uv coefficient is $31/108 - \pi^2/36$.

Some joint-entropy and mutual-information coefficients are:

Budgets	Protocols	$\lim H(\mathcal{T}_0, \mathcal{T}_1)/L$	$\lim I(\mathcal{T}_0; \mathcal{T}_1)/L$
$(1,1)$	all	$5/4$	$1/4$
$(2,1)$	$(+,+)$ or $(+,-)$	$11/8$	$3/8$
$(2,1)$	$(-,+)$ or $(-,-)$	$17/12$	$11/24$
$(2,2)$	$(+,+)$	$29/20$	$11/20$
$(2,2)$	mixed	$47/32$	$21/32$
$(2,2)$	$(-,-)$	$77/72 + \pi^2/24$	$85/72 - \pi^2/24$

Each limit is taken as $M \to \infty$ with the displayed budgets fixed. In particular,

$$\lim_{M\to\infty} \frac{I(\mathcal{T}_r^+; \mathcal{T}_1)}{L} = \frac{3r}{4(r+2)}, \qquad \lim_{M\to\infty} \frac{I(\mathcal{T}_r^-; \mathcal{T}_1)}{L} = \frac{3}{4} - \frac{3}{2^{r+1}} + \frac{3}{4\cdot 3^r}.$$

Taking $r \to \infty$ afterwards recovers the coefficient $3/4$ of the asymmetric experiment. This agreement uses successive limits, not a uniform statement for $r = r(M)$.

The entropy comparison identifies the logarithmic size revealed by the joint congruence information: its leading value is $\mathbb{E}\log\min(n, D_0D_1)$. Mutual information is the expected positive logarithmic excess of the combined divisor over the origin, up to $O(\log L)$. The explicit limits above require only the small-modulus measures already proved.

Part II
Predicting a sign from two labels

12 Predicting a sign on an observable confidence region

We now observe one prime from each of the anchors n and $an+1$, where $a \in \{1,3\}$, with the two draws independent conditional on n. The first draw has the same law with or without replacement. Write $L = \log M$, $d = pq$, and let $z = z_a(p,q)$ be the least positive solution of

$$z \equiv 0 \pmod{p}, \qquad az \equiv -1 \pmod{q}.$$

Every pair that can be observed has $p \neq q$ and $q \nmid a$; for such a pair, $1 \le z < pq$. The least solution is a natural candidate for the hidden integer. The following estimate identifies observations on which this candidate has a uniformly high conditional probability of being correct.

Theorem 12.1 (Observable confidence and universal sign prediction)**.** *For sufficiently large M, define*

$$\mathcal{A}_M = \{(p,q) : z_a(p,q) \in I_M, \quad pq > z_a(p,q)L^4\}. \tag{104}$$

The decoder $\widehat{n}_M$ returns $z_a(p,q)$ on $\mathcal{A}_M$ and abstains otherwise. It satisfies

$$\mathbb{P}(\widehat{n}_M \neq n, \mathcal{A}_M) \le \frac{100}{L}, \tag{105}$$

$$\liminf_{M\to\infty} \mathbb{P}(\widehat{n}_M = n, \mathcal{A}_M) \ge \frac{1}{4}. \tag{106}$$

More precisely, every accepted observation satisfies

$$\mathbb{P}\{n = z_a(p,q) \mid P = p, Q = q\} \ge \frac{1}{1 + 100/L}. \tag{107}$$

Consequently, for every deterministic sign $W_M : I_M \to \{-1, 1\}$, which may depend on M,

$$\liminf_{M\to\infty} \big\|\mathbb{E}[W_M(n) \mid P, Q]\big\|_1 \geq \frac{1}{4}, \tag{108}$$

$$\liminf_{M\to\infty} \sup_{\widehat{W}} \mathbb{P}\{\widehat{W} = W_M(n)\} \geq \frac{5}{8}. \tag{109}$$

If a termination bit is observed at either or both anchors, also require each observed bit to agree with the candidate $z = z_a(p, q)$: the two candidate bits are $\mathbf{1}_{\{p=z\}}$ and $\mathbf{1}_{\{q=az+1\}}$, respectively. With this acceptance rule the same bounds hold, including (107) conditional on the full observed transcript.

For labels alone the acceptance region is $\mathcal{A}_M$. When bits are used, a disagreement rules out the candidate and the decoder abstains. On either acceptance region the conditional error tends uniformly to zero. The zero-error problem, treated in Section 21, has a different answer.

12.1 Marginals of the reconstructible region

Put

$$X_M = \frac{\log P}{\log n}, \qquad Y_M = \frac{\log Q}{\log n}, \qquad G_M = \{PQ > n\}, \qquad B_M = \{PQ \leq n\}.$$

For every fixed $t \in [0, 1]$,

$$\mathbb{P}(G_M, X_M \leq t) \longrightarrow \frac{t^2}{2}, \qquad \mathbb{P}(G_M, Y_M \leq t) \longrightarrow \frac{t^2}{2}. \tag{110}$$

Indeed, Lemma 9.2 with $k = 1$, together with the valuation and denominator comparison (64), gives the uniform law on $[0, 1]$ for each marginal. For the affine anchor, the restriction $q \leq n$ already has mass tending to one, so the possible labels $q > n$ have vanishing mass. With $k = 2$, the same simplex lemma gives Lebesgue measure on $\{x, y \geq 0 : x + y \leq 1\}$ for the restriction to B_M. Its first marginal has density $1 - x$; subtracting

$$\mathbb{P}(B_M, X_M \leq t) \longrightarrow \int_0^t (1 - x)\, dx = t - \frac{t^2}{2}$$

from $\mathbb{P}(X_M \leq t) \to t$ proves the first formula. The second follows in the same way. In particular, $\mathbb{P}(G_M) \to 1/2$.

The formulas remain valid for $t_M \to t \in (0, 1)$: bound t_M between $t-\varepsilon$ and $t+\varepsilon$, take $M \to \infty$, and then let $\varepsilon \downarrow 0$. This argument uses no rate of convergence in the simplex lemma.

12.2 A pointwise posterior estimate

Split the label law into two subprobability measures:

$$h(p, q) = \mathbb{P}(P = p, Q = q, G_M), \qquad \ell(p, q) = \mathbb{P}(P = p, Q = q, B_M).$$

If $z \in I_M$, the only compatible origin in G_M is $n = z$. Thus

$$h(p, q) = \frac{1}{zT_M} \frac{v_p(z) \log p}{\log z} \frac{v_q(az + 1) \log q}{\log(az + 1)}. \tag{111}$$

Since the valuations are at least one, $\log z \leq 2L$, and $\log(az + 1) \leq 4L$, we have

$$h(p, q) \geq \frac{(\log 2)^2}{8zT_M L^2}. \tag{112}$$

The compatible origins in B_M are $z + kd$, $k \geq 1$. Bounding their observation probabilities by one gives

$$\ell(p,q) \leq \frac{1}{T_M} \sum_{\substack{k\geq 1 \\ z+kd<M^2}} \frac{1}{z+kd} \leq \frac{1+\log(M^2/d)}{dT_M} \leq \frac{3}{d} \tag{113}$$

when $d < M^2$; if $d \geq M^2$, then $\ell(p,q) = 0$. Using $T_M \leq 2L$, we obtain

$$\frac{\ell(p,q)}{h(p,q)} \leq \frac{48}{(\log 2)^2} L^3 \frac{z}{d} < 100 L^3 \frac{z}{d}. \tag{114}$$

On $\mathcal{A}_M$, this is less than $100/L$. The conditional probability of $n = z$ is $h/(h+\ell)$, which proves (107). If termination bits are observed and agree with z, the mass at z is still $h(p,q)$, since its bits are deterministic. The competing mass for this transcript is at most $\ell(p,q)$. This proves the conditional bound for every accepted full transcript. Requiring agreement retains every correct acceptance and can only remove false acceptances. Summing $\ell \leq (100/L)h$ over $\mathcal{A}_M$ proves (105).

12.3 The elementary lower bound for acceptance

On G_M, we have $z = n$. An origin in G_M that is not accepted therefore satisfies

$$X_M + Y_M \leq 1 + \frac{4\log L}{\log n} \leq 1 + \frac{4\log L}{L}.$$

Set $t_M = 1/2 + 2\log L/L$. This event is contained in $\{X_M \leq t_M\} \cup \{Y_M \leq t_M\}$. Formula (110) consequently gives

$$\limsup_M \mathbb{P}(G_M, \mathcal{A}_M^c) \leq \frac{1}{8} + \frac{1}{8} = \frac{1}{4}.$$

Together with $\mathbb{P}(G_M) \to 1/2$, this proves (106).

For an accepted pair, the signed mass of the origin z is $W_M(z)h(p,q)$, while the absolute signed mass of all other origins is at most $\ell(p,q)$. Hence

$$\begin{aligned} \big\|\mathbb{E}[W_M(n) \mid P,Q]\big\|_1 &\geq \sum_{(p,q)\in\mathcal{A}_M} (h(p,q) - \ell(p,q)) \\ &\geq \mathbb{P}(\widehat{n}_M = n, \mathcal{A}_M) - 100/L. \end{aligned}$$

The binary decision identity

$$\sup_{\widehat{W}} \mathbb{P}(\widehat{W} = W_M) = \frac{1}{2} + \frac{1}{2}\big\|\mathbb{E}[W_M \mid P,Q]\big\|_1$$

completes the proof of Theorem 12.1. An explicit randomized rule achieves the same lower bound: return $W_M(z)$ on $\mathcal{A}_M$, and an independent fair sign elsewhere. An optimal deterministic rule performs at least as well.

12.4 The posterior outside the confidence region

For an admissible pair with positive observation probability, put

$$\rho_{p,q}(m) = \frac{v_p(m)v_q(am+1)}{m\log m\log(am+1)}.$$

The common factor $\log p \log q/T_M$ cancels from the posterior:

$$\mathbb{E}[W_M(n) \mid P=p, Q=q] = \frac{\sum\limits_{\substack{m\in I_M \\ m\equiv z \ (pq)}} W_M(m)\rho_{p,q}(m)}{\sum\limits_{\substack{m\in I_M \\ m\equiv z \ (pq)}} \rho_{p,q}(m)}. \tag{115}$$

For $W_M(m) = \lambda(m)$, writing $m = pk$ turns the numerator into a weighted sum of $-\lambda(k)$ over

$$k \equiv -(ap)^{-1} \pmod q.$$

For $\lambda(am+1)$, write $am + 1 = qk'$; the sign becomes $-\lambda(k')$. Their product gives $\lambda(k)\lambda(k')$ subject to $apk - qk' = -1$. Thus the undecoded part of the experiment leads to sums over progressions whose moduli and lengths vary together with M. Cancellation for a fixed modulus does not control this posterior.

13 The joint law of the logarithmic prime sizes

The preceding lower bound used only the two marginals and the small-product simplex. To recognize asymptotically the whole reconstructible half, we need a joint law also above that simplex. The following theorem transfers Teräväinen's logarithmic independence of large-prime-factor counts [14, Theorem 1.11] to the sampled prime labels. Its inputs are his binary-correlation theorem for real multiplicative functions and the uniformity calculation in Section 4 of the same paper. We give the affine reduction, the moving-threshold argument, and the passage to weighted prime labels explicitly.

Theorem 13.1 (One prime from each anchor)**.** *For $a \in \{1,3\}$, with the draws independent conditional on n, put*

$$S_M = \frac{\log n}{L}, \qquad X_M = \frac{\log P}{\log n}, \qquad Y_M = \frac{\log Q}{\log n}.$$

Then

$$(S_M, X_M, Y_M) \implies \mathbf{1}_{[1,2]}(s)\,ds \otimes \mathbf{1}_{[0,1]}(x)\,dx \otimes \mathbf{1}_{[0,1]}(y)\,dy. \tag{116}$$

This is weak convergence of logarithmic sizes. It does not assert independence of the integer labels P, Q, or convergence in total variation of their discrete laws.

13.1 The external input and the anchor $3n+1$

For $0 < \alpha < 1$, $u \in [0,1]$, and $x \to \infty$, define

$$g_{x,\alpha,u}(m) = u^{\omega_{>x^\alpha}(m)}, \qquad \omega_{>y}(m) = \#\{p > y : p \mid m\}, \qquad 0^0 = 1.$$

These functions are multiplicative and take values in $[0,1]$. Equations (4.5)–(4.8) of [14] give a polynomial $A_\alpha(u)$ such that, for every residue c modulo a fixed integer q, including nonreduced residues,

$$\frac{1}{x} \sum_{\substack{x\le m\le 2x \\ m\equiv c \ (q)}} g_{x,\alpha,u}(m) \longrightarrow \frac{A_\alpha(u)}{q}. \tag{117}$$

Theorem 1.4 there permits the real multiplicative functions to depend on x. When one factor is uniform in the residue classes of every fixed modulus, its logarithmic correlation with the other

factor at a fixed nonzero shift, on $[x/\log x, x]$, approaches the product of their means on $[x, 2x]$. Applying this with $g_1 = g_{x,\alpha,u}$ and $g_2 = g_{x,\beta,v}$ gives

$$\frac{1}{\log\log x} \sum_{x/\log x < m \le x} \frac{g_1(m) g_2(m+1)}{m} \longrightarrow A_\alpha(u) A_\beta(v). \tag{118}$$

To obtain $3n+1$, let $\chi_{0,3}(m) = \mathbf{1}_{3\nmid m}$. The real multiplicative function $g_1\chi_{0,3}$ has mean $\frac{2}{3}A_\alpha(u) + o(1)$, by (117), but is not uniform modulo three. We therefore use g_2 as the uniform factor, $g_1\chi_{0,3}$ as the other factor, and shift -1. Changing the index and replacing $1/(m+1)$ by $1/m$ costs $o(1)$. The resulting limit is

$$\frac{1}{\log\log x} \sum_{x/\log x < m \le x} \frac{g_1(m)\chi_{0,3}(m) g_2(m+1)}{m} \longrightarrow \frac{2}{3} A_\alpha(u) A_\beta(v).$$

Subtract this from (118). The remaining sum is over $3 \mid m$. For $m = 3n$, we have $g_1(3n) = g_1(n)$ once $x^\alpha > 3$; the factor $1/3$ in the harmonic weight cancels the factor $1/3$ in the limit. Thus, for $a = 1$ and $a = 3$,

$$\frac{1}{\log\log x} \sum_{x/(a\log x) < n \le x/a} \frac{g_{x,\alpha,u}(n) g_{x,\beta,v}(an+1)}{n} \longrightarrow A_\alpha(u) A_\beta(v). \tag{119}$$

This proves the required affine case from the integer-shift theorem.

13.2 Replacing fixed thresholds by thresholds relative to n

On the interval in (119), changing x^α to n^α can alter a count only when a prime in

$$\left[\left(\frac{x}{a\log x}\right)^\alpha, x^\alpha\right]$$

divides the relevant anchor. Its normalized harmonic probability is at most

$$C_a \sum_{(x/(a\log x))^\alpha \le p \le x^\alpha} \frac{1}{p} + O_a\left(\frac{x^\alpha}{[x/(a\log x)]\log\log x}\right) = o(1).$$

Mertens' formula handles the first term, and $\alpha < 1$ handles the second. Counting one residue class modulo p suffices for either anchor; primes dividing a are absent from this band for large x. The same argument applies to β.

Put

$$N_\alpha(n) = \omega_{>n^\alpha}(n), \qquad N_{\beta,a}(n) = \omega_{>n^\beta}(an+1).$$

For fixed α, β, the degrees of $u^{N_\alpha(n)} v^{N_{\beta,a}(n)}$ are bounded uniformly for large n; any integer larger than $2/\alpha + 2/\beta$ is a common bound. Evaluating at finitely many distinct values of u, v and inverting the two Vandermonde matrices converts (119) into convergence of every coefficient to the corresponding coefficient of $A_\alpha(u)A_\beta(v)$.

These local limits imply harmonic limits on $[2, X]$. Writing $y = x/a$, successively concatenate intervals $[y/\log(ay), y]$ until reaching a fixed lower threshold Y. They cover $[Y, X]$ up to negligible endpoints. Above Y, every local error is small by the definition of the local limit, and the total harmonic weight is $\log X + O_Y(1)$. The initial segment below Y disappears after normalization; then let $Y \to \infty$.

The marginal first moment equals $\log(1/\alpha)$, since

$$\frac{1}{x} \sum_{x \le m \le 2x} \omega_{>x^\alpha}(m) = \sum_{x^\alpha < p \le 2x} \frac{1}{p} + O\left(\frac{\pi(2x)}{x}\right) = \log(1/\alpha) + o(1).$$

Bounded degree identifies this moment with $A'_\alpha(1)$. The coefficient limits therefore imply

$$\frac{1}{\log X}\sum_{2\le n\le X}\frac{N_\alpha(n)N_{\beta,a}(n)}{n}\longrightarrow \log(1/\alpha)\log(1/\beta). \tag{120}$$

Subtracting the sums at $X=M^s$ and $X=M^t$ gives, for $1\le s<t\le 2$, the limit $(t-s)\log(1/\alpha)\log(1/\beta)$ on $[M^s,M^t)$ after division by T_M.

13.3 From factor counts to the sampled labels

Consider the counting measures of distinct prime factors

$$\nu_{0,n}=\sum_{p\mid n}\delta_{\log p/\log n},\qquad \nu_{1,n}=\sum_{q\mid an+1}\delta_{\log q/\log n}.$$

Their tails are the preceding counts. Four differences of (120) show that the averaged measure $\nu_{0,n}\otimes\nu_{1,n}$, restricted to $S_M\in[s,t)$, converges on compact rectangles in $(0,1)^2$ to

$$(t-s)\frac{dx\,dy}{xy}.$$

Away from zero, the numbers of factors are uniformly bounded. Step-function approximation therefore permits multiplication by xy and by an arbitrary bounded continuous test function. After removing excess valuations, the conditional weight of a prime pair is

$$\frac{\log p\log q}{\log n\log(an+1)}=xy\,\frac{\log n}{\log(an+1)}.$$

The last ratio tends uniformly to one on I_M, and (64) bounds the total cost of removing the excess valuations by $O_a(1/L)$. Conditional independence is used precisely to multiply the two sampling weights.

Each logarithmic marginal tends to the uniform law on $[0,1]$. Thus the limiting upper bound for the mass within distance ε of an endpoint, or outside the square, is $O(\varepsilon)$. Remove these edges, apply the preceding compact argument, and let $\varepsilon\downarrow 0$. Approximation first in the scale variable s, then in (x,y), proves (116).

13.4 Recognizing the whole reconstructible half

Corollary 13.2 (Confidence selection and separation)**.** *For the acceptance region* (104)*,*

$$\mathbb{P}(\widehat{n}_M=n,\mathcal{A}_M)\longrightarrow\frac12,\qquad \mathbb{P}(\widehat{n}_M\neq n,\mathcal{A}_M)=O(1/L), \tag{121}$$

$$J_M:=\sum_{p,q}\min\{h(p,q),\ell(p,q)\}\longrightarrow 0. \tag{122}$$

Uniformly over deterministic signs $W_M:I_M\to\{-1,1\}$,

$$\liminf_M\|\mathbb{E}[W_M(n)\mid P,Q]\|_1\ge\frac12,\qquad \liminf_M\sup_{\widehat{W}}\mathbb{P}(\widehat{W}=W_M(n))\ge\frac34. \tag{123}$$

With termination bits, use the bit-agreement rule of Theorem 12.1. The selection and sign bounds then hold for the enriched transcript. Its overlap, defined by summing the minimum of the two subprobability masses over full transcripts, also tends to zero.

Proof. The rejected part of G_M lies in

$$1 < X_M + Y_M \le 1 + 4\log L/L.$$

For every fixed $\varepsilon > 0$, this strip eventually lies in $\{|x + y - 1| \le \varepsilon\}$. The joint limit (116) puts no mass on the line $x + y = 1$, so the rejected mass tends to zero. Now $\mathbb{P}(G_M) \to 1/2$, while the false-acceptance probability is at most $100/L$. Moreover,

$$J_M \le \sum_{\mathcal{A}_M} \ell + \sum_{\mathcal{A}_M^c} h = o(1).$$

The signed contribution on $\mathcal{A}_M$ is bounded below by $\sum_{\mathcal{A}_M}(h - \ell)$, uniformly in W_M. The binary decision identity gives the sign bound. With bits, true acceptance is unchanged and false acceptance can only decrease. Splitting each pair by its bits can only decrease the overlap, since $\sum_b \min(h_b, \ell_b) \le \min(\sum_b h_b, \sum_b \ell_b)$. The same sign rule therefore attains the stated lower bound. □

The constant 3/4 thus follows from the binary correlation input after the affine reduction and the logarithmic weighting. To identify the optimal success probability for a particular sign, we must still estimate its signed mass outside the decoded region.

14 The signed remainder beyond the universal guarantee

Retain the subprobability measures h, ℓ from Section 12. For a deterministic sign $W_M : I_M \to \{-1, 1\}$, define

$$b_{M,W}(p, q) = \mathbb{E}\big[W_M(n)\mathbf{1}_{\{P=p,Q=q,PQ\le n\}}\big], \tag{124}$$

$$\Lambda_M(W) = \sum_{p,q} |b_{M,W}(p, q)|. \tag{125}$$

Always $|b_{M,W}(p, q)| \le \ell(p, q)$, so that $0 \le \Lambda_M(W) \le \mathbb{P}(B_M)$. The following finite inequality isolates the role of the overlap between the two label submeasures.

Proposition 14.1 (Uniform reduction of sign prediction)**.** *If*

$$J_M := \sum_{p,q} \min\{h(p, q), \ell(p, q)\} \longrightarrow 0, \tag{126}$$

then, uniformly over deterministic signs W_M,

$$\big\|\mathbb{E}[W_M(n) \mid P, Q]\big\|_1 = \mathbb{P}(G_M) + \Lambda_M(W) + O(J_M), \tag{127}$$

$$\sup_{\widehat{W}} \mathbb{P}\{\widehat{W} = W_M(n)\} = \frac{3}{4} + \frac{1}{2}\Lambda_M(W) + o(1). \tag{128}$$

The error in the first line lies between $-2J_M$ and zero; the $o(1)$ in the second does not depend on W_M. The same formulas, with the same $\Lambda_M(W)$, hold if the two termination bits are included.

Proof. The compatible origin in G_M is unique. Its signed contribution is $W_M(z)h(p, q)$, and hence

$$\big\|\mathbb{E}[W_M(n) \mid P, Q]\big\|_1 = \sum_{p,q} |W_M(z)h(p, q) + b_{M,W}(p, q)|.$$

Choose $W_M(z)$ arbitrarily when $h(p, q) = 0$. For $u \ge 0$, $|s| = 1$, and $b \in \mathbb{R}$,

$$0 \le u + |b| - |su + b| \le 2\min\{u, |b|\}.$$

Apply this with $u = h(p,q)$, $s = W_M(z)$, and $b = b_{M,W}(p,q)$, use $|b_{M,W}| \le \ell$, and sum. The binary decision identity and $\mathbb{P}(G_M) = 1/2 + O_a(1/L)$ prove the second formula.

Both termination bits are zero on B_M: termination at the first anchor would require $P = n$, and termination at the second would require $Q = an + 1$; either condition implies $PQ > n$. Thus the lower signed submeasure is simply assigned to the extended transcript $(p, q, 0, 0)$, without a change of norm. The upper mass carries bits determined by its unique origin. Its overlap with the lower mass is no greater than J_M, so the same argument applies. □

The acceptance set gives the finite bound

$$J_M \le \sum_{(p,q)\in\mathcal{A}_M} \ell(p,q) + \sum_{(p,q)\notin\mathcal{A}_M} h(p,q). \tag{129}$$

The first term is $O(1/L)$, and the joint law makes the second $o(1)$, as proved in Corollary 13.2. The unsigned separation is therefore already available; the remaining problem for a prescribed sign is $\Lambda_M(W)$.

14.1 An exact remaining estimate for Liouville

For $W_M(n) = \lambda(n)$, the positive comparison (64) removes excess valuations and replaces $\log n \log(an+1)$ by $(\log n)^2$, at total cost $O_a(1/L)$. Since the ℓ^1-norm is Lipschitz with constant one,

$$\Lambda_M(\lambda) = \frac{1}{T_M} \sum_{\substack{p,q \text{ prime} \\ p\neq q,\ q\nmid a}} \log p \log q \left| \sum_{\substack{m\in I_M \\ m\equiv z_a(p,q)\ (pq) \\ pq\le m}} \frac{\lambda(m)}{m(\log m)^2} \right| + O_a(1/L). \tag{130}$$

Writing $m = pk$ and using $\lambda(pk) = -\lambda(k)$ gives

$$\Lambda_M(\lambda) = \frac{1}{T_M} \sum_{\substack{p,q \text{ prime} \\ p\neq q,\ q\nmid a}} \frac{\log p \log q}{p} \left| \sum_{\substack{M/p\le k<M^2/p \\ k\ge q \\ k\equiv -(ap)^{-1}\ (q)}} \frac{\lambda(k)}{k(\log(pk))^2} \right| + O_a(1/L). \tag{131}$$

Only $p < M^2$ and $q < 3M^2 + 1$ can contribute. Consequently, the statement that the optimal success probability for Liouville is $3/4 + o(1)$ is equivalent to

$$\Lambda_M(\lambda) \longrightarrow 0. \tag{132}$$

This is a sum of absolute values of progression sums. Cancellation only after a signed sum over the labels does not suffice. The moduli and interval endpoints vary simultaneously. The analogous reduction for $\lambda(an+1)$, or for a product of affine signs, uses the corresponding numerator in (130); it does not by itself establish cancellation.

For example, $W_M \equiv 1$ gives $\Lambda_M(W_M) = \mathbb{P}(B_M)$ and perfect prediction. Thus a universal lower bound for prediction need not be an exact profile for an individual sign.

14.2 The minimax value over deterministic signs

For $W : I_M \to \{-1, 1\}$, let

$$\mathfrak{A}_M(W) = \big\|\mathbb{E}[W(n) \mid P, Q]\big\|_1, \qquad \mathfrak{S}_M(W) = \sup_{\widehat{W}} \mathbb{P}\{\widehat{W} = W(n)\} = \frac{1 + \mathfrak{A}_M(W)}{2}.$$

The decoder knows the sign function to be predicted. The following infimum ranges over deterministic tables chosen before n, P, Q are drawn; the table may depend on M.

Theorem 14.2 (Minimax optimality of the universal guarantee)**.** *Under* (126)*,*

$$\inf_{W:I_M\to\{-1,1\}} \mathfrak{A}_M(W) \longrightarrow \frac{1}{2}, \tag{133}$$

$$\inf_{W:I_M\to\{-1,1\}} \mathfrak{S}_M(W) \longrightarrow \frac{3}{4}. \tag{134}$$

The same limits hold with the termination bits observed.

Proof. Write $o = (p, q)$ and

$$w_B(n, o) = \mathbb{P}\{n, (P, Q) = o, PQ \le n\}.$$

Then $\ell(o) = \sum_n w_B(n, o)$. Maximizing separately for each transcript gives

$$R_{B,M} := \sum_o \max_{n\in I_M} w_B(n, o) = \sup_{\widehat{n}} \mathbb{P}\{\widehat{n} = n, B_M\} = O_a(1/L), \tag{135}$$

where the last bound is (61) with budgets $(1, 1)$.

For the existence argument only, take independent fair signs $(\varepsilon_n)_{n\in I_M}$. Vanishing cross terms and Cauchy–Schwarz imply

$$\mathbb{E}_\varepsilon \left|\sum_n \varepsilon_n w_B(n, o)\right| \le \left(\sum_n w_B(n, o)^2\right)^{1/2} \le \Big(\ell(o) \max_n w_B(n, o)\Big)^{1/2}.$$

Summing over o and applying Cauchy–Schwarz once more gives

$$\mathbb{E}_\varepsilon \Lambda_M(\varepsilon) \le \sqrt{\mathbb{P}(B_M) R_{B,M}} = O_a(L^{-1/2}). \tag{136}$$

At least one deterministic sign table satisfies the same bound. For every W, the proof of Proposition 14.1 gives

$$\mathbb{P}(G_M) + \Lambda_M(W) - 2J_M \le \mathfrak{A}_M(W) \le \mathbb{P}(G_M) + \Lambda_M(W).$$

Therefore

$$\mathbb{P}(G_M) - 2J_M \le \inf_W \mathfrak{A}_M(W) \le \mathbb{P}(G_M) + \sqrt{\mathbb{P}(B_M) R_{B,M}}. \tag{137}$$

Now use $\mathbb{P}(G_M) = 1/2 + O_a(1/L)$, $J_M \to 0$, and $R_{B,M} = O_a(1/L)$. With termination bits, all lower mass still has bits $(0, 0)$; neither $\Lambda_M(W)$ nor $R_{B,M}$ changes, and the upper and lower overlap can only decrease. □

This existence argument allows the minimizing table to change with M. It does not identify the value for Liouville. Section 20 will give a fixed periodic sign that attains the same minimax limit.

15 Large-sieve cancellation on a region of mass 3/8

For a deterministic sign W, write

$$\mathrm{Succ}^*_M(W) = \sup_{\widehat{W}} \mathbb{P}\{\widehat{W} = W(n)\},$$

using the two prime labels, or their extension by termination bits when specified. This section bounds the remainder for $\lambda(n)$ using the qualitative prime number theorem

$$S(t) := \sum_{k \le t} \lambda(k) = o(t). \tag{138}$$

For the second anchor we also use this cancellation in the residue classes of the fixed modulus a. Both follow, for example, from the uniform Davenport estimate recalled after Lemma 1.4 of [15], by taking frequency zero and the rational frequencies j/a. No rate is required.

15.1 An elementary additive large sieve

We use $e(t) = \exp(2\pi i t)$.

Lemma 15.1 (Sampling at separated points)**.** *If $0 < \Delta \le 1$ and the points of $\mathcal{F} \subset \mathbb{R}/\mathbb{Z}$ are separated modulo one by at least Δ, then*

$$\sum_{\alpha \in \mathcal{F}} \left| \sum_{1 \le k \le H} c_k e(k\alpha) \right|^2 \le (\Delta^{-1} + 4\pi H) \sum_{1 \le k \le H} |c_k|^2. \tag{139}$$

Proof. Let $A(\theta) = \sum_{1 \le k \le H} c_k e(k\theta)$ and $F(\theta) = |A(\theta)|^2$. The arcs I_α, centred at α and of length Δ, are disjoint except possibly at endpoints. The fundamental theorem of calculus, averaged over I_α, gives

$$F(\alpha) \le \Delta^{-1} \int_{I_\alpha} F(\theta)\, d\theta + \int_{I_\alpha} |F'(\theta)|\, d\theta.$$

Sum and apply Parseval and Cauchy–Schwarz:

$$\sum_{\alpha} F(\alpha) \le \Delta^{-1} \|A\|_2^2 + 2\|A'\|_2 \|A\|_2 \le (\Delta^{-1} + 4\pi H) \sum_{k} |c_k|^2.$$

□

15.2 A maximal estimate in prime residue classes

For a prime q and $r \bmod q$, define

$$E_q(r,t) = \sum_{\substack{k \le t \\ k \equiv r \ (q)}} \lambda(k) - \frac{S(t)}{q}, \qquad D_q(r;H) = \max_{0 \le t \le H} |E_q(r,t)|.$$

The centre is $S(t)/q$, because all q residue classes, including zero, are included.

Lemma 15.2 (Maximal large sieve for residue classes)**.** *For $H, Q \ge 2$,*

$$\sum_{\substack{Q \le q < 2Q \\ q \text{ prime}}} \sum_{r \bmod q} D_q(r;H)^2 \ll (\log(2H))^2 \left(\frac{H^2}{Q} + HQ \right). \tag{140}$$

The same absolute implied constant works for arbitrary complex coefficients $|c_k| \le 1$, with $\lambda(k)$ replaced by c_k and $S(t)$ by $\sum_{k \le t} c_k$.

Proof. We prove the general version. For an integer interval $J \subset [1, H]$, additive orthogonality gives

$$\sum_{r \bmod q} \left| \sum_{\substack{k \in J \\ k \equiv r \ (q)}} c_k - \frac{1}{q} \sum_{k \in J} c_k \right|^2 = \frac{1}{q} \sum_{b=1}^{q-1} \left| \sum_{k \in J} c_k e(bk/q) \right|^2.$$

The fractions b/q, for primes $Q \le q < 2Q$ and $1 \le b < q$, are distinct, reduced, and separated modulo one by at least $(4Q^2)^{-1}$. Applying Lemma 15.1 to $c_k \mathbf{1}_J(k)$ therefore yields

$$\sum_{q\sim Q} \sum_{r \bmod q} \left| \sum_{\substack{k\in J\\ k\equiv r \ (q)}} c_k - \frac{1}{q}\sum_{k\in J} c_k \right|^2 \ll \left(\frac{H}{Q} + Q\right)|J|. \tag{141}$$

Here and below a sum over $q \sim Q$ is over prime q. Extend the coefficients by zero to a power of two $H_0 \in [H, 2H]$. Every prefix $[1, t]$ is a disjoint union of at most $1 + \log_2 H_0$ canonical dyadic intervals. Cauchy–Schwarz bounds the squared maximum by $O(\log(2H))$ times the sum of squared discrepancies over all such intervals. At each level their useful lengths sum to at most H. Summing (141) over the $O(\log(2H))$ levels proves the lemma. □

15.3 Transferring the estimate to the sampled labels

Fix $0 < \varepsilon < 1/6$ and define the source event

$$\mathcal{C}_{\varepsilon,M} = \{n^{\varepsilon} \le P, Q, \quad Q \le n^{1/2-\varepsilon}, \quad PQ \le n^{1-\varepsilon}\}. \tag{142}$$

This event is contained in B_M. It need not be observable: we use it to split a signed submeasure, rather than to define a decoder. Put

$$\Lambda^{\mathrm{int}}_{\varepsilon,M} = \sum_{p,q} \left| \mathbb{E}[\lambda(n)\mathbf{1}_{\{P=p,Q=q\}}\mathbf{1}_{\mathcal{C}_{\varepsilon,M}}] \right|.$$

Proposition 15.3 (Cancellation in the interior region)**.** *For every fixed* $0 < \varepsilon < 1/6$*,*

$$\Lambda^{\mathrm{int}}_{\varepsilon,M} \longrightarrow 0. \tag{143}$$

Proof. The positive-law comparison (64) is uniform for bounded tests of the origin and labels. It therefore remains valid after restriction to (142), at cost $O_a(1/L)$. We are left to bound

$$\frac{1}{T_M} \sum_{\substack{p,q \text{ prime}\\ p\neq q,\ q\nmid a}} \frac{\log p \log q}{p} \left| \sum_{\substack{M\le pk<M^2\\ (pk,p,q)\in\mathcal{C}_{\varepsilon,M}\\ k\equiv r_{p,q}\ (q)}} \frac{\lambda(k)}{k(\log(pk))^2} \right|, \qquad r_{p,q} = -(ap)^{-1} \pmod q. \tag{144}$$

The sign from $\lambda(pk) = -\lambda(k)$ disappears inside the absolute value.

Partition the origin range into $O(L)$ intervals $[N, 2N) \cap I_M$, with $N = M2^j$. For fixed p, q, the restrictions are

$$n \le p^{1/\varepsilon}, \quad n \le q^{1/\varepsilon}, \quad n \ge q^{1/(1/2-\varepsilon)}, \quad n \ge (pq)^{1/(1-\varepsilon)}.$$

Their intersection with the block is an interval, possibly empty. After $n = pk$, denote the resulting interval by $J_{N,p,q}$.

Also split the prime sizes into $p \in [P, 2P)$ and $q \in [Q, 2Q)$, with dyadic P, Q, and put $K = N/P$. Then $J_{N,p,q} \subset [K/2, 2K]$. The weight $w_p(u) = 1/(u(\log(pu))^2)$ is positive and decreasing. Partial summation gives

$$\left| \sum_{k\in J_{N,p,q}} w_p(k) \left(\lambda(k)\mathbf{1}_{k\equiv r_{p,q}\ (q)} - \frac{\lambda(k)}{q} \right) \right| \ll \frac{D_q(r_{p,q}; 2K)}{KL^2}. \tag{145}$$

Both endpoints may depend on p, q. For fixed q, the residue $r_{p,q}$ determines one residue class of p mod q, of multiplicity at most $1+P/Q$ in $[P, 2P)$. There are $O(PQ)$ pairs. By Cauchy–Schwarz and Lemma 15.2,

$$\sum_{\substack{p\sim P, q\sim Q\\ p\neq q,\ q\nmid a}} D_q(r_{p,q}; 2K) \ll \log(2K)\sqrt{PQ\left(1+\frac{P}{Q}\right)\left(\frac{K^2}{Q}+KQ\right)}. \tag{146}$$

Using $\log p, \log q \ll L$, $T_M \asymp L$, and $PK = N$, the centred contribution of a size rectangle is bounded by

$$\frac{C\log(2K)}{T_M N}\sqrt{PQ\left(1+\frac{P}{Q}\right)\left(\frac{K^2}{Q}+KQ\right)} \ll \sqrt{\frac{1}{P}+\frac{1}{Q}+\frac{Q^2}{N}+\frac{Q}{K}}. \tag{147}$$

The identity behind this bound is

$$\frac{PQ(1+P/Q)(K^2/Q+KQ)}{(PK)^2} = \frac{1}{P}+\frac{1}{Q}+\frac{Q^2}{PK}+\frac{Q}{K}.$$

Whenever the rectangle contains an admissible triple,

$$P, Q \gg_\varepsilon N^\varepsilon, \qquad Q \ll_\varepsilon N^{1/2-\varepsilon}, \qquad PQ \ll_\varepsilon N^{1-\varepsilon}.$$

The bound (147) is therefore $O_\varepsilon(N^{-\varepsilon/2})$. There are $O(L^3)$ triples of blocks (N, P, Q), so their total centred contribution is

$$O_\varepsilon(M^{-\varepsilon/2}L^3) = o(1). \tag{148}$$

For the mean term $S(t)/q$, sum directly within each origin block before splitting the prime sizes. If $J_{N,p,q} \neq \varnothing$, then $p \le (2N)^{1-2\varepsilon}$, and therefore $X = N/p \gg_\varepsilon N^{2\varepsilon}$. Set

$$\eta_\varepsilon(M) = \sup_{X\ge c_\varepsilon M^{2\varepsilon}} \frac{1}{X}\max_{0\le t\le 2X}|S(t)|.$$

By (138), $\eta_\varepsilon(M) \to 0$. Partial summation now gives

$$\frac{1}{q}\left|\sum_{k\in J_{N,p,q}} \frac{\lambda(k)}{k(\log(pk))^2}\right| \ll \frac{\eta_\varepsilon(M)}{qL^2}.$$

Chebyshev's estimate and partial summation imply $\sum_{p\le z}(\log p)/p \ll \log(2z)$. Extending to all primes $p, q \le 2N$, the mean contribution of the block is thus at most

$$\frac{C\eta_\varepsilon(M)}{T_M L^2}\left(\sum_{p\le 2N}\frac{\log p}{p}\right)\left(\sum_{q\le 2N}\frac{\log q}{q}\right) \ll \frac{\eta_\varepsilon(M)}{L}.$$

Summing the $O(L)$ blocks proves, more precisely,

$$\Lambda^{\mathrm{int}}_{\varepsilon,M} \ll_{a,\varepsilon} L^{-1} + M^{-\varepsilon/2}L^3 + \eta_\varepsilon(M) = o(1).$$

□

15.4 The first upper bound for Liouville prediction

Corollary 15.4. *For one exact prime label from each anchor,*

$$\limsup_{M\to\infty} \Lambda_M(\lambda) \le \frac{1}{8}. \tag{149}$$

Consequently, with or without termination bits,

$$\frac{3}{4} \le \liminf_{M} \mathrm{Succ}^*_M(\lambda) \le \limsup_{M} \mathrm{Succ}^*_M(\lambda) \le \frac{13}{16}. \tag{150}$$

Proof. Split the lower signed submeasure on $\mathcal{C}_{\varepsilon,M}$ and its complement:

$$\Lambda_M(\lambda) \le \Lambda^{\mathrm{int}}_{\varepsilon,M} + \mathbb{P}(B_M \setminus \mathcal{C}_{\varepsilon,M}).$$

In logarithmic coordinates the interior region is

$$\Omega_\varepsilon = \{x, y \ge \varepsilon,\ y \le 1/2 - \varepsilon,\ x + y \le 1 - \varepsilon\}.$$

It has a null boundary and area

$$|\Omega_\varepsilon| = \int_\varepsilon^{1/2-\varepsilon} (1 - 2\varepsilon - y)\, dy = \frac{3}{8} - \frac{5}{2}\varepsilon + 4\varepsilon^2.$$

The joint law and $\mathbb{P}(B_M) \to 1/2$ give

$$\limsup_{M} \Lambda_M(\lambda) \le \frac{1}{8} + \frac{5}{2}\varepsilon - 4\varepsilon^2.$$

First take $M \to \infty$ for fixed ε, then $\varepsilon \downarrow 0$, and use Proposition 14.1. □

15.5 Prediction at the second anchor

Corollary 15.5. *For $a \in \{1, 3\}$, both (149) and (150) hold with $\lambda(n)$ replaced by $\lambda(an + 1)$.*

Proof. Use the transposed region

$$\mathcal{C}'_{\varepsilon,M} = \{n^\varepsilon \le P, Q, \quad P \le n^{1/2-\varepsilon}, \quad PQ \le n^{1-\varepsilon}\}. \tag{151}$$

Writing $an + 1 = qk$ gives $n = (qk - 1)/a$, $\lambda(an + 1) = -\lambda(k)$, and $qk \equiv 1 \pmod{ap}$. For large M, $p > a$ on this region, so the congruence is equivalent to the two conditions

$$k \equiv q^{-1} \pmod{p}, \qquad k \equiv q^{-1} \pmod{a}. \tag{152}$$

For each reduced class s mod a, set

$$c_s(k) = \lambda(k)\mathbf{1}_{\{k\equiv s\ (a)\}}, \qquad S_s(t) = \sum_{k\le t} c_s(k).$$

The Davenport estimate cited above gives

$$S_s(t) = o(t) \qquad (s \bmod a \text{ fixed}), \tag{153}$$

because

$$S_s(t) = \frac{1}{a}\sum_{j=0}^{a-1} e(-js/a) \sum_{k\le t} \lambda(k) e(jk/a).$$

For a given q, use $s = q^{-1} \bmod a$. There are only finitely many such classes.

On an origin block $N \le n < 2N$, with $p \sim P$, $q \sim Q$, and $K = N/Q$, the allowed k's form an interval in $[c_a K, C_a K]$. The exact transformed weight in the simplified law is

$$v_q(k) = \frac{1}{((qk-1)/a)(\log((qk-1)/a))^2},$$

positive and decreasing, with maximum and total variation $O_a(1/(NL^2))$. Apply Lemma 15.2 to $c_s(k)$, now with prime modulus p and centre $S_s(t)/p$. For fixed p, each residue $q^{-1} \bmod p$ has multiplicity at most $1 + Q/P$. Restricting q modulo a cannot increase it. Abel summation and Cauchy–Schwarz give a centred contribution at most

$$\frac{C_a \log(2K)}{T_M N}\sqrt{PQ\left(1+\frac{Q}{P}\right)\left(\frac{K^2}{P}+KP\right)} \ll_a \sqrt{\frac{1}{P}+\frac{1}{Q}+\frac{P^2}{N}+\frac{P}{K}}.$$

Here $P, Q \gg N^{\varepsilon}$, $P^2/N \ll N^{-2\varepsilon}$, and $P/K = PQ/N \ll N^{-\varepsilon}$. The total is again $O_{a,\varepsilon}(M^{-\varepsilon/2}L^3)$.

For the mean term, $q \le n^{1-2\varepsilon}$, so $N/q \gg N^{2\varepsilon}$. Define the analogue of $\eta_\varepsilon(M)$ using $\max_{s \bmod a} |S_s(t)|$, $t \le C_a X$. It tends to zero by (153). For each pair, partial summation gives a mean term $O_a(\eta_\varepsilon(M)/(pqL^2))$, before multiplication by $\log p \log q/T_M$: the weight is $O_a(1/(NL^2))$, the prefix mean is $o(N/q)$, and the centre contributes $1/p$. The two prime-weight sums then give $O_a(\eta_\varepsilon(M))$ over all origin blocks. The law comparison adds $O_a(1/L)$. The region therefore has signed norm $o(1)$; its limiting domain is the transpose of Ω_ε, with the same area. The preceding argument proves the claim. □

For $\lambda(n)$, the untreated lower region is $\{y > 1/2,\ x + y < 1\}$, of area $1/8$. For $\lambda(an+1)$, it is the transpose. These two cancellations concern different signs; combining their regions would not cancel the lower measure for either one of the signs.

16 A further gain from smooth cofactors

The preceding large sieve leaves the region $\log Q/\log n > 1/2$. A distribution theorem for multiplicative functions on smooth numbers cancels an additional part of this region. Its mass is strictly positive, although we do not assign it a numerical value.

16.1 The distribution theorem and its prime-modulus form

Let $P^+(m)$ be the largest prime factor of m, with $P^+(1) = 1$, and write $S(X,Y) = \{m \le X : P^+(m) \le Y\}$. Theorem 2.2 of Drappeau–Granville–Shao [16] states that, for every $\epsilon_0 > 0$, there are $C, \delta > 0$ such that

$$\sum_{\substack{q \le X^{3/5-\epsilon_0} \\ (q, a_1 a_2)=1}} \left| \sum_{m \in S(X,Y)} f(m) u_D(m\overline{a_1}a_2; q) \right| \ll_{\epsilon_0} D^{-1/2} X (\log X)^{10}, \tag{154}$$

uniformly for

$$1 \le D \le X^{\delta}, \qquad (\log X)^C \le Y \le X^{\delta}, \qquad 0 < |a_1|, |a_2| \le X^{\delta}.$$

Here an overline denotes a modular inverse, and

$$u_D(v; q) = \mathbf{1}_{v \equiv 1 \ (q)} - \frac{1}{\varphi(q)} \sum_{\substack{\chi \bmod q \\ \mathrm{cond}(\chi) \le D}} \chi(v).$$

The theorem applies to the class $\mathcal{C}$ of multiplicative functions whose logarithmic-derivative coefficients are bounded in modulus by those of $-\zeta'/\zeta$. In particular it applies to every completely multiplicative function of modulus at most one, including 1 and λ.

If q is prime and $q > \max(D, Y)$, the only character modulo q of conductor at most D is the principal character, and every $m \in S(X, Y)$ is coprime to q. Thus the correction term becomes exactly

$$\frac{1}{q-1} \sum_{m \in S(X,Y)} f(m).$$

This prime-modulus specialization is unconditional.

Pascadi's later distribution theorem [17, Theorem 1.7], in the published numbering, reaches moduli of size $X^{66/107-\epsilon_0}$ for 1-bounded completely multiplicative functions supported on Y-smooth integers, under its Siegel–Walfisz criterion. For its associated constant $\delta_{\mathrm{P}} > 0$, the range includes $\exp(\sqrt{\log X \log\log X}) \le Y \le X^{\delta_{\mathrm{P}}}$ and rational classes with $0 < |a_i| \le X^{\delta_{\mathrm{P}}}$; the summed error has arbitrary logarithmic saving relative to the smooth-number count. The conductor-truncated input is Theorem 4.2 of that paper. We use (154), whose short-conductor subtraction specializes directly as above. Improving the gain by the later theorem requires checking the conductor terms, the uniformity in the smoothness threshold and rational classes, and the fixed-character twists at the second anchor. That additional transfer is not used in the bounds below.

Fix $\epsilon_0 = 1/100$, decrease the admissible δ in (154) if necessary so that $0 < \delta \le 1/100$, and set

$$\eta = \delta/10, \qquad c = 51/100, \qquad d = 11/20. \tag{155}$$

Then $d - c = 1/25$. For an origin block $N \le n < bN$, $1 < b \le 2$, put

$$\mathcal{P}_N = \{p \text{ prime} : N^{\eta} < p \le N^{2\eta}\}, \quad \mathcal{Q}_N = \{q \text{ prime} : N^c < q \le N^d\}, \quad Y = N^{2\eta}.$$

For $p \in \mathcal{P}_N$, write $K = N/p$ and $r_{p,q} = -(ap)^{-1} \bmod q$. Uniformly for $t \in [K, 2K]$, once N is large,

$$ap \le t^{\delta}, \quad (\log t)^C \le Y \le t^{\delta}, \quad q \le t^{59/100}, \quad q > \max(Y, t^{\delta/2}).$$

Indeed, $\eta \le 10^{-3}$, $2\eta < \delta(1-2\eta)$, and $d < (59/100)(1-2\eta)$. Apply (154) with $a_1 = -1$, $a_2 = ap$, and $D = t^{\delta/2}$. For $f = 1, \lambda$, this gives

$$\sum_{q \in \mathcal{Q}_N} |E_{p,q,f}(t)| \ll t^{1-\delta/4} (\log t)^{10}, \tag{156}$$

where

$$E_{p,q,f}(t) = \sum_{\substack{k \le t \\ P^+(k) \le Y \\ k \equiv r_{p,q} \ (q)}} f(k) - \frac{1}{q-1} \sum_{\substack{k \le t \\ P^+(k) \le Y}} f(k).$$

The constant is independent of p, t, N and $a \in \{1, 3\}$. In these ranges $q > p$ and $q \nmid a$.

16.2 The mean of the smoothly truncated Liouville function

The mean of λ on all integers does not directly give its mean after a smoothness restriction. The next convolution argument supplies that implication.

Lemma 16.1 (A fixed-power smoothness threshold)**.** *For each fixed $0 < \theta \le 1$,*

$$\sup_{X^{\theta} \le Y \le X} \frac{1}{X} \left| \sum_{\substack{k \le X \\ P^+(k) \le Y}} \lambda(k) \right| \longrightarrow 0. \tag{157}$$

Proof. The function $\lambda_Y(k) = \lambda(k)\mathbf{1}_{P^+(k)\le Y}$ is completely multiplicative. Put

$$h_Y(d) = \mu(d)^2\mathbf{1}_{\{p\mid d \implies p>Y\}}.$$

A check at prime powers gives the exact identity

$$\lambda_Y = \lambda * h_Y. \tag{158}$$

Let $S_\lambda(u) = \sum_{k\le u}\lambda(k)$. For fixed $H \ge 2$, the convolution implies

$$\frac{1}{X}\left|\sum_{k\le X}\lambda_Y(k)\right| \le \left(\sup_{u\ge H}\frac{|S_\lambda(u)|}{u}\right)\sum_{d\le X}\frac{h_Y(d)}{d} + \sum_{X/H<d\le X}\frac{h_Y(d)}{d}.$$

Mertens' formula gives, uniformly for $Y \ge X^\theta$,

$$\sum_{d\le X}\frac{h_Y(d)}{d} \le \prod_{Y<p\le X}(1+1/p) \ll_\theta 1.$$

Fix an integer $B \ge 2$. Once $Y > B$, every contributing d is coprime to $\prod_{p\le B}p$, so

$$\sum_{X/H<d\le X}\frac{h_Y(d)}{d} \le \frac{H}{X}\#\{d\le X : (d,\prod_{p\le B}p)=1\} \le H\prod_{p\le B}(1-1/p) + o_{B,H}(1).$$

For fixed H, take $X\to\infty$, then $B\to\infty$. Finally let $H\to\infty$ and use $S_\lambda(u) = o(u)$. Every bound was uniform in $X^\theta \le Y \le X$. □

16.3 The smooth event and its mass

Partition I_M into disjoint blocks $[N_j, b_jN_j)$, where $N_j = M2^j$ and $1 < b_j \le 2$. Only the final block may have $b_j < 2$. Define

$$\mathcal{F}_M = \bigcup_j\left\{N_j \le n < b_jN_j,\quad P\in\mathcal{P}_{N_j},\quad Q\in\mathcal{Q}_{N_j},\quad P^+(n/P)\le N_j^{2\eta}\right\}. \tag{159}$$

For large M, $\mathcal{F}_M \subset \{PQ\le n,\ Q>\sqrt{n}\}$: indeed $PQ \le N_j^{2\eta+d} < N_j \le n$ and $Q > N_j^c > \sqrt{2N_j} \ge \sqrt{n}$. As before, this source event is used only to split a signed measure; it need not be observable from the labels.

Proposition 16.2 (The smooth submeasure)**.** *For η fixed by (155),*

$$\sum_{p,q}\left|\mathbb{E}[\lambda(n)\mathbf{1}_{\{P=p,Q=q\}}\mathbf{1}_{\mathcal{F}_M}]\right| = o(1), \tag{160}$$

$$\mathbb{P}(\mathcal{F}_M) \longrightarrow \gamma_\eta := \frac{1}{25}\int_\eta^{2\eta}\rho\left(\frac{1-x}{2\eta}\right)dx > 0. \tag{161}$$

Here ρ is the Dickman function: $\rho(u) = 1$ for $0\le u\le 1$, and $u\rho'(u) = -\rho(u-1)$ for $u>1$.

Proof. Apply (64) after restricting to $\mathcal{F}_M$, at cost $O_a(1/L)$. On a block $[N, bN)$, with $K = N/p$, the resulting sum is

$$\frac{1}{T_M}\sum_{p\in\mathcal{P}_N}\frac{\log p}{p}\sum_{q\in\mathcal{Q}_N}\log q\sum_{\substack{K\le k<bK\\ P^+(k)\le Y\\ k\equiv r_{p,q}\ (q)}}\frac{f(k)}{k(\log(pk))^2}.$$

For the signed norm take $f = \lambda$, insert absolute values around the inner sum, and use the triangle inequality over origin blocks. For the mass take $f = 1$. The factor $\lambda(p) = -1$ has no effect on the signed norm.

For $w_p(t) = 1/(t(\log(pt))^2)$,

$$\sup_{K \le t \le 2K} w_p(t) + \int_K^{2K} |w_p'(t)|\, dt \ll \frac{1}{K(\log N)^2}.$$

Partial summation using (156) before the sum over p shows that replacing every progression sum by its principal term costs at most

$$O\left(\frac{1}{T_M} \sum_{p \in \mathcal{P}_N} \frac{\log p}{p} K^{-\delta/4} (\log N)^9\right) = O\left(\frac{N^{-(1-2\eta)\delta/4} (\log N)^{10}}{T_M}\right).$$

Its sum over the origin blocks is $o(1)$. The estimate is used with $\sum_q |E_{p,q,f}(t)|$ inside the Abel integral; it does not require a simultaneous maximum over all t.

For $f = \lambda$, Lemma 16.1 and another partial summation give

$$\sum_{\substack{K \le k < bK \\ P^+(k) \le Y}} \frac{\lambda(k)}{k(\log(pk))^2} = o((\log N)^{-2}),$$

uniformly in p and $b \in [1, 2]$, since $Y \ge (2K)^\eta$ for large N. Both prime-weight sums

$$\sum_{p \in \mathcal{P}_N} \frac{\log p}{p}, \qquad \sum_{q \in \mathcal{Q}_N} \frac{\log q}{q - 1}$$

are $O(\log N)$. The signed principal term is therefore $o(1/T_M)$ per block, uniformly for $N \ge M$. There are $O(\log M)$ blocks and $T_M \sim \log M$, which proves (160).

For $f = 1$, the Dickman–de Bruijn estimate [22, Section 1, p. 51, equations (1.1)–(1.4)],

$$\Psi(X, Y) = X\rho(\log X / \log Y) + o(X),$$

is uniform when the quotient stays in a fixed compact subset of $(0, \infty)$. Here ρ is defined by equation (1.2) of that paper. The compact uniformity follows from combining its estimates (1.3) and (1.4): after division by X, their errors are bounded by a constant depending on the compact set times $R(Y) + 1/\log Y + 1/X$, where $R(Y) \to 0$. The proof of (1.4) is in Section 4 of [22]. Partial summation gives, uniformly in $b \in [1, 2]$,

$$\sum_{\substack{K \le k < bK \\ P^+(k) \le Y}} \frac{1}{k(\log(pk))^2} = \frac{\log b}{(\log N)^2} \rho\left(\frac{1 - \log p / \log N}{2\eta}\right) + o((\log N)^{-2}).$$

Mertens' formula and Stieltjes summation in $x = \log p / \log N$ then yield

$$\sum_{q \in \mathcal{Q}_N} \frac{\log q}{q - 1} = \frac{\log N}{25} + O(1),$$

$$\sum_{p \in \mathcal{P}_N} \frac{\log p}{p} \rho\left(\frac{1 - \log p / \log N}{2\eta}\right) = \log N \int_\eta^{2\eta} \rho\left(\frac{1 - x}{2\eta}\right) dx + o(\log N).$$

The block mass is $(\log b)\gamma_\eta / T_M + o(1/T_M)$. Since $\sum_j \log b_j = \log M$, this proves (161). Positivity follows from $\rho(u) > 0$ for every finite u. □

16.4 A strict upper bound and an exact localization

Corollary 16.3 (A strict improvement on 13/16)**.** *We have*

$$\limsup_M \Lambda_M(\lambda) \le \frac{1}{8} - \gamma_\eta, \tag{162}$$

and hence

$$\frac{3}{4} \le \liminf_M \mathrm{Succ}^*_M(\lambda) \le \limsup_M \mathrm{Succ}^*_M(\lambda) \le \frac{13}{16} - \frac{\gamma_\eta}{2} < \frac{13}{16}. \tag{163}$$

Proof. The smooth event and $\mathcal{C}_{\varepsilon,M}$ are disjoint for large M: the former has $Q > \sqrt{n}$, the latter $Q \le n^{1/2-\varepsilon}$. Their signed norms tend to zero, while their limiting masses are respectively γ_η and $|\Omega_\varepsilon|$. The triangle inequality bounds the remaining signed norm by $\mathbb{P}(B_M \setminus (\mathcal{C}_{\varepsilon,M} \cup \mathcal{F}_M))$. Take $M \to \infty$, then $\varepsilon \downarrow 0$, and use (128). □

For any source event E, write

$$b_E(p,q) = \mathbb{E}[\lambda(n)\mathbf{1}_{\{P=p,Q=q\}}\mathbf{1}_E], \qquad \|b_E\|_1 = \sum_{p,q} |b_E(p,q)|.$$

Define

$$\mathcal{D}_M = \{PQ \le n,\ Q > \sqrt{n}\}, \qquad \mathfrak{R}_M = \mathcal{D}_M \setminus \mathcal{F}_M. \tag{164}$$

The symbol $\mathfrak{R}_M$ denotes an event, distinct from the reconstruction probabilities $\mathcal{R}_M$.

Proposition 16.4 (Localization up to $o(1)$)**.** *The signed remainder satisfies*

$$\|b_{B_M \setminus \mathcal{D}_M}\|_1 \longrightarrow 0, \tag{165}$$

$$\Lambda_M(\lambda) = \|b_{\mathfrak{R}_M}\|_1 + o(1), \tag{166}$$

$$\mathbb{P}(\mathfrak{R}_M) \longrightarrow \frac{1}{8} - \gamma_\eta. \tag{167}$$

Consequently,

$$\mathrm{Succ}^*_M(\lambda) = \frac{3}{4} + \frac{1}{2}\|b_{\mathfrak{R}_M}\|_1 + o(1). \tag{168}$$

Proof. The region $\mathcal{C}_{\varepsilon,M}$ is contained in $B_M \setminus \mathcal{D}_M$, and

$$\|b_{B_M \setminus \mathcal{D}_M}\|_1 \le \Lambda^{\mathrm{int}}_{\varepsilon,M} + \mathbb{P}((B_M \setminus \mathcal{D}_M) \setminus \mathcal{C}_{\varepsilon,M}).$$

The first term tends to zero, while the limiting mass in the second is $3/8 - |\Omega_\varepsilon|$. Letting $\varepsilon \downarrow 0$ proves (165). The disjoint decomposition

$$B_M = (B_M \setminus \mathcal{D}_M) \,\dot\cup\, \mathcal{F}_M \,\dot\cup\, \mathfrak{R}_M$$

and the reverse triangle inequality give

$$|\Lambda_M(\lambda) - \|b_{\mathfrak{R}_M}\|_1| \le \|b_{B_M \setminus \mathcal{D}_M}\|_1 + \|b_{\mathcal{F}_M}\|_1 = o(1).$$

Finally, $\mathbb{P}(\mathcal{D}_M) \to 1/8$ and $\mathbb{P}(\mathcal{F}_M) \to \gamma_\eta$. □

The limiting mass of $\mathfrak{R}_M$ bounds its signed norm; it is not a formula for that norm. The external theorem supplies an admissible $\delta > 0$, but no particular numerical value is extracted here. The constant γ_η is nevertheless positive, with

$$\gamma_\eta \ge \frac{\eta}{25}\rho\left(\frac{1-\eta}{2\eta}\right) > 0.$$

The transfer to the second anchor requires its own fixed-modulus constraint and its own weights; we establish it next.

16.5 From prediction advantage to binary information

The comparison between Bayes error and conditional entropy is classical; see Hellman–Raviv [27]. We give the elementary binary argument in the form needed for our observable confidence region. Entropies and mutual informations are in nats. Division by $\log 2$ converts them to bits. Let

$$h(u) = -u \log u - (1-u)\log(1-u), \qquad 0 \log 0 = 0.$$

Proposition 16.5 (Binary information bounds)**.** *Let $Z_M \in \{-1, 1\}$ be a deterministic sign of the origin, and let T_M be the observed transcript. Suppose $\mathbb{E}Z_M = o(1)$, and suppose there is a T_M-measurable event A_M of probability $1/2 + o(1)$ and a predictor $\widehat{Z}_M(\mathsf{T}_M)$ such that $\mathbb{P}(A_M, \widehat{Z}_M \neq Z_M) = o(1)$. If, for a fixed $0 \le \beta \le 1/2$,*

$$\|\mathbb{E}[Z_M \mid \mathsf{T}_M]\|_1 \le \frac{1}{2} + \beta + o(1),$$

then

$$\left(\frac{1}{2} - \beta\right) \log 2 - o(1) \le H(Z_M \mid \mathsf{T}_M) \le \frac{1}{2} \log 2 + o(1), \tag{169}$$

$$\frac{1}{2} \log 2 - o(1) \le I(Z_M; \mathsf{T}_M) \le \left(\frac{1}{2} + \beta\right) \log 2 + o(1). \tag{170}$$

Existence of limits is not assumed.

Proof. Set $m_M = \mathbb{E}[Z_M \mid \mathsf{T}_M]$. Concavity of h on the two halves of $[0, 1]$ gives

$$h\left(\frac{1+t}{2}\right) \ge (1 - |t|) \log 2 \qquad (-1 \le t \le 1).$$

Hence

$$H(Z_M \mid \mathsf{T}_M) = \mathbb{E}h((1 + m_M)/2) \ge (1 - \mathbb{E}|m_M|) \log 2.$$

For the other direction, let $E_M = \mathbf{1}_{\{\widehat{Z}_M \neq Z_M\}}$. Given the transcript, this error bit determines Z_M. On A_M, its conditional error probability is $e_M = \mathbb{P}(E_M = 1 \mid A_M) = o(1)$. Since A_M is observable,

$$H(Z_M \mid \mathsf{T}_M) \le \mathbb{P}(A_M) h(e_M) + \mathbb{P}(A_M^c) \log 2 = \frac{1}{2} \log 2 + o(1).$$

The marginal mean hypothesis gives $H(Z_M) = \log 2 + o(1)$. Subtracting the conditional entropy bounds proves the mutual information bounds. □

For $\lambda(n)$, marginal balance follows from the prime number theorem and partial summation on I_M. For $\lambda(an + 1)$, use the fixed class 1 mod a. Corollary 13.2 supplies the observable half on which either sign is recovered with error $o(1)$. Taking $\beta = 1/8$ gives, for either sign Z_M,

$$\frac{3}{8} \le \liminf_M \frac{H(Z_M \mid \mathsf{T}_M)}{\log 2} \le \limsup_M \frac{H(Z_M \mid \mathsf{T}_M)}{\log 2} \le \frac{1}{2},$$

and

$$\frac{1}{2} \le \liminf_M \frac{I(Z_M; \mathsf{T}_M)}{\log 2} \le \limsup_M \frac{I(Z_M; \mathsf{T}_M)}{\log 2} \le \frac{5}{8}.$$

For $\lambda(n)$, the smooth improvement permits $\beta = 1/8 - \gamma_\eta$, so

$$\liminf_M \frac{H(\lambda(n) \mid \mathsf{T}_M)}{\log 2} \ge \frac{3}{8} + \gamma_\eta, \qquad \limsup_M \frac{I(\lambda(n); \mathsf{T}_M)}{\log 2} \le \frac{5}{8} - \gamma_\eta. \tag{171}$$

All these bounds remain valid with the termination bits, because the signed reduction retains the same lower remainder.

17 The smooth gain at the second anchor

We keep the parameters δ, η, c, d from (155). To obtain the same positive gain for $\lambda(an+1)$, the congruence modulo a must be handled before transferring the observation weights.

17.1 Smooth truncation of a completely multiplicative function

The convolution argument has a general form that will also apply to character twists.

Lemma 17.1 (Preservation of mean zero under smooth truncation)**.** *Let $f : \mathbb{N} \to \mathbb{C}$ be a fixed completely multiplicative function satisfying $f(1) = 1$, $|f(n)| \le 1$, and*

$$S_f(X) := \sum_{n \le X} f(n) = o(X).$$

For every fixed $0 < \theta \le 1$,

$$\sup_{X^\theta \le Y \le X} \frac{1}{X} \left| \sum_{\substack{n \le X \\ P^+(n) \le Y}} f(n) \right| \longrightarrow 0. \tag{172}$$

The error may depend on f and θ.

Proof. Set $f_Y(n) = f(n)\mathbf{1}_{\{P^+(n) \le Y\}}$ and

$$h_{f,Y}(d) = \mu(d) f(d) \mathbf{1}_{\{p \mid d \implies p > Y\}}.$$

At a power p^j with $p > Y$, convolution gives $f(p)^j - f(p)f(p)^{j-1} = 0$; at primes $p \le Y$ it leaves f unchanged. Hence

$$f_Y = f * h_{f,Y}, \qquad |h_{f,Y}(d)| \le \mu(d)^2 \mathbf{1}_{\{p \mid d \implies p > Y\}}. \tag{173}$$

For fixed $H \ge 2$,

$$\frac{|S_{f_Y}(X)|}{X} \le \sup_{u \ge H} \frac{|S_f(u)|}{u} \sum_{d \le X} \frac{|h_{f,Y}(d)|}{d} + \sum_{X/H < d \le X} \frac{|h_{f,Y}(d)|}{d}.$$

The first weight sum is bounded by $\prod_{Y < p \le X}(1 + 1/p) \ll_\theta 1$. For fixed $B \ge 2$, put $A_B = \prod_{p \le B} p$. Once $Y > B$, the second sum is at most

$$\frac{H}{X} \#\{d \le X : (d, A_B) = 1\} \le H \frac{\varphi(A_B)}{A_B} + o_{B,H}(1).$$

Take successively $X \to \infty$, $B \to \infty$, and $H \to \infty$. All estimates are uniform for $Y \ge X^\theta$. □

Lemma 17.2 (Smooth numbers in fixed residue classes)**.** *Fix $a \ge 1$ and a reduced class $s \bmod a$. If $Y \le X$ and $\log X / \log Y$ remains in a fixed compact subset of $(0, \infty)$, then*

$$\sum_{\substack{k \le X \\ P^+(k) \le Y \\ k \equiv s \ (a)}} 1 = \frac{X}{a} \rho\left(\frac{\log X}{\log Y}\right) + o_a(X), \tag{174}$$

$$\sum_{\substack{k \le X \\ P^+(k) \le Y \\ k \equiv s \ (a)}} \lambda(k) = o_a(X). \tag{175}$$

The errors are uniform over the residue classes and over each such compact set of quotients.

Proof. For every character $\chi \bmod a$, both χ and $\lambda\chi$ are completely multiplicative and bounded by one in modulus. A nonprincipal character has partial sums $O_a(1)$. The Davenport estimate, projected onto the finitely many classes modulo a, gives $\sum_{k\le X} \lambda(k)\chi(k) = o_a(X)$, including for the principal character. Lemma 17.1 applies to these functions.

For the unsigned principal character $\chi_{0,a}$, once Y exceeds all prime factors of a, inclusion–exclusion gives

$$\sum_{\substack{k\le X\\ P^+(k)\le Y}} \chi_{0,a}(k) = \sum_{e\mid \mathrm{rad}(a)} \mu(e)\Psi(X/e,Y) = X\rho\left(\frac{\log X}{\log Y}\right)\frac{\varphi(a)}{a} + o_a(X),$$

using the Dickman–de Bruijn estimate and continuity of ρ. Finally use

$$\mathbf{1}_{\{k\equiv s\ (a)\}} = \frac{1}{\varphi(a)} \sum_{\chi \bmod a} \overline{\chi(s)}\,\chi(k).$$

For $a = 1$, use the constant character and $\varphi(1) = 1$. □

The density in (174) is $1/a$. The factor $\varphi(a)/a$ from coprimality is already included before character orthogonality divides by $\varphi(a)$.

17.2 The transposed smooth event

On a block $N \le n < bN$, put

$$\mathcal{Q}_N^- = \{q \text{ prime} : N^\eta < q \le N^{2\eta}\}, \quad \mathcal{P}_N^+ = \{p \text{ prime} : N^c < p \le N^d\}, \quad Y = N^{2\eta}.$$

With the same disjoint origin blocks as before, define

$$\mathcal{F}_M^{(1)} = \bigcup_j \left\{ N_j \le n < b_j N_j, \quad Q \in \mathcal{Q}_{N_j}^-, \quad P \in \mathcal{P}_{N_j}^+, \quad P^+((an+1)/Q) \le N_j^{2\eta} \right\}. \tag{176}$$

For large M, $\mathcal{F}_M^{(1)} \subset \{PQ \le n,\ P > \sqrt{n}\}$. For $q \in \mathcal{Q}_N^-$, write

$$an + 1 = qk, \qquad A_q = \frac{aN+1}{q}, \qquad B_q = \frac{abN+1}{q}, \qquad K = \frac{N}{q}.$$

The exact interval is $A_q \le k < B_q$, and the congruences are

$$k \equiv q^{-1} \pmod p, \qquad k \equiv s_q := q^{-1} \pmod a. \tag{177}$$

For sufficiently large N, we have $p, q > a$ and $p \ne q$. Thus these two conditions are equivalent to $qk \equiv 1 \pmod{ap}$.

For $f = 1, \lambda$, define

$$F_{f,s}(t,Y) = \sum_{\substack{k\le t\\ P^+(k)\le Y\\ k\equiv s\ (a)}} f(k)$$

and

$$E_{q,p,f,s}(t) = \sum_{\substack{k\le t\\ P^+(k)\le Y\\ k\equiv q^{-1}\ (p)\\ k\equiv s\ (a)}} f(k) - \frac{F_{f,s}(t,Y)}{p-1}.$$

Expand the fixed-class indicator in characters modulo a. For each character, apply (154) to $f\chi$ with $a_1 = 1$, $a_2 = q$, prime modulus p, and $D = t^{\delta/2}$. Uniformly for $t \in [A_q, B_q]$, $t \asymp_a K$, and

$$q, Y \le t^{\delta}, \qquad (\log t)^C \le Y, \qquad p \le t^{59/100}, \qquad p > \max(Y, D).$$

Only the principal character modulo p occurs in the correction, and the smooth cofactors are coprime to p. Projection onto s mod a therefore yields

$$\sum_{p \in \mathcal{P}_N^+} |E_{q,p,f,s}(t)| \ll_a t^{1-\delta/4} (\log t)^{10}, \tag{178}$$

uniformly in q, s, t, N in these ranges. Although D varies with t, the specialized correction remains $F_{f,s}(t, Y)/(p-1)$.

17.3 Transferring the weights through the affine change

Proposition 17.3 (The transposed smooth submeasure)**.** *For $a \in \{1, 3\}$,*

$$\sum_{p,q} \left| \mathbb{E}[\lambda(an+1) \mathbf{1}_{\{P=p, Q=q\}} \mathbf{1}_{\mathcal{F}_M^{(1)}}] \right| = o(1), \tag{179}$$

$$\mathbb{P}(\mathcal{F}_M^{(1)}) \longrightarrow \gamma_\eta = \frac{1}{25} \int_\eta^{2\eta} \rho\left(\frac{1-x}{2\eta}\right) dx > 0. \tag{180}$$

Proof. By (64), it suffices, with error $O_a(1/L)$, to use the mass

$$\frac{\log p \log q}{T_M n (\log n)^2} \mathbf{1}_{\{p \mid n,\ q \mid an+1\}}.$$

After $an + 1 = qk$, its weight in k is

$$v_q(t) = \frac{1}{((qt-1)/a)\{\log((qt-1)/a)\}^2}.$$

It is positive and decreasing on $[A_q, B_q]$, with

$$\sup v_q + \int_{A_q}^{B_q} |v_q'(t)|\, dt \ll_a \frac{1}{N(\log N)^2}.$$

The block mass is

$$\frac{1}{T_M} \sum_{q \in \mathcal{Q}_N^-} \log q \sum_{p \in \mathcal{P}_N^+} \log p \sum_{\substack{A_q \le k < B_q \\ P^+(k) \le Y \\ k \equiv q^{-1} \ (p) \\ k \equiv s_q \ (a)}} v_q(k).$$

For the signed norm insert $\lambda(k)$, take the absolute value after the k-sum, and sum the block bounds. The factor $\lambda(q) = -1$ does not affect this norm. Abel summation using (178), with the sum over p inside the integral, bounds the total progression error on the block by

$$\frac{C_a}{T_M} \sum_{q \in \mathcal{Q}_N^-} \frac{\log q}{q} K^{-\delta/4} (\log N)^9 \ll_a \frac{N^{-(1-2\eta)\delta/4} (\log N)^{10}}{T_M}.$$

The sum of these errors is $o(1)$.

For the signed principal term, Lemma 17.2 gives $F_{\lambda, s_q}(t, Y) = o_a(t)$ uniformly in the block; indeed $Y \ge t^\eta$ for large N. Thus

$$\sum_{\substack{A_q \le k < B_q \\ P^+(k) \le Y \\ k \equiv s_q \ (a)}} \lambda(k) v_q(k) = o_a\left(\frac{1}{q(\log N)^2}\right).$$

The estimates

$$\sum_{q\in\mathcal{Q}_N^-}\frac{\log q}{q}=O(\log N),\qquad \sum_{p\in\mathcal{P}_N^+}\frac{\log p}{p-1}=O(\log N)$$

give $o_a(1/T_M)$ per block and $o(1)$ in total. This proves (179).

For the mass, set $x_q=\log q/\log N$. Uniformly on $[A_q,B_q]$, the same lemma gives

$$F_{1,s_q}(t,Y)=\frac{t}{a}\,\rho\left(\frac{1-x_q}{2\eta}\right)+o_a(t).$$

The dependence of $\log t$ on a,b, and the additive terms 1, is $O_a(1)$, hence disappears after division by $\log Y=2\eta\log N$. Partial summation yields

$$\sum_{\substack{A_q\le k<B_q\\ P^+(k)\le Y\\ k\equiv s_q\ (a)}} v_q(k)=\frac{1}{a}\,\rho\left(\frac{1-x_q}{2\eta}\right)\int_{A_q}^{B_q}v_q(t)\,dt+o_a\left(\frac{1}{q(\log N)^2}\right). \tag{181}$$

Endpoint atoms cost $O_a(1/(N(\log N)^2))$, which the displayed error absorbs. The exact substitution $u=(qt-1)/a$ gives

$$\frac{1}{a}\int_{A_q}^{B_q}v_q(t)\,dt=\frac{1}{q}\int_N^{bN}\frac{du}{u(\log u)^2}=\frac{\log b}{q(\log N)^2}+O\left(\frac{1}{q(\log N)^3}\right). \tag{182}$$

Thus the fixed-class density $1/a$ cancels the Jacobian factor a. Mertens' formula and Stieltjes summation give

$$\sum_{p\in\mathcal{P}_N^+}\frac{\log p}{p-1}=\frac{\log N}{25}+O(1),$$

$$\sum_{q\in\mathcal{Q}_N^-}\frac{\log q}{q}\rho\left(\frac{1-x_q}{2\eta}\right)=\log N\int_\eta^{2\eta}\rho\left(\frac{1-x}{2\eta}\right)\,dx+o(\log N).$$

The block mass is consequently $(\log b)\gamma_\eta/T_M+o_a(1/T_M)$, uniformly for $b\in[1,2]$. Sum over the blocks to obtain (180). □

17.4 The two signed remainders

Corollary 17.4 (The same smooth gain for both anchors)**.** *For $a=1,3$ and either $W(n)=\lambda(n)$ or $W(n)=\lambda(an+1)$,*

$$\limsup_M \Lambda_M(W)\le\frac{1}{8}-\gamma_\eta,$$

and

$$\frac{3}{4}\le\liminf_M \mathrm{Succ}_M^*(W)\le\limsup_M \mathrm{Succ}_M^*(W)\le\frac{13}{16}-\frac{\gamma_\eta}{2}<\frac{13}{16}. \tag{183}$$

The conclusion includes observations with termination bits.

Proof. For $W(n)=\lambda(an+1)$, the large sieve cancels $\mathcal{C}'_{\varepsilon,M}$, whose mass tends to $3/8$ as $M\to\infty$ and then $\varepsilon\downarrow 0$. This event has $P\le n^{1/2-\varepsilon}$ and is disjoint from $\mathcal{F}_M^{(1)}$. Bound the signed norm on the remaining part of B_M by its mass, and apply (128). The first-anchor statement was proved above. □

The precise localization also transposes. Define

$$\mathcal{D}_M^{(1)} = \{PQ \le n,\ P > \sqrt{n}\}, \qquad \mathfrak{R}_M^{(1)} = \mathcal{D}_M^{(1)} \setminus \mathcal{F}_M^{(1)}, \qquad b_E^{(1)}(p,q) = \mathbb{E}[\lambda(an+1)\mathbf{1}_{\{P=p,Q=q\}}\mathbf{1}_E].$$

Exhausting the transposed large-sieve region gives $\|b^{(1)}_{B_M \setminus \mathcal{D}_M^{(1)}}\|_1 = o(1)$. The reverse triangle inequality and (179) then give

$$\Lambda_M(\lambda(an+1)) = \|b^{(1)}_{\mathfrak{R}_M^{(1)}}\|_1 + o(1), \qquad \mathbb{P}(\mathfrak{R}_M^{(1)}) \longrightarrow \frac{1}{8} - \gamma_\eta. \tag{184}$$

The smooth-truncation lemma applies to any fixed completely multiplicative f, bounded by one, with mean zero. For the transfer to the second anchor, it is enough that f have mean zero in every reduced class modulo the fixed a; orthogonality then gives mean zero for each $f\chi$. A global mean-zero assumption alone does not replace this fixed-class hypothesis. The errors may depend on f, a; no uniformity for a varying family $f = f_M$ is asserted.

18 An explicit region without a smoothness restriction

A small explicit part of the remaining region can be treated without requiring smooth cofactors. The required estimate adapts Green's dispersion argument to rational residue classes, keeping track of the cost of their denominator.

18.1 Distribution in rational residue classes

For $Q \ge 2$, write $\mathcal{Q}(Q) = \{q \text{ prime} : Q \le q < 2Q\}$, and let $e_t(v) = \exp(2\pi iv/t)$. We use Green's Proposition 2.2 [18], a general reduction by Ramaré weights, and Theorem 1 of Bettin–Chandee [20], applied to bilinear forms with Kloosterman fractions. Green's theorem for a fixed integer residue is not itself a theorem about the varying rational classes below.

The exponent 20/39 is the one furnished by the particular Green–Bettin–Chandee argument adapted here. Fouvry–Radziwiłł [19, Corollary 1.3 and Theorem 1.2] reach moduli up to $X^{17/33-\epsilon}$ for a common integer residue. In the cited arXiv numbering, Corollary 1.3 assumes a Siegel–Walfisz condition on the values at primes and gives summed error $O_\epsilon(X/(\log X)^{1-\epsilon})$. Theorem 1.2 gives a convolution estimate over prime moduli without that condition. A transfer to the present problem would still have to supply uniformity for the growing denominator in $c/d \bmod q$ and the $o(X/\log X)$ norm required for the weighted summation. We establish that precise contract below.

Proposition 18.1 (Rational classes with a small denominator)**.** *Fix $a_0 \ge 1$, $c \in \{-1, 1\}$, and real numbers*

$$0 < \kappa_0 \le \kappa, \qquad \frac{1}{2} \le \vartheta, \qquad 39\vartheta + 22\kappa < 20.$$

Let r be prime, $d = a_0 r$, and suppose

$$X^{\kappa_0} \le r \le X^{\kappa}, \qquad X^{1/2} \le Q \le X^{\vartheta}.$$

Uniformly in these parameters and in every multiplicative function f with $|f| \le 1$,

$$\sum_{q \in \mathcal{Q}(Q)} \left| \sum_{\substack{n \le X \\ dn \equiv c \ (q)}} f(n) - \frac{1}{q} \sum_{n \le X} f(n) \right| = o_{\kappa_0,\kappa,\vartheta,a_0}(X/\log X). \tag{185}$$

For sufficiently large X, all these q are coprime to d.

Proof. Put

$$\Delta = 1 - \frac{39}{20}\vartheta - \frac{11}{10}\kappa > 0.$$

In particular, $\kappa < 1/44$, $\vartheta < 20/39$, and $\Delta < 1/40$. Choose

$$0 < \zeta < \min\{\kappa_0/2, \Delta/20, 1/100\}, \qquad s = \Delta/20, \qquad \epsilon_{\mathrm{BC}} = \Delta/100.$$

These choices are independent of the tolerance introduced next. For fixed $0 < \varepsilon < 1/2$, set

$$Y = X^{\varepsilon\zeta}, \qquad Z = X^{\zeta}.$$

The choice $\zeta < 1/100$ gives $1 < Y < Z < X^{1/16}$, as required by the corrected form of [18, Proposition 2.2]; also $\log Z/\log Y = \varepsilon^{-1} > 2$. For large X, Y exceeds every prime factor of a_0 and $Z < r$. Hence every auxiliary prime in $[Y, Z)$ is coprime to d.

Take complex coefficients $|\xi_q| \le 1$, supported on $\mathcal{Q}(Q)$, and define

$$F(n) = \sum_{q \in \mathcal{Q}(Q)} \xi_q \left(\mathbf{1}_{dn \equiv c \ (q)} - \frac{1}{q} \right) \qquad (|n| \le 2ZX),$$

extending F by zero elsewhere. This extension covers every value used in the original sum, the auxiliary correlations, and their smoothing. For an integer n in this range, $dn - c \neq 0$ for large X. It cannot have three distinct prime divisors in $[Q, 2Q)$: that would give $X^{3/2} \le Q^3 \le |dn - c| \ll X^{1+\kappa+\zeta}$. Also $\sum_{q \in \mathcal{Q}(Q)} 1/q \ll 1$. Thus

$$|F(n)| \ll 1 \tag{186}$$

uniformly in d, Q, ξ.

Green's Proposition 2.2 gives

$$\left| \frac{1}{X} \sum_{n \le X} f(n) F(n) \right| \ll E_{\mathrm{triv}} + E_{\mathrm{sieve}} + E_{\mathrm{bil}}, \tag{187}$$

where

$$E_{\mathrm{triv}} = Y^{-1/2} \|F\|_\infty, \qquad E_{\mathrm{sieve}} = \frac{1}{X} \sum_{\substack{n \le X \\ (n, \prod_{Y \le \ell < Z} \ell) = 1}} |F(n)|.$$

The product is over primes. For the bilinear term, write $L = \max I$ for the right endpoint of the interval I. Green's corrected proposition normalizes the correlation by $1/\max I$. Thus E_{bil}^2 is the supremum of

$$\frac{1}{L} \left| \sum_{\substack{m \in I \\ (m, uv) = 1}} F(um) \overline{F(vm)} \right|$$

over distinct primes $Y \le u < v < Z$ and intervals $I \subset [0, X]$ with $X/(10YZ) < L = \max I \le X$. In particular $I \subset [0, L]$, even when its length is much smaller than L. All the following smoothing and Poisson estimates retain the normalization $1/L$. We have $E_{\mathrm{triv}} \ll X^{-\varepsilon\zeta/2}$.

Green's sieve estimate (3.2) is uniform in the residue class. Its separate conditions $Z < X^{1/10}$ and $q < X^{3/4}$ follow from $\zeta < 1/100$ and $q < 2Q \le 2X^{\vartheta} < X^{3/4}$ for sufficiently large X. The triangle inequality gives

$$E_{\mathrm{sieve}} \ll \frac{\log Y}{\log Z} \sum_{q \in \mathcal{Q}(Q)} \frac{|\xi_q|}{q} \ll \frac{\varepsilon}{Q} \sum_{q \in \mathcal{Q}(Q)} |\xi_q|. \tag{188}$$

The implied constant is independent of ε, d, ξ. It remains to obtain a power saving in E_{bil} for each fixed ε.

Smoothing and Poisson summation. By (186), removing $(m, uv) = 1$ costs $O(L/Y + 1)$. Replace $\mathbf{1}_I(m)$ by $W(m/L)$, where W is the convolution of $\mathbf{1}_{I/L}$ with a positive smooth kernel of width X^{-s}. The cost is $O(LX^{-s} + 1)$. We may arrange $0 \le W \le 1$, support in $[-1, 2]$, and

$$|\widehat{W}(t)| \ll_A |t|^{-A} X^{s(A-1)} \qquad (A \ge 2).$$

These estimates are uniform in I, u, v.

Expand $F(um)\overline{F(vm)}$. The terms with one progression and one mean contribute an error $O(Q)$: bounded variation gives an $O(1)$ discrepancy between a progression sum of $W(m/L)$ and its $1/q$ share of the unrestricted sum, and $\sum_q 1/q \ll 1$. The diagonal $q = q'$ costs $O(L/Q + 1)$. Indeed, $dum \equiv c \pmod q$ and $dvm \equiv c \pmod q$ cannot both hold, since $q > v - u$ and $(duv, q) = 1$.

For $q \ne q'$, let $R(q, q')$ mod qq' be determined by

$$duR(q, q') \equiv c \pmod q, \qquad dvR(q, q') \equiv c \pmod{q'}.$$

Poisson summation reduces the normalized off-diagonal term to

$$\sum_{h \ne 0} \sum_{\substack{q, q' \in \mathcal{Q}(Q) \\ q \ne q'}} \frac{\xi_q \overline{\xi_{q'}}}{qq'} \widehat{W}\left(\frac{Lh}{qq'}\right) e_{qq'}(hR(q, q')). \tag{189}$$

Replacing the unrestricted sum of $W(m/L)$ by its integral in the mean term costs an arbitrarily large power saving, by the same Fourier decay. Truncate at $0 < |h| \le H$, where

$$H = \left\lceil (Q^2/L) X^{2s} \right\rceil.$$

Since $Q^2/L \ge 1$, the tail is

$$\ll_A (Q^2/L)^A X^{s(A-1)} H^{1-A} \ll (Q^2/L) X^{-s(A-1)}.$$

Also $Q^2/L \ll X^{2\vartheta - 1 + 2\zeta}$. Choose A, depending only on the exponent margin, so that this tail is $O(X^{-s})$.

For each remaining h, use

$$\widehat{W}(Lh/(qq')) = q \int_{\mathbb{R}} W(qt) e(-Lht/q')\, dt.$$

The integral has support of length $O(1/Q)$, and the remaining $1/q'$ is absorbed into a bounded coefficient with a factor Q^{-1}. Thus the inner sum in (189) is at most CQ^{-2} times the supremum of

$$|S_h(\alpha, \beta)|, \qquad S_h(\alpha, \beta) = \sum_{\substack{q, q' \in \mathcal{Q}(Q) \\ q \ne q'}} \alpha(q)\beta(q') e_{qq'}(hR(q, q')), \qquad |\alpha|, |\beta| \le 1. \tag{190}$$

This separation incurs no derivative loss.

Rational reciprocity. The four primes u, v, q, q' are distinct and coprime to d. The Chinese remainder theorem and reciprocity give

$$e_{qq'}(hR(q, q')) = e_{vq}\left(ch(v - u)\overline{duq'}\right) e_d\left(-ch\,\overline{vqq'}\right) e\left(\frac{ch}{dvqq'}\right). \tag{191}$$

For completeness, start from

$$e_{qq'}(hR) = e_q(ch\,\overline{duq'}) e_{q'}(ch\,\overline{dvq}).$$

The first factor is $e_{vq}(chv\,\overline{duq'})$; the second is $e_{duq'}(chu\,\overline{vq})e_d(-ch\,\overline{vqq'})$. Apply $\overline{A}/B+\overline{B}/A\equiv 1/(AB) \pmod 1$ with $A=vq$, $B=duq'$, obtaining (191).

Expand the extra factor on $(\mathbb{Z}/d\mathbb{Z})^\times$:

$$e_d(-ch\,\overline{v}\,\overline{t})=\sum_{\chi \bmod d} b_\chi \chi(t).$$

Parseval gives $\sum_\chi |b_\chi|^2=1$, hence

$$\sum_\chi |b_\chi| \le \sqrt{\varphi(d)} \le \sqrt{d}. \tag{192}$$

The factor modulo d is consequently absorbed into $\alpha(q)\chi(q)$ and $\beta(q')\chi(q')$, at total cost at most $\sqrt{d}$. This step requires no equidistribution of primes modulo d. The last factor in (191) differs from one by $O(|h|/(dvQ^2))$, so its contribution after normalization by Q^2 is $O(|h|/Q^2)$.

The Kloosterman-fraction estimate. For each character, put $m=duq'$, $n=vq$, and $b=ch(v-u)\neq 0$. The remaining form is

$$\sum_{\substack{m\asymp duQ,\ n\asymp vQ\\(m,n)=1}} A_m B_n e_n(b\overline{m}).$$

Its coefficients are supported only where $m/(du)$ and n/v are primes in $[Q,2Q)$. In particular, $\|A\|_2\|B\|_2\ll Q$. Theorem 1 of Bettin–Chandee, with its parameters $A=1$ and $\vartheta=b$, bounds this form by

$$\ll_{\epsilon_{\mathrm{BC}}} Q\left(1+\frac{|b|}{duvQ^2}\right)^{1/2}\left[(duvQ^2)^{7/20+\epsilon_{\mathrm{BC}}}((du+v)Q)^{1/4}+(duvQ^2)^{3/8+\epsilon_{\mathrm{BC}}}((du+v)Q)^{1/8}\right].$$

The differing dyadic interval conventions change only constants. Since $|b|\le |h|Z\ll Q^2X^{2s}Z/L$, the factor $|b|/(duvQ^2)$ is bounded throughout the truncated range. Using (192), $u,v\le Z$, and then dividing by Q^2, gives

$$Q^{-2}|S_h(\alpha,\beta)| \ll_{\epsilon_{\mathrm{BC}}} (dZ^2Q^2)^{\epsilon_{\mathrm{BC}}}\left[d^{11/10}Z^{19/20}Q^{-1/20}+dZ^{7/8}Q^{-1/8}\right]+\frac{|h|}{Q^2}. \tag{193}$$

The second term in brackets is no greater than the first. Thus the right side is

$$\ll X^{3\epsilon_{\mathrm{BC}}}d^{11/10}ZQ^{-1/20}+|h|/Q^2.$$

Keeping the actual sparse coefficient norms is essential for this estimate; their ambient interval lengths would produce an unnecessary extra loss.

Closing the exponent inequalities. Sum over $0<|h|\le H$. Since $L\gg X/Z^2$, the first term is

$$\ll Q^{39/20}d^{11/10}Z^3X^{-1+2s+3\epsilon_{\mathrm{BC}}}\ll X^{-\Delta+3\zeta+2s+3\epsilon_{\mathrm{BC}}}\ll X^{-\Delta/2}.$$

The fixed factor $a_0^{11/10}$ is absorbed into the constant. The reciprocity error is

$$\ll H^2/Q^2\ll X^{2\vartheta-2+4\zeta+4s},$$

also a power saving. The terms Q/L, $1/Q$, $1/L$, the Poisson tail, and the smoothing error are power savings as well. Including the coprimality cost Y^{-1} gives

$$E_{\mathrm{bil}}\ll_\varepsilon X^{-\nu_\varepsilon}$$

for some $\nu_\varepsilon > 0$. The exponent parameters were fixed independently of ε. Combining (187) and (188), we obtain

$$\left|\sum_{n\le X} f(n)F(n)\right| \le C\varepsilon\,\frac{X}{Q}\sum_{q\in\mathcal{Q}(Q)}|\xi_q| + O_\varepsilon(X^{1-\nu'_\varepsilon}), \tag{194}$$

where C is independent of ε, d, f, ξ. Choose ξ_q to be the dual phases of the discrepancies in (185). Since $\#\mathcal{Q}(Q) \ll Q/\log Q$ and $\log Q \asymp \log X$, the normalized limiting upper bound is at most $C'\varepsilon$, uniformly in the stated ranges. Let $\varepsilon \downarrow 0$ to prove the proposition. □

The rational-class adaptation has two explicit denominator costs: character expansion contributes $d^{1/2}$, and the substitution $m = duq'$ contributes $d^{3/5}$ in the first Bettin–Chandee term. Together they give $d^{11/10}$. After Poisson summation, the needed power margin is $Q^{39/20}d^{11/10} < X$, which is exactly $39\vartheta + 22\kappa < 20$.

Wright's refinement [21, Theorem 2.1] treats Kloosterman fractions with a prescribed factor in the denominator. The structured variables $m = duq'$ and $n = vq$ in the present argument make this a relevant comparison. Applying that estimate here would require matching the orientation of the reciprocal phase, the coprimality restrictions and the sparse coefficient norms, and then summing over the truncated frequencies with their full parameter dependence. The bound just proved uses Bettin–Chandee's stated estimate; no improvement from Wright's theorem is incorporated in its exponent region.

18.2 Applying the estimate to the Liouville sign

For $f = \lambda$, the centre in (185) has total norm $o(X/\log X)$ by the prime number theorem. Thus

$$\sum_{q\in\mathcal{Q}(Q)}\left|\sum_{\substack{k\le X\\ apk\equiv -1\ (q)}}\lambda(k)\right| = o(X/\log X), \tag{195}$$

uniformly on every compact subset of the admissible exponent domain, for $a = 1, 3$ and prime p.

A maximal version on comparable lengths is available in the strict interior. Fix $0 < h < 1$ and choose $O(1/h)$ mesh points of spacing at most hX in $[X, 2X]$. At these finitely many points, (195) gives $o_h(X/\log X)$. Between adjacent points, the variation of a sum in a class modulo q is at most $hX/q + 1$. Hence

$$\sum_{q\in\mathcal{Q}(Q)}\max_{X\le t\le 2X}\left|\sum_{\substack{k\le t\\ apk\equiv -1\ (q)}}\lambda(k)\right| \le o_h(X/\log X) + O(hX/\log Q + Q/\log Q).$$

Since $Q = o(X)$, taking $X \to \infty$ and then $h \downarrow 0$ proves

$$\sum_{q\in\mathcal{Q}(Q)}\max_{X\le t\le 2X}\left|\sum_{\substack{k\le t\\ apk\equiv -1\ (q)}}\lambda(k)\right| = o(X/\log X). \tag{196}$$

The same proof applies to the centred discrepancies, whose variation is at most $2hX/q + 1$.

On an origin block $N \le n < 2N$, write $n = pk$. If p has size N^x and q has size N^y, the cofactor scale $X \asymp N/p$ gives

$$\kappa = \frac{x}{1-x}, \qquad \vartheta = \frac{y}{1-x},$$

up to $o(1)$ changes. The exponent condition becomes

$$42x + 39y < 20. \tag{197}$$

It therefore cancels every cofactor in the open triangle

$$\mathcal{G} = \{(x, y) : x > 0,\ y > 1/2,\ 42x + 39y < 20\}.$$

This triangle has $x + y < 1$, so lies in the lower region, and is disjoint from the large-sieve region $y < 1/2$.

To transfer the estimate with all its weights, first fix a finite union of closed rectangles contained in $\mathcal{G}$. In a block $[N, 2N)$, use bounds $N^{x_0} \le p \le N^{x_1}$ and $N^{y_0} \le q \le N^{y_1}$. Then the endpoints for $k = n/p$ are independent of q. A finite covering by exponent neighborhoods supplies uniform power margins for every dyadic block. For $p \asymp P$, $q \asymp Q$, $K = N/p$, partial summation with $w_p(t) = 1/(t(\log(pt))^2)$ and (196) gives

$$\sum_{q \asymp Q} \left| \sum_{\substack{K \le k < bK \\ apk \equiv -1 \ (q)}} \lambda(k) w_p(k) \right| = o((\log N)^{-3}), \qquad 1 \le b \le 2.$$

Equivalently, one may keep the sum of absolute discrepancies inside the Abel integral and use (195) directly. After the factor $(\log p \log q)/(pT_M)$, the estimate $\sum_{p \asymp P} (\log p)/p \ll 1$ gives $o(1/(T_M(\log N)^2))$ for each size block. There are $O((\log N)^2)$ size blocks and $O(\log M)$ origin blocks, with $T_M \sim \log M$, so the final signed norm is $o(1)$.

The coordinates using $\log N$ and $\log n$ differ by $O(1/\log N)$. The corresponding events therefore differ only near the rectangle boundaries, where the positive joint limit gives mass tending to zero. Finally exhaust $\mathcal{G}$ from within. The precision $o(X/\log X)$, rather than merely $o(X)$, is what permits the weighted summation over size blocks.

Theorem 18.2 (An explicit numerical upper bound)**.** *For $a \in \{1, 3\}$,*

$$\limsup_M \Lambda_M(\lambda) \le \frac{1}{8} - \frac{1}{13104} = \frac{1637}{13104}.$$

Consequently,

$$\limsup_M \mathrm{Succ}^*_M(\lambda) \le \frac{13}{16} - \frac{1}{26208} = \frac{21293}{26208} = 0.812461843711\ldots. \tag{198}$$

Proof. The base of $\mathcal{G}$ has length $1/84$, and its height is $1/78$, so

$$|\mathcal{G}| = \frac{1}{2}\frac{1}{84}\frac{1}{78} = \frac{1}{13104}.$$

This cancelled mass is disjoint from the large-sieve mass $3/8$. Bound the signed norm on the complement of these two regions by its mass, and exhaust $\mathcal{G}$ by compact subsets. This yields the remainder bound. Equation (128) gives the success bound. □

The explicit improvement is approximately 3.8156×10^{-5} in the success probability. This number records the region reached by the preceding rational-class argument. No optimality of the analytic bound, or existence of a limiting Liouville success probability, is asserted. The smooth gain may also be used, provided its overlap with the rational region is removed. The next section carries this out for both anchors.

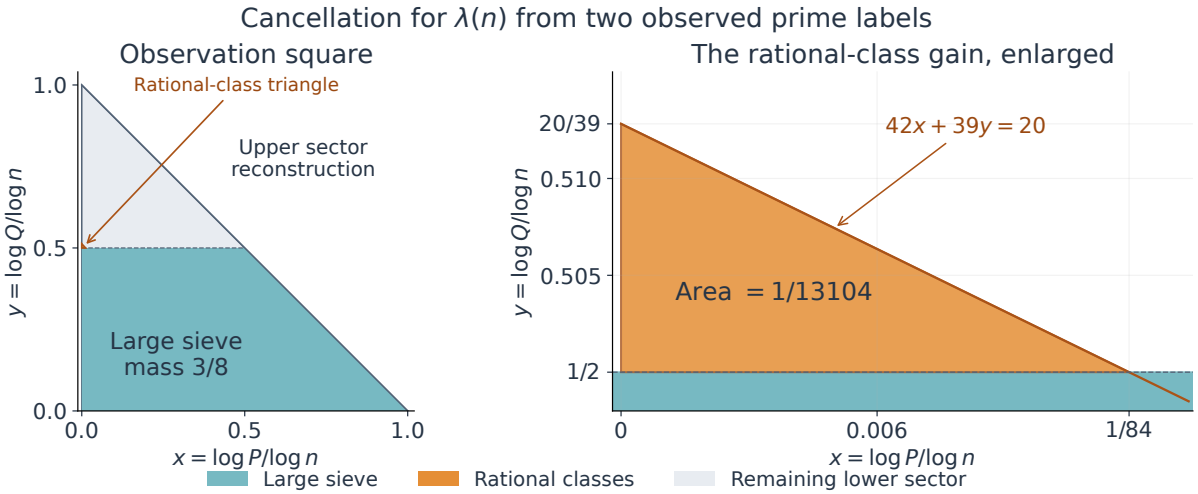

Figure 1: Cancellation regions for predicting $\lambda(n)$. The rational triangle adds mass $1/13104$ to the large-sieve region of mass $3/8$. The smooth event also depends on the cofactor and is not represented by its logarithmic coordinates alone. For $\lambda(an+1)$, the coordinates are interchanged.

19 Combining the rational and smooth regions

The rational estimate transfers to the second anchor through characters of a fixed modulus. A restricted smooth event then provides a disjoint additional gain. Write $W_0(n) = \lambda(n)$, $W_1(n) = \lambda(an+1)$, and

$$x_n = \frac{\log P}{\log n}, \qquad y_n = \frac{\log Q}{\log n}.$$

Define

$$\mathcal{G}_M^{(0)} = \{y_n > 1/2,\ 42x_n + 39y_n < 20\},$$
$$\mathcal{G}_M^{(1)} = \{x_n > 1/2,\ 39x_n + 42y_n < 20\}.$$

Both events lie in B_M and have limiting mass $1/13104$.

19.1 The rational estimate at the second anchor

On $N \le n < bN$, writing $an+1 = qk$ gives

$$A_q = \frac{aN+1}{q} \le k < B_q = \frac{abN+1}{q}, \qquad qk \equiv 1 \pmod p, \qquad k \equiv q^{-1} \pmod a.$$

Work first with finitely many closed rectangles inside $\{(x,y) : y > 0,\ x > 1/2,\ 39x + 42y < 20\}$, using powers of N as prime bounds. The endpoints A_q, B_q are independent of the modulus p. For q of size N^y, p of size N^x, and $t \asymp_a N/q$, the exponents in Proposition 18.1 are

$$\kappa = \frac{y}{1-y}, \qquad \vartheta = \frac{x}{1-y},$$

up to $o(1)$. Thus its condition is precisely $39x + 42y < 20$. The compact exponent margins give uniformity for the relevant prefixes and dyadic p-blocks.

For each $\chi \bmod a$, apply the proposition with $a_0 = 1$, $r = q$, $c = 1$, and $f = \lambda\chi$. The centre $p^{-1}\sum_{k\le t}\lambda(k)\chi(k)$ has norm $o_a(t/\log t)$ summed over $p \asymp P_0$, by fixed-modulus cancellation.

Projecting onto $k \equiv q^{-1} \pmod a$ gives

$$\sum_{p \asymp P_0} \left| \sum_{\substack{k \le t \\ qk \equiv 1 \ (p) \\ k \equiv q^{-1} \ (a)}} \lambda(k) \right| = o_a(t/\log t). \tag{199}$$

The number of characters is fixed; no growing-modulus uniformity is used in this projection. With

$$v_q(t) = \frac{1}{((qt-1)/a)\{\log((qt-1)/a)\}^2}, \qquad \sup v_q + \mathrm{Var}(v_q) \ll_a \frac{1}{N(\log N)^2},$$

Abel summation, keeping the sum of absolute values inside the integral, yields

$$\sum_{p \asymp P_0} \left| \sum_{\substack{A_q \le k < B_q \\ qk \equiv 1 \ (p) \\ k \equiv q^{-1} \ (a)}} \lambda(k) v_q(k) \right| = o_a\left(\frac{1}{q(\log N)^3} \right).$$

Multiplication by $\log p \log q / T_M$ and summation over $q \asymp Q_0$ cost $o_a(1/(T_M(\log N)^2))$ per size block. Summing the $O((\log N)^2)$ size blocks and $O(\log M)$ origin blocks gives $o(1)$. As in the first-anchor transfer, replace $\log N$ by $\log n$ using the null boundary of each rectangle, then exhaust the triangle. We obtain, for $j = 0, 1$,

$$\sum_{p,q} \left| \mathbb{E}[W_j(n) \mathbf{1}_{\{P=p, Q=q\}} \mathbf{1}_{\mathcal{G}_M^{(j)}}] \right| = o(1). \tag{200}$$

The order of choices is: a compact exponent set and positive margin; bilinear parameters fixed independently of the tolerance; $M \to \infty$; tolerance tending to zero; finally exhaustion of the triangle.

19.2 A disjoint portion of the smooth event

In (159) and (176), replace the interval for the large prime by

$$(N^{13/25}, N^{11/20}] = (N^{0.52}, N^{0.55}].$$

Call the resulting events $\mathcal{F}_M^{(0,+)}$, $\mathcal{F}_M^{(1,+)}$, and put

$$b_E^{(j)}(p,q) = \mathbb{E}[W_j(n) \mathbf{1}_{\{P=p, Q=q\}} \mathbf{1}_E].$$

The DGS estimates remain valid when the prime-modulus sums are restricted to these intervals. Applying the smooth proofs with this restriction gives

$$\| b^{(j)}_{\mathcal{F}_M^{(j,+)}} \|_1 = o(1), \tag{201}$$

$$\mathbb{P}(\mathcal{F}_M^{(j,+)}) \longrightarrow \frac{3}{4}\gamma_\eta, \qquad j = 0, 1. \tag{202}$$

The factor $3/4$ is the ratio of the new logarithmic width $3/100$ to the old width $1/25$. This restricted cancellation follows from the restricted distribution estimate, not from cancellation on the larger event alone.

The large prime has logarithmic coordinate $> 13/25 + o(1)$ on the restricted smooth event, whereas that coordinate is $< 20/39$ in the corresponding rational triangle. Since $13/25 > 20/39$, the two events are disjoint for large M, and both are disjoint from the relevant large-sieve region.

Corollary 19.1 (Combined gains for both anchors)**.** *For $a = 1, 3$, $j = 0, 1$, with or without termination bits,*

$$\limsup_M \Lambda_M(W_j) \le \frac{1}{8} - \frac{1}{13104} - \frac{3}{4}\gamma_\eta, \tag{203}$$

$$\limsup_M \mathrm{Succ}^*_M(W_j) \le \frac{13}{16} - \frac{1}{26208} - \frac{3}{8}\gamma_\eta < \frac{21293}{26208}. \tag{204}$$

Proof. Within the lower region of mass $1/2$, the three disjoint cancelled parts have limiting masses $3/8$, $1/13104$, and $3\gamma_\eta/4$. The remaining signed norm is at most its remaining mass. The reduction (128) gives the success bound. □

An exact localization remains available. Define

$$\mathcal{D}^{(0)}_M = \{PQ \le n,\ Q > \sqrt{n}\}, \qquad \mathcal{D}^{(1)}_M = \{PQ \le n,\ P > \sqrt{n}\},$$

and

$$\mathfrak{S}^{(j)}_M = \mathcal{D}^{(j)}_M \setminus (\mathcal{G}^{(j)}_M \cup \mathcal{F}^{(j,+)}_M).$$

The cancellations and reverse triangle inequality give

$$\begin{aligned} \Lambda_M(W_j) &= \|b^{(j)}_{\mathfrak{S}^{(j)}_M}\|_1 + o(1), \\ \mathbb{P}(\mathfrak{S}^{(j)}_M) &\longrightarrow \frac{1}{8} - \frac{1}{13104} - \frac{3}{4}\gamma_\eta. \end{aligned} \tag{205}$$

The explicit success gain $1/26208$ and the positive smooth gain $3\gamma_\eta/8$ are separate quantities. No numerical value for the latter is claimed. Each remainder belongs to one sign; these estimates do not control the product of the two signs.

19.3 Information bounds for each Liouville sign

Let T_M consist of the labels, optionally with their ordinary termination bits.

Corollary 19.2 (Combined binary information bounds)**.** *For $a = 1, 3$ and $j = 0, 1$,*

$$\frac{3}{8} + \frac{1}{13104} + \frac{3}{4}\gamma_\eta \le \liminf_M \frac{H(W_j \mid \mathsf{T}_M)}{\log 2} \le \limsup_M \frac{H(W_j \mid \mathsf{T}_M)}{\log 2} \le \frac{1}{2}, \tag{206}$$

$$\frac{1}{2} \le \liminf_M \frac{I(W_j; \mathsf{T}_M)}{\log 2} \le \limsup_M \frac{I(W_j; \mathsf{T}_M)}{\log 2} \le \frac{5}{8} - \frac{1}{13104} - \frac{3}{4}\gamma_\eta. \tag{207}$$

No limiting entropy or information is asserted to exist.

Proof. Fixed-class cancellation and partial summation give $\mathbb{E}W_j = o(1)$. The confidence selector recovers the origin on an observable event of probability $1/2 + o(1)$ with error $o(1)$, and hence recovers either sign there. Finally the signed reduction and (203) give

$$\|\mathbb{E}[W_j \mid \mathsf{T}_M]\|_1 \le \frac{1}{2} + \beta + o(1), \qquad \beta = \frac{1}{8} - \frac{1}{13104} - \frac{3}{4}\gamma_\eta.$$

Apply Proposition 16.5 and divide by $\log 2$. □

20 A fixed sign attaining the universal guarantee

The minimax argument allowed the sign table to change with the window. A direct calculation on the lower fibres gives examples defined once on all integers.

Proposition 20.1 (Periodic equidistribution in the lower fibres)**.** *Let $g : \mathbb{N} \to \mathbb{C}$ be bounded with a fixed period $m \geq 1$, and let $\overline{g} = m^{-1} \sum_{r=1}^{m} g(r)$. Then*

$$\sum_{p,q} \left| \mathbb{E}[(g(n) - \overline{g}) \mathbf{1}_{\{P=p, Q=q, PQ \leq n\}}] \right| \longrightarrow 0. \tag{208}$$

Both g and its period are fixed before $M \to \infty$.

Proof. By scaling assume $\|g\|_\infty \leq 1$. The probability that a sampled prime divides m is $O_{a,m}(1/L)$, since the expected valuation of each fixed prime is bounded and the sampling denominator is at least $L + O_a(1)$. Remove these labels, then apply the positive comparison (64). This reduces the weights to

$$\frac{\log p \log q}{T_M n (\log n)^2} \mathbf{1}_{\{p|n,\ q|an+1\}}.$$

Fix $0 < \varepsilon < 1$, and first impose $pq \leq n^{1-\varepsilon}$. In an origin block $N \leq n < 2N$, intersected with I_M, the admissible origins form an interval in a single progression of step pq. When $(pq, m) = 1$, the sequence $g(z + jpq) - \overline{g}$ is periodic of period m and has zero sum over a period. Every interval sum has absolute value at most $2m$. Abel summation with $v(t) = 1/(t(\log t)^2)$ gives

$$\left| \sum_{\substack{n \text{ in the block} \\ p|n,\ q|an+1 \\ pq \leq n^{1-\varepsilon}}} \frac{g(n) - \overline{g}}{n(\log n)^2} \right| \ll \frac{m}{N(\log N)^2}.$$

The exterior logarithms cancel $(\log N)^2$ up to a constant. The number of admissible prime pairs is bounded by the number of positive integer pairs of product at most $(2N)^{1-\varepsilon}$, namely $O_\varepsilon(N^{1-\varepsilon} \log N)$. Thus the block costs

$$\ll_\varepsilon \frac{m N^{-\varepsilon} \log N}{T_M}.$$

Its sum over $O(L)$ blocks tends to zero. The discarded strip $n^{1-\varepsilon} < PQ \leq n$ has limiting mass $\varepsilon - \varepsilon^2/2$, by the lower-triangle law. Its norm costs at most twice this mass. First let $M \to \infty$, then $\varepsilon \downarrow 0$. □

Corollary 20.2 (Exact profile of a periodic sign)**.** *For a fixed periodic sign $W : \mathbb{N} \to \{-1, 1\}$ of mean $\overline{W}$,*

$$\Lambda_M(W) \longrightarrow \frac{|\overline{W}|}{2}, \qquad \mathrm{Succ}^*_M(W) \longrightarrow \frac{3}{4} + \frac{|\overline{W}|}{4}. \tag{209}$$

The limits also hold with termination bits. In particular, $W(n) = (-1)^n$ has optimal limiting success probability $3/4$.

Proof. The proposition gives $\|b_{B,W} - \overline{W}\ell\|_1 \to 0$. Since $\sum \ell = \mathbb{P}(B_M) \to 1/2$, the reverse triangle inequality gives the first limit. The signed reduction gives the second. Both bits are zero on B_M, so adding them changes neither calculation. □

There is also a completely multiplicative example showing why fixed-class mean-zero hypotheses are needed at the second anchor. Define

$$f(n) = \chi_3(n/3^{v_3(n)}),$$

where χ_3 is the nonprincipal real character modulo three. Then f is completely multiplicative, takes values ± 1, and has $f(3) = 1$. It is the modified character λ_3 of Borwein, Choi and Coons [34]. The same function, denoted by ψ_+ in [35, Appendix E], enters a mixture with the constant function attaining the dyadic cocycle bound 2/3. For fixed $J \geq 0$,

$$f_J(n) = f(n)\mathbf{1}_{3^{J+1}\nmid n}$$

is periodic of period 3^{J+1} and mean zero: each valuation stratum $0 \leq j \leq J$ splits equally between the two reduced classes after division by 3^j. Moreover,

$$|f - f_J| = \mathbf{1}_{3^{J+1}|n}, \qquad \lim_M \mathbb{E}|f - f_J| = 3^{-J-1}.$$

Proposition 20.1 and $\|b_{B,f} - b_{B,f_J}\|_1 \leq \mathbb{E}|f - f_J|$, followed by $J \to \infty$, give $\Lambda_M(f) \to 0$. Thus $f(n)$ also has optimal success tending to 3/4. The same truncation proves $X^{-1}\sum_{n\leq X} f(n) \to 0$. Yet $f(3n+1) = 1$ for every n, so that sign is predictable with probability one. A global mean-zero condition cannot replace the fixed-class conditions used for the second anchor.

21 Zero-error certificates and dependence of the labels

Keep one logarithmically weighted prime from each anchor, with the draws independent conditional on n. The joint logarithmic law proved above is

$$(S_M, X_M, Y_M) \Longrightarrow (S, U, V), \qquad (S, U, V) \sim \mathrm{Unif}[1, 2] \otimes \mathrm{Unif}[0, 1]^{\otimes 2}. \tag{210}$$

It determines a zero-error certificate probability and also reveals a large total-variation dependence between the exact labels.

21.1 The optimal zero-error reconstruction probability

A decoder with abstention is zero-error if $\mathbb{P}(\widehat{n} \notin \{n, \perp\}) = 0$ for each M. Its success is $\mathbb{P}(\widehat{n} = n)$, so abstention does not count as success. Let R_M^{cert} be the supremum of these success probabilities for observation (P, Q).

Theorem 21.1 (The optimal certificate profile)**.** *For $a = 1, 3$,*

$$R_M^{\mathrm{cert}} \longrightarrow 3 - 4\log 2 = 0.2274112777\ldots. \tag{211}$$

The observable certificate

$$\mathcal{C}_M^{\times} = \{(p, q) : pq > M^2,\ \exists n \in I_M : p \mid n,\ q \mid an + 1\} \tag{212}$$

attains this limit by returning the unique compatible origin and abstaining elsewhere.

Proof. Under the true law a compatible origin exists. The Chinese remainder theorem and $pq > M^2$ guarantee uniqueness in I_M. Since $\log(PQ)/L = S_M(X_M + Y_M)$, the joint law gives

$$\begin{aligned}\mathbb{P}(\mathcal{C}_M^{\times}) &\longrightarrow \mathbb{P}(S(U + V) > 2) \\ &= \int_1^2 \frac{(2 - 2/s)^2}{2}\, ds = 3 - 4\log 2.\end{aligned}$$

The boundary $s(u+v)=2$ has limiting measure zero.

Conversely, every compatible origin for an admissible pair (p,q) has strictly positive likelihood. A zero-error decoder can therefore answer only if the origin is unique in I_M. The CRT class is a progression of step $d=pq$. A half-open interval of length $W_M=M^2-M$ contains at least $\lfloor W_M/d\rfloor$ points of any such progression. Uniqueness forces $d>W_M/2$. Hence

$$R_M^{\mathrm{cert}} \le \mathbb{P}(PQ > (M^2-M)/2).$$

The logarithmic threshold again tends to two. The absence of mass on $s(u+v)=2$ gives the matching upper limit. □

The simpler sufficient condition $P>M$, $Q>M$ has exact limiting mass

$$\mathbb{P}(P>M, Q>M) \longrightarrow \int_1^2 (1-1/s)^2\,ds = \frac{3}{2} - 2\log 2 = 0.1137056389\ldots. \tag{213}$$

The product certificate therefore has twice this mass. The elementary lower bound $3/2-\sqrt{2}$ for the two-threshold event is consistent with this exact value.

The limit (211) also holds if ordinary termination bits after the first removal are observed at either or both anchors. A bit equal to one forces $n=p$ or $an+1=q$. The probability of either event is $O_a(1/L)$: Chebyshev's estimate and Abel summation give harmonic mass $O_a(1)$ for origins whose relevant anchor is prime, before division by T_M. When the observed bits are zero, they can exclude at most the two origins p and $(q-1)/a$ from the CRT class. If only one candidate remains, the original class contained at most three points, forcing $d>W_M/4$. This event has the same limiting mass $3-4\log 2$. This candidate-count argument also accounts for the information carried by observing a zero bit.

21.2 A total-variation separation greater than one half

Write

$$\mathsf{P}_M = \mathcal{L}(P,Q), \qquad \mathsf{Q}_M = \mathcal{L}(P)\otimes\mathcal{L}(Q).$$

The confidence selector is

$$\mathcal{A}_M = \{(p,q) : z_a(p,q)\in I_M,\ pq > z_a(p,q)L^4\}. \tag{214}$$

We already know $\mathsf{P}_M(\mathcal{A}_M)\to 1/2$.

Corollary 21.2 (Dependence of the exact prime labels)**.** *We have*

$$\mathsf{Q}_M(\mathcal{A}_M) = O_a(1/L), \qquad \liminf_M d_{\mathrm{TV}}(\mathsf{P}_M,\mathsf{Q}_M) \ge \frac{3}{4} - \frac{(\log 2)^2}{2} = 0.5097734930\ldots. \tag{215}$$

The same total-variation lower bound holds when termination bits are added to both label observations.

Proof. The pointwise marginal bounds give, for sufficiently large M, $\mathbb{P}(P=p)\le 42/p$ and $\mathbb{P}(Q=q)\le 42/q$. For each $z\in I_M$, there are at most $\omega(z)\omega(az+1)\ll_a L^2$ compatible prime pairs. On $\mathcal{A}_M$, $1/(pq)<1/(zL^4)$, so

$$\mathsf{Q}_M(\mathcal{A}_M) \le 42^2 \sum_{z\in I_M} \sum_{\substack{p\mid z,\ q\mid az+1\\ pq>zL^4}} \frac{1}{pq} \ll_a \frac{1}{L^2}\sum_{z\in I_M}\frac{1}{z} = O_a(1/L).$$

For a stronger test, let $\mathcal{D}_M=\{(p,q): pq\le M\}$. It is disjoint from $\mathcal{A}_M$ for large M. Under the joint law,

$$\mathsf{P}_M(\mathcal{D}_M) \longrightarrow \mathbb{P}(S(U+V)\le 1) = \int_1^2 \frac{1}{2s^2}\,ds = \frac{1}{4}.$$

Each marginal logarithmic size $\log P/L$, $\log Q/L$ converges to SU. On $0 \le t \le 1$,

$$\mathbb{P}(SU \le t) = \int_1^2 \frac{t}{s}\, ds = t \log 2.$$

Under Q_M, these sizes are independent. Their limit therefore has density $(\log 2)^2$ on the triangle $u, v \ge 0,\ u + v \le 1$, yielding

$$\mathsf{Q}_M(\mathcal{D}_M) \longrightarrow \frac{(\log 2)^2}{2}.$$

The test $\mathcal{A}_M \cup \mathcal{D}_M$ has limiting probability $3/4$ under P_M and $(\log 2)^2/2$ under Q_M. Projection onto the labels shows that adding bits cannot reduce total variation. □

The zero-error certificate recovers a limiting proportion $3 - 4\log 2$. The confidence selector accepts a limiting proportion $1/2$, with conditional error tending uniformly to zero but possibly positive at finite M. These are distinct reconstruction guarantees.

22 Reconstruction, stability, and the quartic mean

Recovering the origin makes every prescribed sign observable, but does not determine the mean of that sign. We record exact covariance identities for the quartic expression that motivated the observation problem, then show that successful reconstruction can make the joint observation maximally unstable under an affine change of origin.

22.1 Covariance after certified decoding

Let $\mathcal{O}$ be the observation of Section 5, with two diverging budgets, and let e_M be the abstention probability of its zero-error decoder. For the anchors $n, 3n+1$, put

$$X = \lambda(3n+1), \qquad Y = \lambda(n)\lambda(n+1)\lambda(3n+2), \qquad C = XY, \qquad c_M = \mathbb{E}_{\mu_M} C.$$

Write $q_X = \mathbb{E}[X \mid \mathcal{O}]$ and $q_Y = \mathbb{E}[Y \mid \mathcal{O}]$. On the observable success event, the origin and both signs are determined, so their conditional covariance is zero. Elsewhere its absolute value is at most one. Therefore

$$\mathbb{E}|\operatorname{Cov}(X, Y \mid \mathcal{O})| \le e_M \longrightarrow 0. \tag{216}$$

The full identity is

$$c_M = \mathbb{E}[q_X q_Y] + \mathbb{E}\operatorname{Cov}(X, Y \mid \mathcal{O}), \qquad |c_M - \mathbb{E}[q_X q_Y]| \le e_M. \tag{217}$$

The unknown quartic mean is transferred into the product of the conditional means. No limit for c_M is assumed, and (216) does not imply its cancellation.

22.2 Covariance after decoding with possible errors

The asymmetric decoder has no observable certificate of success. It still gives the following sharp general covariance bound.

Proposition 22.1 (A covariance bound from reconstruction error)**.** *Let $\widehat{n}$ be a deterministic decoder of an observation $\mathcal{O}$, and let $e = \mathbb{P}(\widehat{n} \ne n)$, counting abstentions as errors. For any deterministic signs $X = X(n)$ and $Y = Y(n)$,*

$$\mathbb{E}|\operatorname{Cov}(X, Y \mid \mathcal{O})| \le \gamma(e) \le 4e, \qquad \gamma(u) = \begin{cases} 4u(1-u), & 0 \le u \le 1/2, \\ 1, & 1/2 \le u \le 1. \end{cases} \tag{218}$$

Among general probability experiments, this is the best bound that depends only on the reconstruction error.

Proof. Replace any abstention by a fixed candidate origin to obtain a decoder $\overline{n}$ with error at most e. Set

$$\epsilon(\mathcal{O}) = \mathbb{P}(n \neq \overline{n} \mid \mathcal{O}), \qquad r_X(\mathcal{O}) = \mathbb{P}(X(n) \neq X(\overline{n}) \mid \mathcal{O}).$$

Then $0 \leq r_X \leq \epsilon$, and $\operatorname{Var}(X \mid \mathcal{O}) = 4r_X(1 - r_X)$. The function $4r(1-r)$ increases on $[0, 1/2]$ and is bounded by one on $[0, 1]$, so

$$\operatorname{Var}(X \mid \mathcal{O}) \leq \gamma(\epsilon), \qquad \operatorname{Var}(Y \mid \mathcal{O}) \leq \gamma(\epsilon).$$

Conditional Cauchy–Schwarz gives $|\operatorname{Cov}(X, Y \mid \mathcal{O})| \leq \gamma(\epsilon)$. The function γ is increasing and concave; Jensen's inequality yields

$$\mathbb{E}\gamma(\epsilon) \leq \gamma(\mathbb{E}\epsilon) \leq \gamma(e).$$

For sharpness when $e \leq 1/2$, take a constant observation, two origins with probabilities $1 - e, e$, a decoder choosing the first, and $X = Y$ with opposite signs on the two origins. Their covariance is $4e(1 - e)$. For $e \geq 1/2$, use origins of masses $1 - e$, $e - 1/2$, and $1/2$, carrying signs $+1, +1, -1$. The decoder chooses the first origin; $X = Y$ is balanced, with covariance one. Omit zero masses at endpoints. At $e = 1$, the decoder may choose a candidate outside the support. These examples attain the bound. □

Applied to the quartic signs and the asymmetric decoder of Section 6, the proposition gives

$$\big|c_M - \mathbb{E}\big[\mathbb{E}(X \mid \mathcal{O})\mathbb{E}(Y \mid \mathcal{O})\big]\big| \leq \gamma(e_M) \leq 4e_M \longrightarrow 0. \tag{219}$$

One diverging budget and one auxiliary prime suffice, under any admissible coupling. The bound for the quartic harmonic mean in [35, Theorem 1.1] also applies to μ_M, since $T_M \sim \log M \to \infty$. Thus in either decoding regime above,

$$\limsup_{M \to \infty} |\mathbb{E}[q_X q_Y]| \leq \frac{196567331}{402653184}.$$

Here the expectation is taken before the absolute value.

22.3 An explicit instability of the joint observation

First use the zero-error decoder from Section 5, with mean abstention probability e_M. Fix a prime $p \equiv 1 \pmod 3$, and define

$$g_p(n) = pn + (p-1)/3, \qquad 3g_p(n) + 1 = p(3n+1).$$

For every origin $k \geq 1$, let $\mathsf{J}_M(k)$ be the joint observation kernel, using the same budgets and the same decoding window I_M. The decoder $\mathcal{D}_M$ takes values in $I_M \cup \{\perp\}$. Put

$$\varepsilon_M(k) = \mathbb{P}_{\mathsf{J}_M(k)}(\mathcal{D}_M \neq k).$$

It equals one outside I_M, and $\mathbb{E}_{\mu_M}\varepsilon_M(n) = e_M$. When $g_p(n) \in I_M$, the transported harmonic mass divided by $\mu_M(g_p(n))$ is $g_p(n)/n \leq 2p$. The outgoing boundary has mass $b_{M,p} = O_p(1/L)$, so

$$\mathbb{E}_{\mu_M}\varepsilon_M(g_p(n)) \leq 2pe_M + b_{M,p}.$$

The test $\{\mathcal{D}_M = n\}$ distinguishes the two origins, giving

$$\mathbb{E}_{\mu_M} d_{\mathrm{TV}}(\mathsf{J}_M(n), \mathsf{J}_M(g_p(n))) \geq 1 - (1+2p)e_M - b_{M,p} \longrightarrow 1. \tag{220}$$

A single global sign table also witnesses the instability. Define $\Psi_M(t) = \lambda(3\mathcal{D}_M(t) + 1)$ when the decoder answers, and $\Psi_M(t) = 0$ on abstention. Let $a_M(k) = \mathbb{E}_{\mathsf{J}_M(k)}\Psi_M$. For $k \in I_M$, zero-error decoding gives

$$a_M(k) = (1 - \varepsilon_M(k))\lambda(3k + 1).$$

Complete multiplicativity and $\lambda(p) = -1$ imply

$$\mathbb{E}_{\mu_M}|a_M(g_p(n)) - a_M(n)| \geq 2 - (1 + 2p)e_M - 2b_{M,p} \longrightarrow 2. \tag{221}$$

Thus common reconstruction yields maximal separation of the observation kernels and maximal instability of a bounded signed test. The prime p is fixed; no uniformity in a moving prime is asserted.

For the asymmetric decoder, let e_M now count both wrong answers and abstentions. Under origin n, output n has probability at least $1 - \varepsilon_M(n)$, whereas under the distinct origin $g_p(n)$ it has probability at most $\varepsilon_M(g_p(n))$. The same transport argument proves (220). For the global sign table, use instead

$$|a_M(k) - \lambda(3k + 1)| \leq 2\varepsilon_M(k).$$

This holds within the window and also trivially outside, where $\varepsilon_M(k) = 1$. Consequently

$$\mathbb{E}_{\mu_M}|a_M(g_p(n)) - a_M(n)| \geq 2 - 2(1 + 2p)e_M - 2b_{M,p} \longrightarrow 2. \tag{222}$$

The extra factor two accounts for possible wrong answers. One diverging budget and one auxiliary prime therefore also suffice for maximal instability.

Part III
Complete small-prime information and communication

23 Reconstruction from a cofactor and three small-prime parts

The preceding experiments reveal a prescribed number of prime labels. We now change the information supplied: one observer knows the entire rough cofactor of the origin, while another knows the complete small-prime parts of three neighbouring forms. The reconstruction threshold is then set by the amount of information hidden in the central small-prime part and by the congruences supplied by the neighbours.

For $w \geq 2$ and $m \geq 1$, write

$$s_w(m) = \prod_{p \leq w} p^{v_p(m)}, \qquad r_w(m) = m/s_w(m).$$

Throughout this part, n has harmonic distribution on $I = (U, V] \cap \mathbb{N}$, where $1 \leq U < V$,

$$H_I = \sum_{n \in I} \frac{1}{n}, \quad \ell = \log(V/U) \longrightarrow \infty, \quad y, z \longrightarrow \infty, \quad \log(2\max(y,z)) = o(\ell). \tag{223}$$

Alice observes $v = r_y(n)$. Bob observes the ordered triple

$$S = (S_1, S_2, S_3), \qquad S_j = s_z(L_j(n)), \qquad (L_1(n), L_2(n), L_3(n)) = (n+1, 3n+1, 3n+2).$$

All valuations, including their multiplicities, are included in these integers. Local computation is unrestricted. Public randomness is independent of n. A budget of b bits means at most 2^b possible messages in a one-message protocol. For interaction, terminal incoming bit strings are prefix-free; the stopping rule conveys no extra uncounted information.

Let D have the Dickman probability distribution, normalised by

$$\mathbb{E}e^{-sD} = \exp\left(\int_0^1 (e^{-su} - 1)\frac{du}{u}\right), \qquad s \geq 0. \tag{224}$$

Thus its density is $e^{-\gamma}\rho(x)$ on $x > 0$, where ρ is the Dickman–de Bruijn function and γ is Euler's constant. In particular ρ itself is not the probability density. For independent copies D_0, D_1, D_2, D_3, put

$$P(t, \tau) = \mathbb{P}\{D_0 < t,\ D_0 < \tau(D_1 + D_2 + D_3)\}, \qquad t, \tau \in [0, \infty]. \tag{225}$$

An inequality against $+\infty$ imposes no restriction; if $t = 0$ or $\tau = 0$, this probability is zero.

Theorem 23.1 (Optimal reconstruction with complete small-prime data)**.** *Suppose that*

$$\frac{\log z}{\log y} \longrightarrow \tau, \qquad \frac{b}{\log_2 y} \longrightarrow t, \qquad t, \tau \in [0, \infty].$$

The supremum of $\mathbb{P}(\widehat{n} = n)$ over protocols in which Bob sends at most b bits to Alice tends to $P(t, \tau)$. This remains true if Alice may send arbitrarily many bits in the opposite direction. One message from Bob to Alice already attains the limit.

The attaining protocol can allow an erasure and satisfy

$$\mathbb{P}(\widehat{n} = n) \longrightarrow P(t, \tau), \qquad \mathbb{P}(\widehat{n} \notin \{n, \perp\}) \longrightarrow 0.$$

When the two inputs are pooled, the optimal success probability tends to $P(\infty, \tau)$, with the same vanishing-error selection property.

The theorem is about identifying the complete integer. Its upper bound uses a counting restriction on possible outputs, and therefore also covers interaction. The corresponding upper bound for predicting only $\lambda(n)$ will require a different argument and will be proved for one message in Section 25.

23.1 A product model for the harmonic origin

The first step removes the moving boundary $av \in I$ without conditioning away the information we are trying to measure. Set

$$m_y = \prod_{p \le y} (1 - p^{-1}), \qquad \pi_y(a) = \frac{m_y}{a} \mathbf{1}_{P^+(a) \le y}.$$

Here $P^+(1) = 1$. Let $\nu_{I,y}$ be harmonic measure on the integers $v \in I$ with no prime factor at most y.

The reciprocal-weight model π_y, its independent geometric valuations, and its marginal Dickman limit are classical; see [8, Proposition 1 and Theorem 1]. For the representation of a harmonic sample on $[1, N]$ by geometric valuations conditioned on their product, see [9, Proposition 4.1 and Section 4.3]. The comparison below concerns the joint smooth part and rough cofactor on a moving harmonic window, without retaining a conditioning on their product.

Lemma 23.2 (Harmonic factorisation with a moving cutoff)**.** *The joint law of $(s_y(n), r_y(n))$ is at total variation distance $O(\log(2y)/\ell)$ from $\pi_y \otimes \nu_{I,y}$. The constant is absolute, and the estimate holds uniformly over all bounded functions of the two entries, including functions of their product.*

Proof. Write $S(J) = \sum_{v \in J,\, P^-(v) > y} 1/v$ for an interval J of logarithmic width w. Under π_y,

$$\mathbb{E} \log a = \sum_{p \le y} \frac{\log p}{p - 1} \ll \log(2y).$$

Consequently $\sum_{a \le y^C,\, P^+(a) \le y} 1/a \ge (2m_y)^{-1}$ for a sufficiently large absolute C. Products of these a with rough $v \in J$ are all distinct, by unique factorisation, and lie in an interval of logarithmic width at most $w + C \log y$. Comparing harmonic masses gives

$$S(J) \ll m_y(1 + w + \log(2y)). \tag{226}$$

Unique factorisation also gives $m_y H_I = \mathbb{E}_{\pi_y} S(I/a)$. The symmetric difference of I and I/a consists of at most two intervals of logarithmic width at most $\log a$. Hence

$$\mathbb{E}_{\pi_y} S(I \triangle I/a) \ll m_y \log(2y), \qquad S(I) = m_y H_I + O(m_y \log(2y)).$$

The exact joint weights are $(avH_I)^{-1} \mathbf{1}_{av \in I}$ on smooth a and rough v. The product model replaces this support by $v \in I$ and its normalisation by $S(I)/m_y$. The preceding symmetric-difference bound and normalisation estimate prove the result, since $H_I = \ell + O(1)$. □

23.2 The four logarithmic sizes

Lemma 23.3 (Joint Dickman limit)**.** *In the physical harmonic experiment,*

$$\left(\frac{\log s_y(n)}{\log y}, \frac{\log S_1}{\log z}, \frac{\log S_2}{\log z}, \frac{\log S_3}{\log z} \right) \implies (D_0, D_1, D_2, D_3). \tag{227}$$

The four limiting variables are independent, continuous, and positive almost surely.

Proof. On a dyadic interval $[X, 2X]$, first keep only primes in $(w^\delta, w]$ at a site with threshold w, with $\delta > 0$ fixed. Partition $[\delta, 1]$ into intervals and count primes in the corresponding ranges of $\log p/\log w$. Joint factorial moments of every fixed order k converge to those of independent Poisson counts with means $\int du/u$. Indeed, distinct assigned primes impose one residue class by the Chinese remainder theorem, with density $1/\prod p$ and counting error $O(1/X)$. The sum of these errors is $O_k(\max(y,z)^k/X)$. A prime greater than 3 cannot divide two of the four forms $n, n+1, 3n+1, 3n+2$; the unwanted independent-model collisions have total weight $O_k(\sum_{p>\min(y,z)^\delta} p^{-2}) = o(1)$. Mertens' theorem gives the claimed means. Poisson laws are determined by their moments.

Discard dyadic intervals with $\log X \le \sqrt{\ell \log(2\max(y,z))}$; their harmonic mass is $o(1)$. On the remaining intervals the errors above are uniform. Partial summation supplies the harmonic weight; boundary intervals cost $O(1/\ell)$. Repeated prime powers contribute $O(1/\log w)$ to each normalised logarithm, and primes at most w^δ contribute $O(\delta) + o(1)$, using $\sum_{p\le u} \log p/p \ll \log(2u)$. The primes $2, 3$ can be treated directly from their valuations. Refining the partition and then letting $\delta \downarrow 0$ gives four independent integrals of u against Poisson processes of intensity du/u on $(0, 1]$. Their Laplace transform is (224).

The largest point of one such process is uniform on $(0, 1)$. Conditional on it, the smaller points are an independent rescaled copy. Thus $D \overset{d}{=} W(1 + D')$, where W is uniform and independent of $D' \overset{d}{=} D$. This also proves continuity and positivity. □

23.3 Two necessary conditions for reconstruction

Set $a = s_y(n)$ and $P = S_1S_2S_3$. Fix v and the public randomness. At most 2^b possible incoming terminal strings give at most 2^b possible outputs. This remains true with interaction, since Alice can reconstruct her own outgoing messages from v and the received history. Under the product model of Lemma 23.2, each $a > y^{t+\epsilon}$ has mass at most $m_y/y^{t+\epsilon}$. For $t < \infty$, the total success mass on that set is therefore $o(1)$, uniformly in v.

Pooling both inputs leaves a second obstruction. For fixed v and a fixed triple (s_1, s_2, s_3) a decoder can identify at most one a. Furthermore

$$\#\{(s_1, s_2, s_3) \in \mathbb{N}^3 : s_1s_2s_3 \le H\} \le H(1 + \log H)^2.$$

In the slice $y^{j\eta} \le P < y^{(j+1)\eta}$, successful origins with $a > Py^{2\eta}$ therefore have total mass

$$O\left(m_y(1 + (j+1)\eta \log y)^2 y^{-\eta}\right).$$

First restrict to $a \le y^K$, with fixed K, and use $O(K/\eta+1)$ slices; then let $K \to \infty$ by tightness in (227). Up to $o(1)$ mass, success requires both $\log_y a \le t+\epsilon$ and $\log_y a \le \log_y P+2\eta$. Lemma 23.3, followed by $\epsilon, \eta \downarrow 0$, proves the upper bound of Theorem 23.1, including $t = 0$ or $\tau = 0$.

23.4 Congruences, a hash, and a reliable answer

Remove the factors $2, 3$ from S_j and denote the results by S_j°. These are pairwise coprime and each is coprime to v and to the slope of its form. This follows from the determinants of the four affine forms. Put $P^\circ = \prod_j S_j^\circ$. If $L_j(n) = h_jn + k_j$, the congruences

$$h_jva + k_j \equiv 0 \pmod{S_j^\circ}, \qquad 1 \le j \le 3,$$

specify one residue $a_*(v, S) \in \{1, \dots, P^\circ\}$. When $a \le P^\circ$, it equals a. The removed factors have logarithm $O_{\mathbb{P}}(1)$, so they do not change the normalised limit. For $\tau = \infty$, $\log P^\circ/\log y \to \infty$ in probability.

Fix $0 < t' < t$. Alice enumerates integers $a' \le y^{t'}$, computes $S(va')$, rejects $va' \notin I$, and retains only $a' = a_*(v, S(va'))$. The triples in her resulting list are distinct. Bob sends a public pairwise-independent uniform b-bit hash of his triple, using the universal-hashing principle of [26]. On $a \le y^{t'}$ and $a \le P^\circ$, the true candidate is in the list, and a false hash collision has probability at most $y^{t'}2^{-b} = o(1)$. The joint limit and $t' \uparrow t$ give the lower bound. At $t = \infty$ use any finite t' and then let it increase.

For the vanishing-error assertion, also require

$$a' \le y^K, \qquad P^\circ(S(va')) \ge a'y^\eta$$

with fixed $K < \infty$ and $\eta > 0$, and output an erasure unless the hashed list has one member. Suppose the true $a \le y^K$ and a false $a' \ne a$ has exactly the same triple. Then

$$d = P^\circ(S(va')) \mid |a - a'|, \qquad d \ge a'y^\eta.$$

In the independent product model, rough $v < y^{K+1}$ have mass $O_K(\log(2y)/\ell)$. For the remaining v, any residue class modulo $d \le y^K$ has probability $O(\log(2y)/d)$: on a dyadic interval $[X, 2X]$ with $X \ge y^{K+1}/2$, omit roughness for the numerator and divide by the harmonic rough mass $m_y(\ell + O(1 + \log y))$. For fixed a, a', distributing the prime powers of d among three sites gives at most $3^{\omega(d)}$ classes for v. There are no classes if a prime of d divides a. The probability of any false candidate with the same triple is consequently

$$\ll \log(2y) \sum_{\substack{1 \le a' \le y^K \\ a' \ne a}} \frac{1}{a'y^\eta} \sum_{d \mid |a-a'|} 3^{\omega(d)} \le y^{-\eta + o_K(1)}.$$

Here $\sum_{d|m} 3^{\omega(d)} = \prod_{p^e \| m}(1 + 3e) \le d_2(m)^2 = m^{o(1)}$, uniformly for $m \le y^K$. Distinct-triple hash collisions cost the previous $y^{t'}2^{-b}$. Take the limit in the window first, then $K \uparrow \infty$, $t' \uparrow t$, and $\eta \downarrow 0$, and use a slow diagonal choice. All limiting boundaries have probability zero. This proves the theorem.

24 Cancellation when the congruence information is too small

Reconstruction alone gives a lower bound for sign prediction: on a reliably decoded origin, return its sign; on an erasure, return a fair sign. An upper bound must rule out predicting a parity without learning the integer. We first do this when all inputs are pooled and the neighbouring small-prime parts give a modulus smaller than the hidden central part.

The proof works for a fixed collection $L_j(n) = h_j n + k_j$, $1 \le j \le m$, with positive integer slopes and intercepts, pairwise nonproportional. In this section suppose $z \ge 2$, $\log(2z) = O(\log y)$, and retain $\log(2\max(y, z)) = o(\ell)$. Write

$$a = s_y(n), \quad v = r_y(n), \quad s_j = s_z(L_j(n)), \quad q(\boldsymbol{s}) = \operatorname{lcm}(s_1, \dots, s_m).$$

Proposition 24.1 (Cancellation below the reconstruction threshold)**.** *Fix $0 < u_1 < u_2 < \infty$ and $0 < \eta < 1$. For*

$$E = \{y^{u_1} < a \le y^{u_2}, \quad q(\boldsymbol{s}) \le a^{1-\eta}\},$$

one has

$$\sup_{|F| \le 1} |\mathbb{E}_I \lambda(n) \mathbf{1}_E F(v, \boldsymbol{s})| = o(1). \tag{228}$$

The supremum includes functions depending on the window and on the entire exact valuation tables.

24.1 The signed smooth sequence and its progression estimate

Lemma 23.2 permits independent $a \sim \pi_y$ and $v \sim \nu_{I,y}$. Partition the permitted values of a into intervals $J \subset [A, 2A]$ with $A \asymp y^u$, $u \in [u_1, u_2]$. We prove an $o(1)$ bound after division by A, uniformly over these intervals. There are $O(\log y)$ of them, and the Euler weight $m_y \ll 1/\log y$ pays this number. Absorb $\lambda(v)$ into F.

On one such interval set $f(a) = \lambda(a)\mathbf{1}_{P^+(a)\le y}$ and use the exact convolution

$$f = \lambda * (\mu^2 \mathbf{1}_{P^->y}).$$

Retain terms $a = bd$ with $b > Y_0 = \exp(\sqrt{\log y})$; call the result $g(a)$. Since $a \le 2y^{u_2}$, $|g(a)| \le 2^{u_2+2}$. The upper-bound sieve for rough d gives

$$\frac{1}{A}\sum_{a\sim A} |f(a) - g(a)| \ll_{u_2} \frac{1+\log Y_0}{\log y} = o(1). \tag{229}$$

Indeed an omitted term has $d > 1$, and, when $A/b \ge y/2$, the number of rough $d \in (A/b, 2A/b]$ is $O(A/(b\log y))$; when $2A/b < y$ there are no such terms. This error is removed before any quadratic estimate.

We shall use the following form of dispersion. For fixed $\delta > 0$, $C, H > 0$ and $T \le A^{1-\delta}$,

$$\sum_{Q\le T} d_2(Q)^C \sum_{r \bmod Q} \max_{A\le x\le 2A} \left| \sum_{\substack{A<a\le x \\ a\equiv r \pmod Q}} g(a) \right|^2 \ll_{\delta,C,H,u_1,u_2} A^2(\log A)^{-H}. \tag{230}$$

Here and below $a \sim A$ means $A < a \le 2A$. For completeness, first omit the divisor weight and the maximum. Additive Parseval, grouping fractions by their reduced denominators, bounds the sum by

$$O(\log(2A)) \sum_{r\le T} \frac{1}{r} \sideset{}{^*}\sum_{h \bmod r} \left| \sum_{A<a\le x} g(a) e(ha/r) \right|^2.$$

For $r \le (\log A)^G$, open the convolution defining g. Siegel–Walfisz for λ in each residue class, followed by partial summation, bounds each inner sum by $O_{G,B}(A(\log A)^{-B})$ for any B. The remaining Liouville factor has argument above Y_0, so its logarithm is at least a constant times $\sqrt{\log A}$; the modulus remains polylogarithmic at this scale. The sum of $1/d$ over squarefree rough $d \le 2A$ is $O_{u_2}(1)$. The classical estimate for λ follows from $L(2s,\chi^2)/L(s,\chi)$ and the usual zero-free region; see [24, Chapter 5].

On a remaining denominator interval $R < r \le 2R$, the additive large sieve [24, Chapter 7] gives

$$O_{u_2}\big(\log(2A)(A + 4R^2)A/R\big).$$

For $(\log A)^G \le R \le A^{1-\delta}$ this saves any prescribed logarithmic power when G is large enough. Cauchy–Schwarz reintroduces $d_2(Q)^C$, using the trivial bound $O_{u_2}(A^2/Q + A)$ and the classical divisor-moment estimates. Finally discretise x with spacing $A/(\log A)^J$. The grid costs $O((\log A)^J)$; the discretisation error in each residue class is $O_{u_2}(1 + A/(Q(\log A)^J))$. Its squared, weighted total is $O((T^2 + A^2/(\log A)^{2J})(\log A)^{C'})$. Choose J first, then G and the Siegel–Walfisz powers. This proves (230), also for bounded-variation interval weights by partial summation.

24.2 Exact cells and a uniform sieve transfer

We first isolate the comparison needed here and in Section 25. The remainder is summed over the exact valuation cells before any inverse probability is introduced.

Lemma 24.2 (Uniform sieve transfer for exact valuation cells)**.** *Fix an integer $r \geq 1$, positive integers $c_1, \dots, c_r$, and a modulus $Q_0 \geq 1$. Let $H_1, \dots, H_r$ be arbitrary positive integers, let $\xi \bmod Q_0$ be fixed, and put*

$$\mathsf{S}_{\boldsymbol{H}}(v) = \big(s_z(H_i v + c_i)\big)_{i=1}^{r}, \qquad w = 2\max(y, z), \qquad m_y = \prod_{p \leq y}(1 - p^{-1}),$$

where $y \geq 1$ and $z \geq 2$; for $1 \leq y < 2$ the product is one. Let $X_0 \geq 3$, $1 \leq B \leq X_0^{1/4}$, and let $\mathcal{T}$ be any collection of z-smooth tuples $\boldsymbol{d} = (d_1, \dots, d_r)$ with

$$q(\boldsymbol{d}) = \operatorname{lcm}(d_1, \dots, d_r) \leq B.$$

Write $P_{\boldsymbol{d}}$ for the probability of $v \equiv \xi \pmod{Q_0}$ and $\mathsf{S}_{\boldsymbol{H}}(v) = \boldsymbol{d}$ when the local coordinates of v are independent Haar units at $p \leq y$ and Haar integers at $p > y$. These probabilities may have total mass less than one because of the fixed congruence and the truncation.

There are constants $c, C, s_0 > 0$, depending only on $r, \boldsymbol{c}, Q_0$, such that, if $\log X_0 / \log w \geq s_0$, then for every $X \geq X_0$, with $\epsilon_X = \exp(-c \log X / \log w)$,

$$\left| \frac{N_{\boldsymbol{d}}(X)}{X} - m_y P_{\boldsymbol{d}} \right| \leq C \epsilon_X m_y P_{\boldsymbol{d}} + R_{\boldsymbol{d}}(X), \tag{231}$$

where $N_{\boldsymbol{d}}(X)$ counts positive integers $v \leq X$ satisfying the cell, the fixed congruence, and $P^-(v) > y$. The remainders can be chosen nonnegative with

$$\sum_{\boldsymbol{d} \in \mathcal{T}} R_{\boldsymbol{d}}(X) \leq C B X^{-11/12} (\log(2X))^C. \tag{232}$$

All constants are independent of the moving slopes H_i, the particular cells, and their collection $\mathcal{T}$. A cell of zero local mass has zero physical count.

For the harmonic application, suppose $\log w = o(\log X_0)$ and $\log X_0 = o(\ell)$, where $I = (U, V]$, $1 \leq U < V$, and $\ell = \log(V/U) \to \infty$. Set

$$T = \max(U, X_0), \qquad \mathcal{N}_{I,y} = \sum_{\substack{v \in I \\ P^-(v) > y}} \frac{1}{v}, \qquad \vartheta_I = \frac{m_y \log(V/T)}{\mathcal{N}_{I,y}}.$$

Then $\vartheta_I = 1 + O((\log X_0 + \log(2y))/\ell) = 1 + o(1)$, independently of the slopes and cells. Write

$$p_{\boldsymbol{d}}^I = \nu_{I,y}\{v > T,\ v \equiv \xi \pmod{Q_0},\ \mathsf{S}_{\boldsymbol{H}}(v) = \boldsymbol{d}\}.$$

For arbitrary complex weights $h_{\boldsymbol{d}}$ one has

$$\left| \sum_{\boldsymbol{d} \in \mathcal{T}} h_{\boldsymbol{d}} (p_{\boldsymbol{d}}^I - \vartheta_I P_{\boldsymbol{d}}) \right| \leq C \epsilon_{X_0} \sum_{\boldsymbol{d} \in \mathcal{T}} |h_{\boldsymbol{d}}| P_{\boldsymbol{d}} + C \|h\|_\infty \frac{B X_0^{-11/12} (\log(2X_0))^C}{m_y \ell}. \tag{233}$$

The probability on the left is not renormalised after discarding $v \leq X_0$. A fixed congruence of positive local mass can instead be conditioned on, at the cost of a constant depending on Q_0.

Finally, in the local rough model,

$$\mathbb{E}\left(\prod_{i=1}^{r} s_z(H_i v + c_i) \right)^{1/(4r \log(2z))} \ll_{r, \boldsymbol{c}} 1. \tag{234}$$

Consequently a product cutoff $\prod_i d_i \leq X_0^\kappa$, $0 < \kappa \leq 1/4$ fixed, discards $o(1)$ mass in both the local and harmonic physical models. After controlling this cutoff, their full table laws are at total variation distance $o(1)$ after the negligible deletion of $v \leq X_0$. These assertions include $y = 1$, when no roughness is imposed.

Proof. Fix a tuple and put $q = q(\boldsymbol{d})$ and $Q = \operatorname{lcm}(q, Q_0)$. If $d_i \mid H_i v + c_i$ is soluble, then $\gcd(H_i, d_i) \mid c_i$. There are therefore at most $\prod_i c_i$ simultaneous residue classes for v modulo q. Intersecting with the prescribed class modulo Q_0 gives at most the same number of classes $v = b + Qt$. This bound does not depend on any H_i.

On such a class the quotient forms

$$F_i(t) = \frac{H_i(b + Qt) + c_i}{d_i}$$

have integer coefficients. The exact table condition is $p \nmid F_i(t)$ for every $p \le z$, and roughness additionally requires $p \nmid b + Qt$ for $p \le y$. At a prime, a quotient with both coefficients divisible by p makes the class impossible; a quotient with zero slope and nonzero constant modulo p forbids no residue. Every remaining quotient forbids one residue. Let Ω_p be the union of these forbidden residues and write $\rho(p) = |\Omega_p|$. If $\rho(p) = p$ for some p, the class has zero mass in both models and is discarded. In all other classes,

$$0 \le \rho(p) < p, \qquad \rho(p) \le r + 1.$$

This treatment includes the contents of the quotient forms. It also includes the extra p-adic digit needed for an exact valuation: after the division by d_i, exclusion of a root of F_i modulo p is precisely exclusion of divisibility by the next power of p. When other forced divisibilities already fix that digit, the quotient is constant modulo p and the preceding two cases apply.

For $2 \le u \le t$, Mertens' product estimate gives, uniformly in the surviving class,

$$\prod_{u < p \le t} (1 - \rho(p)/p)^{-1} \le \left(\frac{\log t}{\log u}\right)^{r+1} \left(1 + \frac{C_r}{\log u}\right).$$

Indeed, above $2r + 2$ each factor is at most $(1 - 1/p)^{-(r+1)}(1 + O_r(p^{-2}))$. The finitely many smaller primes are absorbed by increasing C_r, since each surviving factor is at most p. Thus the dimension and the constant in this sieve condition are independent of the moving slopes, their common factors, and the forced modulus.

Apply the fundamental lemma of the sieve in this fixed dimension, in its upper- and lower-weight form [24, Section 6.4, pp. 158–160]. More explicitly, at level $D = X^{1/12}$ it supplies weights λ_d^- and λ_d^+, of modulus at most one, supported on squarefree $d \le D$ with primes at most w, which bracket the indicator of avoiding every Ω_p. Their local expectations both equal

$$V_b(1 + O_r(e^{-c \log D / \log w})), \qquad V_b = \prod_{p \le w} (1 - \rho(p)/p).$$

The parameter $\log D / \log w = \log X / (12 \log w)$ exceeds the required fixed threshold by our choice of s_0. For each d, the Chinese remainder theorem gives $\rho(d) = \prod_{p \mid d} \rho(p)$ residue classes for $t \bmod d$. Each becomes a progression of modulus Qd in v and has count $X/(Qd) + O(1)$. Hence the count on the base class is

$$\frac{X}{Q} V_b(1 + O(\epsilon_X)) + O\left(\sum_{d \le D} (r + 1)^{\omega(d)}\right),$$

with the error interpreted through the two bounds. The final sum is $O_r(D(\log(2D))^C)$.

The exact profinite mass of this base class and all its exclusions is V_b/Q. Summing the base classes gives exactly $m_y P_{\boldsymbol{d}}$, since conditioning the ordinary local integer on roughness changes its mass by m_y. This proves (231). For each q there are at most $d_2(q)^r$ tuples with $\operatorname{lcm}(d_1, \dots, d_r) = q$. Thus the sum of all absolute progression-count errors is at most

$$C \sum_{q \le B} d_2(q)^r \sum_{d \le D} (r + 1)^{\omega(d)} \ll BD(\log(2BD))^C.$$

Equivalently, on regrouping by the final modulus Qd the multiplicity is bounded by a fixed divisor power. Division by X, using $B \le X^{1/4}$, gives (232). No primorial enters this count.

For the harmonic assertion, apply Abel summation to $N_{\boldsymbol{d}}(x) - m_y P_{\boldsymbol{d}} x$ on $(T, V]$. The relative part of (231), including the two boundary terms, costs at most

$$C\epsilon_{X_0} m_y P_{\boldsymbol{d}}(2 + \log(V/T)).$$

For the absolute part, sum over cells first. Its two boundary terms and integral are bounded by

$$CBT^{-11/12}(\log(2T))^C \ll BX_0^{-11/12}(\log(2X_0))^C.$$

Here and below the last comparison is for sufficiently large X_0. Lemma 23.2 gives $\mathcal{N}_{I,y} = m_y \ell(1 + O(\log(2y)/\ell))$; for $y = 1$ this is the ordinary harmonic-sum estimate. Dividing by this mass proves (233). Also $\log(V/T) = \ell + O(\log X_0)$, proving $\vartheta_I = 1 + o(1)$. These estimates remain uniform when $\log U$ is much larger than ℓ.

To prove (234), put $\alpha = 1/(4r\log(2z))$ and use Hölder across the r forms. At a prime not dividing c_i, the tail probability of a positive valuation of $H_i v + c_i$ is at most $2p^{-e}$ at level e, uniformly in H_i and in whether v_p is Haar or a Haar unit. At primes dividing c_i, the same bound holds after shifting the valuation by $v_p(c_i)$. Consequently the local moment with exponent $r\alpha$ is bounded, apart from these fixed factors, by

$$1 + 2\frac{p^{r\alpha} - 1}{p - p^{r\alpha}} = 1 + O\left(\frac{\alpha \log p}{p}\right).$$

Independence over primes and $\sum_{p \le z} \log p/p \ll \log(2z)$ give a bounded product. Markov's inequality therefore bounds the local tail beyond a product cutoff B_0 by $C\exp(-\log B_0/(4r\log(2z)))$.

Apply (233) with $|h| \le 1$ on the retained product cells, which all have $q \le B_0$. The same counting argument with only the roughness condition estimates the total retained-v mass. Subtracting the small-cell mass from that total bounds the physical table tail by its local counterpart plus the same errors. With $B_0 = X_0^{\kappa}$, these errors tend to zero: the absolute term is $X_0^{\kappa - 11/12 + o(1)}/\ell$, since $m_y^{-1} \ll \log(2y) = X_0^{o(1)}$. The deleted low-$v$ mass is $O((\log X_0 + \log(2y))/\ell) = o(1)$ by (226). The weighted bound on the small cells and the two tail bounds prove the total variation assertion. A positive fixed congruence has local rough mass bounded below in terms of Q_0, so conditioning preserves these conclusions. □

Put $\beta = 1 - \eta$, $Q_* = (2A)^{\beta}$, and

$$w_q(a) = \frac{A}{a}\mathbf{1}_{a \in J,\ a \ge q^{1/\beta}}, \qquad W_{\boldsymbol{s}}(v) = \frac{1}{A}\sum_{a \sim A} g(a) w_{q(\boldsymbol{s})}(a)\mathbf{1}_{S(av) = \boldsymbol{s}}.$$

Let $\pi_{\boldsymbol{s}}$ be the exact probability of the table $\boldsymbol{s} = (s_z(L_j(n)))_j$ for a Haar-distributed profinite integer n. Keep only cells of positive mass. Cauchy–Schwarz gives

$$\mathbb{E}_v \sum_{\boldsymbol{s}} |W_{\boldsymbol{s}}(v)| \le \left(\mathbb{E}_v \sum_{q(\boldsymbol{s}) \le Q_*} \frac{|W_{\boldsymbol{s}}(v)|^2}{\pi_{\boldsymbol{s}}}\right)^{1/2}. \tag{235}$$

For fixed forms,

$$(q(\boldsymbol{s})\pi_{\boldsymbol{s}})^{-1} \ll (\log(2z))^{C_m}. \tag{236}$$

At a prime not dividing a slope, intercept or determinant, the local empty state has mass $1 - m/p$, and an occupied state of exact valuation e has mass $p^{-e}(1 - 1/p)$. At the finitely many remaining primes, a soluble configuration contains a residue class modulo $p^{v_p(q)+1}$. These descriptions prove (236).

Discard rough $v \le X_0$, where $X_0 = \exp(\sqrt{\ell \log(2\max(y,z))})$, before applying (235). The loss is $o(1)$ by (226). The parameters satisfy

$$\frac{\log X_0}{\log(2\max(y,z))} \longrightarrow \infty, \qquad \frac{\log X_0}{\ell} \longrightarrow 0, \qquad A = X_0^{o(1)}, \qquad Q_* = X_0^{o(1)}.$$

For $a, a' \sim A$, let $P_{\boldsymbol{s}}(a,a')$ be the joint probability of $S(av) = S(a'v) = \boldsymbol{s}$ in the local rough model. Apply Lemma 24.2 to the $2m$ forms with slopes $(h_j a)_j, (h_j a')_j$ and intercepts $(k_j)_j, (k_j)_j$. Here $Q_0 = 1$ and the two tables are identical, so their joint forced modulus is $q(\boldsymbol{s}) \le B = Q_*$, not its square. In particular the slopes are $O(y^{u_2}) = X_0^{o(1)}$, although the lemma does not need this bound for its counting constants.

Let $\mathcal{Q}_A^{\mathrm{phys}}$ denote the squared expression on the right of (235), after this deletion, and let $\mathcal{Q}_A^{\mathrm{loc}}$ be its local counterpart:

$$\mathcal{Q}_A^{\mathrm{loc}} = \frac{1}{A^2} \sum_{a,a' \sim A} g(a)\overline{g(a')} \sum_{q(\boldsymbol{s}) \le Q_*} w_{q(\boldsymbol{s})}(a) w_{q(\boldsymbol{s})}(a') \frac{P_{\boldsymbol{s}}(a,a')}{\pi_{\boldsymbol{s}}}.$$

Since the interval weights and g are bounded, the harmonic estimate (233), averaged over a, a', gives

$$\left| \mathcal{Q}_A^{\mathrm{phys}} - \vartheta_I \mathcal{Q}_A^{\mathrm{loc}} \right| \ll \epsilon_{X_0} \mathcal{U}_A + \frac{Q_*^2 X_0^{-11/12} (\log(2X_0))^C}{m_y \ell}, \tag{237}$$

where

$$\mathcal{U}_A = \frac{1}{A^2} \sum_{a,a' \sim A} \sum_{q(\boldsymbol{s}) \le Q_*} \frac{P_{\boldsymbol{s}}(a,a')}{\pi_{\boldsymbol{s}}} = O(1). \tag{238}$$

The second power of Q_* in (237) comes from the inverse weight $\pi_{\boldsymbol{s}}^{-1} \ll Q_* (\log(2z))^{C_m}$, after the sum of absolute cell remainders has already cost one power of Q_*. The displayed absolute error is $X_0^{-11/12+o(1)}/\ell$, hence in particular $O(X_0^{-3/4})$. Since $|\mathcal{Q}_A^{\mathrm{loc}}| \ll \mathcal{U}_A$, replacing ϑ_I by one costs at most $O((\log X_0 + \log(2y))\mathcal{U}_A/\ell)$. The next two subsections prove $\mathcal{Q}_A^{\mathrm{loc}} = o(1)$ and (238) entirely in the local model. Thus the relative error multiplies an absolute $O(1)$ quantity. This is needed when $\ell/\log\max(y,z)$ tends to infinity arbitrarily slowly.

24.3 A local expansion with a prescribed modulus

Let $\mathcal{B}$ contain all prime divisors of the slopes, intercepts and nonzero pairwise determinants, and all primes at most $2m+2$. Set

$$R(a,a') = aa' \prod_{i,j} (k_i h_j a' - k_j h_i a).$$

For a good prime $p \le z$ not dividing q, the local factor for the joint empty condition has reference value $1 - 2m/d_p$, where $d_p = p-1$ for $p \le y$ and $d_p = p$ otherwise. Its correction is $O_m(1/p)$, depends on $a, a' \bmod p$, and vanishes unless $p \mid R$.

If $p^e \| q$ is good, only one site can be occupied. Its joint exact valuation requires $p \nmid aa'$ and $a \equiv a' \pmod{p^e}$. After multiplication by p^e, the factor is

$$\frac{p}{d_p} \mathbf{1}_{p \nmid aa',\ a \equiv a' \bmod p^e} \left(1 - \frac{2}{p} + \frac{1}{p} \mathbf{1}_{a \equiv a' \bmod p^{e+1}}\right).$$

At a bad prime the normalised factor is $O_{\mathcal{B}}(1)$ and is determined modulo $p^{v_p(q)+1}$. In every nonempty cell, $q \mid C(a-a')$ for a fixed C, for instance a multiple of $\prod_j k_j^2$. Indeed $s_j \mid h_j v(a-a')$ and the common divisors with $h_j v$ are bounded in terms of k_j.

Expand the good local factors into their reference and correction terms. The contribution indexed by the squarefree product d of corrected primes, in the kernel $qP_{\boldsymbol{s}}$, depends only on residues modulo $Q = C_0qd$, with fixed C_0, and satisfies

$$|k_{d,\boldsymbol{s}}(a,a')| \ll (\log(2A))^{C_m}\frac{B^{\omega(d)}}{d}\mathbf{1}_{a\equiv a' \bmod q/(q,C)}\mathbf{1}_{p|d\Rightarrow p|R(a,a')}. \tag{239}$$

The product of the factors p/d_p is at most logarithmic in $\log(3q)$; all remaining reference factors are at most one. In particular no reference involves the full primorial up to z.

Take $D = A^{\eta/2}$. For $R \neq 0$, $|R| \ll A^{m^2+2}$, and there are only $A^{o(1)}$ permitted cells for the pair (a,a'), since $q \mid C(a-a')$. Thus the tail in (239) is at most

$$(\log(2A))^{C_m}D^{-1/2}\prod_{p|R}(1+Bp^{-1/2}) = D^{-1/2}A^{o(1)},$$

also after summing cells. If $R = 0$ and $a \neq a'$, the pair lies on one of finitely many rational lines, giving $O(A)$ pairs and $A^{o(1)}$ cells. The real and truncated kernels give $A^{-1+o(1)}$ after division by A^2. On the exact diagonal the total real or truncated contribution is $\ll Q_*(\log(2A))^{C'}/A = o(1)$, by counting the permitted cells. The weight (236) costs only additional logarithms.

For $d \le D$, the residue matrix modulo $Q = C_0qd$ has $O(d)$ possible entries in each row and column. Write $S(Q,r;q) = \sum_{a\sim A,\ a\equiv r \bmod Q} g(a)w_q(a)$. The inequality $2|uv| \le |u|^2+|v|^2$ bounds a cell's contribution by

$$\ll \frac{(\log(2A))^{C_m}B^{\omega(d)}}{A^2}\sum_{r \bmod Q}|S(Q,r;q)|^2.$$

For each Q, the number of choices of q,d and cell is bounded by a fixed power of $d_2(Q)$. Since $Q \ll A^{1-\eta/2}$, (230) makes the signed total $o(1)$.

24.4 An absolute bound for the sieve transfer

It remains to justify (238). Compare uniform integers $a,a' \sim A$ with two independent Haar profinite integers. The truncation just proved has error $O(A^{-c_\eta})$ up to fixed logarithmic factors. Under Haar measure, a correction prime not dividing q requires one of a bounded number of lines modulo p, and therefore has an additional probability factor $O_m(1/p)$. At $p^e\|q$ a correction requires equality modulo p^{e+1}, giving the same additional factor conditional on equality modulo p^e. Consequently

$$\mathbb{E}_{a,a'\ \mathrm{Haar}}|k_{d,\boldsymbol{s}}(a,a')| \ll (\log(2A))^{C_m}\frac{B_1^{\omega(d)}}{qd^2}.$$

Since $\sum_{\boldsymbol{s}:q\le Q_*}1/q$ is bounded by a fixed power of $\log(2A)$, the tail is negligible under Haar measure too.

For $d \le D$, change the distribution of one argument at a time. Each residue modulo Q has probability error $O(1/A)$, and only $O(d)$ residues are permitted for the second argument at a fixed first argument. Including (236), the error per cell and per d is $O((\log(2A))^{C_m}B^{\omega(d)}/A)$. There are $O(Q_*(\log(2A))^{C_m})$ cells. The total error is $o(1)$ because $DQ_* \ll A^{1-\eta/2}$.

For the full Haar expression, lift $q \le Q_*$ by positivity. At a fixed profinite v, the two tables are independent. If their common law is $\pi_{\boldsymbol{s}}(v)$, the expression becomes $\mathbb{E}_v\sum_{\boldsymbol{s}}\pi_{\boldsymbol{s}}(v)^2/\pi_{\boldsymbol{s}}$. It factors over primes. For $p \le y$, v is a unit and multiplication by v preserves Haar measure, so the local factor is one. For $y < p \le z$, the prime is eventually good: if v is a unit the factor is one, while if $p \mid v$ the table is the empty state, of probability $1-m/p$ under the reference law. The averaged factor is

$$(1-1/p)+\frac{1}{p(1-m/p)} = 1+O_m(p^{-2}).$$

The Euler product is bounded. This proves (238) independently of the physical measure. Equation (237) now transfers the signed local $o(1)$ estimate to the physical quadratic expression, and (235) proves Proposition 24.1.

Corollary 24.3 (The optimal sign with pooled inputs)**.** *For the three specified neighbours and the regime of Theorem 23.1,*

$$\|\mathbb{E}_I[\lambda(n) \mid v, S]\|_1 \longrightarrow P(\infty, \tau).$$

Thus the optimal probability of correctly predicting $\lambda(n)$ tends to $(1 + P(\infty, \tau))/2$.

Proof. For finite positive τ, the product $\prod_j S_j$ and their least common multiple differ by a bounded factor, because their pairwise greatest common divisors divide fixed determinants. Thus $\log q/\log y \Rightarrow \tau(D_1 + D_2 + D_3)$ jointly with $\log a/\log y \Rightarrow D_0$. Apply Proposition 24.1 and let $u_1 \downarrow 0$, $u_2 \uparrow \infty$, $\eta \downarrow 0$. Continuity gives the upper bound. The selective reconstruction theorem gives the lower bound by returning $\lambda(\widehat{n})$ on acceptance and zero on erasure. At $\tau = 0$, the same cancellation argument and tightness give zero. At $\tau = \infty$, selective reconstruction gives one, which is already the maximal possible advantage. □

25 The exact cost of one message for a Liouville sign

Theorem 25.1 (One-message sign prediction)**.** *Under the hypotheses of Theorem 23.1, let $\mathcal{A}_I(b)$ be the supremum of $|\mathbb{E}_I \lambda(n) F(v, \mathsf{M})|$ over $[-1, 1]$-valued outputs and one message M from Bob to Alice of at most b bits. Then*

$$\mathcal{A}_I(b) \longrightarrow P(t, \tau), \qquad \mathrm{Succ}_I^*(b) \longrightarrow \frac{1 + P(t, \tau)}{2}.$$

Public randomness is allowed. No assertion of this sign upper bound is made for interaction with an unrestricted preliminary message from Alice.

The lower bound follows from the selective decoder. For the upper bound, Proposition 24.1 treats the region where the combined modulus is too small. It remains to prove cancellation on each dyadic interval $A < a \le 2A$, with $a \in [y^{u_1}, y^{u_2}]$, for which

$$2^b \le A^{1-\delta} \tag{240}$$

with fixed $\delta, u_1, u_2 > 0$. This argument permits arbitrary $\log z/\log y$, including a limit of infinity.

25.1 From a message to a weighted quadratic expression

Fix the public randomness. The messages partition the exact tables into $B_1, \dots, B_N$, $N \le 2^b$. Remove a fixed bad-prime set $\mathcal{B}$ containing all divisors of slopes, intercepts and determinants, and all $p \le 2m + 2$. Their valuations can first be bounded by a fixed J, losing $o_{J\to\infty}(1)$, uniformly in sufficiently large windows. We may give these coordinates to Alice for free. There are then finitely many classes modulo a fixed Q_0 for n and, after factorisation, for a and v. Include the class of a in g and the class of v in the sieve. All subsequent conductors are coprime to Q_0.

Let π be the law of the good-prime table when its central argument is a Haar unit at each good $p \le z$, and set $\alpha_l = \pi(B_l)$. Discard messages with $\alpha_l < A^{-\delta/4}/N$ before using inverse weights. For all dyadic intervals one may use the common cutoff $y^{-u_1\delta/4}/N$. The ordinary profinite table law is dominated by $C\pi$: an occupied state has density ratio $(p-1)/p$, and the empty state has ratio

$$\frac{1 - m/p}{1 - m/(p-1)} = 1 + O_m(p^{-2}).$$

The total variation conclusion of Lemma 24.2, applied to the fixed forms in the origin n with no roughness condition ($y = 1$ in that lemma), remains valid on each of the fixed classes modulo Q_0. Since the union of the discarded cells has π-mass at most $y^{-u_1\delta/4}$, its physical mass is $o(1)$. For the remaining messages,

$$\alpha_l^{-1} \le NA^{\delta/4} \le A^{1-3\delta/4}. \tag{241}$$

A common cutoff $y^{-\epsilon_0}/N$ with $0 < \epsilon_0 < u_1\delta/4$ gives the same dyadic inequality.

Discard $v \le X_0 = \exp(\sqrt{\ell \log(2\max(y,z))})$. Put $H = X_0^{1/20}$ and truncate the full tables, including their fixed bad-prime coordinates, at $\prod_j s_j \le H$. The low-v deletion has mass $O((\log X_0 + \log(2y))/\ell) = o(1)$. The Rankin estimate (234) and the table-tail conclusion of Lemma 24.2, again applied first to the physical origin with fixed slopes, show that the discarded table mass is $o(1)$. These deletions precede Cauchy–Schwarz and every inverse probability weight. Intersect B_l with the product truncation; the original α_l may be retained, so $\sum_l \alpha_l \le 1$ and the new cell probabilities are no larger.

Use g from (229), now including the bounded weight A/a, the dyadic interval, and the fixed class. Cauchy–Schwarz in the message and then in v bounds the absolute dyadic correlation by the square root of

$$\mathbb{E}_{v \text{ physical}} \sum_l \frac{1}{\alpha_l} \left| \frac{1}{A} \sum_{a\sim A} g(a) \mathbf{1}_{S(av)\in B_l} \right|^2. \tag{242}$$

The freely chosen output function has disappeared. We prove that the corresponding profinite expression is $o(1)$ for the signed g, and $O(1)$ for the unsigned sequence 1.

Here is the precise second application of Lemma 24.2. For fixed $a, a' \sim A$, use the $2m$ forms with slopes $(h_j a)_j, (h_j a')_j$, and retain the prescribed class of v mod Q_0 in both the physical and local measures. The two tables can now be different. Their joint forced modulus satisfies

$$\operatorname{lcm}(s_1, \dots, s_m, s_1', \dots, s_m') \le \prod_j s_j s_j' \le H^2 = X_0^{1/10}.$$

Thus take $B = H^2$, with the extra fixed modulus handled by Q_0 in the lemma. For every $X \ge X_0$, (232) bounds the sum of the normalised absolute remainders, over all pairs of retained tables, by

$$CH^2 X^{-11/12} (\log(2X))^C \le X^{-49/60+o(1)}.$$

This estimate is uniform over a, a'; their slopes are $O(y^{u_2}) = X_0^{o(1)}$.

The test applied to the two exact cells is

$$K(\boldsymbol{s}, \boldsymbol{s}') = \sum_l \frac{\mathbf{1}_{\boldsymbol{s},\boldsymbol{s}'\in B_l}}{\alpha_l}, \qquad 0 \le K \le \alpha_*^{-1}, \qquad \alpha_* = \min_l \alpha_l.$$

Disjointness of the message cells gives the pointwise bound; there is no additional factor N. Let $\mathcal{V}_A^{\mathrm{phys}}$ denote (242) with the low-v deletion and the fixed class retained, and let $\mathcal{V}_A^{\mathrm{loc}}$ be the same quadratic expression in the local rough model. Write $\mathcal{U}_A^{\mathrm{msg}}$ for its unsigned version with $g = 1$. Averaging (233) over a, a' gives

$$\left| \mathcal{V}_A^{\mathrm{phys}} - \vartheta_I \mathcal{V}_A^{\mathrm{loc}} \right| \ll \epsilon_{X_0} \mathcal{U}_A^{\mathrm{msg}} + \frac{\alpha_*^{-1} H^2 X_0^{-11/12} (\log(2X_0))^C}{m_y \ell}. \tag{243}$$

By (241), $\alpha_*^{-1} \le A^{1-3\delta/4} = X_0^{o(1)}$, and $m_y^{-1} \ll \log(2y) = X_0^{o(1)}$. The absolute term in (243) is therefore $X_0^{-49/60+o(1)}/\ell$, and in particular $O(X_0^{-2/3})$. Replacing ϑ_I by one costs at most $O((\log X_0 + \log(2y))\mathcal{U}_A^{\mathrm{msg}}/\ell)$.

The following two subsections prove $\mathcal{V}_A^{\mathrm{loc}} = o(1)$ and $\mathcal{U}_A^{\mathrm{msg}} = O(1)$ using only the local model. Both the relative sieve error and the harmonic normalisation error then tend to zero without a logarithmic loss, even for arbitrarily slow growth of $\ell/\log(2\max(y,z))$.

25.2 Small conductors

In the profinite rough model, condition on the set $\mathcal{J}$ of primes $y < p \le z$ that divide v. At these good primes all forms $h_j av + k_j$ are units, independently of a. Outside $\mathcal{J}$, v is a product of Haar units. The conditional reference law is dominated by $C(\mathcal{J})\pi$, where

$$C(\mathcal{J}) = \prod_{p\in\mathcal{J}} (1 - m/(p-1))^{-1}, \qquad \mathbb{E}_{\mathcal{J}} C(\mathcal{J}) = \prod_{y<p\le z} (1 + O_m(p^{-2})) \ll 1.$$

This is uniform even if $\log z/\log y \to \infty$. Omit the forced coordinates and let ν be the remaining table law.

Let $G = \prod_{p\le z,\, p\notin\mathcal{B}\cup\mathcal{J}} \mathbb{Z}_p^\times$. A continuous character χ has an exact primitive conductor q; extend it by zero when an integer is not a unit at a prime dividing q. For the table S_1 of a Haar element of G, put $R_\chi(s) = \mathbb{E}[\chi(v) \mid S_1 = s]$. If $\pi_p(0)$ is the local empty probability, then $\pi_p(0) = 1 - m/(p-1)$. Define

$$w_a(s) = \prod_{\substack{p\mid a \\ p \text{ indexes } G}} \frac{\mathbf{1}_{s_p=0}}{\pi_p(0)}.$$

The law of $S(av)$ is $w_a\, d\nu$. If $p \mid (a,q)$ the local Fourier coefficient is zero. Otherwise multiplication by a on the unit coordinates gives

$$\int_G \mathbf{1}_{S(av)\in B_l} \overline{\chi(v)}\, dv = \chi(a) \int_{B_l} \overline{R_\chi(s)} w_a(s)\, d\nu(s). \tag{244}$$

For $p^e \| q$, the local squared norm of R_χ is $O_m(p^{-e})$. At exact valuation k, one knows a residue modulo p^k with one child residue excluded. The character average is zero if $k < e-1$, at most $1/(p-1)$ in modulus if $k = e-1 \ge 1$, and at most one if $k \ge e$; the last states have mass $O_m(p^{-e})$. For $e = 1$, the empty state has character average $O_m(1/p)$ by orthogonality after deleting the m roots, and occupied states have mass $O_m(1/p)$. Thus

$$\|R_\chi\|_2^2 \le C^{\omega(q)}/q.$$

On the support of $\chi(a)$, the weight w_a uses no prime dividing q, so these factors integrate independently. With $\theta_p = \pi_p(0)^{-1} - 1 \ll_m 1/p$ and the squarefree multiplicative extension θ_d, one has

$$\int w_a w_{a'}\, d\nu = \prod_{p\mid(a,a')} (1+\theta_p) = \sum_{d\mid(a,a')} \theta_d.$$

Apply Parseval on G, retaining conductors $q \le D = A^{1-\delta/2}$. Cauchy–Schwarz on each message cell, using $\nu(B_l) \le C(\mathcal{J})\alpha_l$, bounds their total by

$$\frac{CC(\mathcal{J})}{A^2} \sum_{d\le 2A} \theta_d \sum_{q\le D} \frac{C_1^{\omega(q)}}{q} \sideset{}{^*}\sum_{\chi \bmod q} \left| \sum_{\substack{a\sim A \\ d\mid a}} g(a)\chi(a) \right|^2. \tag{245}$$

Positivity permits extending the conductor sum to all q. For $R < q \le 2R$, $R \ge (\log A)^G$, the multiplicative large sieve bounds the contribution for a fixed d, before θ_d, by

$$O_\epsilon\left(R^{-1+\epsilon}/d^2 + R^{1+\epsilon}/(Ad)\right).$$

Here $C_1^{\omega(q)} \ll_\epsilon q^\epsilon$ and $\chi(d)$ separates when $(d,q) = 1$; otherwise the summand is zero. Since $\sum_d \theta_d/d \ll 1$, the sum up to D saves any desired logarithmic power by choosing $\epsilon < \delta/4$ and then G large.

For $q \le (\log A)^G$, the terms $d > (\log A)^H$ are negligible when H is large, by the trivial bound and $\theta_d \ll_\epsilon d^{-1+\epsilon}$. For polylogarithmic q, d, open $a = ru$ in the truncated convolution. Since $P^-(r) > y$, $(r, d) = 1$ and $d \mid u$. Write $u = dm$. Siegel–Walfisz for $\lambda(m)\chi(m)$, with the fixed residue condition and partial summation, gives

$$O_{B,G,H}\left(\frac{A}{rd}(\log A)^{-B}\right).$$

Indeed the remaining Liouville factor lies above Y_0/d, and its logarithm is $\gg \sqrt{\log A}$. The rough reciprocal sum in r is bounded. Thus (245) is $o(1)$ for signed g. For the unsigned expression remove the fixed class by positivity before Parseval. Nonprincipal primitive characters have partial sums $O(q)$, and the conductor-one term is $O(\sum_d \theta_d/d^2) = O(1)$. The same calculation yields $O(1)$.

25.3 Large conductors and the exceptional rational lines

At a good prime, let $J_{p,e}(a, a')$ be the joint law of the two tables when two Haar units v, v' are conditioned to agree modulo p^e and are independent within that class. For $e = 0$ they are independent. Put $K_{p,0} = J_{p,0}$ and $K_{p,e} = J_{p,e} - J_{p,e-1}$ for $e \ge 1$. Pairing against functions of the tables, $J_{p,e}$ is projection on characters of conductor dividing p^e; the difference selects exact conductor p^e. Thus $K_q = \bigotimes_p K_{p,v_p(q)}$ is precisely the contribution of all characters of exact conductor q to Parseval at (a, a'). This is first an identity for finitely truncated tables and then follows by convergence of the orthogonal projections.

Use the same polynomial $R(a, a')$ as above. If $p \nmid R$, the $2m$ roots are distinct modulo p. Passing from two independent residues to an identical residue changes their first-level correlation only. Both marginal laws are fixed; the mass of simultaneously occupied states in the independent law is $O_m(p^{-2})$, whereas it is zero in the common-residue law. It follows that

$$\|K_{p,1}\|_{\mathrm{var}} \ll_m p^{-2} \quad (p \nmid R), \qquad \|K_{p,1}\|_{\mathrm{var}} \ll_m p^{-1} \quad (p \mid R).$$

Here $\|\cdot\|_{\mathrm{var}}$ is absolute mass, without a factor $1/2$. If $p \mid a$ or $p \mid a'$, one marginal is deterministic and the difference is actually zero. For $e \ge 2$, a nonzero difference requires a root of one family to agree with a root of the other modulo p^{e-1}. This implies $p^{e-1} \mid R$. The relevant common class has mass $O_m(p^{1-e})$ and the next digit changes a valuation with probability $O_m(1/p)$; equality of marginals then gives

$$\|K_{p,e}\|_{\mathrm{var}} \ll_m p^{-e}.$$

For $R \neq 0$, $|R| \ll A^{m^2+2}$. For fixed $0 < \epsilon < 1$,

$$\begin{aligned}\sum_{q>D} \|K_q(a, a')\|_{\mathrm{var}} &\ll D^{-1+\epsilon} \prod_{p \nmid R}(1 + C_m p^{-1-\epsilon}) \prod_{p \mid R}(1 + C_{m,\epsilon}) \\ &\le D^{-1+\epsilon} A^{o(1)}.\end{aligned}$$

This is uniform in $\mathcal{J}$. The message test $\sum_l \mathbf{1}_{s,s' \in B_l}/\alpha_l$ is bounded by $\max_l \alpha_l^{-1}$. Combining with (241) and $D = A^{1-\delta/2}$ makes this tail $o(1)$ for sufficiently small ϵ.

If $R = 0$, then $a' = a$ or a'/a is one of a fixed finite set of rational numbers, so there are $O(A)$ pairs. The full joint law is a probability, and the sum of variations for conductors at most D is $O((\log A)^C)$ by the local $O_m(p^{-e})$ bounds and Mertens' product estimate. Subtract the finite projector from the full law instead of summing an infinite tail absolutely. These pairs contribute

$$O\left(A^{-1}(\log A)^C \max_l \alpha_l^{-1}\right) = o(1).$$

The same estimates justify passage from truncated valuations to exact tables. Averaging $C(\mathcal{J})$ finishes the signed $o(1)$ and unsigned $O(1)$ estimates for the profinite version of (242), with any

fixed bad-prime class contributing only its bounded mass. In the notation of (243), these are precisely $\mathcal{V}_A^{\mathrm{loc}} = o(1)$ and $\mathcal{U}_A^{\mathrm{msg}} = O(1)$. That transfer estimate now proves cancellation under (240) in the physical experiment.

To finish Theorem 25.1, split according to $\log_y a \le t + \epsilon$ or its complement. On the first set use the pooled-input upper bound and Proposition 24.1; on the second use (240) on the dyadic intervals. Restrict first to $y^{u_1} < a \le y^{u_2}$, then let $u_1 \downarrow 0$, $u_2 \uparrow \infty$ and all fixed boundary margins decrease to zero. The joint Dickman law has null boundaries. When $\tau = \infty$, the neighbouring modulus dominates a in probability and only the communication bound is needed. When $\tau = 0$, the pooled-input bound is zero. At $t = \infty$, Corollary 24.3 and selective decoding suffice. An unrestricted preliminary message depending on v would make the partition B_l depend on v; the proof above does not cover that different protocol.

26 The Dickman curve and a reproducible numerical value

The limiting probability has useful exact and asymptotic forms. For $0 \le t \le \min(1, \tau)$ with $\tau > 0$,

$$P(t, \tau) = e^{-\gamma} t - \frac{e^{-4\gamma} t^4}{24\tau^3}. \tag{246}$$

Indeed the density of D_0 is $e^{-\gamma}$ on $(0, 1)$, and for $Q = D_1 + D_2 + D_3$, $\mathbb{P}(Q < x) = e^{-3\gamma} x^3/6$ on $0 < x \le 1$. Integrate $\mathbb{P}(Q > u/\tau)$ against $e^{-\gamma}\, du$.

The same convolution gives $\mathbb{P}(Q < x) \le e^{-3\gamma} x^3/6$ for every $x > 0$, since the Dickman density is bounded by $e^{-\gamma}$. Its cumulants are $\kappa_j(D) = 1/j$, so $\mathbb{E}D^3 = 17/6$. Hence

$$1 - P(\infty, \tau) \le \frac{17}{36} e^{-3\gamma} \tau^{-3}, \qquad 1 - P(\infty, \tau) \sim \frac{17}{36} e^{-3\gamma} \tau^{-3} \quad (\tau \to \infty). \tag{247}$$

The asymptotic follows by dominated convergence after conditioning on D_0; this is a second limit, after the arithmetic window limit. With a single neighbour $n + 1$ and equal thresholds, the corresponding limiting pooled reconstruction probability equals $1/2$ by symmetry of two independent continuous Dickman variables.

Proposition 26.1 (Equal thresholds)**.** *For three neighbours and equal thresholds,*

$$0.9302864610 < P(\infty, 1) < 0.9302864868.$$

Accordingly, the limiting optimal success probability for the sign in the pooled or unlimited-one-message experiment lies between 0.9651432305 *and* 0.9651432434.

Proof. Let $Z = D_0 - D_1 - D_2 - D_3$ and put

$$A(t) = \int_0^t \frac{\cos u - 1}{u}\, du, \qquad B(t) = \int_0^t \frac{\sin u}{u}\, du.$$

The characteristic function of Z is $\exp(4A(t) - 2iB(t))$. The inversion formula of Gil-Pelaez [28] at zero, justified here either by an exponential regulariser or by the Dirichlet integral, gives

$$P(\infty, 1) = \frac{1}{2} + \int_0^\infty f(t)\, dt, \qquad f(t) = \frac{e^{4A(t)} \sin(2B(t))}{\pi t} = -\frac{1}{\pi} \mathbb{E} \int_0^Z \cos(tu)\, du.$$

In particular $f(0) = 2/\pi$. Since $A(t) = \mathrm{Ci}(t) - \gamma - \log t$, $B(t) = \mathrm{Si}(t)$, integration by parts in the tails gives $|\mathrm{Ci}(t)| \le 2/t$ and $|B(t) - \pi/2| \le 2/t$. For $T > 8$ it follows that

$$\left| \int_T^\infty f(t)\, dt \right| \le \frac{4}{15 T^5 (1 - 8/T)}. \tag{248}$$

We used $e^{-4\gamma} \le 1$, $e^{8/T} \le (1-8/T)^{-1}$ and $\pi > 3$. The fifth moment of $D_0 + D_1 + D_2 + D_3$ is $14024/5$, obtained from cumulants $4/j$ and the recurrence $m_{n+1} = \sum_{i=0}^{n} \binom{n}{i} \kappa_{i+1} m_{n-i}$. Differentiation in the expectation representation of f yields

$$\sup_t |f^{(4)}(t)| \le \frac{\mathbb{E}|Z|^5}{5\pi} \le \frac{14024}{75}.$$

Composite Simpson integration on $[0, 48]$, with spacing $1/256$, therefore has error at most $1753/150994944000$. The tail bound (248) is $1/796262400$.

The accompanying standard-Python program `verification/dickman_certificate.py` evaluates the finite Simpson sum with outward-rounded integer intervals in units of 2^{-192}. At $t = k/256$, it sums the power series for A and B through degree 240 and bounds each alternating tail by $t^{240}/(240 \cdot 240!)$. For e^{4A} and $\sin(2B)$ it first checks that the arguments have modulus at most 24, then sums through degree 160 and bounds the absolute tail by $2 \cdot 24^{161}/161!$. The identity $\pi = 16 \arctan(1/5) - 4 \arctan(1/239)$ supplies an interval for π using consecutive alternating sums. Adding the two analytic errors gives the displayed rational decimal bounds. The execution record is `verification/dickman_output.txt`. Both files accompany this article in the source supplement and in the Zenodo PDF attachment `dickman-certificate.zip`. This calculation certifies the limiting probability; it supplies no rate of arithmetic convergence. □

AI assistance and verification

The author used OpenAI's GPT-6 Astra Pro to assist with derivations, computation, critical reading and preparation of the English text. The accompanying supplement contains the programs for the finite reciprocity checks and the integer-interval Dickman certificate, together with their reproduced outputs. The programs verify the stated finite identities and numerical enclosure. The author is responsible for the mathematical arguments and the final text.

References

[1] P. Erdős and M. Kac, *The Gaussian Law of Errors in the Theory of Additive Number Theoretic Functions*, American Journal of Mathematics **62** (1940), 738–742. doi:10.2307/2371483. https://combinatorica.hu/~p_erdos/1940-12.pdf.

[2] P. Billingsley, *On the distribution of large prime divisors*, Periodica Mathematica Hungarica **2** (1972), 283–289. doi:10.1007/BF02018667.

[3] E. Bach, *How to Generate Factored Random Numbers*, SIAM Journal on Computing **17** (1988), 179–193. doi:10.1137/0217012.

[4] P. Donnelly and G. Grimmett, *On the Asymptotic Distribution of Large Prime Factors*, Journal of the London Mathematical Society (2) **47** (1993), 395–404. doi:10.1112/jlms/s2-47.3.395. Theorem 1 and Corollary 2.

[5] R. Arratia, *On the Amount of Dependence in the Prime Factorization of a Uniform Random Integer*, in *Contemporary Combinatorics*, Bolyai Society Mathematical Studies **10** (2002), 29–91. Archived edition: arXiv:1305.0941v1 (manuscript updated September 2000).

[6] J. Pitman, *Combinatorial Stochastic Processes*, Lecture Notes in Mathematics **1875**, Springer, 2006. doi:10.1007/b11601500. Equations (2.19), (3.3) and Theorem 3.2.

[7] T. Haddad and D. Koukoulopoulos, *On Arratia's coupling and the Dirichlet law for the factors of a random integer*, Journal de l'École polytechnique – Mathématiques **12** (2025), 1565–1604. doi:10.5802/jep.317. Theorem 1; arXiv:2406.09360v4.

[8] R. G. Pinsky, *A Natural Probabilistic Model on the Integers and Its Relation to Dickman-Type Distributions and Buchstab's Function*, in *Probability and Analysis in Interacting Physical Systems*, Springer Proceedings in Mathematics & Statistics **283** (2019), 267–294. doi:10.1007/978-3-030-15338-0_10. Proposition 1 and Theorem 1 in arXiv:1606.02965v6, 1 February 2017; first circulated in 2016.

[9] A. Jaramillo and X. Yang, *Approximation of smooth numbers for harmonic samples: a Stein method approach*, arXiv:2409.16527v2, 3 March 2025; first circulated in 2024. Theorem 3.1, Proposition 4.1 and Section 4.3.

[10] V. Bernal Ramirez and A. Jaramillo, *Additive functionals of Harmonic samples: the conditioned Dickman regime*, arXiv:2512.02926v2, 8 June 2026. Theorem 3.1.

[11] V. Bernal Ramirez, D. Torres-Flores and A. Jaramillo, *Attenuated Poisson Dirichlet approximations for divisibility configurations*, arXiv:2606.30349, 29 June 2026.

[12] O. Gorodetsky, *A Kubilius model for sieve-theoretic sequences*, Analysis Mathematica **52** (2026), 999–1022. doi:10.1007/s10476-026-00178-w. arXiv:2608.14190v1, 14 August 2026.

[13] T. Tao, *Mertens' theorems*, expository note, 11 December 2013. Theorem 1 and Remark 1. https://terrytao.wordpress.com/2013/12/11/mertens-theorems/.

[14] J. Teräväinen, *On binary correlations of multiplicative functions*, Forum of Mathematics, Sigma **6** (2018), e10. doi:10.1017/fms.2018.10. Theorems 1.4 and 1.11, with Definition 1.1, Remark 1.8 and Section 4 in arXiv:1710.01195v2, 21 May 2018.

[15] K. Matomäki, M. Radziwiłł and T. Tao, *An averaged form of Chowla's conjecture*, Algebra & Number Theory **9** (2015), 2167–2196. doi:10.2140/ant.2015.9.2167. The classical Davenport estimate is recalled after Lemma 1.4.

[16] S. Drappeau, A. Granville and X. Shao, *Smooth-supported multiplicative functions in arithmetic progressions beyond the $x^{1/2}$-barrier*, Mathematika **63** (2017), 895–918. doi:10.1112/S0025579317000225. Theorem 2.2; arXiv:1704.04831.

[17] A. Pascadi, *Smooth numbers in arithmetic progressions to large moduli*, Compositio Mathematica **161** (2025), 1923–1974. doi:10.1112/S0010437X2500747X. Theorem 1.7, p. 1926, and Theorem 4.2, pp. 1934–1935, in the published edition; arXiv:2304.11696.

[18] B. Green, *A note on multiplicative functions on progressions to large moduli*, Proceedings of the Royal Society of Edinburgh, Section A: Mathematics **148** (2018), 63–77. doi:10.1017/S0308210517000191. Corrected edition: arXiv:1604.04481v4, 23 April 2018, Proposition 2.2 and equation (3.2).

[19] É. Fouvry and M. Radziwiłł, *Level of distribution of unbalanced convolutions*, Annales Scientifiques de l'École Normale Supérieure (4) **55** (2022), 537–568. doi:10.24033/asens.2501. Theorem 1.2 and Corollary 1.3 in arXiv:1811.08672v1, 21 November 2018.

[20] S. Bettin and V. Chandee, *Trilinear forms with Kloosterman fractions*, Advances in Mathematics **328** (2018), 1234–1262. doi:10.1016/j.aim.2018.01.026. Theorem 1, equation (1.2), in arXiv:1502.00769v1, 3 February 2015.

[21] T. Wright, *Trilinear Kloosterman fractions I: partially fixed moduli and unbalanced convolutions*, arXiv:2604.25177v2, 7 August 2026. Theorem 2.1.

[22] N. G. de Bruijn, *On the number of positive integers $\leq x$ and free of prime factors $> y$*, Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A **54** (1951), 50–60. Section 1, equations (1.1)–(1.4), p. 51; the proof of (1.4) is in Section 4. Original article.

[23] R. de la Bretèche and G. Tenenbaum, *Friable Averages of Oscillating Arithmetic Functions*, in *Number Theory in Memory of Eduard Wirsing*, Springer, 2023, 43–71. doi:10.1007/978-3-031-31617-3_5. The revised preprint is entitled *Friable averages of oscillating multiplicative functions*, arXiv:2207.04777v14, 20 July 2024.

[24] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications **53**, American Mathematical Society, Providence, RI, 2004. doi:10.1090/coll/053. Chapter 5; the fundamental lemma in Section 6.4, pp. 158–160; Sections 7.4–7.5. Authors' corrections, 8 March 2010.

[25] I. Kontoyiannis, *Some information-theoretic computations related to the distribution of prime numbers*, arXiv:0710.4076v2, 6 November 2007. Conference version: *Counting the primes using entropy*, 2008 IEEE Information Theory Workshop, doi:10.1109/ITW.2008.4578665.

[26] J. L. Carter and M. N. Wegman, *Universal Classes of Hash Functions*, Journal of Computer and System Sciences **18** (1979), 143–154. doi:10.1016/0022-0000(79)90044-8.

[27] M. E. Hellman and J. Raviv, *Probability of Error, Equivocation, and the Chernoff Bound*, IEEE Transactions on Information Theory **16** (1970), 368–372. doi:10.1109/TIT.1970.1054466. Section III, equation (18), p. 369. Author's copy.

[28] J. Gil-Pelaez, *Note on the Inversion Theorem*, Biometrika **38** (1951), 481–482. doi:10.1093/biomet/38.3-4.481.

[29] D. Slepian and J. K. Wolf, *Noiseless Coding of Correlated Information Sources*, IEEE Transactions on Information Theory **19** (1973), 471–480. doi:10.1109/TIT.1973.1055037.

[30] I. Delsol, O. Rioul, J. Béguinot, V. Rabiet and A. Souloumiac, *An Information Theoretic Condition for Perfect Reconstruction*, Entropy **26** (2024), 86. doi:10.3390/e26010086. Sections 6.3–6.4.

[31] B. Pouly, *Arithmetic reconstruction and error recovery from three quadratic traces*, 2026. doi:10.5281/zenodo.23317860. hexagon:2610.00353v1.

[32] OpenAI, *The joint Dickman law for consecutive integers*, OpenAI Math Release preprint, 24 September 2026. Repository snapshot, 6 October 2026 (commit `adc7f1241b42`).

[33] OpenAI, *Ordinary two-point correlations of multiplicative functions*, OpenAI Math Release preprint, 24 September 2026. Repository snapshot, 6 October 2026 (commit `adc7f1241b42`).

[34] P. Borwein, S. K. K. Choi and M. Coons, *Completely multiplicative functions taking values in $\{-1, 1\}$*, Transactions of the American Mathematical Society **362** (2010), 6279–6291. doi:10.1090/S0002-9947-2010-05235-3.

[35] B. Pouly, *An explicit bound below one half for a quartic Liouville correlation*, 2026. doi:10.5281/zenodo.24443184.

[36] B. Pouly, *Simultaneous quadratic probes: reconstruction thresholds and marked Poisson limits*, 2026. doi:10.5281/zenodo.24443281.