hexagon:2611.01196v1

Observations and Shift Decoding in Divisor Incidences

Brice Pouly

Research and writing with AI assistance

Abstract

Hall and Tenenbaum's theorem on local divisor maxima gives the Gaussian success profile for guessing a hidden shift when each origin is weighted by its number of divisors. We extend this profile to uniform sampling of quadruples (h, n, a, b) with $X < n \le 2X$, $1 \le h \le H$, $a \mid n$ and $b \mid n + h$, for coprime odd a, b in dyadic intervals of sizes $X^{1/4}$ and $X^{3/5}$. Given $v = n + h$, the best list of Q guesses succeeds with probability tending to $1 - \Phi(t)$ when $\log(H/Q) = (2\log 2 - 1)L + t(\log 2)\sqrt{2L} + o(\sqrt{L})$, where $L = \log\log X$, H is polylogarithmic and Φ is the standard normal distribution function. Ranking trial origins by their number ω_y of distinct prime factors up to $y = X^{1/L^4}$ attains the limit even against the best list given (v, b). On the sampled origin, the positive part of $\log H + L - (\log 2)\omega_y(n)$ approximates both posterior information and log rank with mean absolute error $O(\log(2L))$. At shift two we count incidences with both cofactors prime. Given a and $(n+2)/b$, the discrepancy for b mod q, summed over cells and averaged over primes $S < q \le 2S$, is $o(X)$ exactly when $S \log X = o(X^{7/20})$, for $X^{\varepsilon} \le S \le X^{7/20}$ and fixed $\varepsilon > 0$.

1 Which displacement produced the observed integer?

Hall and Tenenbaum's 1984 theorem on local maxima of the divisor function already determines how often one can guess a hidden displacement from a single observation in an unrestricted divisor experiment [12, Thm. 1, p. 376]. Choose uniformly a triple (h, n, d) with $1 \le h \le H$, $X < n \le 2X$ and d a positive divisor of n. Report only $v = n + h$. A proposed shift j has one possible label for each divisor of $v - j$, provided this origin belongs to the sampling interval. Thus a best guess selects the most highly represented origin in a short interval. We first make this relation exact, since it identifies the part of the transition supplied by the classical theorem.

Let $\tau(n)$ count the positive divisors of n, and set $W_X = \sum_{X<n\le 2X} \tau(n) = X\log X + O(X)$. Write $P_X^{\tau}(H)$ for the optimal success probability with one guess. Counting a largest row entry for each observation gives the exact identity

$$P_X^{\tau}(H) = \frac{1}{HW_X} \sum_{v} \max_{\substack{1\le j\le H \\ X<v-j\le 2X}} \tau(v-j), \tag{1.1}$$

where an empty maximum is zero. Hall and Tenenbaum sum precisely the corresponding unrestricted local maximum. Put $S_H(Y) = \sum_{1\le m\le Y} \max_{1\le i\le H} \tau(m+i)$. With $m = v - H - 1$, the trial origins fill $(m, m + H]$. Hence the numerator in (1.1) differs from $S_H(2X) - S_H(X)$ by $O(H\sqrt{X})$ for $H \le (\log X)^C$. Only $O(H)$ windows meet either endpoint, and $\tau(u) \le 2\sqrt{u}$.

Set $\kappa = 2\log 2 - 1$, $L_Y = \log\log Y$, and

$$a_Y = \frac{\log H - \kappa L_Y}{2(\log 2)\sqrt{L_Y}}.$$

Theorem 1 of [12] at $\alpha = 1$, applied to $F = \tau$, gives

$$\frac{S_H(Y)}{HY\log Y} = \frac{1}{\sqrt{\pi}}\int_{a_Y}^{\infty} e^{-u^2}\,du + O\Big(\frac{\log(2L_Y)}{\sqrt{L_Y}}\Big).$$

The constant multiplying the integral is 1, as also follows from the case $H = 1$ and the mean of τ. Uniformly for $1 \le H \le (\log X)^C$, the difference $a_{2X} - a_X$ is $O_C(1/(\log X\sqrt{L_X}))$. Taking the dyadic difference and writing $L = L_X$, $\sigma_L = (\log 2)\sqrt{2L}$, we obtain

$$P_X^{\tau}(H) = 1 - \Phi\Big(\frac{\log H - \kappa L}{\sigma_L}\Big) + O_C\Big(\frac{\log(2L)}{\sqrt{L}}\Big), \tag{1.2}$$

where Φ is the standard normal distribution function. Thus the one-guess threshold and its Gaussian success profile already hold for sampling from all labelled divisors.

We now restrict the first divisor, require a second divisor at the other end of the shift, and allow a list of guesses. Each successful representation carries its own mass. Even the first restriction selects a rare event. In the experiment above the probability that the *sampled label* is odd and lies in $(A, 2A]$, with $A = X^{1/4}$, is

$$\frac{\sum_{A<a\le 2A,\ 2\nmid a}(X/a + O(1))}{W_X} \sim \frac{\log 2}{2\log X}.$$

Requiring a second label $b \mid v$ and $(a, b) = 1$ further changes the weight of each origin.

Write $d \sim D$ for $D < d \le 2D$. Throughout,

$$A = X^{1/4}, \qquad B = X^{3/5}, \qquad M = X/A, \qquad K = X/B, \qquad U = X/(AB) = X^{3/20}.$$

Let $\mathcal{C}$ consist of the pairs $a \sim A, b \sim B$ with a, b odd and $(a, b) = 1$. Put

$$\mathcal{I}_{ab}(h) = \{X < n \le 2X : a \mid n,\ b \mid n + h\}.$$

The probability space in Sections 1–4 is uniform on quadruples (h, n, a, b) with $1 \le h \le H$, $(a, b) \in \mathcal{C}$, $n \in \mathcal{I}_{ab}(h)$. Write $v = n + h$, $m = n/a$, $k = v/b$. There is no coprimality condition on h. We use natural logarithms and write $x_+ = \max(x, 0)$. For $O = v$ or $O = (v, b)$, let $P_O(Q)$ be the greatest probability of including h in a list of at most Q shifts chosen from O.

For fixed data (v, b) in the support, the unnormalised posterior weight of shift j is

$$w_{v,b}(j) = \mathbf{1}_{X<v-j\le 2X} \sum_{\substack{a\sim A,\ 2\nmid a\\ a\mid v-j,\ (a,b)=1}} 1, \qquad w_v(j) = \sum_{\substack{b\sim B,\ 2\nmid b\\ b\mid v}} w_{v,b}(j). \tag{1.3}$$

The right side defining $w_{v,b}$ applies to every such divisor b, including zero rows. Dividing by the sum over $1 \le j \le H$ gives the conditional probabilities on positive rows. Ordering them in decreasing order maximises success for every list size. The rank in this order is the conditional guessing rank studied, in a general finite setting, by Arıkan [1].

Example 1.1. Take $X = 4096, H = 5, v = 4125$. The only positive rows are $b = 165, 275$. For $j = 3$ the origin $4122 = 9 \cdot 458$ has the allowed divisor 9, coprime to 275 only. For $j = 4$, the divisor 13 of $4121 = 13 \cdot 317$ occurs with either label. The posterior probabilities of $j = 3, 4$ are consequently $1/3, 2/3$ from v, $0, 1$ from $(v, 165)$, and $1/2, 1/2$ from $(v, 275)$. This is the local ambiguity whose aggregate size we estimate. A complete enumeration at $X = 16384, H = 9$ gives 6557 successful single guesses from v and 6563 from (v, b), out of 12193 incidences.

Theorem 1.2 (Shift decoding). *Suppose* $1 \le Q \le H \le (\log X)^C$ *for fixed* C. *Set*

$$L = \log\log X, \quad \kappa = 2\log 2 - 1, \quad \sigma_L = (\log 2)\sqrt{2L}, \quad y = X^{1/L^4}.$$

Let $\omega_y(u) = \#\{p \le y : p \mid u\}$. *Order* $1 \le j \le H$ *by decreasing* $\omega_y(v-j)$, *resolving ties by increasing* j. *Write* $P_{\mathrm{sc}}(Q)$ *for the success probability of the first* Q *entries. Each of* $P_{\mathrm{sc}}(Q), P_v(Q), P_{v,b}(Q)$ *has the following limits.*

1. *If* $\log H = (c + o(1))L$, $\log Q = (s + o(1))L$, *the limit is* 1 *for* $c - s < \kappa$ *and* 0 *for* $c - s > \kappa$.
2. *If* $\log(H/Q) = \kappa L + t\sigma_L + o(\sqrt{L})$, *with* $t \in \mathbb{R}$ *fixed, the limit is* $1 - \Phi(t)$, *where* Φ *is the standard normal distribution function.*

For $\log H = (c + o(1))L$ *the two conditional entropies satisfy*

$$L^{-1}\operatorname{Ent}(h \mid v), \quad L^{-1}\operatorname{Ent}(h \mid v, b) \longrightarrow (c - \kappa)_+.$$

The origin has two sources of small prime factors. One is the sampled divisor a. The other is the unrestricted part of m. The selected origin consequently carries $2L + O_{\mathbb{P}}(\sqrt{L})$ distinct small prime factors, with fluctuations that are asymptotically Gaussian. A wrong origin usually has only one source. The upper-tail estimate for a false shift has rate $I(x) = x\log x - x + 1$. At $x = 2$, the value $I(2) = \kappa$ gives the threshold and the slope $I'(2) = \log 2$ sets the fluctuation scale. In particular, when $H = \lfloor e^{\kappa L} \rfloor$, the best single guess succeeds half the time in the limit, yet the mean remaining information diverges like $(\log 2)\sqrt{L/\pi}$ (Theorem 4.4). For the converse one must control origins with many divisors in the same dyadic interval, since a single guess can collect all their posterior mass.

Comparison with weighted prime-factor laws. For ordinary integers the centring and variance of the prime-factor count are both asymptotic to L in the Erdős–Kac theorem [3]. Under the weight $\tau(n)$ both are asymptotic to $2L$, as in Elliott's result [4, 5] and the moment theorem of Khan, Milinovich and Subedi [15, Thm. 1.1 and Cor. 1.2]. Fan [7] treats a wider class of multiplicative weights by moments. Our incidence law requires the prime-factor fluctuations under both divisor conditions. In addition to the Gaussian law for the selected origin, Lemma 3.1 bounds the multiplicity of its odd dyadic divisors, and Lemma 4.1 averages the prime factors forced onto the competing origins. The losses in these two estimates are polynomial in L, small enough to retain the full $\sqrt{L}$ window.

For a second comparison, Sun [25, Thm. 1.1] obtains at $k = 2$ the increasing profile $\Phi(t)$ for the short sum of the divisor function, relative to its long mean, at almost all ordinary starting points. The scale of the window agrees with (1.2). Its exceptional set has $o(X)$ starting points, which can still account for a substantial fraction of the mass of labelled incidences. Hall's theorem on integers with a divisor in $(y, z]$ also has this transition width [11, Thm. 7.9]. Ford gives the corrected range in *The distribution of integers with a divisor in a given interval* [8, Thm. H, pp. 369–370]. There $z \le x^{1/\log\log x}$, which excludes a divisor dyad of size $x^{1/4}$ for large x. That comparison counts each integer admitting a divisor once. The posterior in (1.3) depends on the number of representations.

What the observation retains. The list theorem gives the distribution of the true rank by varying the list size (Corollary 4.3). Theorem 4.4 identifies the arithmetic quantity that approximates conditional entropy to $O_C(\log(2L))$. Corollary 4.5 controls the absolute error on the same sampled incidence for posterior information and for all three ranks. Revealing b can change the answer on an individual observation, as Example 1.1 shows. Its mean information gain is $O_C(\log(2L))$. The list form of Fano's inequality used at first order is treated in [24].

At displacement two, Theorem 5.1 counts prime $k = (n+2)/b$ when one reports the exact cofactor $m = n/a$ and one of at most $X^{1/5-\varepsilon}$ preassigned groups of labels a. An observer can therefore use the full factorisation of m. Theorem 5.2 instead retains a and the prime k exactly and reports $b \bmod q$. If $T = B/A$ and $X^{\varepsilon} \le S \le T$, the mean discrepancy over primes $S < q \le 2S$ is $o(X)$ exactly when $S \log X/T \to 0$. Its upper bound uses a translated form of Iwaniec's prime-modulus sieve (Lemma C.1). Its converse uses the minimum mass of a nonempty prime cell.

Table 1 compares the sampling measures and recovery targets in related experiments. Sections 6 and 7 then apply short-interval uniformity and the correlation estimates of Matomäki and collaborators [18, 17] and Evans [6] to observations on individual divisor cells. The displacement in the latter estimates is averaged over an interval. Cancellation of the Liouville product at a prescribed displacement remains open in these cells.

Experiment	Source and observation	Target and choice of test
Reconstruction and sign prediction from sampled prime factors [20]	Harmonic integer. Prime labels are sampled with logarithmic weights.	Recover the integer or predict a prescribed sign under the sampling protocol.
Parity information in a rough cofactor and exact reconstruction thresholds [22]	Uniform integer. Read its rough cofactor and its residue modulo a given modulus.	Optimise recovery, certification or sign prediction for that modulus.
Visible prime factors under neighbouring critical divisor weights [23]	Two divisor-weighted priors. Read a marked prime, its cofactor or successive factor labels.	Distinguish the two laws with bounded tests on the revealed data.
Simultaneous quadratic probes: reconstruction thresholds and marked Poisson limits [21]	Uniform integer. Read four Legendre symbols at each prime of a random panel.	Identify the integer or predict a binary label fixed before the panel is drawn.
Present experiment	Uniform divisor incidence. Read v, or v, b.	Recover h with a list chosen after seeing the data.

Table 1: The observation and the prior in five reconstruction experiments.

2 The prime factors of the selected origin

Let $N_{abh} = |\mathcal{I}_{ab}(h)|$ be the number of origins with these three labels. The Chinese remainder theorem gives

$$N_{abh} = \frac{X}{ab} + O(1) \asymp U, \qquad \mathcal{M} = \sum_{h \le H} \sum_{\mathcal{C}} N_{abh} \sim \frac{2(\log 2)^2}{\pi^2} HX. \tag{2.1}$$

Indeed, for two independently chosen odd dyadic integers with respective weights $1/a, 1/b$, coprimality has limiting probability $\prod_{p>2}(1-p^{-2})$. For each odd $p \le 2A$ the probability of a common multiple is $O(p^{-2})$, since the harmonic probability of a multiple of p on either dyad is $O(1/p)$. Finite sets of primes are handled by counting residues. The separate odd harmonic sums tend to $(\log 2)/2$. The same argument with h fixed proves the mass formula per shift. Replacing every N_{abh} by $X/(ab)$ changes the joint law of (h, a, b) by $O(U^{-1})$ in total variation. In particular h has mass $O(1/H)$ at each value.

Write $n = am$. For a fixed triple let

$$W_{ab}(m) = \sum_{\substack{p \le y \\ p \nmid ab}} \mathbf{1}_{p \mid m}.$$

This separation assigns all small divisors of a to the first summand and avoids repeated counting.

Lemma 2.1 (The two Gaussian contributions)**.** *Under the incidence law,*

$$\left(\frac{\omega_y(a)-L}{\sqrt{L}}, \frac{W_{ab}(m)-L}{\sqrt{L}}\right) \Longrightarrow (Z_1, Z_2),$$

where Z_1, Z_2 are independent standard normal variables. Furthermore

$$Z_X := \frac{\omega_y(n)-2L}{\sqrt{2L}} \Longrightarrow N(0,1), \qquad \Omega(n)-\omega_y(n) = O_{\mathbb{P}}(\log(2L)). \tag{2.2}$$

The same loss satisfies the mean estimate

$$0 \le \mathbb{E}[\Omega(n)-\omega_y(n)] \ll_C \log(2L). \tag{2.3}$$

Here Ω counts prime factors with multiplicity and $O_{\mathbb{P}}$ denotes boundedness in probability after division by the indicated scale.

Proof. The sum of the reciprocal primes up to y is $L - 4\log L + O(1)$. If $d \ll X^2$, deleting its prime factors changes this sum by $O(\log(2L))$, uniformly in d. The contribution up to $(\log X)^2$ has this bound by Mertens' theorem. The larger prime divisors of d contribute $O((\log X)^{-2}\log(2d))$.

We first give the moment comparison used twice below. In a progression of step b with $N \asymp U$ terms, the probability that a product d of distinct primes coprime to b divides m is $1/d + O(1/N)$. In the r-th raw moment of a sum of such indicators, at most y^r ordered tuples occur. Repeated indices give the same indicator. Comparison with independent Bernoulli variables of means $1/p$ therefore costs $O_r(y^r/N)$. In a centred moment it costs at most $O_r((1+L)^r y^r/N) = o(1)$ for each fixed r. The independent sum has variance equal to its mean up to $O(1)$, and its cumulants of every fixed positive order are $O_r(L)$, by differentiating $\sum_p \log(1+(e^t-1)/p)$ at zero. Its centred moments divided by the corresponding powers of $\sqrt{L}$ tend to the Gaussian moments. Replacing its mean by L costs $O(\log(2L)/\sqrt{L})$ on this scale.

For W_{ab}, given h, a, b, the m's occupy consecutive terms in one class modulo b. The comparison just proved and the uniform deletion bound apply to every triple, even when the original progression has a common factor with its modulus. By moment determinacy the conditional distribution functions converge to Φ uniformly over triples at each fixed argument. Otherwise a sequence of exceptional triples would have the same limiting moments and a different weak limit.

For the divisor label there is a direct comparison that also identifies the change at each prime caused by coprimality. Start with independent odd harmonic labels and restrict to $(a,b)=1$. For any odd squarefree d composed of at most r primes $\le y$,

$$\mathbb{P}(d \text{ divides } a \,|\, (a,b)=1) = \prod_{p\mid d} \frac{1}{p+1} + O_r(X^{-1/25}). \tag{2.4}$$

To verify the formula expand the coprimality indicator as $\sum_{\ell \mid a,b} \mu(\ell)$, in which all ℓ are odd. The tail $\ell > D_0 = X^{1/20}$ is $O(1/D_0)$ in absolute mean, by the dyadic multiple bound for a, b. For $\ell \le D_0$ the two divisibilities are by $[d,\ell]$ and ℓ. On an odd harmonic dyad of size A the probability of a multiple of an odd e is $1/e + O(1/A)$, by partial summation, and likewise for B. Summing the two endpoint errors over $\ell \le D_0$ and extending the resulting absolutely convergent series gives an error $O(D_0/A + 1/D_0)$, uniformly in d. At a prime dividing d the unconditioned joint density for the two events is $p^{-1}(1-p^{-1})$. Dividing by the coprimality

density $1-p^{-2}$ gives $1/(p+1)$. All other local factors cancel. The same computation with $d=1$ normalises the probability. By (2.1), restoring the exact incidence labels costs $O(1/U)$ per event.

Since $y^r = X^{o(1)}$ for fixed r, (2.4) gives the moments of the independent sum with probabilities $1/(p+1)$ for odd $p \le y$, up to $o(1)$ even after centring. Its mean and variance differ by $O(1)$ from those with probabilities $1/p$. This is the sieve moment method of [10], with the local sampling probabilities computed above. Multiplying the conditional distribution function of W_{ab} by the indicator of any first-coordinate half-line now gives the product of the two limiting distribution functions.

The difference between $\omega_y(n)$ and $\omega_y(a) + W_{ab}(m)$ counts only prime factors of b that divide n, hence divide h. The expectation of $\Omega(h)$ is $O(1+\log\log(3H)) = O(\log(2L))$, by summing the probabilities of divisibility by prime powers and using (2.1). This proves the central limit assertion.

For the remaining estimate put $d=(b,h)$. Then $m'=m/d$ runs through a primitive class modulo b/d, with the same N terms. Indeed $(m,b)=(h,b)$. Put $Z=X^{1/30}$. For m', the expected number of repetitions from primes $p \le Z$ is bounded by

$$\sum_{p\le Z} \sum_{\substack{j\ge 2\\ p^j\le 3X}} (p^{-j} + O(1/N)) \ll 1 + Z\log X/N.$$

Primes dividing the step have zero contribution. The distinct primes with $y < p \le Z$ have total mean $O(\log(2L) + Z/N)$, and at most $O(1)$ prime factors with multiplicity can exceed Z. For a the analogous estimates hold with denominator A in the errors, by domination of its marginal by the uniform dyadic law up to a bounded factor. For each fixed label a, the conditional mean overlap between the small prime divisors of a and m' is at most $\sum_{p\mid a,\, p\le y}(1/p + O(1/N)) \ll \log(2L)$. Finally $\Omega(d) \le \Omega(h)$ has the same mean bound. Adding the nonnegative losses gives the claimed estimate in expectation and, by Markov, in probability. □

3 How much posterior mass can one origin carry?

Let $d_D(u) = \#\{d \sim D : d \mid u\}$ and $d_D^o(u) = \#\{d \sim D : d \mid u,\ 2 \nmid d\}$. The upper bound in (1.3) is $d_A^o(v-j)$ when b is known. The distinction between counting origins and counting incidences makes a second moment of d_A appear. The completely multiplicative weight $z(u) = 2^{-\Omega(u)}$ compensates for an unusually large factor count.

Lemma 3.1 (A weighted local divisor moment)**.** *For $1 \le h \le (\log X)^C$, with τ the divisor function, there is a constant c_1 such that*

$$\sum_{X<n\le 2X} d_A(n)^2 d_B(n+h) z(n) \ll_C X e^{-L}(1+L)^{c_1}\tau(h).$$

Under the incidence law the positive part of $\log d_A^o(n) - (\log 2)\omega_y(n) + L$ is $O_{\mathbb{P}}(\log(2L))$.

Proof. Factor two ordered dyadic divisors of n as gu, gw, where g is their greatest common divisor. Thus $n = guwr$, $(u,w)=1$, and $r \sim R = X/(guw) \gg X^{1/2}$. Switch the divisor of $n+h$ to its cofactor $k=(n+h)/b \asymp K$. For $e=(guw,k)$ there is no solution unless $e \mid h$. Otherwise r lies in one residue class modulo $q = k/e$. We can enlarge the r-range to the full dyad for an upper bound. Its length measured in progression terms satisfies $N_r = R/q + O(1) \gg X^{1/10}$.

Independently mark each $p \le y,\ p \nmid q$ with probability $1/2$. The probability that none of the marked primes divides r majorises $z(r)$. Upper Bonferroni, stopped at the least even integer $j_0 \ge 10L$, compares its progression average with the corresponding independent

prime model. Each squarefree product has divisibility discrepancy $O(1/N_r)$, for a total error $O((2y)^{j_0+1}/N_r) = X^{-1/10+o(1)}$ as an upper bound. The independent truncation differs from the full product by at most $(L/2+O(1))^{j_0+1}/(j_0+1)!$, which is $O(e^{-2L})$. Mertens' theorem and the deletion bound in Lemma 2.1 give

$$\prod_{\substack{p\le y\\ p\nmid q}}(1-1/(2p)) \ll e^{-L/2}(1+L)^{c_2}.$$

This proves the same estimate for the mean of $z(r)$, without a primitivity assumption on the residue of r.

The sum of $1/q = e/k$ over eligible $k \asymp K$ is $O(\tau(h))$, by summing the harmonic weights of multiples of each $e \mid h$. The remaining reciprocal weights have numerator $z(g)z(u)z(w)$. On an ordinary dyad the preceding Bonferroni argument with $q = 1$ and X replaced by Y gives

$$\sum_{t\sim Y}\frac{z(t)}{t} \ll \frac{(1+\log\log(3Y))^{c_3}}{\sqrt{\log(2Y)}} \qquad (Y \ge 1).$$

Partition $1 \le g \le 2A$ dyadically, including $g = 1$ separately. For $g \asymp G$, both $u, w \asymp A/G$, and the bounded endpoint ranges cost a constant. Dropping $(u, w) = 1$ increases this positive sum. After extracting a power of $1+L$, the remaining dyadic sum is bounded, with $J \asymp \log A$, by

$$\sum_{i=0}^{J}\frac{1}{\sqrt{1+i}\,(1+J-i)} \ll \frac{\log(2+J)}{\sqrt{1+J}}.$$

Split at $J/2$ to see the bound. This supplies the second factor $e^{-L/2}$, up to powers of $1+L$. Multiplying by X from $R = X/(guw)$ proves the first assertion.

Returning to incidences, one copy of d_A accounts for all possible a's. Dropping oddness and coprimality only enlarges the positive numerator, so that

$$\mathbb{E}[d_A^o(n)z(n)] \ll e^{-L}(1+L)^{c_4},$$

since $H^{-1}\sum_{h\le H}\tau(h) \le 1+\log H$. Markov's inequality bounds the positive part of $\log d_A^o(n) - (\log 2)\Omega(n) + L$ by $O_{\mathbb{P}}(\log(2L))$. Lemma 2.1 replaces $\Omega(n)$ by $\omega_y(n)$ at the same cost. □

The factor $(1+L)^{c_1}$ above matters. An estimate with a loss $e^{o(L)}$ would locate the first-order threshold, while allowing errors larger than the width $\sqrt{L}$ in its logarithm.

4 False origins, optimal lists and conditional information

Lemma 4.1 (Competitors in one progression)**.** *Condition on any triple h, a, b. For fixed $x > 1$, uniformly for $1 \le j \le H,\ j \ne h$,*

$$\mathbb{P}(\omega_y(v-j) \ge xL \mid h, a, b) \le \exp\{-[I(x)+o(1)]L\}, \quad I(x) = x\log x - x + 1.$$

More precisely, for fixed $u \in \mathbb{R}$, the number N_u of false shifts with score at least $T_u = \lceil 2L + u\sqrt{2L}\rceil$ satisfies

$$\mathbb{E}(N_u \mid h, a, b) \ll_{u,C} H(1+L)^{c_5}\exp(-\kappa L - u\sigma_L). \tag{4.1}$$

Proof. A prime factor of a dividing $v-j$ divides $h-j$, and one from b divides j. Thus at most $F_j = \omega(j|h-j|)$ prime factors of ab can contribute. Uniformly $F_j = O_C(L/\log(2L))$, since the product of the first d primes is at least $d!$. For the remaining small primes let W_j count divisibility of $v-j$. In the selected cell, n runs in a progression of step ab. For each

product of ℓ distinct primes from this remaining set there is exactly one rank class modulo the product. Therefore

$$\mathbb{E}\left[\binom{W_j}{\ell}\,\middle|\,h,a,b\right] \le (L+O(1))^{\ell}/\ell! + O(y^{\ell}/U).$$

For $\ell = \lfloor (x-1)L \rfloor$, divide by $\binom{\lceil xL \rceil - F_j}{\ell}$. Stirling's formula gives $-I(x)L + o(L)$ for the log of the leading ratio. The additive error is $X^{-3/20+o(1)}$.

For T_u, use $\ell = \lfloor L \rfloor$. With $F_j = 0$ the log of the ratio is

$$\ell \log(L+O(1)) + \log((T_u-\ell)!) - \log(T_u!) = -\kappa L - u\sigma_L + O_u(\log(2L)).$$

This follows by Taylor expansion at $T_u = 2L + O(1)$, where the linear coefficient in T_u is $-\log 2 + O(1/L)$. Lowering the threshold by one multiplies the leading ratio by $r/(r-\ell)$ at threshold r, bounded throughout the F_j steps. The aggregate loss is polynomial in L. In fact Cauchy–Schwarz gives

$$\sum_{j \ne h} \exp(c_6 F_j) \ll H(1+\log H)^{c_7}.$$

To see this, factor the upper bound into terms at j and $|h-j|$. For fixed integer $d \ge e^{2c_6}$, each squared term is bounded by the ordered d-fold divisor function τ_d, whose sum up to H is $O_d(H(1+\log H)^{d-1})$ by repeated harmonic summation. This averaging step is needed at Gaussian precision because the uniform bound for F_j alone can exceed $\sqrt{L}$. Summing the negligible progression errors proves (4.1). □

Proof of Theorem 1.2. Always $P_{\mathrm{sc}}(Q) \le P_v(Q) \le P_{v,b}(Q)$. If the true score is at least T and fewer than Q false shifts score at least T, the score order includes the truth. For $c - s < \kappa$ choose $1 < x < 2$ with $I(x) > c - s$. The true score exceeds xL with probability $1 - o(1)$ by Lemma 2.1. Lemma 4.1 and Markov's inequality control the false shifts. In the central window use $T_{t+\delta}$, $\delta > 0$. Dividing (4.1) by Q gives $o(1)$, while the true score passes with limiting probability $1 - \Phi(t+\delta)$.

For the reverse inequality there are only $O(X)$ possible outputs (v,b), since $b \mid v$ and $\sum_{b \sim B}((X+H)/b+1) \ll X$. On incidences with $d_A^o(n) \le D$, a list can account for at most QD incidences per output, by (1.3). Consequently, for every $D > 0$,

$$P_{v,b}(Q) \le \mathbb{P}(d_A^o(n) > D) + O(QD/H). \tag{4.2}$$

For $c - s > \kappa$, take $\log D = (\kappa + \eta)L$ with $0 < \eta < c - s - \kappa$. Lemmas 2.1 and 3.1 give a vanishing tail. In the central window take $D = (H/Q)e^{-\delta\sigma_L}$. The second term vanishes and the same two lemmas give a tail of limsup at most $1 - \Phi(t-\delta)$. Taking $\delta \downarrow 0$ in both directions proves the theorem. □

Corollary 4.2 (Entropy and a sign readout)**.** *If $\log H = (c + o(1))L$, then for both $O = v, (v,b)$,*

$$\frac{\mathrm{Ent}(h \mid O)}{L} \longrightarrow (c-\kappa)_+.$$

For $c < \kappa$ a single sign function $g_X(v)$ satisfies $\mathbb{E}[\lambda(n) g_X(v)] = 1 - o(1)$, where $\lambda(n) = (-1)^{\Omega(n)}$.

Proof. For a list of success probability p, splitting on success bounds the entropy by $\log 2 + p \log Q + (1-p)\log H$. Conversely, outside the Q largest probabilities every posterior mass is at most $1/Q$, giving the lower bound $(1 - P_O(Q))\log Q$. Apply Theorem 1.2 at exponents just above and just below $(c-\kappa)_+$, using nonnegativity at zero. For $c = 0$, $\log H = o(L)$ suffices. Taking $\widehat{h}$ to be the first score-ranked shift gives the sign $g_X(v) = \lambda(v - \widehat{h})$. Its product with the true sign is $+1$ whenever the guess is correct. □

This readout is an arbitrary function of the endpoint. There is no multiplicativity requirement on it. An observer who fails to identify h may still identify the sign shared by several plausible origins.

Corollary 4.3 (Rank, information and the required list size)**.** *Let R be the rank of the true h in the score order or in either posterior order, with a fixed tie rule. Put $d_X = (\log H - \kappa L)/\sigma_L$.*

1. *If $d_X \to +\infty$, then $(\log H - \kappa L - \log R)/\sigma_L \Rightarrow N(0,1)$. For fixed $0 < p < 1$, the least list length Q_p giving success at least p, in any of these three orders, satisfies*

$$\log Q_p = \log H - \kappa L + \Phi^{-1}(p)\sigma_L + o(\sqrt{L}).$$

2. *If $d_X \to d \in \mathbb{R}$, then $\log R/\sigma_L \Rightarrow (d - Z)_+$ for a standard normal Z, and $\mathbb{P}(R = 1) \to 1 - \Phi(d)$.*

3. *For either observation O, both distributional assertions hold with $\log R$ replaced by the realised conditional information $-\log \mathbb{P}(h \mid O)$, with the posterior mass evaluated at the sampled shift. The assertion about $\mathbb{P}(R = 1)$ concerns ranks.*

Proof. An event $R \le Q$ is precisely success of the first Q guesses. In the first regime take $Q = \lfloor H \exp(-\kappa L - t\sigma_L) \rfloor$. For each fixed t this tends to infinity, so rounding is negligible. Theorem 1.2 gives the first distribution, and quantiles follow by bracketing p using the strict increase of Φ. In the second regime use $Q = \lfloor e^{u\sigma_L} \rfloor$, $u > 0$. Then $\mathbb{P}(R \le Q) \to \Phi(u - d)$, the required distribution at every positive continuity point. The choice $Q = 1$ gives the mass at rank one separately.

For a posterior row with sorted positive probabilities $p_1 \ge p_2 \ge \cdots$, $rp_r \le 1$. The mass of ranks for which $-\log(rp_r) > z$ is at most $e^{-z} \sum_{r \le H} 1/r$. Thus the difference between information and posterior log rank is nonnegative and $O_{\mathbb{P}}(\log(2L))$, uniformly over the rows. This is $o_{\mathbb{P}}(\sigma_L)$. □

4.1 The size of the uncertainty in the critical window

The distribution of the rank describes lists of fixed success probability. To average posterior information one also needs to control rare rows. We control these rows by comparing their posterior probabilities with one explicit probability distribution. A positive-part integral already occurs in Hayashi's lower bound for the divergence of a compressed distribution from uniformity [13, Thm. 8, §VI-B, (41)]. Here the quantity to evaluate is the conditional entropy of an arithmetic shift, and the estimate before the Gaussian limit is (4.3). Write $\phi(t) = (2\pi)^{-1/2} e^{-t^2/2}$.

Theorem 4.4 (Conditional entropy at Gaussian precision)**.** *Uniformly for $1 \le H \le (\log X)^C$ and for either observation $O = v, (v, b)$,*

$$\mathrm{Ent}(h \mid O) = \mathbb{E}[(\log H + L - (\log 2)\omega_y(n))_+] + O_C(\log(2L)), \tag{4.3}$$

$$= \sigma_L\{d_X \Phi(d_X) + \phi(d_X)\} + o_C(\sqrt{L}), \quad d_X = \frac{\log H - \kappa L}{\sigma_L}. \tag{4.4}$$

In particular if $\log H = \kappa L + o(\sqrt{L})$, the entropy is asymptotic to $(\log 2)\sqrt{L/\pi}$. For every H in the stated range the entropy saving from revealing b in addition to v is $O_C(\log(2L))$.

Proof. Set $D_* = \lceil 3L \rceil$, $s_j = 2^{\min(\omega_y(v-j), D_*)}$, and $B_h = \sum_{j \ne h} s_j$. Two moment bounds will be used. First,

$$\mathbb{E}[(\omega_y(n) - 2L)^2] \ll_C L, \qquad \mathbb{E}[\Omega(n) - \omega_y(n)] \ll_C \log(2L). \tag{4.5}$$

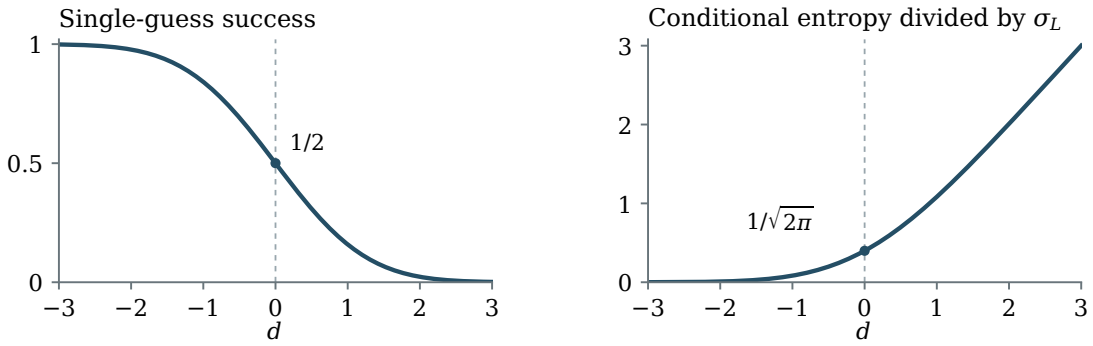

Figure 1: At $\log H = \kappa L + d\sigma_L + o(\sqrt{L})$, the limiting probability of a correct single guess is $1 - \Phi(d)$. The limit of the entropy divided by σ_L is $d\Phi(d) + \phi(d)$. The dots mark the nominal threshold.

For the first, the ordinary dyadic law for a has centred second moment $O(L + (\log(2L))^2)$ by the two-prime count in Lemma 2.1. The label marginal is dominated by a constant multiple of that law. The same centred bound holds conditionally for W_{ab}. The remaining $\omega(h)$ has second moment $O((1 + \log\log(3H))^2)$, by expanding its square and counting multiples up to H with $\mathbb{P}(h = j) \ll 1/H$. The second estimate is (2.3). In particular $\mathbb{E}[(\omega_y(n) - D_*)_+] \ll 1$.

Second, $\mathbb{E}B_h \ll_C He^L(1 + L)^{c_8}$. For a false j, use F_j, W_j of Lemma 4.1 and the deterministic bound

$$s_j \le 2^{F_j} 2^{\min(W_j, D_*)} \le 2^{F_j} \sum_{\ell=0}^{D_*} \binom{W_j}{\ell}.$$

The second inequality holds also when $W_j > D_*$, by comparing the binomial coefficients term by term with those for D_*. On each progression the expected sum on the right, without 2^{F_j}, is at most $\exp(\sum_{p \le y} 1/p) + O((2y)^{D_*+1}/U) \ll e^L$. The aggregate bound on 2^{F_j} in Lemma 4.1 gives the claim. Truncation at $3L$ allows every product in this comparison to be counted within the available progression length.

Use $\pi_j = s_j / \sum_i s_i$ as a candidate conditional distribution given v. Nonnegativity of relative entropy, row by row, gives

$$\mathrm{Ent}(h \mid v) \le \mathbb{E}[-\log \pi_h] = \mathbb{E}\log(1 + B_h/s_h).$$

Pointwise the last logarithm is bounded by $\log 2 + (\log H + L - \log s_h)_+ + \log(1 + B_h/(He^L))$. By Jensen the mean of the last term is $O_C(\log(2L))$. Unclipping $\log s_h$ costs $O(1)$ in expectation by (4.5). This proves the upper bound in (4.3) for both observations.

For the lower bound use the finer observation and let $N(v, b) = \sum_j w_{v,b}(j)$ be the row mass. At most $O(X)$ rows can occur, so

$$\mathbb{P}\{N(v, b) < He^{-u}\} \ll e^{-u} \qquad (u \ge 0)$$

by counting incidences and (2.1). Hence $\mathbb{E}[(\log H - \log N(v, b))_+] = O(1)$. The realised conditional information equals $\log N(v, b) - \log w_{v,b}(h)$ and is nonnegative. From the moment in Lemma 3.1 and Jensen,

$$\mathbb{E}[(\log d_A^o(n) - (\log 2)\Omega(n) + L)_+] \le \log\{1 + e^L \mathbb{E}[d_A^o(n) 2^{-\Omega(n)}]\} \ll_C \log(2L).$$

By (4.5) this also holds with $\omega_y(n)$ in place of $\Omega(n)$. Using $w_{v,b}(h) \le d_A^o(n)$ and $|x_+ - y_+| \le |x - y|$ now gives

$$\mathbb{E}[(\log H + L - (\log 2)\omega_y(n))_+] \le \mathrm{Ent}(h \mid v, b) + O_C(\log(2L)).$$

Conditioning can only decrease entropy, proving (4.3) and the bound on the gain from b.

The expectation in (4.3) is $\sigma_L \mathbb{E}[(d_X - Z_X)_+]$. By (2.2) and (4.5) the Z_X converge to a standard normal Z and are uniformly integrable. For $d_X \to d$ finite this gives convergence of the positive-part averages. If $d_X \to +\infty$, subtract d_X from both expectations and use $\mathbb{E}Z_X \to 0$ and uniform integrability of the positive tails. If $d_X \to -\infty$, both expectations tend to zero by the negative tails. Applying this argument to subsequences gives $\mathbb{E}[(d_X - Z_X)_+] - \mathbb{E}[(d_X - Z)_+] = o_C(1)$ throughout the stated range. Integration of the normal density gives $\mathbb{E}[(d - Z)_+] = d\Phi(d) + \phi(d)$, proving (4.4). □

The same comparison can be made before averaging over the sample. In particular the rankings have a common arithmetic approximation even on the joint probability space where the two posterior orders may disagree.

Corollary 4.5 (Information and rank on the sampled incidence)**.** *Put* $Y_X = (\log H + L - (\log 2)\omega_y(n))_+$ *and* $I_O = -\log \mathbb{P}(h \mid O)$*, evaluated at the sampled shift. For either* $O = v, (v, b)$ *and for each of the three ranks* R *in Corollary 4.3,*

$$\mathbb{E}|I_O - Y_X| + \mathbb{E}|\log R - Y_X| \ll_C \log(2L) \qquad (1 \le H \le (\log X)^C).$$

Proof. Retain the distribution π and the clipped weights from the proof of Theorem 4.4, and put $J_h = -\log \pi_h$. The pointwise upper bound for J_h there gives

$$(J_h - Y_X)_+ \le \log 2 + (\log 2)(\omega_y(n) - D_*)_+ + \log(1 + B_h/(He^L)).$$

Its expectation is $O_C(\log(2L))$. On a row of either observation let $p_j = \mathbb{P}(h = j \mid O)$. The vector π depends only on v, so on this row

$$\mathbb{E}[\exp(I_O - J_h) \mid O] = \sum_{j : p_j > 0} \pi_j \le 1.$$

Integration of the exponential tail gives $\mathbb{E}[(I_O - J_h)_+] \le 1$. In a posterior order, $Rp_h \le 1$. In the score order, every preceding entry has clipped weight at least s_h, so $R\pi_h \le 1$. Consequently each W among the two informations and the three log ranks satisfies $\mathbb{E}[(W - Y_X)_+] \ll_C \log(2L)$.

For the lower bound on the mean of a log rank, condition on the data that specify its order. The shift and its rank are then in bijection. Comparing the conditional law of the rank with the distribution $1/(r \sum_{u \le H} 1/u)$ gives

$$\mathbb{E} \log R \ge \mathrm{Ent}(h \mid O) - \log \sum_{u \le H} 1/u.$$

The last logarithm is $O_C(\log(2L))$. Theorem 4.4 therefore gives $\mathbb{E}W \ge \mathbb{E}Y_X - O_C(\log(2L))$ for each of the five variables. The identity $|z| = 2z_+ - z$ completes the proof. □

5 Prime cofactors at displacement two

We now fix $h = 2$. A label a, a residue of b, and a cofactor give different partitions of the solutions of $am + 2 = bk$. We compare the observed and predicted masses in the sum of absolute differences over those partitions. This bounds the error against any complex test of modulus at most one on the reported data. Put $\vartheta(p) = \log p$ at primes, $\vartheta(u) = 0$ otherwise, and write $\mathcal{L} = \log X$.

Theorem 5.1 (An exact cofactor)**.** *Partition the odd $a \sim A$ into $V \le X^{1/5-\varepsilon}$ sets $\mathcal{A}_j$, chosen before m is known. Put $D_j(m) = \#\{a \in \mathcal{A}_j : X < am \le 2X\}$ and let $C_j(m)$ sum $\vartheta(k)$ over incidences with this j, m. Uniformly in the partition and for each $D > 0$,*

$$\sum_{j,m} \left| C_j(m) - \mathbf{1}_{2\nmid m}(\log 2) D_j(m) \right| \ll_{\varepsilon,D} X\mathcal{L}^{-D}.$$

In particular a reader may use the full factorisation of m. Reporting a exactly would require $\asymp X^{1/4}$ groups.

Proof. In a contributing incidence a, b are odd and the prime $k \asymp K$ exceeds two, forcing m to be odd. Conversely, for odd a, m and a prime $k \mid am + 2$, the integer $b = (am + 2)/k$ is odd. A common factor of a, b would divide two, so the pair is coprime. In a rectangular box of a, m, k write $S_z(\alpha) = \sum_m z_m e(m\alpha)$, where $e(t) = e^{2\pi it}$. Subtract the density $1/k$ from the condition $m \equiv -2\bar{a} \bmod k$. For a set $\mathcal{A}$ of labels the absolute error, summed with prime weights, is bounded by

$$\sum_{k \text{ prime}} \frac{\log k}{k} \sum_{r=1}^{k-1} |S_z(r/k)| \left| \sum_{a \in \mathcal{A}} e(2r\bar{a}/k) \right| \ll \mathcal{L}\sqrt{|\mathcal{A}|(M + K^2)}\, \|z\|_2 . \tag{5.1}$$

Indeed $k > 2A$ and $a \mapsto -2\bar{a}$ is injective. Parseval gives $k|\mathcal{A}|$ as an upper bound for the squared label sum over r. Cauchy–Schwarz, $k \asymp K$, and the additive large sieve [19] with separation $\gg K^{-2}$ give (5.1). By duality choose z_m separately in each group, supported on odd m, $|z_m| \le 1$. Cauchy over groups gives $O(X^{9/10}\sqrt{V}\,\mathcal{L})$.

To impose the two dyadic conditions, divide each coordinate range into $G = \lceil \mathcal{L}^{D+4} \rceil$ comparable subintervals. On boxes meeting a boundary $am = X, 2X$ or $k = (am+2)/(B), (am+2)/(2B)$, the reference mass is $O(X\mathcal{L}/G + X^{1-c})$ for an absolute $c > 0$. For fixed m the first two boundaries allow $O(A/G + 1)$ nearby labels. For fixed a, m the last two allow $O(K/G + 1)$ nearby values of k, each with harmonic weight $O(\mathcal{L}/K)$. The actual boundary mass satisfies the same bound with $O(G^3 X^{9/10}\sqrt{V}\,\mathcal{L})$ added, by positivity and (5.1). All these errors are admissible. For each admissible a, m, the density summed over the remaining k's is $\sum_{(am+2)/(2B) \le k < (am+2)/B} \vartheta(k)/k = \log 2 + O_D(\mathcal{L}^{-D-1})$, uniformly by the prime number theorem and partial summation. □

What a residue of the other divisor reveals. Here a, k are given exactly, k is prime, and a prime q reveals $r = b \bmod q$. Put $T = B/A = M/K = X^{7/20}$ and

$$I_{ak} = \{u \in \mathbb{R} : X < au \le 2X,\ B < (au + 2)/k \le 2B\}.$$

For $q \nmid a$, let P_{akr} be the sum of $\vartheta(m)\vartheta(k)$ in the row with $b \equiv r \bmod q$. The reference and the discrepancy are

$$J_{akr} = \frac{\vartheta(k)|I_{ak}|}{(k-1)(q-1)} \mathbf{1}_{kr \not\equiv 2 \bmod q}, \qquad D(q) = \sum_{\substack{a,k,r \\ q \nmid a}} |P_{akr} - J_{akr}|.$$

The missing class forces $q \mid m$. For large X, each retained prime $k \asymp K$ exceeds $2T$, as does every possible prime $m \asymp M$, so q is different from both. Write $\pi_S = \pi(2S) - \pi(S)$, and use a prime on a sum over q to mean prime moduli.

Theorem 5.2 (Resolution of a modular observation)**.** *For $X^{\varepsilon} \le S \le T$,*

$$\frac{1}{X\pi_S} \sum_{q \sim S}{}' D(q) \to 0 \quad \Longleftrightarrow \quad S \log X/T \to 0.$$

For $T^{2/3} \le S \le T$, the normalised mean is $O_D(\sqrt{S\log X/T} + \mathcal{L}^{-D})$. If $S\log X/T \to \infty$, then $D(q)/(\frac{1}{2}(\log 2)^2 X) \to 2$ uniformly over the prime moduli in the dyad. If P_q, J_q are the arrays of masses and $|P_q|, |J_q|$ their respective totals, then, in this last range, uniformly in q,

$$d_{\mathrm{TV}}\big(P_q/|P_q|,\, J_q/|J_q|\big) \longrightarrow 1, \qquad d_{\mathrm{TV}}(\mu,\nu) = \tfrac{1}{2}\sum_x |\mu(x) - \nu(x)|.$$

Proof. Parameterise the solutions by $m = m_0 + kt,\ b = b_0 + at$ and centre the first admissible rank at 1. Pad $c_{ak}(t) = \vartheta(m_0 + kt)$ with zeros up to $N = \lceil C_0 T\rceil$, for an absolute C_0. Only primes $K/3 < k < 4K$ contribute. We first record the row estimate

$$\sum_{a,k} \vartheta(k) \left| \sum_t c_{ak}(t) - \frac{|I_{ak}|}{k-1} \right| \ll_D X\mathcal{L}^{-D}. \tag{5.2}$$

For a fixed interval I in m, subtract $k^{-1}\sum_{m\in I}\vartheta(m)$ from the progression sum. Parseval across the distinct classes $-2\bar{a}$ mod k bounds the sum of squared errors over a by $k^{-1}\sum_{r=1}^{k-1}|S_I(r/k)|^2$, with S_I formed from ϑ. The large sieve over k, then Cauchy over $O(AK)$ rows, give $O(\mathcal{L}^2\sqrt{A(M+K^2)M})$ including $\vartheta(k)$. Decompose a variable interval using at most two binary subintervals per level of $[M/2, 2M]$. At each level their disjoint supports give a total squared coefficient norm $O(M\mathcal{L}^2)$. Minkowski costs $O(\mathcal{L})$ further. This is $O(X^{9/10}\mathcal{L}^3)$. The prime number theorem for differences of prefixes now proves (5.2). Replacing k by $k-1$ in the denominator costs $O(X/K)$.

By integration in u, the row reference sums to

$$\sum_{a,k} \frac{\vartheta(k)}{k-1}|I_{ak}| \sim (\log 2) \sum_{a\sim A,\ 2\nmid a} X/a \sim \tfrac{1}{2}(\log 2)^2 X.$$

Indeed for each a, u, k traverses a full dyad. Excluding $q \mid a$ costs $O(X\mathcal{L}^2(1/q + 1/A))$ by the trivial row length bound. Thus $\sum_{a,k}\vartheta(k)\sum_t c_{ak}(t)^2 \ll X\mathcal{L}$, while $\sum_{a,k}\vartheta(k) \ll AK = X/T$.

For $q \nmid a$ the rank permutes the residues of b. If V_{akq} denotes the sum of absolute deviations of the rank-class masses from $q^{-1}\sum_t c_{ak}(t)$, then

$$V_{akq}^2 \le \sum_{j=1}^{q-1} \left| \sum_t c_{ak}(t) e(jt/q) \right|^2$$

by Cauchy and Parseval. Switching the uniform row mean to $J_{akr}/\vartheta(k)$ costs $O_D(X\mathcal{L}^{-D} + X/S)$ after weighting, by (5.2) and the one excluded residue. Cauchy over a, k, q and the classical large sieve give the normalised error

$$O\Big(\sqrt{(1/S + S/T)\mathcal{L}\log S}\Big) + O_D(\mathcal{L}^{-D}).$$

This suffices up to $T^{2/3}$.

For $T^{2/3} \le S \le T$ we have $S \le N \le S^2(\log S)^{-8}$ for large X. Use Lemma C.1 with $D_0 = S/\sqrt{N}$. Choose $z = \exp((\log\log S)^5)$, $P = \prod_{p\le z} p$. For large X, $z < k, D_0$. The single forbidden root $t \equiv -m_0\bar{k}$ mod p for each $p \le z$ can be represented by one integer c, by the Chinese remainder theorem. All supported m are primes larger than z. Mertens gives $\Delta = O(1)$ and $\log D_0 \asymp \log S$. Substituting $S^2/\log S$ for $N + S^2$ in the last Cauchy bound proves the claimed rate.

For necessity suppose $S\mathcal{L} \ge c_0 T$ along a subsequence. Choose a small constant $\delta > 0$ depending on c_0, and restrict to $(1/2+\delta)K < k < (1/2+2\delta)K$. Here $|I_{ak}| \asymp \delta M$, since the upper endpoint in au is $2Bk - 2$, just above X. Every nonzero predicted mass divided by $\vartheta(k)$ is $O(\delta T/q) \le \frac{1}{2}\log(M/2)$. Every nonzero observed mass divided by the same weight is

at least $\log(M/2)$. On an occupied cell the excess over its prediction is at least the prediction itself. Empty cells contribute their full prediction. Thus the discrepancy is at least the predicted mass in these rows. Summing the prediction over r, a, k gives $\gg_{c_0} X$ by the prime number theorem. In this range $q > 2A$. Finally if $S\mathcal{L}/T \to \infty$, the predicted mass on all occupied cells is $O(T/(S \log M))$ times the observed mass. Set $M_X = \frac{1}{2}(\log 2)^2 X$. The row estimate and the exclusion bound above give $|P_q|, |J_q| = (1 + o(1))M_X$ uniformly. In the identity

$$D(q) = |P_q| + |J_q| - 2 \sum_{a,k,r} \min(P_{akr}, J_{akr})$$

the overlap is $o(M_X)$, by the estimate on the occupied cells. This proves $D(q)/M_X \to 2$. Replacing $P_q/M_X, J_q/M_X$ by $P_q/|P_q|, J_q/|J_q|$ costs $o(1)$ in ℓ^1 by the same total-mass estimates. Halving the resulting distance gives the claimed probability total variation. □

Proposition 5.3 (A pooled marginal for each modulus)**.** *Let $W = \lceil \exp\sqrt{\mathcal{L}} \rceil$ and partition $K/3 < k < 4K$ into consecutive blocks J of length W, with a possibly shorter last block. Uniformly for every prime $X^\varepsilon \le q \le X^{1/5-\varepsilon}$,*

$$\sum_{\substack{a,J,r \\ q\nmid a}} \left| \sum_{k\in J} P_{akr} - q^{-1} \sum_{k\in J} \frac{\vartheta(k)}{k} |I_{ak}| \right| \ll_{\varepsilon,D} X\mathcal{L}^{-D} \qquad (D > 0).$$

The model retains the actual prime weights in each short block.

Proof. Write $s = (kr-2)\bar{a} \bmod q$. For each exact row subtract $k^{-1} \sum_{m\in I_{ak},\, m\equiv s(q)} \vartheta(m)$ from $P_{akr}/\vartheta(k)$. For a fixed interval I, Parseval over residue pairs modulo k, q bounds the sum of squared errors over a, r by

$$(kq)^{-1} \sum_{\substack{j \bmod kq \\ k\nmid j}} |S_I(j/(kq))|^2.$$

The distinctness of the first residues in a permits this enlargement to all residue pairs. Frequencies for all prime $k \asymp K$ are distinct and spaced by $\gg 1/(qK^2)$, including those for which $q \mid j$. For varying intervals decompose each into at most two binary intervals per level. At any one level sum the coefficient energies over disjoint intervals before applying Cauchy across levels. The large sieve, followed by Cauchy over $O(AKq)$ rows and insertion of $\vartheta(k)$, gives total error

$$\ll X\mathcal{L}^3\{A^{-1/2} + \sqrt{qK^2/X}\}.$$

This saves a fixed power in the asserted range.

Freeze the interval I_{ak} on $G = \lceil \mathcal{L}^{D+5} \rceil$ slices per coordinate in a, k, intersecting k-slices with each block J. The endpoints vary by $O(M/G)$. Summing the empirical model over all r charges disjoint residue classes, so the total boundary error is $O(X\mathcal{L}/G)$, also for the reference. For a frozen interval I and a block piece J', put

$$u_s = \sum_{m\in I,\, m\equiv s(q)} \vartheta(m), \qquad v_t = \sum_{k\in J',\, k\equiv t(q)} \vartheta(k)/k.$$

The mass is $\sum_{xs-rt=-2} u_s v_t$ for $x = a \bmod q$. For uniform $x, r \in \mathbb{F}_q$, each nonzero point (s, t) lies on this line with probability $1/q$. Two noncollinear points give independent incidences, and distinct collinear points are incompatible. The origin has zero weight. Positivity therefore bounds the sum over x, r of squared deviations from $(\sum u)(\sum v)/q$ by $q \|u\|_2^2 \|v\|_2^2$. A slice in a repeats each x at most $O(A/(Gq))$ times. Cauchy over x, r now costs $O((A/G)\sqrt{q}\, \|u\|_2 \|v\|_2)$. Here $\|u\|_2 \ll M\mathcal{L}/\sqrt{q}$ and $\|v\|_2 \ll \sqrt{W}\mathcal{L}/K$ since $W < q$. Summing over $O(G)$ slices in a and $O(K/W + G)$ block pieces gives $O(X\mathcal{L}^2/\sqrt{W})$. The prime number theorem replaces $\sum u$ by $|I|$ with any fixed logarithmic saving in the total. Restoring the endpoints proves the estimate. □

Corollary 5.4 (Two residues can locate the origin)**.** *Fix $7/40 < \eta < 1/5$ and let q_1, q_2 be distinct primes in $(X^\eta, 2X^\eta]$. With a, k given, the mean of $D(q)$ over the primes of this dyad is $o(X)$. For q_1q_2 coprime to a the joint residues determine the origin exactly for all large X. Even if k is reported only through its interval of length $W = \lceil \exp\sqrt{\mathcal{L}} \rceil$, for almost all pairs of these moduli the two residues uniquely determine a $1 - o(1)$ fraction of the origins in any fixed probability measure on incidences independent of the moduli. Each pooled marginal obeys Proposition 5.3 for every modulus in this dyad.*

Proof. For exact k, the rank interval has length $O(T)$ and the two residues determine it modulo $q_1q_2 \gg T$. For pooled k, hold a and its reported k-interval fixed. Each competitor k' allows at most $O(1 + T)$ possible b', because $k'b' \equiv 2 \bmod a$ and $(k', a) = 1$ in any incident row. If $b' = b$, then $a \mid k' - k$, so $k' = k$ since $W < a$. Each nonzero difference $b' - b$ has $O_\eta(1)$ prime factors in the selected dyad. The mean collision probability for a fixed origin is therefore

$$O_\eta(W(1 + T)/\pi^2_{X^\eta}) = X^{7/20-2\eta+o(1)}.$$

Integrate this uniform bound against the given measure and apply Markov. For the prime-weighted marginal, pooling the exact rows contracts the sum of absolute errors. Making each predicted residue uniform costs $O(X/q)$, and the discarded a's cost $o(X)$ uniformly. This proves coexistence of accurate individual marginals and joint recovery. □

5.1 A shared rank difference among prime pairs

Exact divisor labels can also be retained for a rarer event at the fixed displacement two. Write $m_i = m_0 + bi,\ k_i = k_0 + ai$ in each cell, $0 \le i < N_{ab2}$.

Proposition 5.5. *For $T_0 = \lceil C\mathcal{L}^2 \rceil$ with a sufficiently large absolute C, there are $\gg T_0$ even integers $1 \le d < T_0$ for which the following holds. In $\gg |\mathcal{C}|$ primitive odd cells, at least $cU/\mathcal{L}^4$ ranks i have $m_i, k_i, m_{i+d}, k_{i+d}$ all prime and both ranks inside the cell. The constants are absolute. The two progressions have differences bd and ad, with the same label d across these cells.*

Proof. The total number of prime pairs is $\asymp X/\mathcal{L}^2$ by (5.2). For an upper bound in one cell, sieve m_i, k_i in $i \le O(U)$. At primes outside $2ab$ the two forbidden roots are distinct, because $bk_i - am_i = 2$. The upper-bound sieve [9, Chs. 6, 11] at level $U^{1/2}$, sieving up to a sufficiently small fixed power of U, gives $O(UD_c(2ab)/\mathcal{L}^2)$, where $D_c(v) = \prod_{p|v}(1 + c/p)$ with a sufficiently large fixed c. The remainder over $l \le U^{1/2}$ is $O(U^{1/2+o(1)})$ by counting the at most $2^{\omega(l)}$ rank classes. The mean square of $D_c(2ab)$ over a, b is bounded by expanding in squarefree divisors with coefficients $\ll_c \prod_{p|l}(C_c/p)$ and using the dyadic frequency $O(1/l)$ of multiples of l. Hence a positive proportion of cells have at least $c'U/\mathcal{L}^2$ prime pairs, by Cauchy–Schwarz.

In each such cell split the ranks into blocks of length T_0, with a shorter final block allowed. The mean number of prime pairs per block exceeds four for C large. Convexity of $\binom{x}{2}$ now gives $\gg UT_0/\mathcal{L}^4$ pairs of successful ranks in the same block. If V_{abd} counts successes at both $i, i + d$, then its mean over a, b, d is $\gg U/\mathcal{L}^4$.

To turn this into a prevalence assertion, sieve the four linear forms in i. There are four distinct roots outside the prime factors of $D = 2abd(abd - 2)(abd + 2)$, as one sees from the two same-site differences and the cross determinants. Small primes may all be discarded in the upper bound. The same sieve, now with at most $4^{\omega(l)}$ classes, yields $V_{abd} \ll UD_{c'}(|D|)/\mathcal{L}^4$. Its squared local factor has bounded mean. For a, b, d individually use the preceding divisor expansion. For $abd \pm 2$, fix a, d and treat odd primes up to $\mathcal{L}^2$. Each eligible squarefree l prescribes one class of b mod l, of frequency $O(1/l + 1/B)$. A prime common to l, ad makes the congruence impossible. The $1/B$ contribution is $O(B^{-1}\prod_{p\le\mathcal{L}^2}(1 + O(1/p))) = o(1)$.

Larger primes have bounded product since $D = X^{O(1)}$. The estimates also hold for each fixed higher power of the local factor, so Hölder combines the individual means. Cauchy again gives a positive proportion of triples with $V_{abd} \gg U/\mathcal{L}^4$. Counting fibres over d gives the stated conclusion. All four primes exceed two. Since a, b are odd, both $m_{i+d} - m_i = bd$ and $k_{i+d} - k_i = ad$ force d to be even. Each selected difference is shared by a positive proportion of cells. The selection may vary with X, and the argument leaves open a difference prescribed in advance. □

6 Sign words visible inside a fixed cell

Number the points of $\mathcal{I}_{ab}(2)$ in increasing order. Advancing one rank changes m by b and k by a. To compare several observations along the second progression, for $q = u/d \in \mathbb{Q}_{\geq 0}$, $d \geq 1$, define

$$Y_q(n) = \lambda(d)\lambda(d(n+2) + uab) = \lambda(b)\lambda(d)\lambda(dk + ua).$$

Complete multiplicativity makes the definition independent of the chosen representation. At $q = 0$ this is the actual sign of $n + 2$.

Theorem 6.1 (Rational directions and square divisors)**.** *Fix a finite $F \subset \mathbb{Q}_{\geq 0}$ and $C > 0$. In each cell choose an integer $0 \leq t_{ab} \leq CU$. For any specified word $\epsilon \in \{-1, 1\}^F$,*

$$\sum_{\mathcal{C}} \left| \#\{n \in \mathcal{I}_{ab}(2) : Y_{t_{ab}+q}(n) = \epsilon_q \ (q \in F)\} - 2^{-|F|} N_{ab2} \right| = o_{F,C}(X)$$

uniformly in these choices. The estimate also holds with a separate absolute value on every fibre of

$$(s(m), (s(L_q))_{q \in F}), \qquad s(x) = \prod_p p^{\lfloor v_p(x)/2 \rfloor}.$$

Here $q = u_q/d_q$ is in lowest terms and $L_q = d_q(k + t_{ab}a) + u_q a$. Use the actual fibre size in the main term and sum the errors over fibres. On a squarefree L_q the word thus supplies its true Möbius sign, after the known factors $\lambda(b)\lambda(d_q)$ are restored.

Proof. Expand the word indicator in Walsh characters. For a nonempty $S \subset F$ and a common denominator D, the corresponding product, up to a cellwise sign, is $G_a(w) = \prod_{q \in S} \lambda(w + Dqa)$ at $w = D(k + t_{ab}a)$. In the sum over ranks replace each value by the mean of $G_a(w + jDa)$ for $1 \leq j \leq J$. The total endpoint cost is $O(JAB)$. Cauchy with respect to incidence mass reduces the desired normalised square bound to a constant times

$$\frac{1}{AK} \sum_{a \sim A} \sum_{c_1 K \leq w \leq c_2 K} \left| J^{-1} \sum_{j=1}^{J} G_a(w + jDa) \right|^2 + o_J(1).$$

For fixed a, w with $D \mid w$, $b(w/D) \equiv 2 \bmod a$ allows $O(B/a)$ possible b's since a is odd. Each b then determines the original k, also when t_{ab} varies across the cells. The integrand is positive so the divisibility of w can be discarded. The constants $c_1, c_2 > 0$ cover the full translated range.

Expand the square. Each off-diagonal product has slopes in the nonempty symmetric difference of $DS + Dj$ and $DS + Dj'$, as repeated factors square to one. Matomäki, Radziwiłł, Tao, Teräväinen and Ziegler [18, Thm. 1.10] give cancellation after averaging the absolute w-sum over the step variable, for polynomials with nonconstant pairwise differences. Here $A = K^{5/8}$. To fit a prefix of length cK in w, write $a = dv + r$ with fixed d large enough that $2A/d + 1 \leq (cK)^{5/8}$. The polynomials are $p(dv + r)$ at the distinct slopes p. The absolute value per v permits restriction to the desired dyad. Subtract two such prefix estimates. The diagonal contributes $O(1/J)$. First take $X \to \infty$, then $J \to \infty$.

For the fibre assertion, initially reveal only the valuations at $p \le P$ capped at e. This partitions each cell using the rank modulo $D' = \prod_{p \le P} p^e$. Smooth by the rank shifts $D'j$, which preserve the mask exactly. There are only finitely many masks for fixed P, e, the endpoint cost is $O_{P,e}(JAB)$, and after Cauchy the mask can be dropped. The same polynomial argument applies.

To remove the cutoffs, fix $p \le P$. A cell with $p \nmid b$ contributes $O(U/p^e + 1)$ points with $p^e \mid m$. Their total is $O(X/p^e + AB)$. If $p \mid b$ no such point exists. To bound $p^e \mid L_q$, lift $k' = k + t_{ab}a$ to a range of length $O(K)$ with multiplicity $O(B/a)$ as above. For each a its congruence permits at most $(d_q, p^e) \le d_q$ residue classes. This gives $O_{F,C}(X/p^e + B)$. For $p > P > \max d_q$, summing the same bound with exponent two up to $O_{F,C}(\sqrt{K})$ costs $O_{F,C}(X/P + B\sqrt{K})$. For $p^2 \mid m$ split instead at $Z = X^{1/40}$. The smaller primes above P cost $O(X/P + ABZ)$ by the progression count. A larger square divides n. Counting multiples gives $O(X/Z)$ such integers, each carrying at most $\tau(n)\tau(n+2) = X^{o(1)}$ incidences. Outside all these exceptions the exact message is a function of the truncated one. The errors change by at most twice the exceptional mass. Let $X \to \infty$, then $e \to \infty$ and $P \to \infty$. □

A common test cannot have large correlations with several independent directions at once. For real $|x_{ab}(n)| \le 1$, put $D_q(x) = \sum_{\mathcal{C}} |\sum_{n \in \mathcal{I}_{ab}(2)} x_{ab}(n) Y_q(n)|$ and $\mathcal{M}_2 = \sum_{\mathcal{C}} N_{ab2}$. For $c_q \ge 0$,

$$\limsup_{X \to \infty} \frac{\sum_{q \in F} c_q D_q(x)}{\mathcal{M}_2} \le \mathbb{E} \left| \sum_{q \in F} c_q \epsilon_q \right|$$

for independent fair signs ϵ_q. In each cell choose a sign realising each modulus and use Theorem 6.1 on the resulting linear form. The bound is attained asymptotically by $x = \operatorname{sgn}(\sum c_q Y_q)$, with any choice at zero. For $F = \{0, q\}$, $q > 0$, the bound for equal weights is one. It leaves open the value of the single correlation at $q = 0$ with $x = \lambda(n)$.

7 What averaging the displacement permits

The shifts in Theorem 1.2 are hidden. Here the observer knows h as well as the exact divisors a, b, and the error is summed over a long interval of such shifts. Write $\mathcal{H} \subset [1, 2X]$ for any interval of H consecutive integers. Define

$$C_2 = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right), \qquad \mathfrak{S}_{ab}(h) = \begin{cases} 2C_2 \displaystyle\prod_{p>2,\ p \mid abh} \frac{p-1}{p-2}, & 2 \mid h,\ (h, ab) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

All cell sums below range over $n \in \mathcal{I}_{ab}(h)$, with $m = n/a,\ k = (n+h)/b$.

Theorem 7.1 (Three observation ranges)**.** *The following estimates are uniform in the position of $\mathcal{H}$.*

1. *If $H/A \to \infty$, then, uniformly for all arrays $|g_{ab}(k)| \le 1$ common to the different shifts,*

$$\sum_{h \in \mathcal{H}} \sum_{\mathcal{C}} \left| \sum_{n \in \mathcal{I}_{ab}(h)} \lambda(n) g_{ab}(k) \right| = o(HX).$$

In particular one may take $g_{ab}(k) = \lambda(bk)$. The common-test estimate also holds for an interval anywhere in $\mathbb{Z}$, using arrays on all integers and extending a positive-argument test by zero for $k \le 0$.

2. *Let* $P_{ab}(h) = \sum_{n \in \mathcal{I}_{ab}(h)} \vartheta(m)\vartheta(k)$. *For* $H \geq X^{19/44+\varepsilon}$ *and every fixed* $D > 0$,

$$\sum_{h \in \mathcal{H}} \sum_{\mathcal{C}} \left| P_{ab}(h) - \frac{X}{ab} \mathfrak{S}_{ab}(h) \right| \ll_{\varepsilon, D} HX\mathcal{L}^{-D}.$$

3. *Let* $N_{ab}^{(2,1)}(h)$ *count those incidences for which* $\Omega(m) = 2$ *and* k *is prime. If* $H \geq X^{1/4+\varepsilon}$ *and* $b_X = (\log\log M)/(\log M \log K)$, *then*

$$\sum_{h \in \mathcal{H}} \sum_{\mathcal{C}} \left| N_{ab}^{(2,1)}(h) - \frac{X}{ab} \mathfrak{S}_{ab}(h) b_X \right| = o_{\varepsilon}(HXb_X).$$

Prime squares are included. On inadmissible triples the reference is zero, and genuine atypical points can occur in this last population.

The proof is in Appendices A and B. Its structure is easiest to see by fixing a residue of h mod a. A change $h \mapsto h + a\ell$ keeps the admissible class of k the same. For a fixed k, it translates m by $-\ell$. Cauchy–Schwarz then removes the common second-site test, leaving an autocorrelation at the first cofactor. The result of Matomäki, Radziwiłł and Tao [17, Thm. 1.3(i)] supplies the prime autocorrelation for lengths at least $M^{8/33+\delta}$. Requiring H/a to contain that many translates gives $1/4 + (3/4)(8/33) = 19/44$. For the typical semiprime weight, Evans [6, Props. 9.1–9.2] supplies the input for any fixed positive power of M. The smaller range in the third assertion reflects this difference in the inputs. Taking absolute values before summing over the observed divisor labels is the additional step here.

The common-test condition. If H/A stays bounded along a subsequence, choose $H_0 = \min(H, \lfloor A \rfloor)$ consecutive shifts in $\mathcal{H}$. For each a, b, k there is at most one of them, say h_0, with $bk \equiv h_0 \bmod a$. Give $g_{ab}(k)$ the value $\lambda(bk - h_0)$ when $X < bk - h_0 \leq 2X$, and zero if no such origin exists. For each of these H_0 shifts every cell sum in the first assertion equals its occupancy. By (2.1) the resulting lower bound is $\gg H_0 X$. Thus $H/A \to \infty$ is necessary for the whole test class. The constructed array depends on the chosen interval and says nothing about necessity for the sole test $\lambda(bk)$.

On an admissible triple the singular series is at least $2C_2$. Markov's inequality in the second and third estimates gives relative error tending to zero outside $o(HAB)$ triples. Admissible triples have density bounded below. Congruence counts handle primes up to a fixed cutoff. Common larger factors of two labels have total probability at most $\sum_{p>P} O(p^{-2}) + o(1)$. For the pair b, h with a shifted interval one may use $O((B/p)(H/p + 1))$ and sum the last term as $O(B \log X)$. On admissible triples $\log m = \log M + O(1)$, $\log k = \log K + O(1)$ for primes in the box, giving the unweighted prime-pair equivalent with $\mathfrak{S}_{ab}(h)X/(ab \log M \log K)$.

Which sign remains hidden? At polylogarithmic H, Corollary 4.2 constructs the sign of the decoded origin below the single-guess threshold. It uses the factorisations of candidate origins and there is no assumption that this test is multiplicative. Above the threshold, failure to find the exact h leaves open prediction of the one-bit target $\lambda(v - h)$. At fixed shift two, Theorem 6.1 balances signs at the second site, even on square-divisor fibres. Cancellation of $\lambda(n)\lambda(n + 2)$ in the exact divisor cells remains open, as does the prime-pair asymptotic in those same cells at that prescribed shift.

A From cofactor translations to exact divisor cells

We give the dispersion argument for all three claims of Theorem 7.1. It is useful to keep two lengths separate. A rank within a fixed cell has about U possible values. The comparison of shifts in the same residue class uses up to H/a distinct translates.

Write $\mathcal{L} = \log X$ and use the scales and primitive label set $\mathcal{C}$ of Section 1. For an interval $\mathcal{H}$ of H consecutive integers put

$$D_f = \sum_{(a,b)\in\mathcal{C}} \sum_{h\in\mathcal{H}} \left| \sum_{\substack{am+h=bk \\ X<am\le 2X}} f(m) g_{ab}(k) \right|.$$

Lemma A.1 (Dispersion over exact divisor cells)**.** *Suppose $A \le H \le B/2$ and $J \ge 2$ is an integer with $3AJ \le H$. Let $f : \mathbb{Z} \to \mathbb{C}$ be supported on $(M/2, 2M]$, with $|f| \le 1$. For each $(a, b) \in \mathcal{C}$, let $g_{ab} : \mathbb{Z} \to \mathbb{C}$ satisfy $|g_{ab}| \le 1$ and use this same function for every h. Define*

$$\rho = \frac{1}{JM} \sup_{\substack{\alpha\in\mathbb{R} \\ I\subseteq(M/2,2M] \text{ interval}}} \sum_{1\le|\ell|\le J} \left| \sum_{m\in I} f(m)\overline{f(m+\ell)} e(m\alpha) \right|.$$

For $1 < V < X^{3/20}$ and $0 < \xi < 1$,

$$\frac{D_f^2}{(HX)^2} \ll J^{-1} + \xi + U^{-1}(1 + J/B) + \xi^{-1}(V\rho + V^{-1/2} + X^{-1/40}\mathcal{L}^4). \tag{A.1}$$

The implied constant is absolute, uniformly in the location of $\mathcal{H}$ and in the arrays g_{ab}.

Proof. Comparing shifts in a residue class. By duality choose one unit phase per a, b, h to realise the modulus. Write $h = r + as$, $0 \le r < a$, and split the consecutive s's into blocks of size at most J. For a fixed block, all k lie in $k \equiv \bar{b}r \bmod a$, and the union of their ranges contains $O(U + 1 + J/b) = O(U)$ such integers. The number of blocks is $O(AB(H/J + A)) = O(ABH/J)$. Cauchy over block and k thus contributes $O(HX/J)$ to the squared bound and removes g_{ab}. Expanding the other factor gives $O(HX)$ diagonal terms. Two shifts $a\ell$ apart see m and $m + \ell$ at the same k. Consequently

$$\frac{D_f^2}{(HX)^2} \ll \frac{1}{J} + \frac{1}{J} \sum_{0<|\ell|<J} \frac{E_\ell}{HX}, \qquad E_\ell = \sum_{a,b,h} \left| \sum_{\substack{m\in I_{a,\ell} \\ m\equiv -h\bar{a} \ (b)}} f(m)\overline{f(m+\ell)} \right|, \tag{A.2}$$

where $I_{a,\ell} = (X/a, 2X/a] \cap (X/a - \ell, 2X/a - \ell]$. The phases of the two shifts are constant inside the last progression sum. This is the point where using a single second-site test across shifts matters.

How often a residue is used. For a subset of the a's let $w_b(t) = \#\{a, h : h \equiv -at \bmod b,\ (a, b) \in \mathcal{C}\}$ and $\mu_0 = AH/B$. Then

$$\left(\sum_{b\sim B} \sum_{t \bmod b} w_b(t)^2 \right)^{1/2} \ll B\{\mu_0 + (\sqrt{\mu_0} + A/\sqrt{B})\mathcal{L}^3\}. \tag{A.3}$$

Indeed, the zero Fourier mode of the shift interval is $O(\mu_0)$ after summing a. At frequencies $R \le |r| < 2R$ its normalised coefficients are $O(\min(H/b, 1/R))$. Parseval on uniform $t \bmod b$ bounds the norm of this slice by

$$\min(H/b, 1/R)\{AR(1 + AR/b)\}^{1/2} O(\mathcal{L}^{3/2}).$$

For either sign of r, the products $a|r|$ have multiplicities bounded by τ, with $O(1 + AR/b)$ possible products in a given residue. Apply Cauchy to these products and $\sum_{v\le x} \tau(v)^2 \ll x \log^3(2x)$. Summing dyadic R, the square-root term peaks at $R \asymp b/H$, and the other term

is $O(A/\sqrt{b})$ per slice before logarithms. This proves (A.3), uniformly over the chosen subset and the location of $\mathcal{H}$.

The additive frequencies. Split a into $O(\xi^{-1})$ bands of width at most ξA, with $0 < \xi < 1$. In each band freeze the m-interval to one I independent of a, ℓ. The total cost in $E_\ell/(HX)$ is $O(\xi + U^{-1}(1 + J/B))$, by counting progression terms in endpoint intervals of length $O(\xi M + J)$. Put $Z_\ell(\alpha) = \sum_{m\in I} f(m)\overline{f(m+\ell)}e(m\alpha)$. The contribution of reduced denominators $d \asymp P$ to a progression sum is

$$S_b(t) = \frac{1}{b}\sum_{\substack{d\mid b\\ d\asymp P}}\sideset{}{^*}\sum_{j \bmod d} Z_\ell(j/d)e(-jt/d).$$

Here and below $d = 1$ contributes the zero frequency. For $P \le X^{3/20}$, each residue modulo $d \mid b$ is visited $O(H/d)$ times by $-h\bar{a}$. Parseval and $\sum_{b\sim B,\, d\mid b} b^{-1} \ll 1/d$ bound this slice of E_ℓ by

$$\ll AH \sum_{d\asymp P}\frac{1}{d}\left(\sideset{}{^*}\sum_{j \bmod d}|Z_\ell(j/d)|^2\right)^{1/2} \ll HX(P^{-1/2} + \sqrt{P/M}).$$

The last step uses Cauchy and the additive large sieve with spacing $\gg P^{-2}$. For $P > X^{3/20}$, Parseval on the full set of residues and the same sieve give $\sum_{b,t}|S_b(t)|^2 \ll (M + P^2)M/P$. Apply (A.3) and divide by $HX = B\mu_0 M$. The cost is at most

$$O\Big([1 + (\mu_0^{-1/2} + A/(\sqrt{B}\mu_0))\mathcal{L}^3][P^{-1/2} + \sqrt{P/M}]\Big).$$

For $H \ge A$, the first bracket is $O(X^{1/20}\mathcal{L}^3)$ and the second is $O(X^{-3/40})$ up to $2B$. All these large-denominator slices together cost $O(X^{-1/40}\mathcal{L}^4)$ per band.

For $d \le V$ retain absolute values of individual Fourier coefficients. On each band the frozen interval I is independent of ℓ. For each fraction j/d the defining bound for ρ therefore applies before summing over fractions. After the b-sum a denominator d has weight $O(1/d)$, and there are at most d fractions. The contribution per band, averaged in ℓ, is $O(V\rho)$. For $V < d \le X^{3/20}$ the dyadic slices sum to $O(V^{-1/2} + X^{-3/10})$. Substitution in (A.2), followed by the sum over the $O(\xi^{-1})$ bands, proves (A.1). □

Application to the shift windows. A bounded factor depending only on a can be absorbed in the common test. A window longer than $B/2$ splits into intervals of comparable lengths between $B/5$ and $B/2$ for large X. We sum the estimates on these intervals. In the applications, both V and ξ^{-1} are at most polylogarithmic. If the two input sequences are bounded by fixed powers of $\mathcal{L}$, first divide by these bounds. An arbitrary logarithmic saving for the averaged correlations, with polynomial J, gives an arbitrary logarithmic saving for $D_f/(HX)$ by choosing $\xi = \mathcal{L}^{-E}$, $V = \mathcal{L}^F$ and then the precision for ρ, in that order.

For $f = \lambda$ on the cofactor range, use [16, Thm. 1.6, (1.11)–(1.12)], with the shifted multiplicative input λ and the other bounded input $\lambda(m)e(m\alpha)\mathbf{1}_I(m)$. Endpoints of the cofactor range can be put in I, at cost $O(J)$ ordinary terms per correlation. Negative translates follow by a change of variable. The resulting ρ is $o(1)$ uniformly in I, α when $J \to \infty$ and $J \le X^{1/100}$. One may take a constant multiple of $\log\log(3J)/\log(3J) + \mathcal{L}^{-1/3000}$. Choose $J = \lfloor \min(H/(4A), X^{1/100})\rfloor$. Equation (A.1), e.g. with $V \asymp \rho^{-2/3}$, $\xi \asymp \rho^{1/6}$, gives part 1 of Theorem 7.1.

Using an untwisted correlation input. Let F be real, supported on a dyad $Y < m \le 2Y$ with $Y \asymp M$, and $\|F\|_\infty \ll \mathcal{L}^c$. If

$$\frac{1}{J}\sum_{\ell=1}^{J}\left|\sum_n F(n)F(n+\ell)\right| \ll_E Y\mathcal{L}^{-E} \qquad \text{for all } E > 0, \tag{A.4}$$

then for any $\sum_m |v_m|^2 \le Y$ the sum $\sum_{\ell\le J} |\sum_m v_m F(m+\ell)|$ is $O_E(JY\mathcal{L}^{-E})$ when J grows as a positive power of X. Choose a unit phase for each ℓ and apply Cauchy in m to bound the square by

$$YJ\left(\|F\|_2^2 + 2\sum_{t=1}^{J-1}\left|\sum_m F(m)F(m+t)\right|\right).$$

The same holds for negative translates. Taking v_m proportional to $F(m)e(m\alpha)\mathbf{1}_I(m)$ gives ρ with any logarithmic saving. Thus (A.4) is the only first-site correlation input required for the two prime-counting applications.

B Prime and semiprime correlation inputs

On each of the two dyads $Y < m \le 2Y$, $Y = M/2, M$, put

$$R = \exp((\log Y)^{1/10}), \qquad W_R = \prod_{p<R} p, \qquad G_R(m) = \frac{W_R}{\varphi(W_R)}\mathbf{1}_{(m,W_R)=1}.$$

We use a separate R for each dyad. The function is bounded by $O(\log R)$, independently of the size of the full period W_R.

Lemma B.1 (A weighted progression error with variable endpoints)**.** *Let $Q_* = 2X^{3/10}$. For $I \subset [K/3, 5K]$ a half-open interval and $(u,q) = 1$, put*

$$E_q(I,u) = \sum_{\substack{k\in I\\ k\equiv u \bmod q}} \vartheta(k) - |I|/\varphi(q).$$

For every fixed $D > 0$,

$$\sum_{q\le Q_*} \frac{\tau(q)}{q} \sideset{}{^*}\sum_{u \bmod q} \sup_I |E_q(I,u)| \ll_D K\mathcal{L}^{-D}. \tag{B.1}$$

Proof. On one dyad $Q < q \le 2Q$, first remove $\tau(q) > \mathcal{L}^C$, where $C > D + 7$. Their weighted contribution is $O(K\mathcal{L}^{4-C})$. Indeed, $\sum_u^* \sup_I |E_q(I,u)| \ll K\mathcal{L}$ by summing all positive prime masses and the length bounds over the classes, and $\sum_{q\sim Q} \tau(q)^2/q \ll \mathcal{L}^3$.

For the remaining moduli take a grid of spacing $K\mathcal{L}^{-P}$, with $P > C + D + 6$. For one prefix endpoint $z \asymp K$ let $E_q(z,u)$ denote the prime sum up to z minus $z/\varphi(q)$. Barban–Davenport–Halberstam [2, Ch. 29] at upper level $K\mathcal{L}^{-B_0}$ gives

$$\sum_{q\le Q_*} \sideset{}{^*}\sum_u |E_q(z,u)|^2 \ll_E K^2\mathcal{L}^{-E}$$

for any prescribed $E > 0$, choosing B_0 sufficiently large. The enlarged level exceeds Q_* for large X. Replacing the von Mangoldt weight by ϑ costs $O(Q_*K\mathcal{L}^4)$ in the squared norm, since the prime-power mass in one prefix is $O(\sqrt{K}\,\mathcal{L}^2)$. By Cauchy–Schwarz and $\sum_{q\sim Q} \varphi(q)/q^2 \ll 1$, a grid endpoint costs $O_E(K\mathcal{L}^{-E/2})$ with weight $1/q$. Multiplying by $\mathcal{L}^C$ and summing over the $O(\mathcal{L}^P)$ endpoints covers both ends of every interval.

In a segment of length $K\mathcal{L}^{-P}$ each class contains at most $1 + K/(q\mathcal{L}^P)$ integers. Rounding the two endpoints thus costs, over the whole dyad with its weights, at most

$$\ll \mathcal{L}^C\{Q\mathcal{L} + K\mathcal{L}^{1-P}\}.$$

Here $Q_* \ll KX^{-1/10}$. Sum the estimates over $O(\mathcal{L})$ dyads, with $q = 1$ included in the first, and take $E > 2(C + P + D + 3)$. This proves (B.1). □

Computing the reference at a prescribed shift. Fix $h \in [1, 2X]$ even, and restrict to $(ab, h) = 1$. If $\ell_{a,Y}$ is the length of $(X, 2X] \cap (aY, 2aY]$, we claim

$$\sum_{\substack{\mathcal{C} \\ (ab,h)=1}} \left| \sum_{\substack{n \in \mathcal{I}_{ab}(h) \\ m \sim Y}} G_R(m)\vartheta(k) - \frac{\ell_{a,Y}}{ab}\mathfrak{S}_{ab}(h) \right| \ll_D X\mathcal{L}^{-D}, \tag{B.2}$$

uniformly in h. For $d \le D_0 = X^{1/20}$ squarefree, the condition $d \mid m$ is empty if $(d, bh) > 1$. Indeed $p \mid (d, b)$ would divide h, and $p \mid (d, h)$ with $p \nmid b$ would divide the prime $k > D_0$. Otherwise $k \equiv h\bar{b} \bmod ad$ is a reduced class in an interval of length $\ell_{a,Y}/b$ in $[K/3, 5K]$. Its expected prime mass is $\ell_{a,Y}g(d)/(b\varphi(a))$ for multiplicative g on squarefree integers, where

$$g(p) = 0 \quad (p \mid bh), \qquad g(p) = 1/p \quad (p \mid a), \qquad g(p) = 1/(p-1) \quad (p \nmid abh).$$

Group these prime errors by $q = ad \le Q_*$. There are at most $\tau(q)$ ways to write $q = ad$. For each representation, $(b, q) = (h, q) = 1$, so the map $b \mapsto h\bar{b} \bmod q$ permutes the reduced classes. Each class occurs $O(1 + B/q) = O(B/q)$ times on the b-dyad. The interval may depend on b, which accounts for the supremum inside the sum in (B.1). Multiplying that bound by $O(B)$ gives $O_E(BK\mathcal{L}^{-E}) = O_E(X\mathcal{L}^{-E})$ for the sum of all absolute remainders.

Apply the upper and lower fundamental lemma of the sieve [9, Chs. 6, 11] at level D_0, with $s = \log D_0/\log R \gg \mathcal{L}^{9/10}$. Its dimension-one condition follows uniformly from $g(2) = 0$ and $g(p) \le 1/(p-1)$ for $p > 2$. After multiplication by $W_R/\varphi(W_R)$ the local factor is $1-(p-1)^{-2}$ for $p \nmid abh$, and $p/(p-1)$ for $p \mid abh$, including the factor $a/\varphi(a)$ in the reference. Completing the tail changes the product by a factor $1 + O(\mathcal{L}/R)$, by summing $\sum_{p \ge R} p^{-2}$ and the $O(\mathcal{L})$ prime divisors of abh. This gives (B.2).

A true prime at the first site. For $\ell \neq 0$ define

$$S_R(\ell) = \prod_{p<R} \left(1 + \frac{p\mathbf{1}_{p|\ell} - 1}{(p-1)^2}\right), \qquad \mathfrak{S}(\ell) = \lim_{R\to\infty} S_R(\ell).$$

On the common dyad $n, n + \ell \sim Y$, both mixed correlations $\vartheta G_R, G_R\vartheta$ and G_RG_R equal $Y\mathfrak{S}(\ell) + O_E(Y\mathcal{L}^{-E})$, uniformly for $1 \le \ell \le J \le Y^{1-\delta}$. For odd ℓ they vanish. For even ℓ sieve the translate of a prime at level $Y^{1/4}$. The local density is zero for $p \mid \ell$, $1/(p-1)$ otherwise, and Bombieri–Vinogradov with its maximum over reduced classes bounds all the remainders [2, Ch. 28]. For two model weights sieve ordinary integers at the same level, excluding the one or two distinct classes $0, -\ell \bmod p$. The total counting error is $O(\sum_{d \le Y^{1/4}} 2^{\omega(d)})$ before logarithmic factors. The fundamental lemma in dimension at most two yields the displayed S_R after normalisation, with an error smaller than every log power. The endpoints cost $O(J\mathcal{L}^2)$, and the tail beyond R is as in (B.2) (for odd ℓ both series are zero).

For $Y^{8/33+\delta} \le J \le Y^{1-\delta}$, Theorem 1.3(i) of [17] gives this same singular main term for the two von Mangoldt weights, with an arbitrary logarithmic saving off a set of $O_E(J\mathcal{L}^{-E})$ shifts. A trivial $O(Y\mathcal{L}^2)$ bound handles the exceptional set after increasing E. Removing proper prime powers costs $O(\sqrt{Y}\,\mathcal{L}^4)$ uniformly per shift, with $O(J\mathcal{L}^2)$ for the endpoints. Subtraction of the two mixed terms and addition of the model-model term prove (A.4) for $F = (\vartheta - G_R)\mathbf{1}_{(Y,2Y]}$. Take $J = \lfloor X^{2/11+\nu} \rfloor$ with $0 < \nu < \min(\varepsilon/2, 1/100)$. This meets both the input range and $3AJ \le H$ in the original or subdivided window. Appendix A applies with the genuine $\vartheta(k)$ test. On inadmissible triples a prime pair is impossible. Parity excludes odd h, a common factor of a, h would divide $k > 2A$, and one of b, h would divide $m > 2B$. On the other triples (B.2) evaluates the reference. Since $\sum_Y \ell_{a,Y} = X$, this proves part 2.

A typical semiprime and its small factor. Put $u=(\log\log Y)^{-1/2}$, $P_1=\exp((\log Y)^u)$, $P_2=\exp((\log Y)^{1-u})$ and $c_Y=\sum_{P_1\le p\le P_2}1/p\sim\log\log Y$. Let $w(n)=\log r$ on $n=pr\sim Y$ with p,r prime and $P_1\le p\le P_2$, and zero elsewhere. Such a factorisation is unique for large Y. For $J\asymp Y^\delta$, $0<\delta<1/4$ fixed, Evans's estimates imply

$$J^{-1}\sum_{1\le\ell\le J}\left|\sum_n w(n)w(n+\ell)-Yc_Y^2\mathfrak{S}(\ell)\right|\ll_{\delta,E} Y\mathcal{L}^{-E}. \tag{B.3}$$

For clarity, the input is the logarithmically weighted circle-method estimate in [6, §9, Props. 9.1 and 9.2]. Its (9.1) permits this choice of u, and the lengths satisfy $P_2(\log Y)^D\le J$ for every fixed D, as well as $J\ge\exp((\log Y)^{1-\epsilon_0})$ for each fixed $\epsilon_0>0$. Proposition 9.1 bounds the minor-arc energy on each interval of length $1/J$ by $Y(\log Y)^{-B}$. The circle-method Bessel bound [17, Prop. 3.1] and $\sum w^2\ll Y\mathcal{L}^2$ then bound the mean absolute minor-arc coefficient by $O_E(Y\mathcal{L}^{-E})$ upon increasing B. Proposition 9.2 gives the stated main term on the major arcs except for $O(JQ_0^{-1/3})$ shifts, where Q_0 is an arbitrarily large fixed power of $\log Y$. Their trivial contribution is absorbed. This proves (B.3) with the accuracy used in (A.1).

Each mixed correlation of w and G_R equals $Yc_Y\mathfrak{S}(\ell)+O_E(Y\mathcal{L}^{-E})$, for these same ℓ. For even ℓ, fix the smaller factor p, write $n=pr$ and sieve $pr+\ell$ with r prime on a dyad of length $\asymp Y/p=Y^{1-o(1)}$. The $p\mid\ell$ and their main terms have total size $O(Y\mathcal{L}^{O(1)}/P_1)$. For the other p, a sieving prime dividing $p\ell$ excludes no reduced residue of r, and each other sieving prime excludes one. Bombieri–Vinogradov is applicable at $Y^{1/8}$ uniformly in p, with errors $O_E((Y/p)\mathcal{L}^{-E})$. Summing these costs loses only c_Y. Relative to $S_R(\ell)$ the factor at $p<R$ changes by $1+O(1/p)$, negligible after summation. Dyadic endpoint errors contribute $O(J\mathcal{L}^{O(1)}+P_2\mathcal{L}^{O(1)})$. For the reversed order change the sign of ℓ. Both mixed sums vanish at odd ℓ. Thus (B.3) and the model-model computation prove (A.4) for $F=(w-c_YG_R)\mathbf{1}_{(Y,2Y]}$. Take $J=\lfloor X^\gamma\rfloor$, $0<\gamma<\min(\varepsilon/2,1/100)$, and apply Appendix A. This time $H\ge X^{1/4+\varepsilon}$ suffices.

On admissible cells (B.2) supplies $c_Y\ell_{a,Y}\mathfrak{S}_{ab}(h)/(ab)$. For odd h or $(a,h)>1$ there are no typical solutions. If $(b,h)>1$ in a typical solution, some $p\mid(b,h)$ exceeds $P_1/2$ because $p\mid m$. The number of such incidences is $\ll\sum_{P_1/2<p\le2B}AU(B/p)(H/p+1)\ll HX/P_1+X\log X$; any fixed logarithmic weights can be absorbed. Uniformly on typical points $\log r/\log M\to1$, $\log k/\log K\to1$. These replacements preserve the required relative accuracy in the summed norm because the total singular reference is $O(HX)$. Indeed $\mathfrak{S}_{ab}(h)\ll f(a)f(b)f(h)$ with $f(v)=\prod_{p>2,\,p\mid v}(1+1/(p-2))$. Its mean on either dyad, or on $\mathcal{H}$ of polynomial length, is bounded by expansion in squarefree divisors. The endpoint error on the shifted interval is $O(\prod_{2<p\le2X}(1+1/(p-2))/H)=o(1)$. Positivity now allows the removal of both logarithmic weights.

Finally, on either dyad the number of atypical m with $\Omega(m)=2$ is $o(M\log\log M/\log M)$. Sum over the least prime $p\le\sqrt{2Y}$ outside $[P_1,P_2]$. Chebyshev gives $O(Y/(p\log Y))$ choices for the other prime, while the omitted reciprocal sum is $O(1+u\log\log Y)$. This includes squares. To transport this saving to incidences with a genuine second prime, for $0\le z(m)\le\mathcal{L}$ one has uniformly in $1\le h\le2X$

$$\sum_{a,b}\sum_{am+h=bk,\ X<am\le2X}z(m)\vartheta(k)\ll A\,\|z\|_1+\mathcal{L}\sqrt{A(M+K^2)}\,\|z\|_2+AM\mathcal{L}^2/K.$$

For an upper bound discard the remaining size condition on b, keeping $k\asymp K$. For $k\nmid h$ the proof of (5.1) applies with 2 replaced by h, and the reference uses $\sum_{k\asymp K}\vartheta(k)/k=O(1)$. The $O(1)$ primes $k\asymp K$ dividing h force $k\mid m$ and give the last term. For the atypical indicator z, we have $\|z\|_1=o(M\log\log M/\log M)$ and $\|z\|_2\le\sqrt{2M}$. Divide by $\log K+O(1)$ to remove the prime weight. The first term is $o(Xb_X)$ since $AM=X$. Before this division the second term is $O(X^{9/10}\mathcal{L})$ and the third is $O(X^{3/5}\mathcal{L}^2)$, using $K^2\ge M$. Both are $o(Xb_X)$ even before

division, as $b_X \asymp (\log\log X)/\mathcal{L}^2$. The bounds are uniform in h. Summing over the window gives $o(HXb_X)$, completing the proof of Theorem 7.1.

C A translated support in the prime-modulus large sieve

For a row $m = m_0 + kt$ in Theorem 5.2, the prime $p \nmid k$ excludes one residue of t. These residues can be combined into a single integer c whose size has no useful bound in terms of the row length. The following form keeps the length of the original interval in the sieve.

Lemma C.1 (The translated prime-modulus sieve)**.** *Let S be sufficiently large, $S \le N \le S^2(\log S)^{-8}$, and $D = S/\sqrt{N}$. Let $P \ge 1$ and $c \in \mathbb{Z}$. Suppose that $w_t \in \mathbb{C}$ is supported on $1 \le t \le N$ and vanishes when $(t-c, P) > 1$. With $e(x) = e^{2\pi i x}$, put*

$$\Delta = (\log\log S)^{5/2} \prod_{\substack{p \mid P \\ p < D}} (1 - 1/p)^{1/2}.$$

Then, with an absolute implied constant,

$$\sum_{\substack{S < q \le 2S \\ q \text{ prime}}} \sum_{j=1}^{q-1} \left| \sum_t w_t e(tj/q) \right|^2 \ll \frac{(1+\Delta)S^2}{\log D} \sum_t |w_t|^2.$$

In particular the bound is uniform in c and in P.

Proof. For $c = 0$ this is Iwaniec's Theorem 2 in its dyadic form [14, (1.4)–(1.6) and §2]. Write $F(\alpha) = \sum_t w_t e(t\alpha)$. In Iwaniec's reduction to the hybrid integral (5.3) the length parameter is N and the input is on $[1, N]$. Keep F on this interval for that reduction. The support restriction is used to improve (5.3) by his Theorem 5. Only this step needs to be checked on a translated interval.

The integrand in (5.3) is an absolute square. Replace F there by $G(\alpha) = e(-c\alpha)F(\alpha)$. Its nonzero indices $u = t - c$ satisfy $(u, P) = 1$ and lie in $[1-c, N-c]$. In the proof of Theorem 5, the insertion of the divisor weights at (6.2) uses $(u, P) = 1$, an identity valid also for negative indices. Majorise the translated support in that Cauchy inequality by $f_c(u) = f(u + c)$, with f the nonnegative majorant of the original interval in (A.3). Its Fourier transform has the same support and modulus as $\widehat{f}$.

More precisely, in the convention $\widehat{f}(\beta) = \int f(x)e(x\beta)\,dx$ the new multiplier is $e(-c\beta)\widehat{f}(\beta)$. In the Poisson sum (6.3), compact support forces precisely the congruence (6.4). The surviving Fourier argument is $\beta = y - y_1$. Thus the extra phase is independent of both divisors and of the rational residues in (6.5). It factors through the summations leading to (6.8). For each divisor $s \mid P$ used in the residue sums, one still has

$$\sum_{\beta \bmod s}^{*} G(x + \beta/s) = \mu(s)G(x).$$

Indeed each supported coefficient has index u with $(u, s) = 1$, so the inner Ramanujan sum equals $\mu(s)$. After the sieve weights are summed, the remaining kernel is bounded in absolute value by $\widehat{f}(y - y_1)$, which is nonnegative. The integral bound for G has the constant of Theorem 5. Since $|G| = |F|$, inserting it in (5.3) and finishing the reduction of Theorem 2 proves the claim. □

AI assistance and reproducibility

OpenAI's ChatGPT assisted in developing and writing the mathematical arguments. Codex prepared the sources, checked the local-maximum comparison and the translation in the prime-modulus sieve against the cited articles, and ran the exact finite enumerations. The source archive includes two independent incidence enumerations, the all-divisor maximum identity, and code for Example 1.1 and the finite posterior checks. The limiting laws follow from the proofs. The author assumes responsibility for the manuscript.

References

[1] E. Arıkan, *An inequality on guessing and its application to sequential decoding*. IEEE Trans. Inform. Theory **42** (1996), 99–105. DOI 10.1109/18.481781.

[2] H. Davenport, *Multiplicative Number Theory*, 3rd ed., revised by H. L. Montgomery. Graduate Texts in Mathematics **74**, Springer, 2000. ISBN 978-0-387-95097-6.

[3] P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*. Amer. J. Math. **62** (1940), 738–742. DOI 10.2307/2371483.

[4] P. D. T. A. Elliott, *Central limit theorems for classical cusp forms*. Ramanujan J. **36** (2015), 81–98. DOI 10.1007/s11139-013-9516-9.

[5] P. D. T. A. Elliott, *Corrigendum: Central limit theorems for classical cusp forms*. Ramanujan J. **36** (2015), 99–102. DOI 10.1007/s11139-014-9629-9.

[6] N. Evans, *Correlations of almost primes*. Math. Proc. Cambridge Philos. Soc. **174** (2023), 301–344. DOI 10.1017/S0305004122000251.

[7] K. Fan, *Weighted Erdős–Kac theorems via computing moments*. Acta Arith. **217** (2025), 99–158. DOI 10.4064/aa231014-9-8.

[8] K. Ford, *The distribution of integers with a divisor in a given interval*. Ann. of Math. (2) **168** (2008), 367–433. DOI 10.4007/annals.2008.168.367.

[9] J. Friedlander and H. Iwaniec, *Opera de Cribro*. Amer. Math. Soc. Colloq. Publ. **57**, AMS, Providence, 2010. DOI 10.1090/coll/057.

[10] A. Granville and K. Soundararajan, *Sieving and the Erdős–Kac theorem*. In *Equidistribution in Number Theory, an Introduction*, NATO Sci. Ser. II Math. Phys. Chem. **237**, Springer, 2007, 15–27. DOI 10.1007/978-1-4020-5404-4_2.

[11] R. R. Hall, *Sets of Multiples*. Cambridge Tracts in Mathematics **118**, Cambridge University Press, 1996. DOI 10.1017/CBO9780511566011.

[12] R. R. Hall and G. Tenenbaum, *On the local behaviour of some arithmetical functions*. Acta Arith. **43** (1984), 375–390. DOI 10.4064/aa-43-4-375-390.

[13] M. Hayashi, *Second-order asymptotics in fixed-length source coding and intrinsic randomness*. IEEE Trans. Inform. Theory **54** (2008), 4619–4637. DOI 10.1109/TIT.2008.928985. Author's version.

[14] H. Iwaniec, *The large sieve with prime moduli*. Rev. Mat. Iberoam. **38** (2022), 2337–2354. DOI 10.4171/RMI/1381.

[15] R. Khan, M. B. Milinovich and U. Subedi, *A weighted version of the Erdős–Kac theorem*. J. Number Theory **239** (2022), 1–20. DOI 10.1016/j.jnt.2021.10.010.

[16] 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.

[17] K. Matomäki, M. Radziwiłł and T. Tao, *Correlations of the von Mangoldt and higher divisor functions I. Long shift ranges*. Proc. Lond. Math. Soc. (3) **118** (2019), 284–350. DOI 10.1112/plms.12181.

[18] K. Matomäki, M. Radziwiłł, T. Tao, J. Teräväinen and T. Ziegler, *Higher uniformity of bounded multiplicative functions in short intervals on average*. Ann. of Math. (2) **197** (2023), 739–857. DOI 10.4007/annals.2023.197.2.3.

[19] H. L. Montgomery and R. C. Vaughan, *The large sieve*. Mathematika **20** (1973), 119–134. DOI 10.1112/S0025579300004708.

[20] B. Pouly, *Reconstruction and sign prediction from sampled prime factors*. Zenodo, 2026. DOI 10.5281/zenodo.24443266.

[21] B. Pouly, *Simultaneous quadratic probes: reconstruction thresholds and marked Poisson limits*. Zenodo, 2026. DOI 10.5281/zenodo.24443281.

[22] B. Pouly, *Parity information in a rough cofactor and exact reconstruction thresholds*, Zenodo (2026). 10.5281/zenodo.24578849.

[23] B. Pouly, *Visible prime factors under neighbouring critical divisor weights*, Zenodo (2026). 10.5281/zenodo.24578867.

[24] Y. Sakai, *Generalizations of Fano's inequality for conditional information measures via majorization theory*. Entropy **22** (2020), article 288. DOI 10.3390/e22030288.

[25] Y.-C. Sun, *The critical-window profile for the k-fold divisor function in almost all short intervals*. arXiv:2401.08432v3, 3 August 2026. DOI 10.48550/arXiv.2401.08432.