A fixed-profile proof for prime gaps of at most 182

GPT-6 Astra (OpenAI), prompted by Romyar Sharifi

11 September 2026*

Abstract

Recent progress toward smaller explicit bounds for prime gaps draws substantially on Julia Stadlmann's 2023 advance in the distribution of primes to squarefree smooth moduli. Her August 2026 mixed-support preprint lowered the more-than-twelve-year-old bound $H_1 \leq 246$ to $H_1 \leq 240$; independent concurrent work by OpenAI, using her 2023 estimates, reached $H_1 \leq 186$, and the Axiom Math team further developed her framework to reach $H_1 \leq 212$. Building on these developments with additional Type II and Type III inputs and a new fixed profile, this paper proves $H_1 \leq 182$.

The prime-gap argument has two distinct tasks: make a quadratic sieve form exceed its mass, and ensure that its coefficients lie in the arithmetic domains on which the distribution estimates apply. We explain both tasks for a fixed profile in 39 coordinates. The profile is described by 726 rational coefficients. Its larger, easily integrated support gives a surplus greater than 0.0002766919371; all corrections needed to restore the arithmetic support cost at most 0.000211478178. The remaining surplus exceeds 0.000065213759. The resulting assertion is $\mathrm{DHL}[39, 2]$, which gives a consecutive-prime gap at most 182 infinitely often.

We develop the fragment measure, the sharp five-prime minorant, the exceptional-square estimate, the signed hybrid inequality, and the support-restoration argument before describing their finite certificate. The incidence Type II and finite-exceptional Type III estimates are proved in the two companion papers. A Lean formalization verifies the analytic-to-prime-gap implication under explicit numerical and finite-field premises. The local Fourier proposition remains outside the formal proof; Section 13 describes the precise scope.

Contents

*Original version: 4 September 2026. This version expands citations and exposition; the stated results and numerical data are unchanged.

1 Introduction

Let p_n denote the n-th prime. We use the customary notation

$$H_1 = \liminf_{n\to\infty}(p_{n+1} - p_n).$$

Thus an upper bound for H_1 controls infinitely many consecutive prime gaps. The twin-prime conjecture predicts $H_1 = 2$. The present paper proves $H_1 \le 182$, using a fixed sieve profile and the distribution estimates in the two companion papers [10, 11].

1.1 From bounded gaps to the recent explicit bounds

The starting point for the modern theory is the work of Goldston, Pintz and Yıldırım. They proved

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} = 0$$

and showed that a suitable improvement in the distribution of primes in arithmetic progressions would imply bounded gaps [8, Theorems 1 and 2]. In 2013, Yitang Zhang supplied the first unconditional proof that H_1 is finite, obtaining $H_1 < 7 \times 10^7$ [24, Theorem 1]. His decisive input was distribution beyond the Bombieri–Vinogradov range for suitably restricted moduli. The decomposition into Type I, Type II and Type III sums, and the exploitation of the factorization of those moduli, remain part of the argument used here.

Polymath8a refined Zhang's method and reached $H_1 \leq 4680$; its distribution estimates and dense-divisibility framework are recorded in [19, §§1–2 and Theorem 2.4]. Later in 2013, Maynard introduced a multidimensional sieve that gave $H_1 \leq 600$ and bounded intervals containing any prescribed number of primes [15, Theorems 1.1 and 1.3]. Tao independently developed the multidimensional sieve idea; see his contemporaneous account [23] and [19, §1]. This changed the role of the distribution estimates: even a positive level of distribution suffices for bounded gaps, while stronger distribution information improves the numerical bound [15, Propositions 4.2–4.3]. Polymath8b combined and optimized these ingredients to obtain $H_1 \leq 246$ in 2014 [20, Theorem 1.4(i)].

The recent advances share a central analytic foundation in Julia Stadlmann's work, first posted in 2023 and published in 2025. She improved the exponent of distribution for primes to squarefree smooth moduli from the Polymath8a value $1/2 + 7/300$ to every exponent below $1/2 + 1/40$, using a modified q-van der Corput argument that exploits an additional averaging variable [21, Theorem 1]. Her distribution theorem and associated near-full-mass prime minorant supply central inputs to the recent explicit prime-gap bounds discussed below [21, Proposition 2].

In a subsequent preprint posted on 31 August 2026, Stadlmann combined the classical distribution range with stronger estimates for suitable moduli and introduced a mixed sieve support, allowing different parts of the sieve to use different distribution information. She thereby proved $H_1 \leq 240$ [22, Abstract and §6.2], improving Polymath8b's unconditional bound $H_1 \leq 246$, a record that had stood for more than twelve years, since July 2014. The paper of Charton, Hong, Lau, Ono, Remy, Siu, Swaminathan, Thorner and Xie, announced by Axiom Math on 3 September 2026, obtains $H_1 \leq 212$ [3, Theorem 1.1]. It refines Stadlmann's framework by retaining the divisor conditions actually needed in the proof, treating the continuous range of rough factor configurations, and controlling the transition near the half-level; see [3, §1 and §§7–8].

OpenAI's *Improved short gaps between primes*, dated 30 August 2026, proves $\mathrm{DHL}[40, 2]$ and hence $H_1 \leq 186$ [17, Theorem 1.1 and Corollary 1.2]. That paper describes its work as independent and concurrent with Stadlmann's bound 240. It uses both her bilinear distribution estimate and her near-full-mass prime minorant [21, Propositions 1–2], with extensions to the densely divisible modulus families required by its sieve [17, Propositions 2.18 and 2.21–2.22]. It combines these and Polymath8a's estimates with complementary factorizations and triply densely divisible supports. It also introduces the fragment-based profile calculation, nested inner sieve sums, hybrid inequality and quantitative support restoration that form the immediate framework of the present paper; see [17, §1.2, §3 and §§4.3–4.6]. The bound 182 below extends that framework using additional analytic estimates and a new certified profile.

1.2 The result and its analytic inputs

An integer set $\mathcal{H} = \{h_1, \ldots, h_k\}$ is *admissible* if its image modulo each prime p omits at least one residue. This is the local obstruction relevant to simultaneous primality: a translate of an inadmissible set always has one entry divisible by a particular prime. We use the DHL notation of [20, Claim 3.1].

Definition 1.1. The assertion $\mathrm{DHL}[k, 2]$ means that every admissible set of k distinct integer shifts has infinitely many translates containing at least two primes.

Theorem 1.2. *One has* $\mathrm{DHL}[39, 2]$*. Consequently* $H_1 \leq 182$*.*

The first assertion reduces the tuple size from 40 in [17, Theorem 1.1] to 39. The second is a finite consequence of the admissible 39-tuple of diameter 182 in the MIT prime-gaps tables [16], reproduced and checked in Section 12. There is no assertion that all its entries become prime, or that a particular pair of entries is known in advance to be prime infinitely often.

The change in tuple size requires several ingredients to work together. The incidence estimate of [10] gives additional Type II distribution ranges. The local Fourier estimate of [11] gives a Type III range by restricting the exceptional Fourier frequencies for a correlation of Kloosterman sums. The allowable coefficient moduli in Section 3 combine these two inputs with the estimates inherited from Polymath8a and Stadlmann. Thus the analytic results in the companions are substantive parts of the proof of Theorem 1.2.

For the sieve calculation, we retain the outer and inner profile organization of [17]. We develop the five-prime minorant arising from Baker–Irving's decomposition [2], bound its mean deficit sharply, and prove the exceptional-square and signed hybrid estimates needed for this profile. We then bound the cost of returning from a larger integration domain to the actual arithmetic support. The resulting calculation has 726 fixed rational coefficients in 39 coordinates. Its surplus before support restoration exceeds 0.0002766919371, the restoration cost is at most 0.000211478178, and the final surplus exceeds 0.000065213759. These inequalities are the numerical content of the argument; the preceding distribution and sieve results explain why they imply the prime-gap theorem.

1.3 Why a quadratic inequality proves a statement about primes

For a fixed admissible tuple, we construct a real sieve sum $A(n)$. Its square is a nonnegative weight. If, on arbitrarily large intervals,

$$\sum_n \left(\sum_{i=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_i) - 1 \right) A(n)^2 > 0, \tag{1.1}$$

then at least one n in the interval has two prime entries. Indeed, if every translate had at most one prime entry, each summand would be nonpositive. The analytic problem is therefore to compare two weighted sums. The comparison is homogeneous in A: the absolute size of the coefficients has no significance.

The Maynard sieve turns these sums into quadratic forms of a profile F [15, Proposition 4.1]. Its norm $I(F)$ describes the total mass of the sieve. Integrating out a coordinate describes detecting a prime in that coordinate. A ratio greater than one gives (1.1). The present argument uses the individual prime factors of coefficient divisors, as well as their sizes. This extra information allows use of distribution estimates on suitable densely divisible moduli. It also makes the permitted support of F more complicated.

1.4 The arithmetic support and its computable enlargement

The true arithmetic domain will be denoted by O. For calculation we first work in a larger domain $\mathcal{H}_O$, specified by simple radius and largest-fragment bounds. A good quotient on $\mathcal{H}_O$ is only an intermediate result: the coefficients outside O must be removed. We use the support-restoration framework of [17, §§4.4–4.5]. The removal is the orthogonal projection

$$PF = \mathbf{1}_O F.$$

The removal changes both the norm and the erased-coordinate forms. Consequently its cost is not just the mass of F outside O. Sections 8–9 give a quantitative operator argument that controls this change and retains the favorable change in the denominator.

The proof can be read in the following order.

Ingredient	Its role in the argument
Distribution estimates	Evaluate prime and minorant correlations on the actual coefficient moduli; see Section 3.
Sharp minorant and auxiliary sieve	Bound the exceptional contribution that remains when the minorant replaces the prime indicator; see Sections 4–5.
Signed hybrid	Convert the available mixed correlations into a lower quadratic form; see Section 6.
Support restoration	Transfer the computed form from F on $\mathcal{H}_O$ to PF on O; see Section 8.
Finite certificate	Enclose every quantity in that transfer for one fixed rational profile; see Sections 7 and 10.

The two companion papers address the first row. Their results are substantive analytic inputs, not conclusions of the numerical checker. The remaining rows are developed here. The digital supplement supplies the lengthy rational coefficient and source tables. A reader can follow the proof without interpreting the implementation's file names.

Section 13 gives the hypotheses of the Lean formalization and derives its positive margin from the numerical certificate using a convexity bound for the minorant's mean deficit.

1.5 Conventions

Every terminating decimal used as an exact parameter is a rational number. Bounds displayed with $<, \leq, >$, or $\geq$ are directed bounds, not rounded equalities. The parameter x tends to infinity; the tuple, profile, grid, and auxiliary partitions are fixed first. All implied constants may depend on these fixed objects. In prime and minorant sums, the bounded endpoint intervals are trimmed so that every $n + h_i$ lies in $[x, 2x]$. This removes only $O_{\mathcal{H}}(1)$ integers and has negligible cost under the fixed divisor-power coefficient bounds. We suppress this trimming when displaying interval sums.

We distinguish ρ_*, the logarithmic scaling of sieve divisors, from the Dickman function ρ_D and from the minorant ϱ_x. The cap domains use calligraphic $\mathcal{H}_*$, to distinguish them from the prime-gap constant H_1. The domain $\mathcal{C}$ is a set of retained roots; the sans-serif C is an operator. A prime-factor density parameter δ is different from the operator bound δ_{op}.

2 Prime fragments and the quadratic sieve

We use the prime-fragment realization of the Maynard sieve developed in [17, Sections 3.1–3.3]. Its coefficient algebra comes from [15, Section 5]; the fragment variables retain the individual prime factors needed by the enlarged divisor supports.

2.1 What a fragment records

Put $R = x^{\rho_*}$, where $0 < \rho_* < 1$ is fixed. If r is a squarefree coefficient divisor, each prime $p \mid r$ contributes the fragment

$$u = \log_R p = \frac{\log p}{\rho_* \log x}.$$

The sum of its fragments is $\log_R r$. A bound on this sum controls the size of the divisor; a bound on every individual fragment controls its largest prime factor. The two bounds provide different information. For example, two divisors can have the same logarithmic size while having very different collections of large prime factors.

We impose a fixed largest-fragment cutoff $\zeta > 0$. A fragment configuration X is a locally finite multiset in $(0, \zeta]$, with finite total

$$|X| = \sum_{u \in X} u.$$

There may be infinitely many fragments accumulating at zero, but only finitely many above any positive cutoff. The configurations used in finite-dimensional proofs can instead be written

$$X = (t_0; u_1, \ldots, u_m), \qquad t_0 \geq 0, \quad u_j > u_0,$$

where t_0 collects the fragments at most u_0. This description makes every fixed-cutoff integral an ordinary integral in finitely many real variables.

2.2 The measure and its normalization

Let Π_ζ be the Poisson process of intensity du/u on $(0, \zeta]$, and let γ_E be Euler's constant. Define

$$\nu_\zeta = e^{\gamma_E} \zeta \ \mathrm{Law}(\Pi_\zeta). \tag{2.1}$$

Thus ν_ζ is a finite measure, of total mass $e^{\gamma_E}\zeta$. It is not a probability measure. This convention is dictated by the Mertens normalization of the harmonic divisor sums, recorded in [17, equation (3.5)]. Changing to a probability measure in only some of the forms would change their ratio.

For completeness, let ρ_D be the continuous Dickman function:

$$\rho_D(t) = 1 \quad (0 \leq t \leq 1), \qquad t\rho_D'(t) = -\rho_D(t-1) \quad (t > 1).$$

In the description with a cutoff u_0, the measure is

$$\rho_D(t_0/u_0)\, dt_0 \, \frac{1}{m!} \prod_{j=1}^{m} \frac{du_j}{u_j}, \qquad u_0 < u_j \leq \zeta, \tag{2.2}$$

summed over $m \geq 0$. The factorial accounts for unordered fragments. The image of ν_ζ under $X \mapsto |X|$ is

$$\rho_D(t/\zeta)\, dt. \tag{2.3}$$

For intensity du/u, the sum of the points of the process on $(0, 1]$ has density $e^{-\gamma_E}\rho_D$, by [1, equations (2.4)–(2.6)]. Scaling by ζ and multiplying by the mass in (2.1) gives (2.3). The independent restrictions of a Poisson process [14, Theorem 5.2] then give (2.2): the void factor u_0/ζ cancels the change of total mass. The Laplace-functional formulation is [14, Theorem 3.9]; its use for the harmonic sieve measure is [17, Proposition 3.7 and Lemma 3.8].

There is a useful consistency check. For $0 < c \leq \zeta$,

$$\nu_\zeta\big|_{\{\max X \leq c\}} = \nu_c. \tag{2.4}$$

Indeed, the Poisson probability of no fragment in $(c, \zeta]$ is $\exp(-\int_c^\zeta du/u) = c/\zeta$, and the configuration below c remains a Poisson process. Multiplying by the mass $e^{\gamma_E}\zeta$ in (2.1) leaves $e^{\gamma_E}c$. This explains why there is no further c/ζ factor in a largest-fragment cap integral.

If a function depends only on the total $t = |X|$, integration on a cap c uses $\rho_D(t/c)\, dt$. In particular, the measure has nonzero mass at totals greater than c; several small fragments can have a large sum. Replacing this density by $\mathbf{1}_{t \leq c}$ would impose a different restriction.

2.3 Profiles, erasure, and shared coordinates

A profile in k coordinates is a real function $F(X_1,\ldots,X_k)$. In the final construction it is bounded and supported on a domain with $\sum_i |X_i| \le S$. Initially one may use smooth profiles in the finite-cutoff coordinates of (2.2). The fixed discontinuous profile used in the certificate is reached by the approximation argument in Section 11.

Write $d\nu^k$ for the product measure and define

$$I(F) = \int |F|^2 \, d\nu^k, \tag{2.5}$$

$$(E_iF)(\widehat{X}_i) = \int F(X_1,\ldots,X_{i-1},X,X_{i+1},\ldots,X_k)\, d\nu(X), \tag{2.6}$$

$$J_i(F) = \int |E_iF|^2 \, d\nu^{k-1}. \tag{2.7}$$

The hat means that the i-th coordinate has been omitted. Erasure is an integral, not a conditional expectation.

The square in (2.7) is worth writing out:

$$J_i(F) = \int_Y \int_X \int_{X'} F(Y \oplus_i X) F(Y \oplus_i X')\, d\nu(X)\, d\nu(X')\, d\nu^{k-1}(Y). \tag{2.8}$$

Here $Y \oplus_i X$ inserts X into the missing coordinate. The two erased coordinates X, X' are independent integration variables. The retained configuration Y is shared. This distinction is responsible for several cross terms in the numerical calculation. It also explains why one must add the signed components of F before squaring an erased profile.

2.4 The arithmetic sums represented by a profile

Fix an admissible $\mathcal{H}$, and choose a presieving modulus $W = \prod_{p \le w} p$ and a residue $b_0 \bmod W$ for which all $b_0 + h_i$ are coprime to W. Take w large enough to contain every prime dividing a nonzero $h_i - h_j$. As in the source sieve, w grows sufficiently slowly with x. Put

$$B_x = \frac{\varphi(W)}{W} \log x, \qquad B_R = \rho_* B_x.$$

A diagonal array $y_{\mathbf{r}}$ is indexed by tuples for which $r_1 \cdots r_d$ is squarefree, coprime to W, and has all prime factors at most R^ζ. Sampling F on their fragment configurations gives such an array.

The canonical coefficient transform is [17, equation (3.2) and Lemma 3.2]

$$\lambda^{(d)}_{\mathbf{d}}(y) = B_R^{-d} \left(\prod_{j=1}^{d} \mu(d_j) d_j \right) \sum_{\substack{\mathbf{r} \\ d_j \mid r_j \ (1 \le j \le d)}} \frac{y_{\mathbf{r}}}{\prod_j \varphi(r_j)}. \tag{2.9}$$

This is Maynard's diagonal change of variables [15, equations (5.7)–(5.9)], with each d-coordinate coefficient multiplied by B_R^{-d}. The associated sieve sum is

$$A_F(n) = \sum_{\substack{d_i \mid n + h_i \\ 1 \le i \le k}} \lambda^{(k)}_{\mathbf{d}}(y[F]).$$

The coefficients are real and can have either sign. The final weight is $A_F(n)^2$.

A support condition is *hereditary* if deleting prime factors preserves it. Formula (2.9) only replaces roots by divisors, so hereditary conditions pass from the diagonal array to the coefficient array. No hereditary condition is required of the sign pattern, or of the set on which the profile is nonzero inside the permitted domain.

The normalizing constants for unweighted and prime-weighted sums are

$$\mathfrak{C}_x = \frac{x}{W} B_R^{-k}, \qquad \mathfrak{P}_x = \frac{x}{\varphi(W)\log x} B_R^{-(k-1)} = \rho_* \mathfrak{C}_x. \tag{2.10}$$

The unweighted identity is [17, Proposition 3.4], obtained from [15, Lemma 5.1 and equation (5.13)]. For a bounded profile of total support at most S, its counting hypothesis is $2\rho_* S < 1$; the present value is checked in Section 7. It gives

$$\sum_{\substack{x \le n \le 2x \\ n \equiv b_0 \pmod W}} A_F(n)^2 = \mathfrak{C}_x (I(F) + o(1)).$$

The prime-weighted counterpart is controlled in units of $\mathfrak{P}_x$. Consequently the final comparison is ρ_* times a face form versus $I(F)$. The factor ρ_* is supplied by (2.10), not by numerical optimization.

2.5 The exact erased array and its limiting profile

For a k-coordinate array y, define

$$y^{(i)}_{\widehat{\mathbf{r}}_i} = \frac{1}{B_R} \sum_{a \ge 1} \frac{y_{r_1,\dots,r_{i-1},a,r_{i+1},\dots,r_k}}{\varphi(a)}. \tag{2.11}$$

The zero extension of y retains its original squarefree and coprimality conditions; an a sharing a prime with another coordinate contributes zero. Let $A_i^0 = B_{y^{(i)},i}$ be the canonical sum in the remaining $k-1$ coordinates. Setting $d_i = 1$ in (2.9) gives this exact array and its factor B_R^{-1}. This is the exact erasure identity of [17, Lemma 3.3]. Therefore $A_F = A_i^0$ whenever every nontrivial coefficient divisor in the detected coordinate is excluded by roughness.

The arithmetic-to-continuous marginal statement uses the harmonic norm

$$\|y\|_{R,d}^2 = \frac{1}{B_R^d} \sum_{\mathbf{r}} \frac{|y_{\mathbf{r}}|^2}{\prod_j \varphi(r_j)}.$$

For a fixed smooth bounded profile, the realization lemmas of [17, Proposition 3.7 and Lemma 3.8] give

$$\|y[F]\|_{R,k}^2 \longrightarrow I(F), \qquad \|y^{(i)}[F] - y[E_i F]\|_{R,k-1}^2 \longrightarrow 0. \tag{2.12}$$

They also give the corresponding inner-product limits for finite fixed combinations.

The mechanism behind (2.12) is useful to keep in view. Harmonic sums of the small cofactor have the Dickman limit, while each recorded prime contributes du/u; their product is (2.2). Excluding primes shared by different coordinates costs a vanishing bounded-form error after presieving, because their local contribution is $O_k(p^{-2})$. Expanding the square in the second limit gives two independently summed erased coordinates and shared retained coordinates, exactly as in (2.8). Its three terms converge to $J_i(F), -2J_i(F), J_i(F)$. This proves mean-square convergence of the marginal; it does not assert pointwise convergence of the array entries.

These are the realization facts used by the hybrid pairing identities. The passage from smooth profiles to the fixed bounded grid profile is made at the end of the argument, with all supports and approximation choices held fixed before the arithmetic limit.

2.6 A uniform bound for all erasure operators

The following estimate adapts the weighted Cauchy–Schwarz argument of Polymath8b [20, Lemma 6.1 and Corollary 6.4], which bounds the Maynard variational quantity by $k \log k/(k-1)$. Here the total support radius is S, and the Dickman density is at most one. These two observations give the fragment-measure version below, also after compression to a smaller domain.

Lemma 2.1. *On profiles supported in $\sum_i |X_i| \le S$, one has*

$$\sum_{i=1}^{k} \|E_i F\|_2^2 \le \frac{Sk \log k}{k-1} \|F\|_2^2.$$

Proof. Write $t_i = |X_i|$ and $\widehat{s}_i = \sum_{j \ne i} t_j$. For fixed retained coordinates with $\widehat{s}_i < S$, use the positive weight

$$w_i = S - \widehat{s}_i + (k-1)t_i.$$

Since $\rho_D(t_i/\zeta) \le 1$, the integral of w_i^{-1} over the permitted erased configurations is at most

$$\int_0^{S-\widehat{s}_i} \frac{dt}{S - \widehat{s}_i + (k-1)t} = \frac{\log k}{k-1}.$$

Weighted Cauchy–Schwarz bounds $|E_i F|^2$ by this number times $\int w_i |F|^2 \, d\nu(X_i)$. Integrating over the retained coordinates and summing over i proves the claim because $\sum_i w_i = kS$. The boundary $\widehat{s}_i = S$ has zero fiber measure. Extra cap restrictions only decrease the first integral. $\square$

In operator notation, on $L^2(\mathcal{H}_O, \nu^k)$ with functions extended by zero outside $\mathcal{H}_O$, this says

$$\sum_i E_i^* E_i \le C_E \mathrm{Id}, \qquad C_E = \frac{Sk \log k}{k-1}. \tag{2.13}$$

The adjoints are compressed back to $\mathcal{H}_O$. For the present S and $k = 39$, $C_E < 4$. An entirely rational verification uses $\sum_{j=0}^{32} (367/100)^j / j! > 39$, hence $\log 39 < 3.67$.

3 Distribution estimates and the actual domains

3.1 What a source estimate must say

For a sequence f supported on the relevant dyadic interval, define

$$\Delta(f; q, a) = \sum_{n \equiv a \pmod{q}} f(n) - \frac{1}{\varphi(q)} \sum_{(n,q)=1} f(n).$$

A distribution estimate controls a sum of $|\Delta(f; q, a(q))|$ over a specified family of squarefree moduli, with an arbitrarily large logarithmic saving. Here the residues $a(q)$ form a *coherent primitive system*: for a prime dividing two moduli, the prescribed local residue is the same, and is nonzero. The estimates retain the fixed divisor-power weights needed by the coefficient transform; the passage from unweighted discrepancy bounds to these weights is [17, Proposition 2.11].

This is the uniformity required by the sieve. Expanding a mixed coefficient sum produces moduli $W[D, E]$ and residue classes determined by the shifts. Presieving makes these classes primitive. Their coordinate assignments can vary with the coefficient pair, so they do not immediately form a single coherent system. The reduction in [17, Proposition 3.5] averages over global primitive classes whose residue at a prime belongs to $\{h_i - h_j : j \neq i\}$, for the detected coordinate i. It costs at most $[3(k-1)^2]^{\nu(q/W)}$, where $\nu(m)$ counts distinct prime factors. This is one of the fixed divisor-power weights retained in the source estimate. The modulus family is independent of the global residue choice. In this way the coherent-class theorem supplies the required mixed sums without asserting a maximum over unrelated residue choices inside the modulus sum.

For orientation, a Type II estimate treats a convolution $\alpha_M * \beta_N$ with $MN \asymp x$. Its coefficients are divisor bounded and one specified factor satisfies the Siegel–Walfisz condition. A Type III estimate treats a residual divisor-bounded sequence convolved with three smooth sequences. The labels describe which factors can supply cancellation. They do not mean that the three primes detected by a sieve are being counted; only two primes are sought here. The coefficient and smoothness conventions, including the auxiliary coprimality condition in Siegel–Walfisz, are [19, Definition 2.5], restated in [17, Definitions 2.5–2.9].

3.2 Dense divisibility and the source orders

An integer is Y-smooth if all its prime factors are at most Y. The weaker notion used here preserves the ability to choose divisors in prescribed ranges. We use the recursive definition in [19, Definition 2.1].

Definition 3.1. Put $\mathscr{D}^{(0)}(Y) = \mathbb{Z}_{>0}$. For $r \geq 1$, an integer q belongs to $\mathscr{D}^{(r)}(Y)$ if, for every $j, l \geq 0$ with $j + l = r - 1$ and every $1 \leq U \leq Yq$, there is a factorization

$$q = uv, \qquad U/Y \leq v \leq U, \qquad u \in \mathscr{D}^{(j)}(Y), \quad v \in \mathscr{D}^{(l)}(Y).$$

At order one this asks for a divisor near every target size. At order three it asks, in particular, for the allocation $(j, l) = (1, 1)$: both selected factors are singly densely divisible. The definition requires *every* allocation. This strong recursive class is essential in the application of the new Type II estimate.

Every smooth integer belongs to every such class [19, Lemma 2.10(iii)], but a theorem stated only for smooth moduli cannot be applied to a dense class without an additional argument. The incidence companion [10, Corollaries 9.2–9.3] provides that argument for the source family used here. We spell out its required divisor choices below.

3.3 The new analytic estimates

Write the distribution level as $\theta = \frac{1}{2} + 2\omega$. The incidence Type II estimate [10, Theorem 2.1 and Corollary 9.3] is applied to $\alpha_M * \beta_N$, where $MN \asymp x$, the coefficients are divisor bounded, and β_N has the source theorem's Siegel–Walfisz property. With $N = x^\gamma$, its strict conditions are

$$\begin{gathered} 12\omega + 6\delta < \gamma < \tfrac{1}{2} - 2\omega, \qquad 2\gamma + 4\omega + 2\delta < 1, \\ \gamma > 16\omega + 7\delta, \qquad 5\gamma > \tfrac{3}{2} + 40\omega + 16\delta. \end{gathered} \tag{3.1}$$

The companion proves the estimate by an incidence-operator bound inside the original positive dispersion square. The dependence of Fourier coefficients on a short modulus variable must be separated before that bound is applied. The divisor and residue masks also remain in the calculation. These issues explain why the source theorem has more content than a bound for an isolated exponential sum.

The Type III input concerns $\alpha * \psi_1 * \psi_2 * \psi_3$, where α is divisor bounded and each ψ_i is smooth. It requires

$$\begin{gathered} MN_1N_2N_3 \asymp x, \qquad x^{2\sigma} \ll N_i \ll x^{1/2-\sigma}, \\ N_iN_j \gg x^{1/2+\sigma} \quad (i \neq j), \qquad 0 < \sigma < \tfrac{1}{6}, \end{gathered} \tag{3.2}$$

together with

$$\sigma > \tfrac{1}{30} + \tfrac{56}{15}\omega + \tfrac{4}{15}\delta, \qquad 108\omega + 2\delta > 1. \tag{3.3}$$

It applies to squarefree singly densely divisible moduli. The distribution statement is [11, Theorem 1.1]; its new local input is the finite-exceptional Fourier estimate [11, Proposition 3.1]. The restrictions on all three individual lengths and on all pair products in (3.2) are part of the theorem; a bound on their product alone would not suffice.

The old ladder uses genuine-prime distribution rather than minorant distribution. Its source is [17, Corollary 2.19], with the following sufficient conditions on the corresponding strong dense class:

$$\begin{aligned} 108\omega + 30\delta &< 1, \quad \text{order 1},\\ 280\omega + 80\delta &< 3, \quad \text{order 2},\\ 240\omega + 80\delta &< 3, \quad \text{order 3}. \end{aligned} \tag{3.4}$$

The order-one and order-two cases use Polymath8a's bilinear and Type III estimates [19, Theorem 2.8(i),(ii),(iv),(v)]. The order-three case uses Stadlmann's bilinear estimate [21, Proposition 1], extended from smooth to triply densely divisible moduli in [17, Proposition 2.18 and Appendix A.4], together with Polymath8a's Type III estimate. Triple dense divisibility describes the moduli; it does not identify a Type III convolution estimate. The common passage to genuine primes is the Heath–Brown reduction [19, Lemma 2.7]. The resulting statement includes uniformity on subintervals of $[x, 2x]$. The old rows in the digital table satisfy the appropriate line of (3.4). The two additional order-two sources needed for the inner pairings are checked separately in Section 3.7.

The final application fixes

$$a = .41361, \qquad I_0 = .34941, \qquad c = .41364, \qquad \tau = 10^{-10}. \tag{3.5}$$

Here I_0 is a cutoff in the arithmetic decomposition, not the norm $I(F)$. The new order-three Type II band is

$$a - \tau \le \gamma \le .45.$$

The upper bilinear range is supplied by [19, Theorem 2.8(iv)], with the handoff conditions in [10, Section 10]. Smooth Type I uses $I_0 - \tau > 1/4 + 7\omega + 2\delta$, as in [2, Lemmas 3 and 5] and [10, Proposition 10.1(i)]. The complementary coefficient has the Siegel–Walfisz property required by that input. The short smooth-factor estimate on dense moduli is [17, Lemma 2.20 and Appendix A.5]. The lower-order Type II conditions are

$$54\omega + 15\delta + 5\sigma_m < 1, \qquad 56\omega + 16\delta + 4\sigma_m < 1, \qquad \sigma_m = \tfrac{1}{2} - a + \tau,$$

at orders one and two respectively. These are [19, Theorem 2.8(i),(ii)], in its Type I terminology, combined with its part (iv) and the additional condition $68\omega + 14\delta < 1$; see [17, Proposition 2.16]. The Type III choice is $\sigma = \frac{1}{2} - c - \tau$, with $\bar{\omega} = \max(\omega, .01)$ when necessary. The enlargement of ω is used only where all conditions of that Type III source are checked at the enlarged value.

3.4 Why these estimates apply to the dense source family

For a retained order-three row, the geometric conditions give $q \in \mathscr{D}^{(3)}(x^{\delta_g})$, where

$$\delta_g = \frac{\rho_*}{\rho}\delta < \delta.$$

Choose an intermediate δ' and then a small positive ε, both fixed. The allocation $(1, 1)$ in Definition 3.1 supplies $q = r(q/r)$ with both factors singly dense and

$$x^{\gamma - 3\varepsilon - \delta'} \le r \le x^{\gamma - 3\varepsilon}.$$

The two further targets in the dispersion argument are

$$u \mid q/r: \quad U = \frac{x^{1-\gamma-6\varepsilon}}{q}, \qquad d_1 \mid r: \quad D_1 = \frac{r^2 x^{2-\gamma-52\varepsilon}}{q^4}. \tag{3.6}$$

These are the three intervals in [22, Lemma 6 and Section 3.3.6] and [3, Lemma 5.6 and Section 6.4]. Their use with the new incidence estimate is proved in [10, Definition 9.1 and Corollaries 9.2–9.3]. The earlier sources supply the factorization scheme; the new parameter range (3.1) is the conclusion of the incidence companion. It does not require every divisor encountered in the proof to remain densely divisible with an unchanged parameter.

Here are the size checks that make the targets legal. For $q \geq x^{1/2-\varepsilon}$, their ratios to the available factors are at most $x^{-7\varepsilon}$ and $x^{-51\varepsilon}$. For $q \leq x^{1/2+2\omega}$, their lower exponents are at least

$$\tfrac{1}{2} - \gamma - 2\omega - 6\varepsilon, \qquad \gamma - 8\omega - 2\delta' - 58\varepsilon.$$

Both are positive on every actual row, after choosing ε sufficiently small. Thus the targets lie inside the ranges where the two singly dense factors can be used. The companion source-transfer argument retains the same local estimates after these choices.

The strict gaps between ρ_* and ρ, and between $\delta_g, \delta', \delta$, absorb the presieving factor, fixed divisor weights and fixed subdivisions. Below the initial half-level range we use the bilinear Bombieri–Vinogradov theorem in [19, Theorem 2.9]. Polymath8a attributes its underlying principle to Gallagher and Motohashi and cites [4, Theorem 0] for the proof. The prime and convolution versions with the present interval conventions are [17, Proposition 2.15]. There is no need to extend the new dispersion argument to a target of size less than one.

3.5 Writing the arithmetic domains as explicit predicates

All fragments of a multi-coordinate configuration are collected into one multiset for the following tests. Write s for their total. For a cutoff ξ, put

$$M_\xi = \max(\{u : u > \xi\} \cup \{0\}).$$

If ϕ is nonnegative, nondecreasing and vanishes at zero, define

$$H_{\phi,\xi} = \max_{u>\xi} \left(\sum_{v \geq u} v + \phi(u) \right), \tag{3.7}$$

with an empty maximum equal to zero. The sum includes the marked fragment and retains multiplicities. We abbreviate $H_{m,\xi} = H_{\phi,\xi}$ for $\phi(u) = (m-1)u$. Deleting fragments decreases all these quantities.

The meaning of (3.7) is straightforward: a large prime is permitted when enough of the divisor's logarithmic mass lies below it to support the required factorization. Exponentiating a tail inequality converts it into a prime-factor criterion for dense divisibility.

A source row contains $\omega_{\text{prev}}, \omega, \delta$, an order $r \in \{1,2,3\}$, and the two radius bounds S, T. Put

$$B = \frac{1/2 + 2\omega_{\text{prev}}}{\rho}, \quad \xi = \frac{\delta}{\rho}, \quad a_r = B - T, \quad b_r = B - S,$$

$$\eta = \begin{cases} \xi, & r = 1, 2, \\ (\xi + S + T - B)/2, & r = 3, \end{cases} \qquad A_r = a_r + \eta, \quad C_r = b_r + \eta.$$

The subscript on a_r distinguishes this row threshold from the fixed five-prime cutoff a.

At orders one and two the outer predicate is

$$O_r = \{s \le a_r\} \ \cup\ \{H_{2,\xi} \le A_r\}. \tag{3.8}$$

At order one there is no additional inner predicate. At order two it is

$$I_r = \{s \le b_r\} \ \cup\ \{H_{2,\xi} \le C_r\}. \tag{3.9}$$

The first alternative in each formula is important: below the activation size the extra fragment test is unnecessary. The lower-order transfer also has the size guards in [17, Lemma 4.1]; these are checked on each row.

At order three use

$$\phi_D(u) = \min(3u/2, L), \qquad \phi_E(u) = 3u - \phi_D(u).$$

Both functions are nondecreasing and sum to $3u$. The outer and inner predicates become

$$O_r = \{s \le a_r\} \ \cup\ \{H_{\phi_D,\xi} \le A_r,\ \phi_E(M_\xi) \le C_r\}, \tag{3.10}$$

$$I_r = \{s \le b_r\} \ \cup\ \{H_{\phi_E,\xi} \le C_r,\ \phi_D(M_\xi) \le A_r\}. \tag{3.11}$$

The opposite-root conditions involving M_ξ are needed when only one of the two divisors contains a large prime. They are not optional largest-prime simplifications. The budget $A_r + C_r \le B + \xi$ makes [17, Proposition 4.2] applicable.

The old ladder uses genuine-prime sources and the new ladder uses the sharp minorant. Their order-three plateaux differ:

$$\begin{aligned}\Lambda_{\mathrm{old},r} &= \mathrm{clip}\left(\frac{3A_r - C_r}{4}; \frac{3\max(A_r, C_r)}{7}, \frac{3\min(A_r, C_r)}{5}\right),\\ \Lambda_{\mathrm{new},r} &= \frac{23}{40}C_r,\end{aligned}$$

where L is chosen as the corresponding $\Lambda_{\mathrm{old},r}$ or $\Lambda_{\mathrm{new},r}$, and $\mathrm{clip}(z; l, u) = \min(u, \max(l, z))$. All intervals used here have legal endpoints. The active order-three outer largest-fragment bound is $\min(A_r - L, (C_r + L)/3)$; the inner bound is $(C_r + L)/4$. Retaining only one term in the outer minimum would describe a larger domain.

3.6 The four domains used in the proof

Let $O_{\mathrm{old}}, O_{\mathrm{new}}$ be the intersections of the retained outer predicates in the two ladders, with the global radius and prime caps. Let $L_{\mathrm{old}}, L_{\mathrm{new}}$ be the corresponding raw inner domains. The finite grid provides cap domains

$$\mathcal{H}_0 \subset \mathcal{H}_1 \subset \mathcal{H}_C, \qquad \mathcal{H}_O$$

whose exact construction is given in Section 7. Define

$$\begin{aligned} O &= \mathcal{H}_O \cap O_{\mathrm{old}} \cap O_{\mathrm{new}},\\ L_0 &= \mathcal{H}_0 \cap L_{\mathrm{old}} \cap L_{\mathrm{new}},\\ L_1 &= \mathcal{H}_1 \cap L_{\mathrm{new}},\\ \mathcal{C} &= \mathcal{H}_C \cap \mathcal{C}_{14}^{\mathrm{raw}}. \end{aligned} \tag{3.12}$$

The last domain is the subtraction domain. Its common order-two source is the specified row-14 source and its physical radius is

$$.2607860319243637\ldots.$$

The exact row data and inward cap intersections give

$$L_0 \subset L_1 \subset \mathcal{C}.$$

The raw old and new inner domains need not be nested. Their intersection in L_0 is what makes the first inclusion immediate.

The following table states precisely the arithmetic information used later. A pairing refers to all moduli generated by coefficient roots in the two indicated domains.

Pairing	Required sequence	Purpose
$O \times O$	ordinary counting	Mass of the sieve square
$O \times L_0$	prime indicator	Base prime correlation
$O \times L_1$	sharp minorant	Additional mixed correlation
$L_1 \times L_1$	both	Inner prime and exceptional self forms
$\mathcal{C} \times L_1$	both	Canonical subtraction against the correction

The last row includes the fourth because $L_1 \subset \mathcal{C}$. We list the fourth separately to identify the self forms in the hybrid proof. There is no use of a $\mathcal{C} \times \mathcal{C}$ distribution estimate.

There are 43 retained old rows and 55 retained new rows. For each ladder the first level is $B_{\text{first}} = 1/(2\rho)$. If

$$B_r^+ = \frac{1/2 + 2\omega_r}{\rho}$$

is the upper level of a retained row, the next row has $B_{\text{next}} = B_r^+$. Thus there is no gap between successive modulus ranges. At the last row the exact rational data satisfy

$$B_{\text{last}}^+ \geq h\left[(N-1) + \min(\lfloor T/h \rfloor, N-2)\right], \quad T = \begin{cases} T_0, & \text{old ladder,} \\ T_1, & \text{new ladder.} \end{cases} \tag{3.13}$$

The right side is the largest permitted sum of the two whole-cell root bounds on the fine grid in Section 7: the outer total is at most $(N-1)h$, and the retained total is at most the indicated minimum. Hence every mixed modulus above the initial Bombieri–Vinogradov range is covered by a row. The hereditary predicates and the two-root transfer put its least common multiple in that row's dense class. The strict scaling $\rho_* < \rho$ leaves room for the presieving factor W, including at an upper endpoint.

Checking the displayed analytic inequalities and transfer guards on their exact rational parameters is a finite task. The current interface check has 638 scalar source inequalities and 31 explicit order-three extraction witnesses. These counts describe the coverage of the application; they are not substitutes for the distribution theorems or for the transfer arguments just stated.

3.7 The self source and the subtraction domain

We now justify the two inner pairings in the preceding table. This also proves the inclusion $L_1 \subset \mathcal{C}$ used by the hybrid inequality. A largest-fragment cap alone would not establish these statements: the full activated tail predicate must be retained. The transfer criterion is [17, Lemma 4.1(ii)]; its equal-radius case is [17, Lemma 4.3]. The source parameters and domain inclusions checked below belong to the present profile.

For a pooled fragment configuration of total s, abbreviate

$$\mathcal{Q}(c_0, \xi, U) = \{s \le c_0\} \ \cup \ \{H_{2,\xi} \le U\}.$$

There is a simple monotonicity principle:

$$c_0' \ge c_0, \qquad \xi' \ge \xi, \qquad U' \ge U \quad \Longrightarrow \quad \mathcal{Q}(c_0, \xi, U) \subset \mathcal{Q}(c_0', \xi', U'). \tag{3.14}$$

Indeed, the core alternative is preserved by increasing its endpoint. Otherwise, increasing the activation removes possible witnesses from the maximum defining $H_{2,\xi}$, and increasing the threshold weakens the remaining test. This argument applies to configurations, not merely to their largest fragments.

Here are the exact parameters needed for the present application. Use the ρ, ρ_* of (7.1), and put

$$T_1 = \frac{1262625033}{1312494500}, \qquad B_0 = \frac{1}{2\rho}, \qquad \epsilon_r = 10^{-7}, \qquad U_{\mathrm{n}} = T_1 + \frac{\epsilon_r}{\rho}.$$

The new ladder's rows numbered 13 and 14 both have order two. Their inner predicates are $\mathcal{Q}(b_j, \xi_j, U_{\mathrm{n}})$, with

$$b_{13} = \frac{37058221276498005595947498293}{40868109574262446536646110000}, \qquad \xi_{13} = U_{\mathrm{n}} - b_{13}, \tag{3.15}$$

$$b_{14} = \frac{26070004679778661762784174330 3}{286076767019837125756522770000}, \qquad \xi_{14} = U_{\mathrm{n}} - b_{14}. \tag{3.16}$$

The row numbers here are those of the supplied rational ladder, starting at zero. Every element of L_1 satisfies both predicates because $L_1 \subset L_{\mathrm{new}}$.

The common self source. Take

$$\omega_s = \frac{21}{5000}, \qquad \delta_s = \frac{2279}{100000}, \qquad \xi_s = \frac{\delta_s}{\rho} = \frac{22790}{262499}, \qquad c_s = B_0 - T_1.$$

The required order-two predicate on either root of radius T_1 is $\mathcal{Q}(c_s, \xi_s, c_s + \xi_s)$. The following directed bounds follow from the displayed rationals:

difference	strict lower bound	strict upper bound
$c_s - b_{13}$	.035989096083	.035989096084
$\xi_s - \xi_{13}$	.031590949551	.031590949552
$c_s + \xi_s - U_{\mathrm{n}}$	.067580045635	.067580045636

Thus (3.14) shows that new row13 implies the self-source predicate. The lower-order transfer guard and level checks are

$$B_0 + \xi_s > T_1, \qquad 2\rho T_1 < \tfrac{1}{2} + 2\omega_s.$$

The first difference exceeds 1.02958; the second has margin greater than .00334979. The actual physical level has the still larger margin

$$\tfrac{1}{2} + 2\omega_s - 2\rho_* T_1 = .0033499868.$$

Consequently the order-two transfer applies to every pair of roots in L_1, with strict size and density retreats.

The larger subtraction source. Its exact radius and source parameters can be given economically in terms of row14:

$$\begin{aligned} S_C &= B_0 - b_{14} - \frac{\epsilon_r}{\rho_*}, \qquad \omega_C = \frac{\rho(S_C + T_1) - 1/2 + \epsilon_r}{2}, \\ \sigma_0 &= .100001, \qquad \delta_C = \frac{1 - 4\sigma_0}{16} - \frac{7}{2}\omega_C - \epsilon_r, \qquad \xi_C = \frac{\delta_C}{\rho}. \end{aligned} \tag{3.17}$$

These are fixed rational definitions. For orientation,

$$\begin{aligned} .2607860319243 &< \rho_* S_C < .2607860319244, \\ .0066556670361 &< \omega_C < .0066556670362, \\ .0142048153735 &< \delta_C < .0142048153736. \end{aligned}$$

Set

$$a_C = B_0 - T_1 = c_s, \qquad b_C = B_0 - S_C.$$

The raw subtraction domain and its actual inward version are

$$\begin{aligned} \mathcal{C}_{14}^{\mathrm{raw}} &= \{s \le S_C,\ M_0 \le \zeta\} \cap \mathcal{Q}(a_C, \xi_C, a_C + \xi_C), \\ \mathcal{C} &= \mathcal{H}_C \cap \mathcal{C}_{14}^{\mathrm{raw}}. \end{aligned} \tag{3.18}$$

The complementary predicate required of the root from L_1 is $\mathcal{Q}(b_C, \xi_C, b_C + \xi_C)$. It follows from new row14: the three differences in (3.14) satisfy

$$\begin{aligned} b_C - b_{14} &= \frac{1}{2624989} > 0, \\ .003403370303 &< \xi_C - \xi_{14} < .003403370304, \\ .003403751257 &< b_C + \xi_C - U_{\mathrm{n}} < .003403751258. \end{aligned}$$

The required lower-order guard is also strict:

$$B_0 + \xi_C > \max(2S_C - T_1,\ 2T_1 - S_C). \tag{3.19}$$

Its margin exceeds .93393748. The other size and density checks are

$$\begin{aligned} \tfrac{1}{2} + 2\omega_C - \rho(S_C + T_1) &= 10^{-7}, \\ \tfrac{1}{2} + 2\omega_C - \rho_*(S_C + T_1) &> 2.9554 \cdot 10^{-7}, \\ \delta_C - \rho_* \xi_C &> 5.411 \cdot 10^{-9}, \\ \tfrac{1}{2} - \rho_* B_0 &= \frac{1}{5249980} > 0. \end{aligned} \tag{3.20}$$

They allow the presieving factor and all fixed subdivisions in the usual order of limits.

For clarity, the order-two source estimates used for these two pairs of parameters are the following. For genuine primes, combine [17, Propositions 2.16–2.17], which restate [19, Theorem 2.8(ii),(iv),(v)], with the Heath–Brown reduction [19, Lemma 2.7]. The original identity is [9, Lemma 1, equation (6)]; the smooth decomposition and grouping used here are those of Polymath8a. The reduction uses $\sigma_0 > 1/10$, as well as

$$56\omega + 16\delta + 4\sigma_0 < 1, \qquad 68\omega + 14\delta < 1, \qquad \sigma_0 > \frac{1}{18} + \frac{28}{9}\omega + \frac{2}{9}\delta.$$

For the sharp minorant use [10, Theorem 11.1 and Proposition 10.1], followed by the addback in Section 4, and put $\sigma_m = 1/2 - a + \tau$, $\sigma_3 = 1/2 - c - \tau$, and $\bar{\omega} = \max(\omega, .01)$. In addition to the common Harman cutoff conditions, the requirements are

$$56\omega + 16\delta + 4\sigma_m < 1, \qquad 68\omega + 14\delta < 1, \qquad a + \tfrac{1}{2} + 2\omega < 1,$$
$$I_0 - \tau > \tfrac{1}{4} + 7\omega + 2\delta,$$
$$\sigma_3 > \tfrac{1}{30} + \tfrac{56}{15}\bar{\omega} + \tfrac{4}{15}\delta, \qquad 108\bar{\omega} + 2\delta > 1.$$

At the self source the first genuine-prime inequality has margin 39/250000; at the subtraction source it has margin 1/625000. Every other listed inequality has margin greater than .0096 at both sources. These assertions are direct rational substitutions into the displayed formulas, so they can be checked independently of an integration program.

We can now conclude both the support inclusions and the arithmetic pairings. We have $T_1 < S_C < S$, hence $a_C > b_C$. Increasing both the core and threshold from b_C to a_C shows that the complementary predicate also implies the predicate in (3.18). Since L_1 has radius at most T_1, respects the global prime cap, and is contained in $\mathcal{H}_1 \subset \mathcal{H}_C$, this proves

$$L_1 \subset \mathcal{C}.$$

Applying the two-root order-two transfer to (3.18) and its complementary predicate, using (3.19)–(3.20), gives both the genuine-prime and sharp-minorant estimates on $\mathcal{C} \times L_1$. The self-source argument gives the same two estimates on $L_1 \times L_1$. Only these pairings are claimed: no source estimate on $\mathcal{C} \times \mathcal{C}$ is needed.

4 The sharp five-prime minorant

The starting Harman decomposition is [10, Theorem 11.1 and Section 11]. It follows the Buchstab decompositions of [2, Section 3, Lemmas 7–13] and [21, Sections 4.3–4.8]; [3, Proposition 7.2] records a version with variable cutoffs. The addback below identifies the sharp exceptional set for the parameters (3.5). Its coefficient and mass are determined here.

4.1 Its support and its exact meaning

Set

$$\lambda = 1 - 2a = .17278.$$

On $x \le n \le 2x$, let $\mathcal{S}_5(a;x)$ consist of products of five distinct primes such that every pair product is less than x^a. Equivalently, order the primes decreasingly and require only $p_1 p_2 < x^a$. The minorant is the literal arithmetic function

$$\varrho_x(n) = \mathbf{1}_{\mathbb{P}}(n) - b_x(n), \qquad b_x(n) = 241\mathbf{1}_{\mathcal{S}_5(a;x)}(n). \tag{4.1}$$

It is at most the prime indicator pointwise. Its exceptional part b_x is nonnegative and supported on composites.

Every exceptional integer is x^λ-rough. To see this, fix one of its five primes p_i. Partition the other four into two pairs. Their product is less than x^{2a}, so

$$p_i = \frac{n}{\prod_{j \ne i} p_j} > \frac{x}{x^{2a}} = x^\lambda.$$

The inequality uses $n \geq x$. It would not be a statement about all positive integers independently of the interval.

The global coefficient-prime cap is strictly below $x^{.17277}$. Therefore a coefficient prime cannot divide a prime $n+h_i$, or an exceptional $n+h_i$, at the large scales in question. The detected divisor coordinate is exactly one on both supports. This is why the same erased-coordinate identity can be used for the prime and exceptional weights.

4.2 How the coefficient 24 arises

The Harman construction in [10, Section 11, "The first five-prime error" and "Reversal and the sign of the second error"] has two ordered five-prime errors E_1, E_2. The first has the labeling

$$n = p_1p_2p_3p_4p_5, \qquad x^{\lambda} < p_4 < p_3 < p_2 < p_1 < x^a,$$
$$p_1p_2 < x^a, \qquad p_2p_3p_4 > x^{1-a}, \qquad p_5 > p_4.$$

The second has

$$n = p_2p_3p_4p_5p_6, \qquad x^{\lambda} < p_i < x^{8a-3} \quad (2 \leq i \leq 6),$$
$$p_2, p_4 > p_3, \quad p_2p_4 < x^a, \quad p_3p_4p_5 > x^{1-a}, \quad p_6 > p_5.$$

The original minorant is $\mathbf{1}_{\mathbb{P}} - E_1 - E_2$.

Call a pair *eligible* if its product is at least x^a. Partition each error according to the first eligible pair in a fixed lexicographic order on the labels. If there is no such pair, place the tuple in a residual class. Denote the eligible and residual counts by A_j, B_j. Then

$$E_j = A_j + B_j, \qquad \mathbf{1}_{\mathbb{P}} - E_1 - E_2 + A_1 + A_2 = \mathbf{1}_{\mathbb{P}} - B_1 - B_2.$$

The choice of a first pair assigns every tuple once.

The eligible pair is a Type II factor. All five primes exceed x^{λ}, hence its product lies in

$$[x^a,\, 2x^{1-3\lambda}] = [x^a,\, 2x^{6a-2}].$$

Here $6a-2 = .48166 < 1/2$. The endpoint padding and the handoff at .45 in Section 3 put this whole range in the available Type II band. Thus A_1, A_2 can be added back without losing the distribution estimate.

On a residual tuple every pair is below x^a, so every complementary triple exceeds x^{1-a}. In addition,

$$np_i^3 = \prod_{j \neq i}(p_ip_j) < x^{4a}, \qquad p_i < x^{(4a-1)/3}.$$

The inequalities $a > 2/5$ and $3I_0 + 7a < 4$ imply

$$\frac{4a-1}{3} < \min(a, 8a-3, 1-I_0-a).$$

Consequently the original individual caps and triple conditions are automatic on the residual set.

For a fixed set of five distinct primes, B_1 has four labelings: p_4 is the smallest, p_5 can be any of the other four, and the other labels are then ordered. The count B_2 has twenty labelings: choose the three primes at labels $2, 3, 4$ in ten ways, put their minimum at label 3, and assign the other two in either order. The last two primes have a forced order at labels 5 and 6. Therefore

$$(B_1 + B_2)\mathbf{1}_{\text{squarefree}} = 24\mathbf{1}_{\mathcal{S}_5(a;x)}.$$

This proves the coefficient in (4.1); there is no unrecorded factorial.

4.3 Why the addback has the required distribution

The eligible-pair argument must control progression discrepancies, not just the unweighted number of discarded tuples. Here are the relevant details. Partition each prime variable into intervals of relative length $(\log x)^{-B_0}$, with B_0 fixed and sufficiently large. Inside a box lying entirely in a selected-pair region, all order and product indicators are fixed. The pair and its complement give prime-product convolution coefficients. Their positive polynomial scales supply the required Siegel–Walfisz property: prime intervals are covered by [17, Proposition 2.10], and convolution preserves this property by the argument in [19, proof of Lemma 3.4(iii)].

The passage from boxes to the selected-pair region is [17, Proposition 2.13], extending the subdivision in [20, Section 4.5, proof of equation (63)]. Its hypotheses require distribution for the positive boundary majorants as well as the interior boxes. Boundary boxes lie in finitely many thin product or ratio strips. Fixing the other variables confines a variable with nonzero boundary coefficient to a short relative interval. Positive box majorants stay inside the padded Type II range. Applying Type II to both the interior boxes and these majorants controls their summed progression discrepancies. Increasing B_0 and the requested logarithmic saving absorbs the polynomial number of boxes and $\sum_{q\le x^\theta} 1/\varphi(q) \ll \log x$. In particular, a possible pair in $[x^a, 2x^a)$ in E_1 is treated by the same argument. It is not deleted by replacing $\log_x n$ with one.

Repeated-prime tuples have a prime square $p^2 \mid n$ with $p > x^\lambda$. For primitive residues their progression and coprime mean contributions vanish when $p \mid q$; otherwise CRT applies. With the fixed divisor-power factors absorbed into $x^{o(1)}$, the discrepancy sum is at most

$$x^{1-\lambda+o(1)} + x^{\theta+a+o(1)}.$$

Both are power savings on every row, since $\theta + a < 1$. These terms can also be added back. This establishes the literal sharp minorant with the same distribution inputs.

4.4 The mass deficit

Its exact limiting deficit has a simple integral description. Let

$$\mathcal{P}_a = \left\{ (\alpha_1, \ldots, \alpha_4) : \alpha_1 > \cdots > \alpha_5 > 0, \quad \alpha_5 = 1 - \sum_{j=1}^{4} \alpha_j, \quad \alpha_1 + \alpha_2 < a \right\}.$$

Then

$$\kappa_{\text{true}} = 24 \int_{\mathcal{P}_a} \frac{d\alpha_1\, d\alpha_2\, d\alpha_3\, d\alpha_4}{\alpha_1\alpha_2\alpha_3\alpha_4\alpha_5}. \tag{4.2}$$

The inequalities force $\alpha_i \ge \lambda$ on the closure of the region, so there is no singularity at a zero prime exponent. Apply the prime number theorem and partial summation as in [21, proofs of Lemmas 21–22], with the ordered region $\mathcal{P}_a$ in place of the regions there. This gives

$$\sum_{x \le n \le 2x} \varrho_x(n) = (1 - \kappa_{\text{true}} + o(1)) \frac{x}{\log x}.$$

The ordering in $\mathcal{P}_a$ counts a distinct factorization once. The factor 24 is the multiplicity just proved.

Exact polytope integration with outward rational bounds gives

$$.000093307088324061 < \kappa_{\text{true}} < .000093307088324062.$$

The proof uses the convenient upper bound

$$\bar{\kappa} = \frac{46653544162031}{500000000000000000}. \tag{4.3}$$

The same arithmetic function and the same κ_{true} occur at every source row; changing a density parameter does not change the minorant. Since coefficient primes lie below its roughness threshold, the coprime mean on these moduli equals the full mean.

4.5 A reproducible enclosure of the sharp mass

We give the calculation behind (4.3). It is a four-dimensional integral with a uniformly convergent series; it does not use a quadrature error estimate inferred from successive meshes. The exact region is the one derived in Section 4. The enclosure below evaluates its mass for the present cutoff a.

Let $\kappa_{5,1}$ denote the limiting mass of the first labeled error E_1. On the hyperplane $\sum_i \alpha_i = 1$, its two pair constraints are $\alpha_1 + \alpha_2 < a$ and $\alpha_1 + \alpha_5 < a$, the latter being equivalent to its complementary triple constraint. Along with the order of labels 1 through 4, they imply that every pair sum is below a. Thus this limiting region is exactly four labeled copies of the ordered region in (4.2). Permuting five coordinates on $\sum_i \alpha_i = 1$ has absolute Jacobian one in the four independent coordinates, and the density is symmetric. Hence

$$\kappa_{mtrue} = 6\kappa_{5,1}. \tag{4.4}$$

This is an identity of limiting integrals. The preceding addback argument handles the literal interval endpoints of the integer counts; it is not replaced by this hyperplane argument.

Put

$$u = a - \frac{2}{5} = \frac{1361}{100000}, \qquad \alpha_i = \frac{1}{5} + uz_i, \qquad z_5 = -z_1 - z_2 - z_3 - z_4.$$

The closure of the first labeled region becomes the fixed polytope $P_E \subset \mathbb{R}^4$ defined by

$$\begin{gathered} z_i \geq -2 \quad (1 \leq i \leq 5), \qquad z_1 \geq z_2 \geq z_3 \geq z_4, \\ z_1 + z_2 + z_3 + 2z_4 \leq 0, \qquad z_1 + z_2 \leq 1, \qquad z_2 + z_3 + z_4 \geq -1. \end{gathered} \tag{4.5}$$

The original upper cap on α_1 is redundant here; it also follows immediately from the vertices below. The integral is

$$\kappa_{mtrue} = 6u^4 \int_{P_E} \frac{dz_1\, dz_2\, dz_3\, dz_4}{\prod_{i=1}^{5}(1/5 + uz_i)}. \tag{4.6}$$

The factor u^4 is the coordinate Jacobian. There is no additional factor 5!, or Euclidean hyperplane-area factor.

For explicit geometric data, the vertices are

j	v_j			
0	$-1/3$	$-1/3$	$-1/3$	$-1/3$
1	0	0	0	0
2	$1/2$	$1/2$	$-3/4$	$-3/4$
3	$1/2$	$1/2$	$-1/3$	$-1/3$
4	$1/2$	$1/2$	$1/2$	-2
5	$1/2$	$1/2$	$1/2$	$-3/4$
6	$4/3$	$-1/3$	$-1/3$	$-1/3$

and a triangulation is

$$\Sigma_1 = (0, 1, 3, 5, 6), \qquad \Sigma_2 = (0, 2, 3, 5, 6), \qquad \Sigma_3 = (0, 2, 4, 5, 6).$$

One way to check the partition is to cut P_E by $z_3 + 2z_4 = -1$ and $z_2 + 2z_4 = -1$. Since $z_2 \geq z_3$, the three resulting regions are respectively

$$z_3 + 2z_4 \geq -1, \qquad z_3 + 2z_4 \leq -1 \leq z_2 + 2z_4, \qquad z_2 + 2z_4 \leq -1.$$

Their vertex sets are precisely those of $\Sigma_1, \Sigma_2, \Sigma_3$; the interiors are disjoint. Determinants divided by 4! give the volumes

$$V_1 = \frac{125}{5184}, \qquad V_2 = \frac{625}{20736}, \qquad V_3 = \frac{625}{6912}, \qquad \sum_j V_j = \frac{125}{864}. \tag{4.7}$$

Fix a simplex Σ, and write $t_0, \ldots, t_4$ for its barycentric coordinates. Each $z_i(t)$, including $z_5(t)$, is a linear form. The forms $B_i(t) = -12z_i(t)$ have integer coefficients. With $v = 5u/12$, expand

$$\prod_{i=1}^{5}(1/5 + uz_i)^{-1} = 5^5 \prod_{i=1}^{5}(1 - vB_i)^{-1} = 5^5 \sum_{n=0}^{\infty} v^n h_n(B_1, \ldots, B_5), \tag{4.8}$$

where h_n is the complete homogeneous polynomial of degree n. Its coefficients can be computed by the finite recurrence

$$H_{i,n} = H_{i-1,n} + B_i H_{i,n-1}, \qquad H_{i,0} = 1, \quad H_{0,n} = 0 \ (n > 0), \quad h_n = H_{5,n}.$$

For the uniform probability measure on a four-simplex, successive applications of the beta integral give the exact Dirichlet moment formula

$$\mathbb{E}_\Sigma[t_0^{b_0} \cdots t_4^{b_4}] = \frac{4! \prod_{j=0}^{4} b_j!}{(n+4)!}, \qquad \sum_{j=0}^{4} b_j = n. \tag{4.9}$$

Combining the two formulas computes every integrated coefficient in (4.8) by integer and rational arithmetic.

For the error bound put

$$r_\Sigma = 5u \max_{z \text{ a vertex of } \Sigma} \max_i |z_i|.$$

Linearity bounds $|5uz_i(t)|$ by r_Σ everywhere in the simplex. The three values are

$$r_{\Sigma_1} = r_{\Sigma_2} = \frac{1361}{15000}, \qquad r_{\Sigma_3} = \frac{1361}{10000}.$$

In particular all are less than one. There are $\binom{n+4}{4}$ monomials in h_n, counted with multiplicity in the expansion of the product of five geometric series. Thus after degree D its uniform absolute tail, before the factor 5^5, is at most

$$T_D(r) = \frac{1}{(1-r)^5} - \sum_{n=0}^{D} \binom{n+4}{4} r^n. \tag{4.10}$$

This proves absolute uniform convergence and justifies termwise integration. Let

$$A_\Sigma = \sum_{n=0}^{22} v^n \mathbb{E}_\Sigma[h_n(B_1, \ldots, B_5)].$$

Equations (4.6)–(4.10) give the explicit rational enclosure

$$\kappa_{mtrue} \;\in\; 6 \cdot 5^5 u^4 \sum_{j=1}^{3} V_j [A_{\Sigma_j} - T_{22}(r_{\Sigma_j}),\ A_{\Sigma_j} + T_{22}(r_{\Sigma_j})]. \tag{4.11}$$

Its width is less than $3 \cdot 10^{-20}$, and outward comparison with terminating rationals yields

$$.000093307088324061 < \kappa_{mtrue} < .000093307088324062 = \bar{\kappa}.$$

For each of these three simplices, the displayed formulas determine 23 rational moments, their partial sum and tail, and the resulting rational contribution interval. This is an explicit finite calculation. Its normalization, convergence and tail estimate do not depend on trusting a numerical quadrature routine.

5 An upper bound for the exceptional square

Knowing the mean of b_x does not bound its correlation with a sieve square. A separate positive upper-bound sieve is needed. This section explains the bound and specifies the finite kernel actually used in the certificate. The starting method is the auxiliary Selberg sieve of [17, Lemma 3.9 and Propositions 3.10–3.11]. We prove below the variant with the hinge weight and an auxiliary radius depending on the retained root.

5.1 A fixed positive weight on prime pairs

For a fixed $t < 2/5$, with $t \geq 2\lambda$, put

$$w_t(s) = \frac{12}{5} \frac{(s-t)_+}{2/5 - t}.$$

On an exceptional integer $n = p_1 \cdots p_5$, all ten pairs have logarithmic sizes $s_{ij} = \log_x(p_i p_j) < a$, and

$$\sum_{i<j} s_{ij} = 4 \log_x n \geq 4.$$

It follows that

$$\sum_{i<j} (s_{ij} - t)_+ \geq 4 - 10t.$$

Multiplication by $12/(5(2/5 - t))$ gives 24. Off the exceptional set the right-hand side below is nonnegative. Thus, pointwise on the interval,

$$b_x(n) \leq \sum_{\substack{p<q,\ pq|n \\ p,q>x^\lambda,\ pq<x^a}} w_t(\log_x(pq)). \tag{5.1}$$

We use $t = .38$ throughout the profile.

The threshold is fixed. It is not selected separately for different retained roots inside a signed square. Such a selection would change the mixed products of their coefficients and would require a different Gram inequality.

5.2 The auxiliary sieve on a retained root

Fix a detected coordinate i, and put $d = k - 1$. For a retained configuration Y, let

$$q(Y) = \rho_* \sum_{j \neq i} |Y_j|.$$

This is its *physical* logarithmic total, in base x. The largest possible value in this proof is

$$r_c = \rho_* S = .2758746934.$$

For a pair bin with upper endpoint s_j, choose

$$z_j(q) = \min\{\lambda - \varepsilon_0,\ (1 - s_j)/2 - q - \varepsilon_0\}, \qquad \varepsilon_0 = 10^{-4}. \tag{5.2}$$

All these radii are positive. At the worst endpoint,

$$(1 - a)/2 - r_c - \varepsilon_0 = .0172203066 > 0.$$

The reason for the second term in (5.2) is that a marked pair and two auxiliary-augmented roots must fit inside an ordinary CRT counting range:

$$s_j + 2\big(q + z_j(q)\big) \leq 1 - 2\varepsilon_0.$$

The first term ensures that every auxiliary prime lies below the roughness threshold.

We recall the exact identity behind this construction, since a root-dependent auxiliary radius is otherwise easy to misunderstand. For a prime $p \nmid W$, define

$$U_{p,l}(n) = \frac{1 - p\mathbf{1}_{p|n+h_l}}{p - 1}.$$

A retained canonical sum can be written

$$B_y(n) = B_R^{-d} \sum_{\mathbf{r}} y_{\mathbf{r}} \prod_{l \neq i} \prod_{p|r_l} U_{p,l}(n).$$

Let $q_{\mathbf{r}} = \log_x \prod_{l \neq i} r_l$, and put

$$Z_{\mathbf{r}} = x^{z_j(q_{\mathbf{r}})}, \qquad G_{\mathbf{r}} = \sum_{\substack{b \leq Z_{\mathbf{r}} \\ (b,W)=1}} \frac{\mu^2(b)}{\varphi(b)}.$$

The auxiliary sum is

$$T_j(n) = B_R^{-d} \sum_{\mathbf{r}} \frac{y_{\mathbf{r}}}{G_{\mathbf{r}}} \sum_{\substack{b \leq Z_{\mathbf{r}} \\ (b,W)=1}} \mu^2(b) \prod_{l \neq i} \prod_{p|r_l} U_{p,l}(n) \prod_{p|b} U_{p,i}(n). \tag{5.3}$$

If $n + h_i$ is x^λ-rough, every auxiliary contrast in the last product is $1/(p-1)$. The inner sum is then exactly $G_{\mathbf{r}}$, and

$$T_j(n) = B_y(n).$$

This identity holds for signed y and for each root's own auxiliary radius. It also holds when an auxiliary prime divides a retained root. Those shared-prime terms must remain in (5.3).

5.3 Why the resulting bound is a quadratic-form bound

Expand $T_j(n)^2$ after marking a pair $p < q$. The CRT modulus is bounded by $x^{1-2\varepsilon_0}$, so ordinary counting gives a power-saving remainder after the fixed divisor factors are included. No distribution theorem for primes is used in this step.

For a fixed auxiliary radius, the comparison with independent local roots is proved in [17, proof of Proposition 3.10]. The same local kernels apply here; the coefficient array now includes the cutoff and normalization belonging to each retained root. We give the comparison explicitly to keep that dependence in the square. At an unmarked prime put $a_p = p - 1$. Each root's state is one of

$$\varnothing, \quad \{i\}, \quad \{l\}, \quad \{i,l\} \quad (l \neq i).$$

The actual and independent kernels are

$$K_p(S,T) = \frac{1 - |S \cup T| + a_p|S \cap T|}{a_p^{|S|+|T|}}, \qquad D_p(S,T) = \mathbf{1}_{S=T}\, a_p^{-|S|}.$$

The total absolute entrywise discrepancy is $O_k(p^{-2})$. Tensor telescoping after presieving therefore gives a vanishing normalized error, since $\sum_{p>w} p^{-2} \to 0$. The comparison uses a uniform bound for the whole coefficient array, not a factorization of the coefficients into independent functions of the retained and auxiliary roots.

In the independent model the retained indices and the auxiliary indices must match. Its energy is exactly

$$B_R^{-2d} \sum_{\mathbf{r}} \frac{|y_{\mathbf{r}}|^2}{\prod_{l\neq i} \varphi(r_l)} \frac{1}{G_{\mathbf{r}}}.$$

Uniformly on the retained support, $G_{\mathbf{r}} = (z_j(q_{\mathbf{r}}) + o(1))B_x$. This is the harmonic-sum estimate in [17, Lemma 3.9], from [15, Lemma 6.1]; its error is uniform because $z_j(q_{\mathbf{r}})$ lies in a fixed compact subinterval of $(0,\infty)$. After the normalizations in (2.10), the limit is

$$\int \frac{|H(Y)|^2}{z_j(q(Y))}\, d\nu^{k-1}(Y)$$

when H is the retained limiting profile. Weighted L^2 convergence suffices, since $1/z_j(q)$ is bounded on the support. This also covers exact erased arrays and their fixed finite signed combinations with canonical arrays.

The unordered harmonic pair density, as in [17, equation (3.33)], is

$$\frac{1}{s} \log \frac{s-\lambda}{\lambda}\, ds.$$

Combining it with (5.1) gives the continuous upper kernel

$$K_t(q) = \frac{12}{5(2/5-t)} \int_t^a \frac{(s-t)\log((s-\lambda)/\lambda)}{s \min\{\lambda, (1-s)/2 - q\}}\, ds. \tag{5.4}$$

There is no additional factor of two: the density already counts unordered pairs. In this display the auxiliary retreat has been allowed to tend to zero: one first proves the estimate with an arbitrary fixed positive retreat, then takes its decreasing limit. The finite calculation below instead retains the fixed retreat $\epsilon_0 = 10^{-4}$; its kernel is correspondingly larger.

5.4 The finite step kernel used in the calculation

The certificate uses a slightly larger, explicit step kernel. Its bin convention is part of the fixed trial. Partition the *whole* interval $[2\lambda, a]$ into 1024 equal bins, with upper endpoints s_j. Set

$$L_{21}(v) = \sum_{m=1}^{21} (-1)^{m+1} \frac{v^m}{m}, \qquad v_j = \frac{s_j - 2\lambda}{\lambda}.$$

Here $0 \le v_j < 1$, so the odd alternating sum bounds $\log(1 + v_j)$ above. Round the following nonnegative quantity upward to a multiple of 10^{-25}:

$$W_j^+ \ge (s_j - s_{j-1}) \frac{w_t(s_j)}{s_j} L_{21}(v_j). \tag{5.5}$$

The weight is zero below the hinge except possibly on the bin straddling it, where an upper-endpoint bound is used. The pair density is increasing on this interval, as $a < 3\lambda$; the hinge is also nondecreasing. Thus (5.5) is a valid upper bin mass.

For the fine grid in Section 7, put $n_0 = N - k$. Divide retained sum indices into blocks of length $\lceil n_0/512 \rceil = 768$. If the last index in a block is $r_{\max}$, use the physical upper total

$$q^+ = \rho_*(r_{\max} + k - 1)h.$$

The value of K on that entire block is

$$K = \sum_{j=1}^{1024} \frac{W_j^+}{z_j(q^+)}. \tag{5.6}$$

The numerator rounding, auxiliary retreat and upper total all act in the direction of an upper bound. They give

$$0 \le K \le 2.194 \tag{5.7}$$

on the whole retained domain. The same K is used in the cap form and in all restoration terms.

It is useful to summarize the result in the form needed next. For a retained sum B_H, or a fixed finite combination including an exact erased sum,

$$\limsup_{x\to\infty} \frac{1}{\mathfrak{P}_x} \sum_{\substack{x \le n \le 2x \\ n \equiv b_0 \pmod W}} b_x(n + h_i) B_H(n)^2 \le \int K(q(Y)) |H(Y)|^2 \, d\nu^{k-1}(Y). \tag{5.8}$$

All cross terms of the finite combination are included in its square. The bound controls a quadratic form, which is stronger than separately bounding the mean of each summand.

6 The signed hybrid inequality

The purpose of this section is to use the minorant only where its distribution gives extra information. The prime indicator remains the nonnegative weight whose square is being estimated.

6.1 The pairing identities

Fix a detected coordinate i, an arithmetically allowed profile U supported on O, and write

$$V = E_i U, \quad V_0 = \mathbf{1}_{L_0} V, \quad v = \mathbf{1}_{L_1 \setminus L_0} V, \quad V_C = \mathbf{1}_{\mathcal{C}} V, \quad g = \mathbf{1}_{\mathcal{C}^c} V.$$

Complements are taken in the retained-root ambient space. They are not new independently capped domains. Let $A = A_U$, and let B_0, B_C, B_H be the canonical retained sums with profiles V_0, V_C, H, where $\operatorname{supp} H \subset L_1 \setminus L_0$.

In this subsection $\mathcal{M}_x(fQ)$ denotes the sum of $f(n+h_i)Q(n)$ on the presieved interval, divided by $\mathfrak{P}_x$. The table in Section 3.6 supplies the distribution hypotheses for the argument in [17, proof of Proposition 3.5]. The minorant constructed here has the same required roughness property, with its own mean $1-\kappa_{\text{true}}$. That argument and the harmonic limits [17, Proposition 3.7 and Lemma 3.8] give

$$\mathcal{M}_x(\mathbf{1}_{\mathbb{P}} A B_0) \longrightarrow \|V_0\|^2, \qquad \mathcal{M}_x(\mathbf{1}_{\mathbb{P}} B_0^2) \longrightarrow \|V_0\|^2, \tag{6.1}$$

$$\mathcal{M}_x(\mathbf{1}_{\mathbb{P}} B_0 B_H) \longrightarrow 0, \qquad \mathcal{M}_x(\mathbf{1}_{\mathbb{P}} B_H^2) \longrightarrow \|H\|^2, \tag{6.2}$$

$$\mathcal{M}_x(\varrho_x A B_H) \longrightarrow (1-\kappa_{\text{true}})\langle V, H\rangle, \qquad \mathcal{M}_x(b_x B_C B_H) \longrightarrow \kappa_{\text{true}}\langle V, H\rangle. \tag{6.3}$$

Also $\mathcal{M}_x(b_x B_H^2) \to \kappa_{\text{true}}\|H\|^2$. The zero in (6.2) comes from the disjoint supports of V_0 and H. These limits follow from the harmonic marginal limit and the arithmetic pairings; they do not require pointwise convergence of an erased array.

6.2 Subtracting the full available canonical face

The prime/minorant comparison uses the positive-square method of [17, Section 4.3]. Here the evaluated mixed pairing on $\mathcal{C} \times L_1$ permits subtraction of the entire available canonical face. The resulting tail estimate is proved below.

Let A_i^0 be the exact erased sum. The roughness argument in Section 4 gives $A = A_i^0$ on both the prime and exceptional supports. Set

$$T_i = A_i^0 - B_C.$$

Its limiting profile is g. Since $A = B_C + T_i$ on the support of b_x, (6.3) gives

$$\mathcal{M}_x(\mathbf{1}_{\mathbb{P}} A B_H) = \langle v, H\rangle + \mathcal{M}_x(b_x T_i B_H) + o(1). \tag{6.4}$$

The coefficient of $\langle v, H\rangle$ is exactly one: $(1-\kappa_{\text{true}}) + \kappa_{\text{true}} = 1$. Only an upper bound for κ_{true} will be used numerically, but this identity uses its true value.

The evaluated subtraction term is $B_C B_H$. The correction H is supported in L_1; therefore the mixed $\mathcal{C} \times L_1$ source suffices. No square B_C^2 is evaluated by a distribution theorem. Its contribution inside T_i^2 is instead included in the positive exceptional bound (5.8).

Proposition 6.1 (Tail hybrid)**.** *Put*

$$J_{0,i} = \|V_0\|^2, \qquad J_{+,i} = \|v\|^2, \qquad E_i^{\text{tail}} = \int_{\mathcal{C}^c} K|V|^2 \, d\nu^{k-1}.$$

Then

$$\liminf_{x\to\infty} \mathcal{M}_x(\mathbf{1}_{\mathbb{P}} A^2) \geq J_{0,i} + \left(\sqrt{J_{+,i}} - \sqrt{\kappa_{\text{true}} E_i^{\text{tail}}}\right)_+^2. \tag{6.5}$$

Proof. The pointwise inequality $\mathbf{1}_{\mathbb{P}}(A - B_0 - B_H)^2 \geq 0$, together with (6.1)–(6.4), gives

$$\liminf \mathcal{M}_x(\mathbf{1}_{\mathbb{P}} A^2) \geq J_{0,i} + 2\langle v, H\rangle - \|H\|^2 + 2 \liminf \mathcal{M}_x(b_x T_i B_H).$$

Cauchy–Schwarz in the positive measure with weight b_x, followed by (5.8), bounds the last term below by $-2\sqrt{\kappa_{\mathrm{true}} E_i^{\mathrm{tail}}}\|H\|$. Choose $H = \alpha v$, $\alpha \geq 0$. If $J_{+,i} > 0$, maximizing the resulting quadratic in α gives

$$\alpha = \left(1 - \sqrt{\kappa_{\mathrm{true}} E_i^{\mathrm{tail}}/J_{+,i}}\right)_+ .$$

This gives (6.5). If $J_{+,i} = 0$, take $H = 0$. □

6.3 A fixed quadratic form

For $u, v \geq 0$ and $0 < \ell < 1$,

$$(\sqrt{u} - \sqrt{v})_+^2 \geq (1 - \ell)u - (1/\ell - 1)v. \tag{6.6}$$

For $u \geq v$ this follows from $2\sqrt{uv} \leq \ell u + v/\ell$; for $u < v$ the right side is nonpositive. Replace κ_{true} by its upper bound $\bar{\kappa}$ in the negative term and set

$$d_h = \bar{\kappa}(1/\ell - 1).$$

Summing over the detected coordinates yields the computable lower form

$$\mathcal{J}(U) = J_0(U) + (1 - \ell)J_+(U) - d_h E_{\mathcal{C}^c}(U), \tag{6.7}$$

where the three quantities are the sums of $J_{0,i}, J_{+,i}, E_i^{\mathrm{tail}}$.

For the fixed profile we use the rational

$$\ell = \frac{6308693370963335}{576460752303423488}, \qquad .0108 \leq \ell \leq .0112. \tag{6.8}$$

There is no need to prove that this is an optimal choice. Any fixed choice that passes the final strict inequality suffices.

Using $\mathfrak{P}_x = \rho_* \mathfrak{C}_x$, we obtain the following criterion. Its final positivity argument is the one in [15, proof of Proposition 4.2]; the hybrid face form $\mathcal{J}$ is supplied by the preceding calculation.

Corollary 6.2. *If a fixed, nonzero, arithmetically allowed profile U satisfies*

$$\rho_* \mathcal{J}(U) > I(U),$$

and has the approximation properties specified in Section 11, then every admissible k-tuple has infinitely many translates containing at least two primes.

The remaining work is to verify this strict inequality for $U = PF$. The numerical cap profile F by itself need not be arithmetically allowed.

7 The fixed profile

The numerical search is not part of the proof. We fix one profile, give its mathematical definition, and enclose its forms. This makes the claim independent of the path by which the coefficients were found. The trial uses the fragment and power-sum construction of [18, Section 1.1], with the 39-coordinate spline profile and exact data specified here.

7.1 Scales and the original fine grid

The exact values are

$$\begin{gathered}
k=39, \qquad N=393216, \qquad n_0=N-k=393177, \\
S=\frac{1379373467}{1312494500}, \qquad h=\frac{S}{N}=\frac{1379373467}{516093837312000}, \\
\rho_*=\frac{2624989}{10000000}<\rho=\frac{262499}{1000000}.
\end{gathered} \tag{7.1}$$

We can take the ambient fragment cutoff $\zeta=.17277/\rho$. All caps actually used lie below it, so (2.4) identifies these forms with the arithmetic ambient measure without a change of normalization. The physical radii are

domain	radius in base x
$\mathcal{H}_O$	$\rho_*S=.2758746934$
$\mathcal{H}_0$	$\rho_*T_0=.2491228441011$
$\mathcal{H}_1$	$\rho_*T_1=.2525250066.$

In particular $2\rho_*S=.5517493868<1$, which leaves a positive ordinary counting margin for the outer square.

For a configuration X_i, put

$$j_i=\lfloor |X_i|/h\rfloor, \qquad t_i^\circ=(j_i+\tfrac{1}{2})h.$$

The profile samples its smooth factors at these midpoints, but its support tests use the actual fragments and the whole-cell upper total. For d coordinates write

$$r=\sum_{i=1}^{d} j_i, \qquad s^\circ=(r+d/2)h, \qquad s^+=(r+d)h.$$

The distinction between s° and s^+ is deliberate. A midpoint support test could admit part of a cell beyond an allowed arithmetic radius.

7.2 The cap domains as finite mathematical data

For a configuration in d coordinates let M_0 be the largest fragment among all its coordinates, with value zero on the empty configuration. Each row of the following table specifies an inclusive range of the integer r and the bound $M_0\le ch$. Outside the listed ranges the domain is empty.

Domain and dimension	First r	Last r	c
$\mathcal{H}_O,\ d=39$	0	352697	246256
	352698	376449	196608
	376450	393176	186253
$\mathcal{H}_0,\ d=38$	0	334253	246256
	334254	343163	177542
	343164	355047	140306
$\mathcal{H}_1,\ d=38$	0	336274	246256
	336275	343169	179967
	343170	359896	141724
$\mathcal{H}_C,\ d=38$	0	352698	246256
	352699	371671	186491

These integer ranges define the inward caps exactly. They give $\mathcal{H}_0 \subset \mathcal{H}_1 \subset \mathcal{H}_C$. Deleting fragments can only lower r and M_0, and the allowed largest-fragment cap increases or stays the same when r decreases. Hence all four domains are hereditary. Their intersections with the row predicates in (3.12) are hereditary as well.

7.3 The angular functions and the radial splines

Set

$$g(t) = \frac{21/200}{1 + t/100} + \frac{179/200}{1 + (907/5)t}. \tag{7.2}$$

This positive one-coordinate factor is the one in [18, equation (1.7)]. The set of eleven power-sum signatures below is that of [18, equation (1.6)], listed in the order used by the present coefficient table. Define the ordinary power sums

$$P_j(t^\circ) = \sum_{i=1}^{39} (t_i^\circ)^j, \qquad P_\sigma(t^\circ) = \prod_{j \in \sigma} P_j(t^\circ), \qquad P_\varnothing = 1.$$

The eleven signatures, in their fixed order, are

$$\varnothing,\ (2),\ (3),\ (4),\ (2,2),\ (5),\ (2,3),\ (6),\ (2,4),\ (3,3),\ (2,2,2).$$

These are products of power sums, so repeated coordinate assignments are included. For example,

$$P_{(2,2)} = \sum_i (t_i^\circ)^4 + 2\sum_{i<j} (t_i^\circ)^2 (t_j^\circ)^2.$$

This example makes explicit the diagonal contribution that would be lost by interpreting a signature as a sum over distinct coordinates only.

There are 22 cubic B-spline basis functions $B_{d,3}$, $0 \le d \le 21$, with a rational 26-entry knot vector. To specify it, put

$$u_1 = \frac{324809441462533}{344528493755500}, \qquad u_2 = \frac{1626288847188681852468891377825170326 7}{1616192119940869600750361656598394600 0}.$$

The knot vector is

$$(3/20)^{[4]},\ 1/4,\ 7/20,\ 9/20,\ 11/20,\ 13/20,\ 3/4,\ 33/40,\ 7/8,\ 37/40,$$
$$u_1^{[4]}, \quad 49/50, \quad u_2^{[4]}, \quad S^{[4]},$$

where the superscript [4] means four consecutive copies. We use the Cox–de Boor recurrence [5, 6], with the following convention at repeated knots:

$$B_{d,0}(s) = \mathbf{1}_{[\kappa_d, \kappa_{d+1})}(s),$$

$$B_{d,m}(s) = \frac{s - \kappa_d}{\kappa_{d+m} - \kappa_d} B_{d,m-1}(s) + \frac{\kappa_{d+m+1} - s}{\kappa_{d+m+1} - \kappa_{d+1}} B_{d+1,m-1}(s).$$

A term with zero denominator is defined to be zero. Values at internal knots use the right-hand convention. The fourfold internal knots allow jumps; smoothness across them is not assumed. The argument of the spline is s°, the sum of the coordinate midpoints.

7.4 The complete profile

Let $\mathcal{A} = \{141720h, 177528h, 246256h\}$. For each $c \in \mathcal{A}$, fix 242 rational numbers $c_{c,\sigma,d}$. The profile is

$$\begin{aligned} F(X_1, \ldots, X_{39}) = \mathbf{1}_{\mathcal{H}_O}(X) \prod_{i=1}^{39} g(t_i^\circ) \\ \times \sum_{c \in \mathcal{A}} \mathbf{1}_{M_0(X) \le c} \sum_{\sigma} \sum_{d=0}^{21} c_{c,\sigma,d} B_{d,3}(s^\circ) P_\sigma(t^\circ). \end{aligned} \tag{7.3}$$

The total number of coefficients is $3 \cdot 11 \cdot 22 = 726$. The complete rational table is supplied as digital data; Appendix B identifies it unambiguously. It also supplies a flat coefficient table with the cap, signature and spline index as explicit columns.

The three cap components have independent coefficients. They are added before every square. Since the profile can change sign, replacing $(F_1 + F_2 + F_3)^2$ by $F_1^2 + F_2^2 + F_3^2$ would not be a valid evaluation. The exported coefficients already include the angular scaling used to generate them. Formula (7.3) uses them exactly as written, without an additional power of 39.

7.5 The common numerical normalization

For $0 \le j < n_0$, set

$$g_j = g((j + \tfrac{1}{2})h), \qquad Z = \sum_{j=0}^{n_0 - 1} g_j^2, \qquad \mathcal{N} = (hZ)^{39}.$$

Every quadratic form reported by the certificate, including a form with 38 retained coordinates, is divided by the same positive number $\mathcal{N}$. The size of the resulting denominator is therefore unrelated to a numerical underflow of the unscaled sieve mass.

For a largest-fragment cap c, the exact one-coordinate cell mass is

$$w_c(j) = \int_{jh}^{(j+1)h} \rho_D(t/c)\, dt.$$

Thus a normalized root law has weights

$$\frac{g_j^2}{Z} \frac{w_c(j)}{h}.$$

The factor $w_c(j)/h$ is a cell-average density, not an independent probability correction. In a face square there are 38 retained integrations and two erased integrations. After division by $\mathcal{N}$, the sum over symmetric detected coordinates has the prefactor

$$\frac{39h}{Z}.$$

This is the same factor obtained by counting powers of h in (2.8). Outer source integrals have an additional full-root integration; their normalization is recorded explicitly in Section 8.5.

7.6 The cap form and its enclosure

Replace $L_0, L_1, \mathcal{C}$ in (6.7) by $\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_C$, keeping the same F, K, ℓ. Write the resulting face forms as $J_{0,H}, J_{+,H}, E_{\mathcal{H}_C^c}$. The cap quotient is

$$q_H = \rho_* \frac{J_{0,H} + (1-\ell)J_{+,H} - d_h E_{\mathcal{H}_C^c}}{I(F)}. \tag{7.4}$$

It is a computable relaxation of the final problem. The normalized denominator has the rigorous enclosure

$$\begin{aligned} I_- \le \frac{I(F)}{\mathcal{N}} \le I_+, \\ I_- = \frac{990031661365651864280059 3}{6646139978924579364519035301401722 88}, \\ I_+ = \frac{9900316889124784142456785}{664613997892457936451903530140172288}. \end{aligned} \tag{7.5}$$

In particular $I_- > 0$, and both endpoints are about $1.4896341 \cdot 10^{-11}$. The quotient enclosure gives

$$\mu := q_H - 1 \ge \mu_- = \frac{9968877639291}{36028797018963968} > .0002766919371. \tag{7.6}$$

For a positive lower quotient bound the positive numerator terms use their lower endpoints, the negative term uses its upper endpoint, and the denominator is I_+. The lower denominator I_- will instead be used when normalizing error upper bounds.

8 Restoring the arithmetic support

The cap-to-arithmetic operator comparison follows [17, Section 4.4]; a componentwise signed restoration inequality appears in [18, Proposition 1.3]. We prove the form needed for the present multipliers, including the root-dependent exceptional kernel and the subtraction domain.

8.1 The two changes that have to be controlled

The cap calculation relaxes both the outer support and the permitted faces. First the face multipliers must be replaced by their arithmetic versions; then F must be replaced by $PF = \mathbf{1}_O F$. The second change interacts with the first, so we handle them in one operator calculation.

Work on $\mathscr{H} = L^2(\mathcal{H}_O, \nu^{39})$, extending functions by zero outside $\mathcal{H}_O$. On a retained face define

$$a_H = \ell \mathbf{1}_{\mathcal{H}_0} + (1-\ell)\mathbf{1}_{\mathcal{H}_1} + d_h K \mathbf{1}_{\mathcal{H}_C} - d_h K, \tag{8.1}$$
$$a = \ell \mathbf{1}_{L_0} + (1-\ell)\mathbf{1}_{L_1} + d_h K \mathbf{1}_{\mathcal{C}} - d_h K. \tag{8.2}$$

Thus a_H is 1 on $\mathcal{H}_0$, $1-\ell$ on $\mathcal{H}_1 \setminus \mathcal{H}_0$, zero on $\mathcal{H}_C \setminus \mathcal{H}_1$, and $-d_h K$ outside $\mathcal{H}_C$. The analogous four-region description holds for a.

Define the self-adjoint operators

$$A = \sum_i E_i^* a_H E_i, \quad B = \sum_i E_i^* a E_i, \quad \mathsf{C} = \rho_* A, \quad \mathsf{T} = \rho_* B, \quad \mathsf{D} = \mathsf{C} - \mathsf{T}.$$

Because each actual domain is a subset of its cap domain, $a_H - a \ge 0$. Therefore $\mathsf{D} \ge 0$, even though A and B need not be positive operators.

Put

$$\epsilon_h = \bar{\kappa}(1/.0108 - 1)\,2.194,$$
$$\delta_{\mathrm{op}} = 4\rho_* \epsilon_h, \qquad \gamma_{\mathrm{op}} = 1 - \delta_{\mathrm{op}}, \qquad \Gamma = 4\rho_*(1 + \epsilon_h). \tag{8.3}$$

The inequalities $-\epsilon_h \le a_H, a \le 1$ and Lemma 2.1 imply

$$0 \le \mathsf{D} \le \Gamma \mathrm{Id}, \qquad \mathsf{T} \ge -\delta_{\mathrm{op}} \mathrm{Id}. \tag{8.4}$$

For orientation, $\epsilon_h \approx .01875045$, $\delta_{\mathrm{op}} \approx .01968789$, and $\gamma_{\mathrm{op}} \approx .98031211 > 0$. The proof uses the exact rationals in (8.3).

8.2 Partitioning the true outer error

Let E_j be the true outer failure events in the finite source cover. Their union covers $\mathcal{H}_O \setminus O$. Define a soft partition by

$$p_j(X) = \begin{cases} \dfrac{\mathbf{1}_{\mathcal{H}_O \setminus O}(X) \mathbf{1}_{E_j}(X)}{\sum_l \mathbf{1}_{E_l}(X)}, & \sum_l \mathbf{1}_{E_l}(X) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then $p_j \ge 0$, $\sum_j p_j = \mathbf{1}_{\mathcal{H}_O \setminus O}$, and p_j is supported on E_j. If a computable positive kernel w_j majorizes $\mathbf{1}_{E_j}$, then $p_j \le w_j$.

It matters that this partition is made from the *true* events before their kernels are enlarged. A local estimate valid on E_j need not remain valid everywhere on the support of its analytic majorant.

After dividing by $I(F)$, suppose the source integrations give

$$\frac{\int p_j |F|^2}{I(F)} \le A_j, \qquad \frac{\int p_j |\mathsf{C}F|^2}{I(F)} \le M_j, \qquad \frac{\langle F, \mathsf{D}F \rangle}{I(F)} \le b. \tag{8.5}$$

Here A_j measures discarded mass, M_j measures the cap operator on that discarded region, and b measures all face deletions. They are different energies; none can be replaced by the unweighted volume of the bad set.

Theorem 8.1 (Support restoration)**.** *Under* (8.4) *and* (8.5)*, for every $t > \gamma_{\mathrm{op}}$,*

$$\frac{\langle PF, (\mathsf{T} - \mathrm{Id})PF \rangle}{I(F)} \ge \mu - b - \frac{\Gamma b}{t} - 2 \sum_j \sqrt{A_j M_j} - (t - \gamma_{\mathrm{op}}) \sum_j A_j. \tag{8.6}$$

No sign condition on F or commutation of P with the other operators is required.

Proof. Set $e = (1 - P)F$, and normalize

$$u_j = \frac{\int p_j |F|^2}{I(F)}, \qquad b_0 = \frac{\langle F, \mathsf{D}F \rangle}{I(F)}.$$

Then $\sum_j u_j = \|e\|^2 / I(F)$. Because P is an orthogonal projection, $\langle e, F \rangle = \|e\|^2$. Expanding $PF = F - e$ and $\mathsf{T} = \mathsf{C} - \mathsf{D}$ gives the exact identity

$$\begin{aligned} \langle PF, (\mathsf{T} - \mathrm{Id})PF \rangle = {} & \langle F, (\mathsf{C} - \mathrm{Id})F \rangle - \langle F, \mathsf{D}F \rangle - 2\,\mathrm{Re}\langle e, \mathsf{C}F \rangle \\ & + 2\,\mathrm{Re}\langle e, \mathsf{D}F \rangle + \langle e, (\mathsf{T} + \mathrm{Id})e \rangle. \end{aligned} \tag{8.7}$$

In particular the last term contains $+\mathrm{Id}$. That sign retains the beneficial removal of $\|e\|^2$ from the denominator.

Weighted Cauchy–Schwarz on the partition gives

$$\frac{|\langle e, \mathsf{C}F\rangle|}{I(F)} \le \sum_j \sqrt{u_j M_j}.$$

Since $0 \le \mathsf{D} \le \Gamma \mathrm{Id}$, spectral calculus gives $\mathsf{D}^2 \le \Gamma \mathsf{D}$. Consequently

$$\frac{|\langle e, \mathsf{D}F\rangle|}{I(F)} \le \sqrt{\Gamma b_0 \sum_j u_j}.$$

Finally, $\mathsf{T} + \mathrm{Id} \ge \gamma_{\mathrm{op}} \mathrm{Id}$, so the last term of (8.7), divided by $I(F)$, is at least $\gamma_{\mathrm{op}} \sum_j u_j$.

Apply Young's inequality once:

$$2\sqrt{\Gamma b_0 \sum_j u_j} \le \frac{\Gamma b_0}{t} + t \sum_j u_j.$$

There is one shared inner-deletion cross term, so there is one charge $\Gamma b_0/t$. Since $t > \gamma_{\mathrm{op}}$, each function

$$u \longmapsto 2\sqrt{u M_j} + (t - \gamma_{\mathrm{op}})u$$

is nondecreasing for $u \ge 0$. We may therefore use $u_j \le A_j$ and $b_0 \le b$ in the remaining terms. This proves (8.6). $\square$

8.3 Why keeping the individual components helps

The sum $\sum_j \sqrt{A_j M_j}$ is no larger than $\sqrt{(\sum_j A_j)(\sum_j M_j)}$, by Cauchy–Schwarz. The inequality is often strict because a source with substantial discarded mass may have small operator energy, while a different source has the opposite behavior. Keeping their separate ceilings prevents the error estimate from reallocating mass between unrelated events.

The chosen value is simply

$$t = \frac{11}{10} > \gamma_{\mathrm{op}}.$$

No optimizer or numerical stationarity claim is needed to justify this value. It makes all substitutions in the proof monotone in the available upper bounds.

8.4 The local operator count

To obtain M_j, write

$$\mathsf{C}F(X) = \rho_* \sum_i a_H(\widehat{X}_i) V_i(\widehat{X}_i), \qquad V_i = E_i F.$$

The multiplier satisfies

$$|a_H| \le \epsilon_h + (1 - \epsilon_h)\mathbf{1}_{\mathcal{H}_1}.$$

At full-root midpoint total s°, an erased face can belong to $\mathcal{H}_1$ only if

$$t_i^\circ \ge d(s^\circ) := s^\circ + \frac{k-1}{2}h - T_1.$$

When $d(s^\circ) > 0$, at most $\lfloor s^\circ / d(s^\circ) \rfloor$ coordinates can meet this inequality.

There is a sharper event-specific count if marked fragments of total at least Q_0 lie in r coordinates. Those owners consume midpoint mass at least $Q_0 - rh/2$. Among the other coordinates, each additional eligible face consumes $d(s^\circ)$. Hence the eligible count is bounded by the minimum of k, the preceding global count, and

$$r + \left\lfloor \frac{\max(0, s^\circ - Q_0 + rh/2)}{d(s^\circ)} \right\rfloor.$$

If the denominator is nonpositive, use the safe count k. Largest-fragment conditions can only reduce the actual count.

This gives a nonnegative rational $L_j(s^\circ)$ that bounds $\sum_i |a_H(\widehat{X}_i)|$ on the true event E_j. Pointwise weighted Cauchy–Schwarz gives

$$|\mathsf{C}F|^2 \le \rho_*^2 L_j(s^\circ) \sum_i |a_H(\widehat{X}_i)||V_i(\widehat{X}_i)|^2 \quad \text{on } E_j.$$

Use this inequality on the support of p_j first, then enlarge p_j to w_j in the nonnegative integral. The multiplier L_j belongs to the full root and is inserted *before* erasing a coordinate.

8.5 The stored source normalization

The stored outer root ceiling R_j bounds

$$\frac{k}{\mathcal{N}} \int w_j |F|^2 \, d\nu^k.$$

The factor k is a convention in the stored ceiling R_j; it is removed in A_j below. The cap multiplier has absolute value at most one for the parameters here. The stored ceiling $Q_{L,j}$ bounds

$$\frac{1}{\mathcal{N}} \int w_j(X) L_j(s^\circ) \sum_i |a_H(\widehat{X}_i)||V_i(\widehat{X}_i)|^2 \, d\nu^k(X).$$

Thus valid normalized inputs in (8.5) are

$$A_j = \frac{R_j}{39 I_-}, \qquad M_j = \frac{\rho_*^2 Q_{L,j}}{I_-}. \tag{8.8}$$

The coordinate sum and the local count are already in $Q_{L,j}$; they are not multiplied in a second time. The common denominator is the one for the original fine profile (7.3).

For the inner deletion, old-inner failures have coefficient at most .0112, new-inner failures at most one, and subtraction-domain failures at most ϵ_h. A new-inner failure can affect both L_0 and L_1, with total coefficient at most $\ell + (1 - \ell) = 1$. These observations explain the weights used to assemble b. They also explain why both old and new predicates must occur in L_0.

9 From support failures to positive integrals

The restoration theorem reduces the application to upper bounds for A_j, M_j, b. This section explains how the geometric failure sets become the finite, positive source integrals in the certificate. The low-witness, rank-two and higher-rank organization follows [18, Lemma 1.4 and Section 1.5]. The seven families and their coverage inequalities below are those of the present profile.

9.1 The seven source families

An arithmetic support fails when one of its activated row predicates in Section 3.5 fails. The inward caps already handle the largest-fragment and opposite-root conditions. The remaining failures are nonlargest tail witnesses. The order-three allocation reduces them to effective order 5/2, in addition to the order-2 conditions. The reduction uses the checked row inequality $\max(A_r, C_r) \le 7L/3$; it retains any lower-order conditions absorbed into that row.

Here is the reason for the effective order 5/2. For an outer witness, $\phi_D(u) \le 3u/2$, so failure of its tail test implies failure of the $H_{5/2,\xi}$ test with the same threshold. For an inner witness with $u \le 2L/3$, one has $\phi_E(u) = 3u/2$, giving the same conclusion. If an inner nonlargest witness has $u > 2L/3$, at least two fragments are at least u. Its $H_{5/2,\xi}$ tail value is therefore at least

$$2u + \frac{3}{2}u = \frac{7}{2}u > \frac{7}{3}L \ge C_r.$$

These two cases cover the inner failures. Thus the effective-order reduction follows from the explicit allocation functions and the checked row inequality.

The source names in the data have the following mathematical meaning.

Family	Underlying failure	Low	Rank two	High
G1	Outer, effective order 2	8	14	1
G2	Outer, effective order 5/2	22	14	1
G3	Old inner, effective order 2	8	14	1
G4	Old inner, effective order 5/2	21	12	1
G5	New inner, effective order 2	8	14	1
G6	New inner, effective order 5/2	20	13	1
C	Subtraction domain, effective order 2	7	15	1

G1 and G2 together cover the outer failures of *both* ladders; they are not an old/new split. G3–G6 cover the two raw inner domains separately. The C family covers the true row-14 subtraction predicate within $\mathcal{H}_C$. Both inner pairs are imposed on the base domain L_0.

The first two rows give 60 outer components. The last five give 137 inner components, consisting of 64 low intervals, 68 rank-two intervals and five high terms. The number 197 is thus a complete inventory of these upper bounds, not a sample of a larger collection.

9.2 Why the low, rank-two and high cases cover every failure

Consider an effective tail obstruction

$$H_{m,\xi} > U$$

on a row where the radial and largest-fragment conditions are already enforced. Its nonlargest witness is a fragment $w > \xi$. Each group has a split point p_G. The group parameters satisfy

$$\xi < p_G \le z, \qquad z + mp_G \le U, \qquad \max(\xi, p_G) \le U/(m+1),$$

where z is the relevant largest-fragment cap. The split points, in units of h, are

G	G1	G2	G3	G4	G5	G6	C
p_G/h	98304	78643	88771	71017	89983	71986	93245.

They are the inward integer values of $U_G/(2m_G h)$, where m_G is the common effective order of the group.

If $w \le p_G$, it belongs to one of the complete half-open low-witness intervals. If $w > p_G$ and only two fragments exceed p_G, write them $q > w$. Then

$$q + mw > U, \qquad q > \frac{U}{m+1}, \qquad w > \frac{U-q}{m}.$$

This is the rank-two sector. If at least three fragments exceed p_G, the third factorial count covers the configuration. Indeed, the number $\binom{N_{>p_G}}{3}$ is at least one on that event.

The evaluated rank-two majorant may also include configurations with more than two high fragments. Such overlap is harmless because all source integrals are nonnegative. The argument requires coverage, not a disjoint decomposition of the probability space.

When a lower-order row has been absorbed into a group of nominal order 5/2, its own effective order is retained in the low-source radial clipping. The nominal group label alone is not sufficient to determine that clipping.

9.3 Marked fragments and the positive majorants

The Poisson description makes marked-fragment integrals explicit. In one coordinate, for a nonnegative measurable Ψ, the reduced Mecke identity [14, Theorem 4.5, equation (4.15)] is

$$\int \sum_{\mathbf{u}\in X^r_{\neq}} \Psi(\mathbf{u}, X\setminus \mathbf{u})\, d\nu_\zeta(X) = \int_{(0,\zeta]^r} \int \Psi(\mathbf{u}, X)\, d\nu_\zeta(X) \prod_{\ell=1}^{r} \frac{du_\ell}{u_\ell}. \tag{9.1}$$

The marks on the left are ordered distinct occurrences. A factor $1/r!$ is introduced only when converting a symmetric ordered integrand to unordered marks. The source calculations retain the designated witness as a labeled role. The intensity du/u is sigma-finite and the process is proper; multiplying its probability law by $e^{\gamma_E}\zeta$ multiplies both sides by the same constant.

Here is the precise exponential majorant for a low witness. In the pooled configuration X, let

$$T_X(w) = \sum_{\substack{v\in X\\ v\ge w}} v + (m-1)w, \qquad L_{\le l}(X) = \sum_{\substack{v\in X\\ v\le l}} v.$$

Thus $H_{m,\xi}$ is the maximum of $T_X(w)$ over the activated witnesses $w > \xi$. For a witness in a bin $w \in (l, u]$, and $m \ge 1$,

$$T_X(w) \le s - L_{\le l}(X) + (m-1)u.$$

Consequently every $\beta > 0$ gives the pointwise bound

$$\mathbf{1}_{\{T_X(w)>U\}} \le \exp\{\beta[s + (m-1)u - U - L_{\le l}(X)]\}. \tag{9.2}$$

One sums this bound over witnesses in the active part of the bin and applies (9.1). This explains both the upper endpoint u and the negative low-fragment term in the kernel. Replacing the full tail merely by a largest-fragment condition would not give this estimate. The high-fragment and residual-low-fragment laws factor into positive convolution kernels by the Poisson restriction theorem [14, Theorem 5.2]; their Laplace transforms are given by [14, Theorem 3.9]. The corresponding finite kernel construction is [18, Lemmas 2.3–2.4]. Upper endpoints of witness bins and original source windows are used in every outward replacement.

The rank-two and high terms use two and three marked fragments respectively. The marks may belong to the same coordinate or to different coordinates; all owner assignments are included. In an outer-face source the event is still an event of the full 39-coordinate root. It is not redefined after the detected coordinate has been erased.

Thus each component has a positive kernel $w_j \geq \mathbf{1}_{E_j}$ and a source integral of the form required in (8.8), or an inner deletion integral. The finite geometry verification checks the assigned windows, activations and bin endpoints against the row predicates, including their lower-order ancestors.

9.4 Aggregating a fine grid without changing the profile

The source bounds are accelerated by grouping 48 consecutive fine indices. This grouping does not replace (7.3) by a new profile on a coarser grid. Write a fine coordinate index as

$$j = 48q + r, \qquad 0 \leq r \leq 47.$$

For d coordinates with coarse-index sum q_{tot}, the fine-index sum lies in

$$[48q_{\mathrm{tot}},\, 48q_{\mathrm{tot}} + 47d]. \tag{9.3}$$

Each branch is intersected with every relevant original source piece and spline span. Positive measures are aggregated *after* multiplying by the original fine midpoint powers. Therefore the angular functions and the normalizing Z remain those of the same fine profile.

Signed radial coefficients are handled by enclosing the full polynomial or affine expression over (9.3). Bernstein coefficients give a convenient scalar range enclosure on a polynomial interval: the Bernstein basis is nonnegative and sums to one, so its coefficient minimum and maximum enclose every value. For the face forms we use a matrix enclosure that also includes jumps. Let $J = [r_-, r_+] \cap \mathbb{Z}$ be one of the clipped fine windows, let $\ell_J = r_+ - r_-$, and let v_r be the vector of coefficients of the *complete signed sum* of cap affines at fine index r. Put $\Delta v_j = v_{j+1} - v_j$. For any fixed $r_0 \in J$, telescoping and Cauchy–Schwarz give the positive semidefinite inequality

$$(v_r - v_{r_0})(v_r - v_{r_0})^* \preceq \ell_J \sum_{j=r_-}^{r_+-1} \Delta v_j \Delta v_j^* \qquad (r \in J). \tag{9.4}$$

For a singleton window both sides are zero. Otherwise, testing against any vector reduces this statement to the scalar Cauchy–Schwarz inequality for at most ℓ_J increments. A knot or cap jump is simply one of those increments, so it cannot disappear from the bound. Young's inequality now yields

$$v_r v_r^* \preceq (1+\eta) v_{r_0} v_{r_0}^* + (1 + 1/\eta)\ell_J \sum_{j=r_-}^{r_+-1} \Delta v_j \Delta v_j^*, \qquad \eta > 0. \tag{9.5}$$

After contraction with the angular monomials, the right-hand side is the positive Gram majorant used for the entire window. The full signed coefficient vector is formed before either bound is applied. For root components, a separately valid Minkowski bound may instead combine their individual norm bounds.

Multiplying such a positive majorant by an upper bound for a positive source kernel preserves the inequality. By contrast, entrywise enlarging coefficients of an arbitrary signed integrand would not be justified. Positivity is the reason the aggregation is safe.

9.5 Directed recurrences for the low kernels

One of the costly positive kernels has a simple exact recurrence. For integers $1 \le f < c$, define the formal power series

$$H(z) = \sum_{j=f}^{c-1} \frac{z^j}{j}, \qquad T_r(z) = \frac{H(z)^r}{r!}, \qquad E(z) = e^{H(z)}.$$

Differentiating gives, for $n \ge 1$,

$$nE_n = \sum_{j=f}^{c-1} E_{n-j}, \qquad n(T_r)_n = \sum_{j=f}^{c-1} (T_{r-1})_{n-j}. \tag{9.6}$$

Negative subscripts mean zero. There is no extra factor $1/r$ in the second formula: differentiation has already changed $r/r!$ to $1/(r-1)!$.

At dyadic scale $Q = 2^b$, store lower and upper bounds as integers. The moving-window sums are exact integer operations. Division by n is rounded down for the lower bound and up for the upper bound. Starting at $E_0 = (T_0)_0 = 1$, induction proves that all coefficients are enclosed. The exact possible support of T_r is $[rf, r(c-1)]$, so empty ranges can be skipped without approximation.

For a finite count prefix, the remaining positive series is

$$T_{>R} = E - \sum_{r=0}^{R} T_r.$$

Its upper coefficients are obtained by subtracting a *lower* prefix from an *upper* full exponential. The nonnegative carry polynomials have each coefficient at most one, so the entire omitted-count contribution is majorized coefficientwise by

$$\frac{T_{>R}(z)}{1-z}.$$

The prefix subtracted here is the unreplicated count prefix; subtracting an upper or already tilted prefix would not give this bound.

The remaining low-residual coefficients are rounded upward to positive dyadic numbers. Exponentially small tails are bounded above, rather than silently replaced by zero. All these replacements occur inside positive convolutions and positive Gram bounds. The original fine cell geometry, signed marginal and common normalization are retained throughout.

10 The complete numerical certificate

10.1 How to read the tables

The following tables are generated from the frozen exact-rational certificate. All their decimal entries are rounded upward and are in units of 10^{-6}. The full rational values, including individual source bins, are in the digital supplement. The display rounding is for readability; the final sum is assembled from the exact values.

For outer components, the last column is their contribution

$$2\sqrt{A_j M_j} + (t - \gamma_{\mathrm{op}})A_j$$

to the restoration loss. It does not include either of the two shared inner-error terms.

Outer source	Count	$\sum A_j$	$\sum M_j$	Loss
G1: low	8	0.749574898	14.579171091	6.699042427
G1: rank two	14	0.224037660	1.558976746	1.177680761
G1: high	1	0.000019087	0.000031473	0.000051304
G2: low	22	3.755008125	40.782076795	25.186856792
G2: rank two	14	23.449034139	151.904930860	121.397753113
G2: high	1	0.779494639	5.803481072	4.347133100

The inner table already includes ρ_*, the appropriate old/new/C coefficient, and division by I_-. All rank-two and high terms of the subtraction domain occur once.

Inner source	Count	Contribution to b
G3: low	8	0.007585797
G4: low	21	0.017382180
G5: low	8	0.413045470
G6: low	20	2.684023505
C: low	7	9.828292827
All rank-two and high terms	73	13.752470897

Their exact sum satisfies

$$b < .000026702801. \tag{10.1}$$

The restoration calculation, with $t = 11/10$, has the following four terms.

Term	Upper bound in units of 10^{-6}
Inner deletion b	26.702800674
Shared inner cross term $\Gamma b/t$	25.966859017
Outer cross terms $2\sum_j \sqrt{A_j M_j}$	155.342695198
Remaining outer mass $(t - \gamma_{\mathrm{op}})\sum_j A_j$	3.465822296

Every square root in this calculation is rounded outward using integer square roots of rational numbers. No floating optimizer or unproved error tolerance enters the last inequality.

10.2 The final strict comparison

The exact assembly gives

$$b + \frac{\Gamma b}{t} + 2\sum_j \sqrt{A_j M_j} + (t - \gamma_{\mathrm{op}})\sum_j A_j \le \bar{L} = \frac{105739089}{500000000000} = .000211478178. \tag{10.2}$$

Subtracting this from the cap lower bound (7.6) gives

$$\begin{aligned} \frac{\langle PF, (\mathsf{T} - \mathrm{Id})PF\rangle}{I(F)} &\ge \mu_- - \bar{L} \\ &= \frac{57362629154996791985 1}{879609302220800000000000} \\ &> .000065213759 > 0. \end{aligned} \tag{10.3}$$

This includes all 197 source components. No omitted component is assigned zero and no prospective target budget is used in place of an evaluated bound.

In particular $PF \neq 0$. Since $\mathsf{T} = \rho_* B$, the numerator on the left of (10.3) is exactly

$$\rho_* \mathcal{J}(PF) - I(PF).$$

It is the criterion of Corollary 6.2 for the true arithmetic profile.

10.3 Certificate construction

The coordinate-kernel and directed-contraction methods are developed in [18, Sections 2.1–2.3]. The current coefficient vector, integration records and final bounds are supplied by the digital supplement identified in Appendix B. The real-ball arithmetic uses Arb's enclosure operations [12, Sections 2–3]; the supplement pins the library versions and the producers used for these calculations.

The cap integrals use interval arithmetic, exact rational spline pieces, and exact Dickman logarithm and dilogarithm primitives on $0 \le u \le 3$. The largest Dickman argument in this trial is $393177/140306 < 3$. The source bounds use the positive constructions described in Section 9.

There are two independent scalar implementations. The main implementation assembles the stored interval endpoints, checks matching trials, inventories and normalizations, and evaluates the restoration inequality. The second reads the raw source outputs directly and uses exact fractions and integer square roots, without importing the main assembler. Both give the conservative ceiling (10.2) and the margin (10.3).

The source integrations are supplied by their recorded producers and pinned runtime. The fast scalar checks do not repeat those long integrations. File hashes identify which data and producer versions were used; a hash does not establish the mathematical validity of a producer. The underlying integration commands are also provided for complete numerical regeneration.

11 Passing from the fixed profile to primes

It remains to justify applying the arithmetic identities to the discontinuous fixed profile. The bounded-profile and exact-marginal extension is [17, Lemma 3.8]. The relevant limiting procedure is finite before x tends to infinity. We do not assert uniformity as the grid size grows.

Choose a positive fragment cutoff below every active witness threshold and cap. The smaller fragments can be collected into their total as in (2.2). On bounded total support there are only finitely many recorded large fragments. The grid profile, row predicates and upper step kernel have finitely many boundary types: cell totals, spline knots, activation levels, tail bounds and largest-fragment caps. Their nonconstant boundaries have limiting measure zero.

To respect hereditary arithmetic supports, use the fixed-band construction of [17, §4.2.4]. As the bands are refined, the chance that two large fragments occupy the same coordinate and band is bounded by a constant times

$$\sum_{\text{bands } (\alpha,\beta]} \left(\int_\alpha^\beta \frac{du}{u} \right)^2,$$

which tends to zero. Away from these collisions each active block represents one fragment, so the band predicates agree with the desired fragment tests. Zero-band strata are retained, and inward smooth approximations remove small neighborhoods only of the nonconstant boundaries.

Dominated convergence gives convergence of the ordinary norms. Lemma 2.1 gives convergence of the erased-coordinate forms. The fixed K is bounded, so the exceptional energies converge as well. Positively marked source integrals have bounded mark counts above their fixed cutoffs and fixed Chernoff parameters, so their corresponding approximation is also valid.

The order of choices is as follows: fix the positive margin in (10.3); choose bands and inward approximations fine enough to preserve a positive margin and all strict source inequalities; then let $x \to \infty$ with those choices fixed. This is precisely the fixed-profile realization used by the arithmetic pairing lemmas.

For every admissible 39-tuple, the ordinary and prime-weighted identities therefore give, at every sufficiently large scale,

$$\sum_n \sum_{i=1}^{39} \mathbf{1}_{\mathbb{P}}(n+h_i) A_{PF}(n)^2 > \sum_n A_{PF}(n)^2.$$

The right side is nonzero and every weight is nonnegative. Equation (1.1) produces a translate with two primes at arbitrarily large scales. This proves DHL$[39, 2]$.

12 The admissible tuple and the number 182

The following tuple is from the MIT prime gaps table [16]:

$$0, 6, 14, 20, 24, 26, 30, 36, 42, 44, 50, 54, 62, 66, 72, 80, 84, 86, 92, 96,$$
$$104, 110, 114, 120, 126, 132, 134, 140, 146, 150, 152, 156, 162, 164,$$
$$170, 174, 176, 180, 182.$$

It has 39 distinct entries and diameter 182. For each prime at most 39, the following table supplies one omitted residue.

p	2	3	5	7	11	13	17	19	23	29	31	37
omitted residue	1	1	3	4	1	12	5	2	10	3	12	1

For a larger prime, 39 entries cannot occupy all its residue classes. This verifies admissibility.

Every translate supplied by DHL$[39, 2]$ has two primes whose difference is at most 182. Between these primes there is a consecutive-prime pair with no larger gap. The translates occur arbitrarily far out, so these give infinitely many such pairs. Hence $H_1 \leq 182$, completing Theorem 1.2.

13 Formal verification

The public Lean development [25], hosted at `romyar1/prime-gaps-lean`, follows the fixed 39-coordinate profile through the analytic reductions, source geometry, sieve limits and prime-detection argument. It proves the implications from the explicit inputs below to DHL[39, 2] and $H_1 \le 182$. This section specifies those inputs and the deficit enclosure used in the formal proof. The source and reproduction links in this paper refer to commit `9c1ab6648bcd`; Appendix B gives the full revision identifier and commands. The main entry is `formal/PrimeGaps182Analytic.lean`. In the namespace `PrimeGap182`, the theorem `infinite_consecutive_prime_pairs182_of_local_inputs` asserts the infinitude of actual consecutive-prime pairs with gap at most 182. Its hypotheses are listed in `docs/assumptions.md` and summarized below.

13.1 Verified implications and explicit inputs

262 physical integral inequalities.

There are five cap bounds, 60 outer-root bounds, 60 outer-face bounds, and 137 inner-face bounds. They concern the fragment-law integrals of the fixed profile. The five cap bounds enclose the denominator on both sides and the three terms in the cap numerator in the required directions. Each outer record supplies two inequalities, which explains why 197 source records become 262 numerical premises after the cap bounds are included. The normalizations are those of (8.8).

Established finite-field estimates.

These include the normalized rank-three Kloosterman bound 3, from [13, Theorem 4.1.1(1)–(2)]: rank three and weight two give a raw bound $3p$, and the sum used here is divided by p. The raw two-rank-two correlation bound $8p\sqrt{p}$, at nonzero parameters and with its two poles excluded, is [7, Proposition 2]. Each rank-two sum there is divided by $\sqrt{p}$, which accounts for the extra factor p here; see also [17, Lemma A.4]. The incidence argument uses the scalar rank-four bound $4p\sqrt{p}$ for every prime and nonzero argument. It is the rank-four, weight-three specialization of the same theorem of Katz, and is a separate premise in the formal development. It does not assume an incidence matrix bound or a global distribution estimate.

The local Fourier proposition.

This is the finite-exceptional four-cycle estimate in [11, Proposition 3.1], with constants uniform in the prime and all its nonzero parameters. It remains an explicit hypothesis of the final prime-gap theorem. The repository also contains a conditional geometric derivation, described below.

More precisely, the local Fourier proposition requires fixed $C_{\mathrm{loc}} \ge 0$ and integers D, p_0 such that, for every prime $p > p_0$, the actual four-cycle W of the companion satisfies the following bounds for its unnormalized positive-sign two-variable Fourier transform. For distinct row and column indices there is a set $Z \subset \mathbb{F}_p^2$, with $\#Z \le D$, such that

$$|\widehat{W}(h,k)| \le C_{\mathrm{loc}} p^3 \big(1 + \sqrt{p}\,\mathbf{1}_Z(h,k)\big).$$

For a repeated row or column, there is a nonzero polynomial P over $\overline{\mathbb{F}}_p$, of degree at most D, such that

$$|\widehat{W}(h,k)| \le C_{\mathrm{loc}} p^3 \big(1 + \sqrt{p}\,\mathbf{1}_{\{P=0\}}(h,k) + p\,\mathbf{1}_{\{(0,0)\}}(h,k)\big).$$

A curve of exceptions is insufficient in the distinct-index case. The constants cannot be chosen separately for each prime. Small primes are covered by a proved trivial estimate; the arithmetic moduli are not required to omit them.

Subject to these inputs, the formal proof derives the global source-distribution estimates, the sharp minorant and its mean, the true-event covers, the operator bounds, the approximation by admissible profiles, and the actual positive sieve moment. None of those global conclusions is an additional endpoint hypothesis. The prime-gap endpoint concerns ordinary primes with no intervening prime, and the translate endpoint retains the admissibility and cardinality hypotheses of DHL[39, 2].

The endpoint declarations have logical dependency reports containing only the standard Lean principles `propext`, `Classical.choice` and `Quot.sound`. Such reports check the conditional implications, not their explicit hypotheses. In particular they do not prove the 262 integral inequalities or the local Fourier proposition. The companion supplies the ordinary mathematical proof of the latter.

The posted numerical release records a fresh external recomputation completed on 9 September 2026. All 262 rational target comparisons, 14 implementation audits and four deliberate rejection tests passed; the acceptance receipt and comparison inventory are linked from `research/numerical_182/README.md`. This is computational evidence for the integral premises. The Lean theorem continues to take `PhysicalSourceBounds182` as an argument; the external computations do not supply Lean proof terms for its inequalities. Appendix B gives separate commands for checking the release and repeating its integrations.

The supporting Type III entry is `formal/TypeIIILocalProgress.lean`. Its `research/type_iii/published_inputs/README.md` describes deductions of support exclusion and Fourier trace bounds from explicit general geometric laws and uniform data realizing the actual correlation family. These theorems retain the source-family identifications, compatibility of the geometric operations and uniform complexity witnesses as mathematical premises. They therefore leave the local Fourier input explicit in the prime-gap endpoint.

13.2 Two enclosures for the deficit

The rational moment argument of Section 4.5 gives $\kappa_{\text{true}} < \bar{\kappa}$, with $\bar{\kappa}$ as in (4.3). An alternative convexity argument gives

$$\kappa_{\text{true}} < \kappa_{\text{cvx}} := \frac{11696}{125000000} = .000093568.$$

The convexity bound is proved in Lean without numerical-integral, distribution or finite-field premises. For the positive convex density on each simplex in Section 4.5, the integral is at most the simplex volume times the average of its five vertex values. The three rational vertex averages give the displayed upper bound.

Both enclosures apply to the same minorant and fixed trial. Fix ℓ as in (6.8) and $d_h = \bar{\kappa}(1/\ell - 1)$. Put

$$\xi_{\text{add}} = \frac{(1-\ell)^2}{d_h}(\kappa_{\text{cvx}} - \bar{\kappa}) = \ell(1-\ell)\frac{\kappa_{\text{cvx}} - \bar{\kappa}}{\bar{\kappa}}. \tag{13.1}$$

Choosing the auxiliary profile $(1-\ell)v$ in the proof of the tail hybrid and applying Young's inequality with the same coefficient d_h gives the lower form

$$\mathcal{J}_{\text{cvx}}(U) = \mathcal{J}(U) - \xi_{\text{add}} J_+(U).$$

The additional charge to J_+ is exactly $(1-\ell)^2(\kappa_{\mathrm{cvx}}-\bar{\kappa})/d_h$. Thus $\mathcal{J}_{\mathrm{cvx}}$ uses the coefficient d_h and tail kernel of $\mathcal{J}$, together with this unweighted enlargement-square term.

Exact rational arithmetic gives

$$0 < 4\rho_*\xi_{\mathrm{add}} < \frac{32}{10^6}.$$

By Lemma 2.1 and the orthogonal projection, $J_+(PF) \le 4I(PF) \le 4I(F)$. Consequently the same numerical premises and (10.3) imply

$$\frac{\rho_*\mathcal{J}_{\mathrm{cvx}}(PF) - I(PF)}{I(F)} > \frac{33}{10^6} > 0.$$

Both deficit enclosures therefore give a positive sieve moment for the fixed trial, using the same cap and source integral bounds and their stated normalizations. The convexity enclosure supplies the margin used by the formal implication.

The public repository supplies the `./verify.sh` verifier, complete endpoint types and numerical-premise definitions; see `docs/verification.md` and Appendix B. Rebuilding the formal implication does not regenerate the interval integrals. The companion's many-primes exponential-bound theorem is an ordinary analytic result; it is not a formal endpoint of this development.

A Notation at a glance

Symbol	Meaning
$x, R = x^{\rho_*}$	Arithmetic scale and coefficient-divisor scale
ρ_*, ρ	Actual scaling and slightly larger design parameter
ρ_D, ν_ζ	Dickman function and finite fragment measure
$X, \lvert X\rvert, M_\xi$	Fragment configuration, its total, and its largest fragment above ξ
$H_{\phi,\xi}$	Inclusive tail obstruction in (3.7)
$F, U = PF$	Fixed cap profile and its actual arithmetic restriction
E_i, V_i	Erasure of coordinate i, and the resulting face profile
$I(F), \mathcal{N}$	Unscaled profile norm and common numerical normalization
$\mathcal{H}_O, \mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_C$	Outer, base, enlarged and subtraction cap domains
$O, L_0, L_1, \mathcal{C}$	Their actual arithmetic counterparts
ϱ_x, b_x	Sharp minorant and its nonnegative exceptional part
$\kappa_{\mathrm{true}}, \bar{\kappa}$	True mean deficit and a rational upper bound
K, ℓ, d_h	Exceptional step kernel, fixed hybrid parameter, and $\bar{\kappa}(1/\ell - 1)$
$\mathsf{C}, \mathsf{T}, \mathsf{D}$	Scaled cap operator, actual-face operator and their positive difference
A_j, M_j, b	Normalized outer mass, outer operator-energy and complete inner-deletion bounds
I_-, I_+	Lower and upper enclosures of $I(F)/\mathcal{N}$
$\mu_-, \bar{L}$	Certified cap surplus and complete restoration loss ceiling

B Reproduction and the digital supplement

The source code, fixed data and verification instructions are public in `romyar1/prime-gaps-lean` [25]. The numerical programs, runtime and complete accepted outputs are distributed in its `external-numerics-20260909` release. The instructions below distinguish the Lean build, the integrity checks on the numerical evidence, and a fresh numerical recomputation. All source paths are relative to the public repository; commands for the extracted release are identified separately.

B.1 Fixed data and numerical evidence

The literal Lean trial is in `formal/TrialData182.lean`, and the source parameters and rational thresholds are in `formal/SourceData182.lean`. The integral definitions and inequality types are in `formal/PhysicalTrial182.lean`, `formal/SourceKernels182.lean` and `formal/SourceBounds182.lean`. The collection `inputs/source_certificates/` contains the fixed trial and saved cap and source records. Its file identities and numerical payloads are checked by `provenance/check_data.py`.

The trial has 726 coefficients, 26 knots, 11 angular signatures, three component caps and 12 shells. The numerical package checks these against the literal Lean data; see `research/numerical_182/evidence/trial-comparison.json`. For the JSON representation, the full profile is given by the three `independent_components` arrays. The top-level vector is a constructor seed. The spline variable is the uncentered s° of (7.3); its definition does not use the JSON `center` field. The exact trial file in the numerical release has SHA-256 digest

`3ebae41700a7839f2d7a7b3fe068b11ed364`
`a7542aab110173bbcdbd66645952.`

The fresh external run completed on 9 September 2026. Its `research/numerical_182/evidence/verification.json` records acceptance of all 262 target comparisons, 14 implementation audits and four deliberate rejection tests. The `research/numerical_182/evidence/coverage.json` record pairs the comparisons with their exact rational thresholds: five cap inequalities, two for each of 60 outer components, and one for each of 137 inner components. The same run checked 328,936 source geometry equalities and exact agreement of 20 freshly generated affine arrays with the reference arrays. These are external computational checks; the named integral inequalities remain explicit Lean premises.

B.2 Checking and repeating the numerical computation

The release asset is `prime-gaps-external-182-verified-20260909.zip`, with SHA-256 digest

```
2960729efa8996538731820ce6cadc20283
1c8e98b2c0bdb14cfd19f33726f3b.
```

The archive includes the programs, input arrays, tested runtime, fresh production outputs and acceptance evidence. The release guide is `research/numerical_182/README.md`; `research/numerical_182/asset.json` pins the archive identity and the corresponding repository files.

With Python 3.11 or later, the following commands run from the repository root. The second assumes that the downloaded ZIP is in the parent directory; its argument can be replaced by the chosen download location.

```
python3 -B scripts/check_external_numerics.py
python3 -B scripts/check_external_numerics.py \
  --archive ../prime-gaps-external-182-verified-20260909.zip
```

The first checks the repository's verifier copies, acceptance records and five current Lean target files. The second additionally checks the ZIP checksum and every archived file. Neither command repeats an integration. The original source-record collection can be checked by `python3 -B provenance/check_data.py`.

For a fresh computation, extract the release and run from its `prime-gaps-external-182` directory. The tested platform is Apple Silicon macOS with Python 3.12 and NumPy 2.3.5. The package provides python-flint 0.9.0 and FLINT 3.6.0 with the signed-FFT repair; `research/numerical_182/RUNTIME.md` records the exact library identity and its regression check. Use a Python 3.12 environment with NumPy 2.3.5, for example after `python3.12 -m pip install -r requirements.txt`. The work and export directories below must be new, on a local disk outside a synchronized folder.

```
python3.12 -B check_package.py
python3.12 -B audit.py --work ../pg182-audits
python3.12 -B test_rejection.py --work ../pg182-negative
python3.12 -B verify.py --work ../pg182-production --jobs 3
python3.12 -B finish.py --work ../pg182-production \
  --audits ../pg182-audits --negative ../pg182-negative \
  --export ../pg182-evidence
```

Use `-jobs 1` to reduce simultaneous work, and do not invoke Python with `-O`. The production driver regenerates the cap, affine and source outputs; the final driver accepts the full run only after its comparisons, audits and rejection tests succeed. The final `verification.json`, with status `PASS_COMPLETE_EXTERNAL_NUMERICAL_RECOMPUTATION`, records acceptance. The method and dependency chain are explained in `research/numerical_182/METHOD.md`.

The release targets the same five Lean numerical-source files as the repository revision below; the attachment checker verifies those identities. Repeating the numerical programs supplies external evidence for their integral bounds. The mathematical identification of the algorithms with those integrals is distinct from file integrity and from the conditional Lean build.

B.3 Reproducing the Lean verification

The referenced repository revision is

`9c1ab6648bcd901c24c32e0101136b2253bbe9bd.`

It pins Lean `4.34.0-rc2` and Mathlib commit

`bbcd1968ee6950abe88b85dba6995da346c4b2a8.`

With Elan, Git and Python 3.11 or later installed, run

```
git clone https://github.com/romyar1/prime-gaps-lean.git
cd prime-gaps-lean
git checkout --detach 9c1ab6648bcd901c24c32e0101136b2253bbe9bd
lake exe cache get
./verify.sh
```

The verifier checks the pinned dependencies, builds the project, checks source and numerical-data integrity, and prints theorem types and axiom dependencies. For a fresh build of the project sources and vendored baseline, use `./verify.sh -fresh`. `docs/verification.md` specifies the complete audit policy and the generated verification receipts. These commands check all formal modules, including the incidence and Type III supporting work.

The Lean verification workflow and the separate external-package integrity workflow both report successful runs for this revision. The Lean endpoint is the conditional theorem described in Section 13; its numerical and finite-field hypotheses remain visible in the printed type.

References

[1] R. Arratia, A. D. Barbour and S. Tavaré, *The Poisson–Dirichlet distribution and the scale-invariant Poisson process*, Combin. Probab. Comput. **8** (1999), 407–416. Author's copy.

[2] R. C. Baker and A. J. Irving, *Bounded intervals containing many primes*, Math. Z. **286** (2017), 821–841. Theorem and section numbers refer to arXiv:1505.01815v1.

[3] F. Charton, L. Hong, K. Lau, K. Ono, G. Remy, H. C. Siu, A. A. Swaminathan, J. Thorner and Y. Xie, *A new bound for small gaps between primes*, 3 September 2026. Full text.

[4] E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), 203–251. Published article.

[5] M. G. Cox, *The numerical evaluation of B-splines*, J. Inst. Math. Appl. **10** (1972), 134–149. Published article.

[6] C. de Boor, *On calculating with B-splines*, J. Approx. Theory **6** (1972), 50–62. Published article.

[7] É. Fouvry, E. Kowalski and P. Michel, *The Friedlander–Iwaniec character sum*, 14 June 2013. Full text.

[8] D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples I*, Ann. of Math. (2) **170** (2009), 819–862. Full text.

[9] D. R. Heath-Brown, *Prime numbers in short intervals and a generalized Vaughan identity*, Canad. J. Math. **34** (1982), 1365–1377. Published article.

[10] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Incidence operators and intervals containing many primes*, companion paper, 2026.

[11] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Kloosterman correlations with finitely many exceptional Fourier frequencies*, companion paper, 2026.

[12] F. Johansson, *Arb: Efficient arbitrary-precision midpoint-radius interval arithmetic*, IEEE Trans. Comput. **66** (2017), 1281–1292. arXiv:1611.02831v1.

[13] N. M. Katz, *Gauss Sums, Kloosterman Sums, and Monodromy Groups*, Annals of Mathematics Studies, vol. 116, Princeton University Press, 1988. Author's copy.

[14] G. Last and M. Penrose, *Lectures on the Poisson Process*, Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, 2018. Theorem numbers refer to the authors' text of 21 August 2017. Authors' copy.

[15] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), 383–413. Published article.

[16] MIT prime gaps project, *Admissible 39-tuple of diameter 182*, online tuple table. Tuple data.

[17] OpenAI, *Improved short gaps between primes*, 30 August 2026. Full text.

[18] OpenAI, *Numerical certificate for prime gaps at most 186*, accompanying manuscript and code, 2026. Numerical companion; PrimeGaps186 repository.

[19] D. H. J. Polymath, *New equidistribution estimates of Zhang type*, Algebra Number Theory **8** (2014), 2067–2199. Theorem and section numbers refer to arXiv:1402.0811v3.

[20] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), article 12; erratum, **2** (2015), article 15. Theorem and section numbers refer to arXiv:1407.4897v4.

[21] J. Stadlmann, *On primes in arithmetic progressions and bounded gaps between many primes*, Adv. Math. **468** (2025), article 110190. Theorem and section numbers refer to arXiv:2309.00425v3.

[22] J. Stadlmann, *Bounded gaps between primes*, 31 August 2026. arXiv:2608.31126v1.

[23] T. Tao, *Polymath8b: Bounded intervals with many primes, after Maynard*, What's New, 19 November 2013. Online article.

[24] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), 1121–1174. Published article.

[25] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Prime gaps in Lean*, source code, numerical certificates and reproduction instructions, 2026. Public repository, commit `9c1ab6648bcd` (accessed 11 September 2026). External numerical release `external-numerics-20260909`.

Incidence operators and intervals containing many primes

GPT-6 Astra (OpenAI), prompted by Romyar Sharifi

11 September 2026*

Abstract

We retain a positive quadratic form in the dispersion method for bilinear convolutions. An integer change of variables expresses this form as an incidence operator acting on a weighted set of fractions. Its complete Fourier modes reduce to a circulant operator whose eigenvalues are rank-four hyper-Kloosterman sums. This gives the operator bound needed for completion in two smooth, possibly sheared, windows. An elementary weighted Farey estimate then yields the Type II condition

$$\gamma > \frac{3}{10} + 8\omega + \frac{16}{5}\delta,$$

together with the preliminary inequalities stated in the paper. We prove the transfer with the original Fourier weights, congruence conditions, and divisor sectors retained, and state the precise extension to densely divisible moduli. Combining this result with the Type III theorem in the companion paper and a Harman minorant gives $H_m \ll \exp(3.7877m)$. The constant implicit in this asymptotic statement is not evaluated.

Contents

*Original version: 4 September 2026. This version expands citations and exposition; the stated results and numerical data are unchanged.

1 The result and the structure of the argument

Let p_n be the nth prime and write

$$H_m = \liminf_{n\to\infty}(p_{n+m} - p_n).$$

Thus H_1 measures gaps between consecutive primes, whereas H_m measures the length needed to contain $m+1$ consecutive primes infinitely often. There are two related quantitative questions: an explicit numerical bound for H_1, and a bound for the growth of H_m as m increases. This paper concerns the second question, as well as a convolution estimate that can be used for the first.

The sieve of Goldston, Pintz, and Yıldırım [3, Theorems 1–2] connected small prime gaps with the distribution of primes in arithmetic progressions. Yitang Zhang's 2013 breakthrough proved $H_1 < 70\,000\,000$ by obtaining the necessary distribution beyond the Bombieri–Vinogradov range for restricted moduli [14, Theorem 1]. Polymath8a improved this bound to 4680 and developed the smooth-modulus distribution estimates used throughout this subject [7, Section 1 and Theorem 1.1]. Later in 2013, Maynard introduced a multidimensional sieve, independently developed by Tao [7, Section 1], that gave $H_m < \infty$ for every fixed m; Maynard's numerical bound was $H_1 \le 600$ [6, Theorems 1.1 and 1.3]. Polymath8b subsequently obtained $H_1 \le 246$ [8, Theorem 1.4(i)]. The bounds 600 and 246 use the classical Bombieri–Vinogradov theorem; their proofs do not require Zhang's stronger distribution estimates.

The many-primes estimates follow a different numerical sequence. Maynard proved

$$H_m \ll m^3 e^{4m},$$

and Polymath8b used stronger distribution information to obtain

$$H_m \ll m \exp\left(\left(4 - \frac{28}{157}\right) m\right) \qquad \text{[8, Theorem 1.4(vi)]}.$$

Baker and Irving combined the multidimensional sieve with Harman's prime minorants to reach $H_m \ll e^{3.815m}$ [1, Theorem 1.1]. Stadlmann then improved this to $H_m \ll e^{3.8075m}$ [9, Theorem 2]. Her underlying arithmetic-progression theorem gives exponent $1/2 + 1/40$ for squarefree smooth moduli, with an arbitrary fixed retreat in the exponent [9, Theorem 1]. A minorant is useful here because its improved distribution can more than compensate for the prime density that it discards. The transfer depends on both quantities, as made explicit in Section 12.

Recent work has also lowered the explicit bound for H_1. Stadlmann obtained $H_1 \le 240$ by combining the Bombieri–Vinogradov range with stronger estimates for moduli having suitable factorizations [10, Theorem 1 and Sections 2–4]. Charton, Hong, Lau, Ono, Remy, Siu, Swaminathan, Thorner, and Xie, in the Axiom Math paper, built on her construction and obtained $H_1 \le 212$ [11, Theorem 1.1]. OpenAI obtained $H_1 \le 186$ from complementary factorization conditions and a larger sieve support [12, Theorem 1.1 and Corollary 1.2]. That paper describes its work and Stadlmann's 240 result as independent and concurrent. These finite bounds have separate variational certificates; they do not follow by substituting $m = 1$ into a many-primes estimate with an unspecified constant.

Our many-primes application is the following.

Theorem 1.1. *There is an absolute constant C such that*

$$H_m \le C \exp(3.7877m) \qquad (m \ge 1).$$

The main new estimate proved here concerns a bilinear convolution $\alpha * \beta$ at total scale x. If the shorter factor has length $N = x^\gamma$, we obtain distribution to certain squarefree moduli of size x^θ, where $\theta = 1/2 + 2\omega$. The significant lower bound for γ is

$$5\gamma > \frac{3}{2} + 40\omega + 16\delta. \tag{1.1}$$

The complete list of hypotheses appears in Theorem 2.1. The quantity δ measures either smoothness of the original modulus or the multiplicative width within which its required divisors can be chosen. The starting dispersion reduction is Stadlmann's [9, Section 3, especially Lemmas 3–6], which in turn uses Polymath8a's bilinear reductions [7, Section 5]. For the dense extension we retain the divisor conditions of Stadlmann [10, Lemma 6], in the explicit form developed in [11, Lemma 5.6 and Section 6.4]. The new step is the operator estimate for the positive form inside that reduction. Neither the dispersion framework nor the Harman minorant construction is new here: the latter follows the line of Baker–Irving [1, Sections 2–4] and Stadlmann [9, Section 4], with the parameter geometry and coefficient conditions proved below.

The application also uses a new Type III estimate. That estimate is proved in the separate paper [15, Theorem 1.1]; its full scale hypotheses are recalled in Proposition 10.1. It is an estimate for a convolution with three smooth factors, not a distribution theorem for primes to all moduli up to $x^{0.5284}$. Harman's sieve combines the different convolution estimates into a particular prime minorant. Its mean and its distribution level, together, give Theorem 1.1.

1.1 Why an incidence operator appears

After the usual dispersion reductions, one must bound a positive sum of squares with an inner expression of the form

$$\sum_{\lambda} c_\lambda \sum_{n \equiv j\lambda \pmod{e}} \psi_N(n) e_q\left(\frac{A\lambda}{e(n+Be)}\right). \tag{1.2}$$

For this explanation suppose that e and λ are units modulo q, and temporarily suppress the additional arithmetic masks. Writing

$$n = j\lambda + ek, \qquad b = (k+B)/\lambda \pmod{q}$$

turns the phase into $e_q(A/(ej+e^2b))$. Thus the dependence on the two row variables (e,j) is carried by the matrix

$$R_{e,j;b} = \mathbf{1}_{(ej+e^2b,q)=1} e_q\left(\frac{A}{ej+e^2b}\right).$$

The input vector counts fractions $(k+B)/\lambda$ with the original weights c_λ. This representation preserves the signed sum over λ inside each square. It also preserves the exact integer congruence: k is an integer, not a real substitute for the quotient $(n-j\lambda)/e$.

There are three distinct estimates in the argument. First, the complete Fourier modes of R^*R have norm $O(q^{3/2+o(1)})$, except for a controlled factor when both frequencies vanish at a prime divisor of q. Second, completion transfers this bound to the actual row windows. Third, the input vector has small enough squared norm by an elementary count of equal fractions. These steps occupy Sections 3–5.

The arithmetic transfer is not just a substitution into this model. The coefficients in (1.2) initially depend on e; some primes divide λ; the original congruence conditions couple e and n; and the numerator interval depends on λ. Sections 6–8 treat these points in the order needed to preserve positivity. In particular, an input-dependent mask is never removed from a signed sum by asserting an inequality.

1.2 What the many-primes conclusion uses

The final choice has

$$\theta = 0.5284, \qquad a = 0.41361, \qquad I = 0.34941, \qquad c = 0.41364.$$

Here I is the threshold for a long smooth Type I factor, $[a, 1-a]$ is the Type II band, and c describes the three Type III factors. The minorant is exactly

$$\rho = \mathbf{1}_{\mathbb{P}} - E_1 - E_2,$$

where E_1, E_2 count specified labeled products of five primes. Its mass loss is an ordinary four-dimensional prime-density integral. An explicit upper bound for that loss, followed by small fixed retreats, gives an exponential constant below 3.7877. No numerical value for H_1 follows merely by setting $m = 1$ in Theorem 1.1, since its absolute constant is unevaluated. The fixed-profile application to a specific gap belongs to [16].

2 Arithmetic conventions and the Type II statement

All asymptotic statements concern $x \to \infty$. Exponents such as ω, δ, γ are fixed. When several strict inequalities are imposed, the estimates are uniform on closed parameter sets lying a fixed positive distance inside the permitted region. We write $e_q(t) = \exp(2\pi i t/q)$ and $e(t) = \exp(2\pi i t)$. A quotient in $e_q(\cdot)$ denotes inversion modulo q and is used only when its denominator is a unit. The symbol $*$ denotes Dirichlet convolution, $(\alpha * \beta)(n) = \sum_{uv=n} \alpha(u)\beta(v)$. Parentheses (u, v) denote a greatest common divisor, and brackets $[u, v]$ a least common multiple.

For a finitely supported sequence f and a primitive residue $a \pmod q$, define its discrepancy by

$$\mathcal{E}_q(f; a) = \sum_{n \equiv a \pmod q} f(n) - \frac{1}{\varphi(q)} \sum_{(n,q)=1} f(n). \tag{2.1}$$

The second term is the coprime mean. In particular, we do not replace it by the unrestricted mean unless rough support later justifies that replacement.

2.1 Coefficient hypotheses

We use the coefficient classes of [7, Definition 2.5]. A coefficient sequence at scale N is supported in $[cN, CN]$, for fixed $0 < c < C$, and satisfies

$$|\alpha(n)| \ll \tau(n)^B (\log x)^B$$

for a fixed B. Complex sequences are allowed by separating their real and imaginary parts. The bounds for a family of sequences are uniform. The Siegel–Walfisz property means that, for every fixed $A > 0$, all integers $q, r \geq 1$, and all primitive $a \pmod q$,

$$\left|\mathcal{E}_q(\alpha \mathbf{1}_{(\cdot, r)=1}; a)\right| \ll_A \tau(qr)^{O(1)} N (\log x)^{-A}. \tag{2.2}$$

The divisor exponent is fixed. The extra coprimality condition in (2.2) is part of the hypothesis.

A smooth coefficient at scale N has the form $\psi(n/N)$, where ψ is smooth, supported in a fixed compact subinterval of $(0, \infty)$, and, for every fixed j, $\|\psi^{(j)}\|_\infty \ll_j (\log x)^{O_j(1)}$. Translated smooth windows will also occur, with their centers specified separately. Fixed smooth subdivisions and the separation of product inequalities are permitted with the same uniform coefficient hypotheses.

2.2 Residues and modulus classes

A *coherent primitive residue system* assigns a nonzero residue $a_p \pmod p$ to each relevant prime. For a squarefree modulus q, the class $a(q)$ is the Chinese-remainder combination of the a_p for $p \mid q$. Equivalently, on the divisors of any fixed squarefree product P, it is the restriction of one primitive class modulo P. Our estimates are uniform in this system, as in [7, Theorem 1.1 and Claim 2.3]. This convention fixes the compatibility between residue classes at different divisors of the same squarefree product.

An integer is y-smooth if every prime divisor is at most y. Initially the moduli divide a squarefree product P of primes below x^δ. Such integers are multiply x^δ-densely divisible in every fixed order [7, Lemma 2.10(iii)]. Section 9 states the more general factorization class on which the same proof works.

Theorem 2.1 (Incidence Type II estimate)**.** *Fix $\omega, \delta > 0$ and put $\theta = 1/2 + 2\omega$. Suppose*

$$12\omega + 6\delta < \gamma < \frac{1}{2} - 2\omega, \qquad 2\gamma + 4\omega + 2\delta < 1, \qquad \gamma > 16\omega + 7\delta, \qquad 5\gamma > \frac{3}{2} + 40\omega + 16\delta. \tag{2.3}$$

Let α and β be coefficient sequences at scales M, N, where $MN \asymp x$ and $N = x^\gamma$, and suppose that β has the Siegel–Walfisz property. For every fixed $A > 0$,

$$\sum_{\substack{q \mid P \\ q \ll x^\theta}} |\mathcal{E}_q(\alpha * \beta; a(q))| \ll_A x(\log x)^{-A}, \tag{2.4}$$

uniformly over squarefree P whose prime divisors are below x^δ, and over coherent primitive residue systems. The implied constant may depend on the coefficient bounds and the strict margins in (2.3)*.*

The role of the four conditions is worth recording before the proof. The first is a preliminary dispersion range. The next controls the volume of the Farey input. The third ensures adequate divided row lengths in the source transfer. The last makes the principal oscillatory error small and implies the two remaining error inequalities. At $\delta = 0$, its boundary is $\gamma = 3/10 + 8\omega$.

We will use ε for the small, fixed retreat in the dispersion proof and η for arbitrarily small divisor-bound losses. The order of choice is important: first fix the strict margins, then choose ε, then choose the finitely many η's much smaller than ε. A displayed $O(\varepsilon)$ in an exponent is not an $o(1)$ term when ε is fixed.

3 Complete modes of the incidence operator

This section is independent of the sieve. Let q be squarefree and $(A, q) = 1$. On the space $\ell^2(\mathbb{Z}/q\mathbb{Z})$ with counting measure, set

$$F_q(u) = \mathbf{1}_{(u,q)=1} e_q(A/u), \qquad R_{e,j;b} = F_q(ej + e^2 b). \tag{3.1}$$

Nonunit e give zero rows. The complete matrices are

$$D_q^{h,\nu}(b, b') = \sum_{e,j \bmod q} e_q(he + \nu j)\overline{R_{e,j;b}} R_{e,j;b'}. \tag{3.2}$$

These Fourier transforms are unnormalized: there is no q^{-1} or q^{-2} in (3.2). Matrix norms below are Euclidean operator norms.

Proposition 3.1. *For every fixed $\eta > 0$,*

$$0 \le D_q^{0,0} \le q^2 \,\mathrm{Id}, \qquad \left\| D_q^{h,\nu} \right\|_{\mathrm{op}} \ll_\eta q^{3/2+\eta} (q, h, \nu)^{1/2}. \tag{3.3}$$

At $q = 1$ the operator has one constant column and the first inequality is interpreted directly.

The exact constant one on the zero mode will matter in the arithmetic argument. The oscillatory estimate alone, with a q^η factor everywhere, would obscure the available retreat in the main term.

3.1 Prime-modulus identities

For an odd prime p, use the normalization

$$\mathrm{Kl}_r(t; p) = p^{-(r-1)/2} \sum_{x_1, \ldots, x_{r-1} \in \mathbf{F}_p^*} e_p\left(x_1 + \cdots + x_{r-1} + \frac{t}{x_1 \cdots x_{r-1}} \right). \tag{3.4}$$

Thus $\mathrm{Kl}_3(0; p) = 1/p$. For $t \ne 0$, the standard hyper-Kloosterman bound gives $|\mathrm{Kl}_r(t; p)| \le r$; we need $r = 4$ only. In Katz's construction the unnormalized sum is, up to sign, the trace of a rank-r sheaf pure of weight $r - 1$ [5, Theorem 4.1.1(1)–(2)]. The factor $p^{-(r-1)/2}$ in (3.4) gives the stated bound, in exactly the normalization of [7, Proposition 6.9 and Remark 6.10]. The nonzero-argument hypothesis is retained throughout its use.

Fix $b \ne b'$ and let $d = b' - b$. On a nonzero row put

$$u = j + eb, \qquad v = j + eb'.$$

Then $e = (v - u)/d$. The two denominators are units precisely when $u, v \ne 0$, and the omitted $e = 0$ rows are the line $u = v$. The complete phase becomes

$$-\frac{Ad}{uv} + \frac{\nu b' - h}{d} u + \frac{h - \nu b}{d} v. \tag{3.5}$$

For $C \ne 0$, a useful elementary identity is

$$\sum_{u, v \ne 0} e_p(C/(uv) + \alpha u + \beta v) = p\, \mathrm{Kl}_3(C\alpha\beta; p) - p \mathbf{1}_{\alpha = \beta = 0}. \tag{3.6}$$

If $\alpha\beta \ne 0$, rescale u, v. If exactly one coefficient vanishes, sum first in the variable without a linear term; the result is 1. If both vanish, the sum is $-(p - 1)$. These are exactly the three cases in (3.6).

Subtracting the line $u = v$ from (3.5) gives

$$\begin{aligned} D_p^{h,\nu}(b, b') &= p\, \mathrm{Kl}_3\left(\frac{A(h - \nu b)(h - \nu b')}{b' - b}; p \right) - S_\nu(b' - b) - p \mathbf{1}_{h = \nu = 0}, \\ S_\nu(d) &= \sum_{u \ne 0} e_p(-Ad/u^2 + \nu u). \end{aligned} \tag{3.7}$$

On the diagonal the original phase cancels. Direct summation over $e \ne 0$ and $j \ne -eb$ gives

$$D_p^{h,\nu}(b, b) = \begin{cases} 1 - p \mathbf{1}_{h = \nu b}, & \nu \ne 0, \\ (p - 1)(p \mathbf{1}_{h = 0} - 1), & \nu = 0. \end{cases} \tag{3.8}$$

This calculation keeps all zero linear coefficients. In particular, the last term of (3.7) cannot be omitted when $h = \nu = 0$.

3.2 A circulant operator with hyper-Kloosterman eigenvalues

For $a \in \mathbf{F}_p^*$ define

$$L_a(x,y) = \mathbf{1}_{x \neq y} \operatorname{Kl}_3(a/(x-y);p). \tag{3.9}$$

It depends only on $x-y$, so the additive characters $p^{-1/2}e_p(sx)$ diagonalize it. Their eigenvalues are

$$\lambda_s = \begin{cases} \sqrt{p}\operatorname{Kl}_4(-as;p), & s \neq 0, \\ -1/p, & s = 0. \end{cases} \tag{3.10}$$

Indeed, for $s \neq 0$ expand (3.4) and put $z = -st$ in the sum over $t = x-y \neq 0$:

$$\lambda_s = p^{-1} \sum_{t,u,v \neq 0} e_p\Big(u+v+\frac{a}{tuv} - st\Big) = p^{-1} \sum_{z,u,v \neq 0} e_p\Big(u+v+z-\frac{as}{zuv}\Big).$$

This is the first case of (3.10). When $s = 0$, the sum over t is -1, giving the second case. Consequently

$$\|L_a\|_{\mathrm{op}} \leq 4\sqrt{p}. \tag{3.11}$$

Suppose next that $\nu \neq 0$. The projective transformation

$$g(b) = \frac{1}{\nu(h-\nu b)} \tag{3.12}$$

is represented by $\left(\begin{smallmatrix} 0 & \nu^{-1} \\ -\nu & h \end{smallmatrix}\right)$, of determinant one. Away from its pole,

$$-\frac{A}{g(b)-g(b')} = \frac{A(h-\nu b)(h-\nu b')}{b'-b}.$$

Extend L_{-A} by a zero row and column at infinity. The principal matrix in (3.7) is therefore a compression of a permutation conjugate of pL_{-A}, except for the off-diagonal entries in the pole row and column. Those entries are all 1, because $p\operatorname{Kl}_3(0;p) = 1$. Their symmetric correction has norm $\sqrt{p-1}$.

The matrix with entries $S_\nu(b'-b)$ is also circulant. Its eigenvalues, up to a harmless Fourier sign, are

$$p \sum_{\substack{u \neq 0 \\ -A/u^2 = \xi}} e_p(\nu u).$$

There are at most two terms, so its norm is at most $2p$. Since $S_\nu(0) = -1$ when $\nu \neq 0$, the remaining diagonal correction in (3.8) has norm p. We have proved

$$\left\| D_p^{h,\nu} \right\|_{\mathrm{op}} \leq 4p^{3/2} + 3p + \sqrt{p-1} \qquad (\nu \neq 0). \tag{3.13}$$

For $\nu = 0$ and $h \neq 0$, the principal part is already circulant, with entries $p\operatorname{Kl}_3(Ah^2/(b'-b);p)$ off the diagonal. Equations (3.8) and (3.11) give the bound $4p^{3/2} + 2p$.

3.3 The zero mode and squarefree moduli

Write $\mathbf{J}$ for the all-ones matrix and S_0 for the full circulant with entries $S_0(b'-b)$, including its diagonal. The preceding exact identities give

$$D_p^{0,0} = (p^2-1)\,\mathrm{Id} - (p-1)\mathbf{J} - S_0. \tag{3.14}$$

The Fourier eigenvalues of S_0 are nonnegative: they equal p times the number of solutions of $-A/u^2 = \xi$. Hence $S_0 \geq 0$. The matrix $D_p^{0,0}$ is a Gram matrix, so (3.14) proves $0 \leq D_p^{0,0} \leq p^2\,\mathrm{Id}$. More explicitly its eigenvalues are $p-1$ on the constants and $p^2-p-1-p\chi_p(-As)$ for $s \neq 0$. At $p=2$ the two nonzero rows are a signed permutation matrix; each complete-mode norm is 1. This also gives the required zero-mode inequality.

For squarefree q, the Chinese remainder theorem identifies the input space with the tensor product of the prime input spaces. The complete matrices tensor as well. In the local factor at p, the parameters A, h, ν are multiplied by the unit $(q/p)^{-1} \pmod p$ in the appropriate additive characters. The estimates above are uniform in these unit changes. At a prime for which $(h,\nu) \not\equiv (0,0)$ the factor is $O(p^{3/2})$; at a common zero it is at most p^2. Multiplying gives

$$\left\| D_q^{h,\nu} \right\|_{\mathrm{op}} \ll C^{\omega(q)} q^{3/2} (q,h,\nu)^{1/2} \ll_\eta q^{3/2+\eta} (q,h,\nu)^{1/2}.$$

The exact zero-mode bounds multiply without $C^{\omega(q)}$. This proves Proposition 3.1.

4 Completion in smooth and sheared windows

The row windows in the dispersion argument are not generally rectangles. Their natural coordinates are a divisor and a ratio of two integers. A uniform shear in the following lemma lets us retain that geometry. The completion argument extends the one-dimensional smooth completion in [7, Section 4.4, Lemma 4.9 and its proof]; the shear and the treatment of all positive window scales are proved below.

Lemma 4.1 (Smooth-window bound). *Let $w \geq 0$ be a smooth function of compact support in $\mathbb{R}^2$. For arbitrary real e_0, j_0, τ and positive E_1, E_2, put*

$$W(e,j) = w\left(\frac{e-e_0}{E_1}, \frac{j-\tau e - j_0}{E_2} \right).$$

For squarefree q, $(A,q)=1$, and every input vector r,

$$\sum_{e,j \in \mathbb{Z}} W(e,j) |(Rr)_{e,j}|^2 \ll_{\eta,w} \{E_1E_2 + q^{3/2+\eta} + q^{1/2+\eta}(E_1+E_2)\} \|r\|_2^2 . \tag{4.1}$$

The bound is uniform in the size and arithmetic nature of τ. It also holds after an integer affine change of the row coordinates whose linear determinant is coprime to q.

Proof. Periodize W modulo q and apply finite Fourier inversion. The factor preceding each complete matrix is q^{-2}. The zero residue of the Fourier transform contributes exactly

$$q^{-2} \Big(\sum_{e,j \in \mathbb{Z}} W(e,j) \Big) D_q^{0,0}.$$

Its norm is at most the *unrestricted* lattice mass of W by Proposition 3.1. The restriction to unit rows is already inside R; it is not inserted a second time in this mass.

Poisson summation and repeated integration by parts give, for every fixed sufficiently large J,

$$|\widehat{W}(h,\nu)| \ll_{J,w} E_1E_2\left(1+\frac{|h+\tau\nu|}{T_1}\right)^{-J}\left(1+\frac{|\nu|}{T_2}\right)^{-J}, \qquad T_i = q/E_i. \tag{4.2}$$

Here the complete periodized coefficient is the sum over integer lifts of its residue. In estimating nonzero residues, those lifts avoid $q\mathbb{Z}^2$; enlarging to all integer pairs other than $(0,0)$ gives an upper bound.

For each $d \mid q$ the weighted lattice sum satisfies

$$\sum_{\substack{(h,\nu)\in\mathbb{Z}^2\setminus\{(0,0)\}\\ d|h,d|\nu}}\left(1+\frac{|h+\tau\nu|}{T_1}\right)^{-J}\left(1+\frac{|\nu|}{T_2}\right)^{-J} \ll_J \frac{T_1T_2}{d^2}+\frac{T_1+T_2}{d}. \tag{4.3}$$

To see this uniformly in τ, first take $\nu=0$ and sum over nonzero multiples of d. This costs $O(T_1/d)$. For $\nu \neq 0$, the h-sum is $O(1+T_1/d)$ uniformly in its center $-\tau\nu$; the remaining nonzero ν-sum is $O(T_2/d)$.

Use $(q,h,\nu)^{1/2} \le \sum_{d|(q,h,\nu)} d^{1/2}$, then (4.3). With the factor $q^{-2}E_1E_2$, the nonzero modes contribute at most

$$q^{3/2+\eta}+q^{1/2+\eta}(E_1+E_2)$$

times $\|r\|_2^2$, after a harmless adjustment of η. Finally

$$\sum_{e,j} W(e,j) \ll_w (1+E_1)(1+E_2)$$

by summing first in j, uniformly in its center. The terms $1+E_1+E_2$ are absorbed by the displayed nonzero-mode bound. This proves (4.1), including subunit E_i.

For an integer affine change of rows, complete Fourier frequencies are carried by the transpose of its linear part. When its determinant is a unit modulo q, it preserves (q,h,ν) and permutes the complete row classes. Translation only introduces unit-modulus factors. The same proof applies in the new integer coordinates. □

The exact zero-mode statement can be useful separately: before replacing the lattice mass by an area, its contribution is bounded with constant one by $\sum W$. In the application below the divided row lengths are polynomially large, so ordinary area estimates also apply. Retaining the lattice mass also covers intervals shorter than one in intermediate divisor sectors.

5 The weighted Farey input

The local estimate controls an operator on all residues $b \pmod q$. The input coming from the sieve occupies these residues through fractions with short numerator and denominator ranges. The next lemma requires no cancellation in their weights.

Lemma 5.1 (Weighted fraction energy)**.** *Let $\mathcal{A}$ be a set of P integers of one sign, with $|u| \asymp U$ and $(u,q)=1$. For each $u \in \mathcal{A}$, let I_u be an interval of length $O(V)$, where $V \ge 1$; its center may depend arbitrarily on u. For arbitrary $|\beta_{u,k}| \le 1$ and complex c_u, define*

$$r(b) = \sum_{u\in\mathcal{A}} c_u \sum_{k\in I_u\cap\mathbb{Z}} \beta_{u,k}\mathbf{1}_{b=(k+B)/u \pmod q}.$$

If the integer heights are bounded by a fixed power of x, then

$$\|r\|_2^2 \ll_\eta x^\eta (1 + UV/q)(V + P) \sum_{u \in \mathcal{A}} |c_u|^2. \tag{5.1}$$

There is no coprimality condition on $k + B$.

Proof. For fixed u, u', equality of the two residue fractions means

$$u'k - uk' \equiv B(u - u') \pmod{q}.$$

As (k, k') ranges in $I_u \times I_{u'}$, its determinant $u'k - uk'$ lies in an interval of length $O(UV)$. There are $O(1 + UV/q)$ possible values in the prescribed residue class, independently of the locations of the intervals. For any fixed determinant value, integer solutions form a one-dimensional progression with step $(u/(u, u'), u'/(u, u'))$. The number in the two intervals is at most

$$O\left(1 + \frac{V(u, u')}{U}\right).$$

Taking absolute values of the bounded weights produces a symmetric nonnegative majorant for the input Gram matrix. Its row sum is at most

$$(1 + UV/q) \left(P + \frac{V}{U} \sum_{|u'| \asymp U} (u, u')\right).$$

The elementary divisor estimate

$$\sum_{|u'| \asymp U} (u, u') \le \sum_{d | u} d\, O(U/d) \ll U \tau(|u|)$$

and Schur's test prove the result, since $\tau(|u|) \ll_\eta x^\eta$ in the stated height range. □

Two features will be used. The endpoint contribution is P, the number of denominators actually represented, rather than the length U of their ambient interval. Also, the numerator intervals may move with u. The latter permits angular localization of the source row variable without lengthening every numerator interval to the full denominator range. The divisor loss in (5.1) depends on the integer heights, which are bounded in terms of x; this explains the factor x^η even for small residual q.

6 The dispersion form and its coefficients

We now connect the preceding operator lemmas to the bilinear convolution in Theorem 2.1. The starting reduction is the one in [9, Definitions 3–6 and Lemmas 3–6]. For dense moduli we use the factorization conditions in [10, Lemma 6 and Section 3.3.6], in the explicit form of [11, Lemma 5.6 and Section 6.4]. We stop immediately before the entrywise absolute-value estimate in the proof of Stadlmann's Lemma 6. The positive expression at that point is the object estimated here.

For ease of comparison, most source symbols are retained. There are four deliberate changes: the original progression modulus is denoted q_{orig} when needed; the completed exponential modulus called m in the source is $\mathfrak{m}$ here; the row residue called γ there is j here; and the divisor called c in the mask expansion below is f. Thus γ continues to mean $\log_x N$, and c remains available for the Harman cutoff. The residue on a q_0-line is denoted ζ_0 to distinguish it from the later Harman cutoff ζ.

6.1 Source ranges and the smaller divisor target

Let Q_s, R_s be the lengths of the two initial modulus factors. After the preliminary reductions in [9, Definition 3, equations (3.2)–(3.4), and Lemma 3], we may assume

$$x^{-4\varepsilon-\delta}N \ll R_s \ll x^{-2\varepsilon}N, \qquad x^{1/2-\varepsilon} \ll Q_sR_s \ll x^{1/2+2\omega+\varepsilon}. \tag{6.1}$$

The omitted smaller moduli are covered by the bilinear Bombieri–Vinogradov theorem [7, Theorem 2.9]. Let q_0 be the common divisor arising in the dispersion step and set

$$H = \frac{x^{\varepsilon}R_sQ_s^2}{q_0M}. \tag{6.2}$$

The nontrivial part has $H \geq 1$. A signed dyadic block of Fourier indices has length $H_* \leq H$.

The source extracts u_1, v_1, v_2, q_2 at lengths $U_s, V_s, V_s, Q_s/q_0$, respectively. We distinguish the published smooth source, its enlarged formal source interface, and the dense-modulus source. The published source in [9, Definition 3, equations (3.5)–(3.6)] has the stronger upper bound $U_s \ll q_0^{-1}x^{-5\varepsilon}Q_s/H$ and lower bound $V_s \gg x^{5\varepsilon}H$. We prove the estimate also for the larger source intervals used in the public formalization [13]: the declaration `sourceSigmaTwo_uniform_secondary_reduction` records the interval interface below. Only that interval interface is being compared here; the improved parameter range is proved by the estimates of this paper.

$$\begin{gathered} q_0^{-1}x^{-\delta-5\varepsilon}\frac{Q_s}{H} \ll U_s \ll x^{-5\varepsilon}\frac{Q_s}{H}, \\ q_0^{-1}x^{5\varepsilon}H \ll V_s \ll x^{\delta+5\varepsilon}H. \end{gathered} \tag{6.3}$$

In the explicit dense-modulus extraction of [11, Section 6.4], one first splits $q_0q_1 = uv$ and then removes the factors of q_0 from u, v. Its ranges are

$$\begin{gathered} q_0^{-2}x^{-\delta-5\varepsilon}\frac{Q_s}{H} \ll U_s \ll q_0^{-1}x^{-5\varepsilon}\frac{Q_s}{H}, \\ x^{5\varepsilon}H \ll V_s \ll q_0x^{\delta+5\varepsilon}H. \end{gathered} \tag{6.4}$$

The relation $U_sV_s \asymp Q_s/q_0$ holds in both cases. Neither displayed pair of intervals contains the other. For every estimate after extraction we use the common envelope

$$\frac{x^{5\varepsilon}H}{q_0} \ll V_s \ll q_0x^{\delta+5\varepsilon}H, \qquad U_sV_s \asymp Q_s/q_0. \tag{6.5}$$

Both powers of q_0 in this envelope matter. The weaker lower bound is needed for the enlarged public smooth-source interface; the larger upper bound is needed for the dense extension. The stronger q_0-free lower bound in (6.4) cannot be used when proving the common statement.

Choose the common divisor-extraction target

$$D = \frac{N}{q_0^2x^{50\varepsilon}H^2}, \qquad D/x^{\delta} \ll \Delta \ll D, \qquad \Delta_1 = x^{-5\varepsilon}\Delta. \tag{6.6}$$

We extract a divisor $d \asymp \Delta$ of the factor of size R_s, write that factor as dr_1, and subdivide the d-range into windows of width Δ_1 centered at $d_0 \asymp \Delta$. The target (6.6) is smaller by q_0^2 than the target in [9, Definition 5]; it is the target used for the dense extraction in [11, Section 6.4, p. 23]. Here we use (6.6) for both source classes.

Here is its legality in the smooth case. From (6.1), (6.2), and $MN \asymp x$,

$$q_0 H \gg (N/R_s)x^{-\varepsilon}.$$

It follows that

$$\frac{D}{R_s} \ll \frac{R_s}{N} x^{-48\varepsilon} \ll x^{-50\varepsilon}. \tag{6.7}$$

The lower-scale inequalities in (2.3) give $D > 1$ for sufficiently small ε. A smooth integer has a divisor in the required interval [7, Definition 2.1 and Lemma 2.10(iii)]. For the dense case the same interval is one of the explicit witnesses verified in Section 9. Notice that $q_0^2 H^2$ in (6.6) is independent of q_0 after substitution of (6.2).

Put

$$\mathfrak{m} = r_1 q_0 u_1 [v_1, v_2] q_2, \qquad v_0 = (v_1, v_2), \qquad v_i^* = v_i / v_0. \tag{6.8}$$

This $\mathfrak{m}$ is squarefree. The CRT reduction in [9, Lemma 5] gives $A, B \pmod{\mathfrak{m}}$, independent of d, n and the Fourier indices, such that

$$(A, \mathfrak{m}) = 1, \qquad B \equiv 0 \pmod{r_1 q_0 u_1 [v_1, v_2]}, \qquad B \equiv \ell r_1 \pmod{q_2}. \tag{6.9}$$

Here ℓ is the integer shift from the dispersion argument. Choose a fixed integer representative of B throughout the changes of variables below.

6.2 Separating the divisor-dependent Fourier weights

The Fourier coefficients in [9, Definition 4] are

$$\varphi_i(h; d) = \frac{1}{M} \sum_t \psi_M(t) e\Big(-\frac{th}{d r_1 q_0 u_1 v_i q_2}\Big) \qquad (i = 1, 2). \tag{6.10}$$

The profiles ψ_M, ψ_N here are nonnegative smooth majorants introduced by the preceding Cauchy reductions. They are not the original coefficient sequences α, β; their positivity adds no hypothesis to Theorem 2.1. The product of the two Fourier coefficients depends on d. An operator estimate for fixed c_λ is applicable only after separating that dependence, before taking the square.

Put $R_i = r_1 q_0 u_1 v_i q_2$ and $F_{\boldsymbol{h}}(d) = \varphi_1(h_1; d)\overline{\varphi_2(h_2; d)}$. Fix a support bound $t/M \le T_M$ for ψ_M and a bound L_M for its absolute value. If $|h_i| \ll H$, the useful dimensionless Taylor parameter is

$$S_T = \frac{\Delta_1}{d_0}\left(1 + \frac{T_M M H}{d_0}(R_1^{-1} + R_2^{-1})\right) \ll x^{-4\varepsilon}. \tag{6.11}$$

Indeed, $d_0 r_1 \asymp R_s$, $q_0 u_1 v_i \asymp Q_s$, and $q_2 \asymp Q_s/q_0$. Hence $MH/(d_0 R_i) \ll x^{\varepsilon}$, while $\Delta_1/d_0 \ll x^{-5\varepsilon}$. This argument uses the product relation in (6.5) and works for both extractions.

For a fixed integer $J \ge 1$, Taylor's theorem gives

$$\begin{aligned} F_{\boldsymbol{h}}(d) &= \sum_{j=0}^{J-1} \left(\frac{d - d_0}{\Delta_1}\right)^j a_j(\boldsymbol{h}) + r_J(d, \boldsymbol{h}), \\ a_j(\boldsymbol{h}) &= \frac{\Delta_1^j}{j!} F_{\boldsymbol{h}}^{(j)}(d_0). \end{aligned} \tag{6.12}$$

On $|d - d_0| \le C_D \Delta_1$, differentiation of the finite Fourier sums shows

$$|a_j(\boldsymbol{h})| \le C_J (T_M L_M)^2 S_T^j, \qquad |r_J(d, \boldsymbol{h})| \le E_J := C_J (T_M L_M)^2 (C_D S_T)^J.$$

All constants and profile seminorms are fixed before the arithmetic source parameters vary. In particular, a_j is independent of d, n and the row residue, and has a uniform logarithmic bound when $S_T \le 1$.

Here is the positivity step explicitly. For nonnegative row weights ρ_r, any finite features $A_{r,h}$, and an expansion $F_h(d_r) = \sum_{j<J} u_{r,j} a_j(h) + r_{r,h}$ with $|r_{r,h}| \le E_J$, one has

$$\begin{aligned} \sum_r \rho_r \left| \sum_h A_{r,h} F_h(d_r) \right|^2 &\le 2J \sum_{j<J} \sum_r \rho_r |u_{r,j}|^2 \left| \sum_h A_{r,h} a_j(h) \right|^2 \\ &\quad + 2E_J^2 \sum_r \rho_r \left(\sum_h |A_{r,h}| \right)^2 . \end{aligned} \tag{6.13}$$

This is just the triangle inequality followed by finite Cauchy–Schwarz. It leaves a positive Gram form for each fixed coefficient vector. No assertion about positivity of a matrix of four-factor Taylor coefficients is required.

For the source rows, let $\mathcal{F}$ be the original Fourier pair set, $\mathcal{I}$ the integer support of ψ_N, and $D_{\max}$ an upper bound for d. If $|\psi_N| \le L_N$ and the divisor-window weight is at most one, the last line of (6.13) is at most

$$2E_J^2 \big(|\mathcal{F}|\, |\mathcal{I}| L_N\big)^2 (D_{\max}/w_1)^2. \tag{6.14}$$

The final factor bounds the number of pairs (e, j) with $1 \le e \le D_{\max}/w_1$ and $0 \le j < e$. All three finite cardinality factors are at most x^{100} for sufficiently large x, uniformly in the source: the original supports and all extracted factors have fixed polynomial height. Once ε is chosen, choose J with $6\varepsilon J > 1000 + 48\varepsilon$. The fixed constants and logarithmic profile bounds can then be absorbed after increasing a single threshold for x, and (6.14) is at most $x^{-48\varepsilon}$. For example, bounding each remaining fixed or logarithmic factor by x and $C_D S_T$ by $x^{-3\varepsilon}$ gives the larger upper bound $x^{613-6\varepsilon J}$. The threshold is chosen after J and the profile bounds, but before the arithmetic source data. This gives the needed uniformity rather than a pointwise Taylor approximation.

For a fixed Taylor term, aggregate by

$$y = h_1 v_2^* - h_2 v_1^*.$$

In a source sector write

$$d = w_1 e, \qquad y = w_1 \lambda, \qquad \lambda = w_2 a_0, \qquad \Lambda = |Y|/w_1. \tag{6.15}$$

Here λ lies in a signed dyadic interval of size Λ, w_1 is squarefree and coprime to $\mathfrak{m}$, and w_2 is the full $\mathfrak{m}$-supported part of λ. Thus $(a_0, \mathfrak{m}) = 1$; w_2 may contain prime powers. The vectors are

$$c_\lambda = \sum_{h_1 v_2^* - h_2 v_1^* = w_1 \lambda} a_j(h_1, h_2).$$

For a fixed represented y, the pairs (h_1, h_2) form an arithmetic progression of step (v_1^*, v_2^*). If $|h_i| \le H_b \ll H$, its cardinality is at most

$$1 + \frac{2v_0 H_b}{\max(v_1, v_2)} \ll q_0 v_0.$$

The last inequality uses the weaker lower bound in (6.5). Dividing y by w_1 on the exact divisibility block does not change a fiber. If $|a_j| \le B_J$, Cauchy within each fiber, followed by summing the original Fourier pairs only once, gives

$$\begin{gathered} \|c\|_\infty \ll q_0 v_0 B_J, \qquad \|c\|_2^2 \ll q_0 v_0 B_J^2 H_*^2, \\ P := \#\{\lambda : c_\lambda \ne 0\} \ll H_*^2 \le H^2. \end{gathered} \tag{6.16}$$

Thus the pointwise square costs q_0^2, whereas the total energy needs only q_0. We shall use the convenient common upper budget q_0^2 for both. This deliberate weakening still suffices. The total-energy bound must nevertheless use the number of Fourier pairs, rather than the ambient length of the Y-interval; the latter replacement would spoil the mean.

6.3 The positive form with its full masks

Let b_1, b_2 be the primitive source residues. Define

$$C_d(n) = \mathbf{1}_{b_1/n = b_2/(n+\ell d r_1) \pmod{q_0}}, \tag{6.17}$$

$$\chi_e(n) = C_{w_1 e}(n)\mathbf{1}_{(n, w_1 r_1 q_0 u_1 v_1 v_2)=1}\mathbf{1}_{(n+\ell w_1 e r_1, q_0 q_2)=1}. \tag{6.18}$$

The inverses in (6.17) are restricted to units. By (6.9), all remaining prime-to-$\mathfrak{m}$ conditions are exactly the unit-pole conditions for the rational phase. The nonnegative row weight is $\mu_e = \psi_{\Delta_1}(w_1 e - d_0)$. The form to be bounded is

$$\mathcal{U} = \sum_{\substack{e \\ (e, \mathfrak{m} w_1)=1}} \mu_e \sum_{0 \le j < e} \left| \sum_{\substack{\lambda \\ (\lambda, e)=1}} c_\lambda \sum_{n \equiv j\lambda \pmod{e}} \psi_N(n)\chi_e(n) e_{\mathfrak{m}}\left(\frac{A\lambda}{e(n + B w_1 e)}\right) \right|^2. \tag{6.19}$$

The sum includes every $j \pmod{e}$, including zero and nonunit classes. This is a positive enlargement already made in the source before its last Cauchy inequality. No squarefreeness condition on e is needed in this enlarged form.

To specify precisely how the source is being used, its estimate before the entrywise step is

$$\Sigma_2 \ll_\varepsilon x^{5\varepsilon/2} \Delta \sup \sqrt{\mathcal{U}_{\mathrm{raw}}}. \tag{6.20}$$

The supremum includes the fixed frequency and divisor sectors just described. If $\kappa_0 = (q_0, \ell)$, the required bound for that stage is $\Sigma_2 \ll q_0 \kappa_0 \Delta N v_0 x^{-10\varepsilon}$. It is therefore sufficient to prove

$$\mathcal{U}_{\mathrm{raw}} \ll x^{-40\varepsilon} q_0^2 \kappa_0^2 v_0^2 N^2. \tag{6.21}$$

Here $\mathcal{U}_{\mathrm{raw}}$ retains the actual two-factor Fourier amplitude inside the square; $\mathcal{U}$ in (6.19) is one fixed-coefficient form after (6.13). The exponent 40 leaves a reserve: the displayed Cauchy step yields $x^{-35\varepsilon/2}$, stronger than the required $x^{-10\varepsilon}$. Equation (6.20) is the Cauchy step in [9, proof of Lemma 6, p. 16], with its Υ_3 denoted by $\mathcal{U}_{\mathrm{raw}}$ here. The sufficient target (6.21) includes extra slack beyond the Σ_2 bound required in [9, Lemma 4]. The preceding diagonal and coefficient reductions are those of Lemmas 3–4 there; small moduli use [7, Theorem 2.9]. Its later entrywise exponential-sum estimate is replaced by Sections 7–8 below.

7 Congruence masks and the exact incidence coordinates

Several transformations of (6.19) are needed before applying Lemma 4.1. They are recorded explicitly because changing their order can remove a condition inside a signed sum without justification.

7.1 The outside-modulus condition

Expand the factor $(n, w_1) = 1$ as

$$\mathbf{1}_{(n, w_1)=1} = \sum_{\substack{f \mid w_1 \\ f \mid n}} \mu(f).$$

Apply Cauchy to this finite divisor sum at each row. The cost is a divisor function, hence $x^{o(1)}$. In the f-sector set $n = fn_1$. The original restriction $(e, w_1) = 1$ ensures $(f, e) = 1$, so all row classes $j \pmod e$ can be relabeled by multiplication by f^{-1}. The phase numerator is divided by f, and the fixed integer Bw_1 is replaced by

$$B_f = Bw_1/f. \tag{7.1}$$

All original restrictions $(e, \mathfrak{m}w_1) = 1$ are retained at this stage.

7.2 The common modulus and its residue lines

Fix a unit row class $e \equiv e_0 \pmod{q_0}$. The condition (6.17), after the preceding substitution, reads

$$(b_2 - b_1)fn_1 \equiv b_1\ell r_1 w_1 e \pmod{q_0}. \tag{7.2}$$

Its allowed solutions have the form $n_1 \equiv \zeta_0 e \pmod{q_0}$ for a set of at most κ_0 unit residues ζ_0. In fact the number of solutions of the unrestricted linear congruence, when nonempty, is $(b_2 - b_1, q_0)$, which must divide ℓ; the required unit conditions can only reduce the count.

For one such line write the integer identity

$$n_1 = \zeta_0 e + q_0 n_2, \qquad \widehat{B} = q_0^{-1}(B_f + \zeta_0) \pmod{\mathfrak{m}/q_0}. \tag{7.3}$$

The definition of $\widehat{B}$ is modular: $B_f + \zeta_0$ need not be divisible by q_0 as an integer. Relabeling j by $q_0^{-1} \pmod e$ gives the exact parametrization

$$n_2 = j\lambda + ek, \qquad 0 \le j < e, \qquad k \in \mathbb{Z}. \tag{7.4}$$

The part of the phase modulo q_0 is

$$e_{q_0}\left(\frac{A\lambda}{(\mathfrak{m}/q_0)f(B_f + \zeta_0)e_0^2}\right). \tag{7.5}$$

It is a fixed multiplier of c_λ of modulus one. It is independent of k and of the remaining row variables. Thus no partition of λ into residue classes is needed.

There are at most q_0 row classes. For each, Cauchy over the allowed ζ_0's costs at most κ_0, and there are at most κ_0 such terms. The resulting factor is $q_0\kappa_0^2$. This factor will be displayed through the normalization in Section 8.

7.3 Primes dividing the frequency numerator

Recall $\lambda = w_2 a_0$ and $(a_0, \mathfrak{m}) = 1$. Put

$$g = (w_2, \mathfrak{m}/q_0), \qquad q = \mathfrak{m}/(q_0 g). \tag{7.6}$$

At a prime $p \mid g$, the original row still has $e \neq 0 \pmod p$, whereas $\lambda = 0 \pmod p$. Equation (7.4) then gives

$$n_2 + \widehat{B}e = e(k + \widehat{B}) \pmod p.$$

The local exponential is 1 on its pole complement, and the remaining condition is exactly $\mathbf{1}_{(k+\widehat{B},g)=1}$. This is a bounded input weight, independent of both row variables. We move it into the numerator weight in Lemma 5.1 and remove g from the oscillatory modulus.

This step is necessary: the unit-numerator zero-mode estimate for p cannot be applied to a prime at which the numerator vanishes. Keeping its pole mask as an input weight avoids any extra factor g in the mean.

7.4 The frequency–row coprimality condition

On the *original* unit-e rows, $(w_2, e) = 1$, so $(\lambda, e) = 1$ is equivalent to $(a_0, e) = 1$. Use the exact expansion

$$\mathbf{1}_{(a_0,e)=1} = \sum_{t|(a_0,e)} \mu(t)$$

and apply Cauchy at the fixed row. Its cost is at most $\tau(e) = x^{o(1)}$. Every occurring t is coprime to $\mathfrak{m}w_1$. Write

$$e = tE, \qquad \lambda = tw_2u, \qquad E \equiv e_0t^{-1} \pmod{q_0}. \tag{7.7}$$

Only now, inside each resulting positive square, may the remaining row restrictions outside q be dropped. In particular, the condition $(e, g) = 1$ is dropped after the expansion, not before it. Otherwise the equivalence of $(\lambda, e) = 1$ and $(a_0, e) = 1$ would no longer hold on the enlarged row set.

The residual phase in this sector is

$$e_q\left(\frac{A/(fq_0^2gt)}{Ej + E^2(k + \widehat{B})/(w_2u)}\right). \tag{7.8}$$

Its numerator is a unit modulo q. The factor g^{-1} is the CRT scalar in the surviving local characters; it has no size cost. The remaining unit-pole condition is exactly the one in (3.1). The input residue is

$$b = (k + \widehat{B})/(w_2u) \pmod{q}.$$

Multiplication of all input residues by the unit w_2 preserves their squared norm, so the fraction-energy estimate uses the denominator u and its length, not the larger physical length of λ.

The row range remains $0 \le j < tE$. There is no change to $j \pmod{E}$. The progression in E in (7.7) is an integer affine row change of determinant q_0, a unit modulo q. It is therefore permitted in Lemma 4.1.

8 Angular localization and summation of the source errors

8.1 Moving numerator intervals and sheared rows

For a fixed sector, define

$$e_{\mathrm{sc}} = \Delta/w_1, \qquad V = \frac{Nw_1}{fq_0\Delta}, \qquad U = \frac{|Y|}{w_1w_2t}, \qquad J_0 \asymp \max(1, \Lambda/V). \tag{8.1}$$

The real ratio $z = j/e = j/(tE)$ lies in $[0, 1]$. Cover this interval by smooth, bounded-overlap angular cells of length $O(1/J_0)$. In a cell centered at τ, the support of ψ_N implies

$$k = -\tau\lambda - \zeta_0/q_0 + O(V). \tag{8.2}$$

Thus its center depends on λ, exactly as allowed in Lemma 5.1. A common numerator interval of length Λ is unnecessary.

The smooth factor can be expressed on fixed compact normalized ranges as a smooth function of

$$\frac{e}{e_{\mathrm{sc}}}, \qquad \frac{(z-\tau)\Lambda}{V}, \qquad \frac{\lambda}{\Lambda}, \qquad \frac{k + \tau\lambda + \zeta_0/q_0}{V}. \tag{8.3}$$

Indeed its argument is proportional to $e\{k + z\lambda + \zeta_0/q_0\}$, and the product of the middle two variables accounts for the mixed term. Fourier separation on these compact ranges gives absolutely summable combinations of a row factor and a bounded joint (u,k)-weight. The smooth seminorms cost only powers of $\log x$, or an arbitrarily small x^η after the finite subdivisions. The triangular row cutoff can be enlarged by nonnegative smooth cutoffs in these same angular cells, including the two end cells. Their overlap is bounded, so rowwise Cauchy has no factor J_0.

After the progression in E is parametrized, the two row lengths are

$$E_1 = \frac{\Delta_1}{w_1 t q_0}, \qquad E_2 \asymp \frac{\Delta}{w_1 J_0}. \tag{8.4}$$

The shear may be of size $q_0 t$, which is harmless by Lemma 4.1. The area contains $1/J_0$; this cancels the number of angular cells. The nonzero-mode term, which has no area factor, pays at most J_0.

With all the bounded separated factors included, put $c_t(u) = c_{tw_2u}$. The coefficient bounds give

$$\begin{gathered} P_t := \#\operatorname{supp}(c_t) \ll \min(H^2, U), \qquad \|c_t\|_2^2 \ll x^{o(1)} q_0^2 v_0^2 P_t, \\ \sum_t \|c_t\|_2^2 \ll x^{o(1)} q_0 v_0 H^2 \le x^{o(1)} q_0^2 v_0 H^2. \end{gathered} \tag{8.5}$$

For the last bound, each original index occurs at most $\tau(|a_0|)$ times in the t-sum. The first and last estimates are used for different terms; neither is discarded in favor of the other.

8.2 The source envelopes

We now keep the small powers of x explicit. Fix $C \ge 2$ large enough for all the dyadic size comparisons in (6.1)–(6.8). This constant does not depend on the arithmetic parameters. For sufficiently large x, we may assume $C \le x^\varepsilon$ and put

$$L = C^{10} x^{8\varepsilon}, \qquad Z = x^{54\varepsilon}, \qquad A_\Delta = x^{5\varepsilon}. \tag{8.6}$$

The following simultaneous envelope is valid for both source extractions:

$$\begin{gathered} \frac{x}{L} \le MN \le Lx, \qquad N = x^\gamma, \qquad 1 \le H \le \frac{Lx^{4\omega+\delta}}{q_0}, \\ \frac{N}{Lq_0^2 x^\delta H^2 Z} \le \Delta_1 \le \frac{N}{q_0^2 H^2 Z}, \qquad \Delta = A_\Delta \Delta_1, \qquad 1 \le A_\Delta \le Z, \\ |Y| \le \frac{Lq_0 x^\delta H^2}{v_0}, \qquad \frac{MH^2}{Lq_0 v_0 \Delta_1} \le \mathfrak{m} \le \frac{Lq_0 x^\delta MH^2}{v_0 \Delta_1}. \end{gathered} \tag{8.7}$$

Here $q_0, v_0, w_1 \ge 1$. The upper bound for Δ_1 uses $Cx^{-55\varepsilon} \le x^{-54\varepsilon}$; its lower bound retains the same extraction retreat within LZ. The estimate for H follows from $H = x^\varepsilon (Q_s R_s)^2/(q_0 M R_s)$ and (6.1). The bound for Y follows by counting the original Fourier ranges and using the upper bound in (6.5).

It is useful to derive the modulus bounds separately. The exact identity

$$\mathfrak{m} v_0 = r_1 q_0 u_1 v_1 v_2 q_2$$

and the six size comparisons give

$$\mathfrak{m} \asymp \frac{R_s Q_s^2 V_s}{q_0 \Delta v_0} = \frac{MHV_s}{x^\varepsilon \Delta v_0}. \tag{8.8}$$

Inserting the common lower bound $V_s \gg x^{5\varepsilon} H/q_0$ gives the lower bound in (8.7), including its q_0 in the denominator. Using the original scale Δ in (8.8) and only then substituting $\Delta = x^{5\varepsilon}\Delta_1$ is essential. The resulting bounds are

$$\frac{q_0 ZMH^4}{Lv_0 N} \le \mathfrak{m} \le \frac{L^2 q_0^3 x^{2\delta} ZMH^4}{v_0 N}. \tag{8.9}$$

No smoothness assumption on $\mathfrak{m}$ has entered these calculations.

For the geometric quantities of (8.1) we have

$$J_0 \ll Lx^{\delta}, \qquad P_t \ll H^2 \le V. \tag{8.10}$$

Fixed comparison constants can be included in the actual lengths V_f and J_f used below. The incidence density is

$$\begin{aligned} \frac{UV}{q} &= \frac{g}{w_2 t}\frac{|Y|N}{f\Delta\mathfrak{m}} \\ &\le \frac{L^2 q_0^2 x^{\delta} N}{A_\Delta M} \le q_0 L^4 x^{2\gamma+4\omega+2\delta-1}. \end{aligned} \tag{8.11}$$

For the last step use $N/M \le Lx^{2\gamma-1}$ and $q_0 \le Lx^{4\omega+\delta}$, a consequence of $H \ge 1$. The strict density condition makes the factor multiplying q_0 less than one, even after the fixed row comparison constants. Accordingly the valid uniform density allowance is

$$UV \le D_{\mathrm{dens}} q, \qquad D_{\mathrm{dens}} = q_0. \tag{8.12}$$

It need not be $UV \le q$. In particular q_0 is retained as an arithmetic factor, not absorbed into x^{η}. Moreover a nonempty t-sector has $w_1 t \ll |Y|/w_2$, so both lengths in (8.4) exceed

$$x^{\gamma-16\omega-7\delta-O(\varepsilon)}. \tag{8.13}$$

For example, the lower bound for E_1 follows by using $w_1 t \ll |Y|/w_2$ and then $(q_0 H)^4 \le L^4 x^{16\omega+4\delta}$; it is at least $x^{\gamma-16\omega-6\delta}/(L^6 Z)$ up to fixed constants. The lower bound for E_2 also uses $J_0 \ll Lx^{\delta}$. This explains the third condition of the theorem.

8.3 The four terms and the outside normalization

There are two levels of normalization: the output for an individual progression sector, and the original masked energy. We record both. This also shows exactly where $D_{\mathrm{dens}} = q_0$ is paid.

Let $P_{\mathrm{sm}} \ge 1$ bound the fixed and subpower losses, including the relevant divisor counts, finite harmonic sums, Fourier-separation costs, and fixed profile bounds. These losses will be quantified below, separately from L. Write

$$B_c = \sup_{\lambda} |c_\lambda|, \qquad E_c = \sum_{\lambda} |c_\lambda|^2, \qquad \mathcal{E}_c = \tau_c E_c, \qquad \mathcal{B}_c = B_c^2 \sum_t \frac{1}{t},$$

where τ_c bounds the multiplicity from the t-divisor sum. By (6.16)–(8.5) we may arrange

$$\mathcal{E}_c \le P_{\mathrm{sm}}^2 q_0^2 v_0 H^2, \qquad \mathcal{B}_c \le P_{\mathrm{sm}}^2 q_0^2 v_0^2. \tag{8.14}$$

For each $f \mid w_1$, take the actual numerator length V_f and angular count J_f, with fixed comparison constants included. Put $E_{\mathrm{base}} \asymp \Delta_1/(w_1 q_0)$ and $e_{\mathrm{sc}} \asymp \Delta/w_1$. The four contributions, after the t and angular sums, are bounded by $V_f \mathcal{T}_f$, where

$$\begin{aligned} \mathcal{T}_f = {} & (2E_{\mathrm{base}} e_{\mathrm{sc}} + q^{1/2+\eta} J_f E_{\mathrm{base}} + 2q^{1/2+\eta} e_{\mathrm{sc}})\mathcal{E}_c \\ & + 6J_f q^{3/2+\eta} \frac{\Lambda}{w_2} \mathcal{B}_c. \end{aligned} \tag{8.15}$$

The first term comes from the exact zero-mode inequality $D_q^{0,0} \le q^2\,\mathrm{Id}$. The next two come from the two row lengths in Lemma 4.1; the last comes from its $q^{3/2+\eta}$ term. For the first three terms use the summed coefficient energy. For the last, use $P_t \ll \Lambda/(w_2 t)$ and the pointwise square bound, which give the displayed harmonic sum. In particular, the zero mode does not pay the ambient frequency height or an angular-count factor. The $1/J_f$ in its row area cancels the number of cells.

Here is a precise way to retain the outside factors. Let τ_e, τ_a bound the row-divisor and divided-frequency divisor counts, and let C_{card} satisfy $P_t \le C_{\mathrm{card}} V_f$. Let K be the fixed window constant, L_{row} a bound for the nonnegative row weight, and $W = 1 + x^{2\eta}$ the subpower Fourier-separation factor. The mask expansions, Cauchy inequalities, and Lemma 5.1 give

$$\mathcal{U} \le \tau(w_1)\kappa_0\tau_e\big[q_0\kappa_0 W K L_{\mathrm{row}}\, 9D_{\mathrm{dens}}(C_{\mathrm{card}} + 8\tau_a)\big] \sum_{f \mid w_1} V_f \mathcal{T}_f. \tag{8.16}$$

To see the density factor, the fraction-energy estimate is proportional to $1 + UV_f/q$, and the denominator row count is bounded by $C_{\mathrm{card}} V_f$. Thus it is bounded by a constant times $D_{\mathrm{dens}}(C_{\mathrm{card}} + \tau_a)V_f$. The factor $q_0\kappa_0^2$ outside it is precisely the residue and compatibility-line factor derived in Section 7. The two additional divisor Cauchy steps contribute $\tau(w_1)\tau_e$.

With (8.12), the arithmetic part of the outside factor is $q_0^2\kappa_0^2$. It cancels against that part of the raw normalization $(q_0\kappa_0 v_0 N)^2$. Thus the inner expressions to bound are $V_f\mathcal{T}_f/(v_0^2N^2)$, not those expressions divided a second time by $q_0^2\kappa_0^2$. Neither q_0 nor κ_0 is required to be subpower.

8.4 The normalized monomials

Choose $L_Y, L_J \ge 1$ with

$$\Lambda \le \frac{L_Y q_0 x^\delta H^2}{w_1 v_0}, \qquad J_f \le L_J x^\delta.$$

We can take $L_Y = L$ and $L_J = 2CL$. All fixed comparisons in $V_f, E_{\mathrm{base}}, e_{\mathrm{sc}}$ are included in P_{sm}; in particular

$$V_f \le \frac{P_{\mathrm{sm}} N w_1}{q_0\Delta}, \quad E_{\mathrm{base}} \le \frac{P_{\mathrm{sm}}\Delta_1}{w_1 q_0}, \quad e_{\mathrm{sc}} \le \frac{P_{\mathrm{sm}}\Delta}{w_1}, \quad q^\eta \le P_{\mathrm{sm}}.$$

Using $q \le \mathfrak{m}/q_0$, (8.14), and $\Delta_1 \le \Delta$ in (8.15) gives

$$\frac{V_f\mathcal{T}_f}{v_0^2N^2} \le P_{\mathrm{sm}}^5(2B_0 + L_J B_1 + 2B_2) + 6P_{\mathrm{sm}}^4 L_Y L_J B_{\mathrm{osc}}, \tag{8.17}$$

where

$$B_0 = \frac{\Delta_1 H^2}{w_1 v_0 N} \le \frac{1}{q_0^2 w_1 v_0 Z}, \tag{8.18}$$

$$B_{\mathrm{osc}} = \frac{q_0^{1/2} x^{2\delta}\mathfrak{m}^{3/2} H^2}{v_0 N\Delta}, \tag{8.19}$$

$$B_2 = \frac{q_0^{1/2}\sqrt{\mathfrak{m}}H^2}{v_0 N}, \qquad B_1 = \frac{x^\delta\sqrt{\mathfrak{m}}H^2}{q_0^{1/2} v_0 N}. \tag{8.20}$$

For example, the area contribution is bounded by

$$\frac{P_{\mathrm{sm}} N w_1}{q_0\Delta}\,\frac{2P_{\mathrm{sm}}^2\Delta_1\Delta}{w_1^2 q_0}\,\frac{P_{\mathrm{sm}}^2 q_0^2 v_0 H^2}{v_0^2 N^2} = 2P_{\mathrm{sm}}^5 B_0.$$

This calculation displays the surviving q_0^2 coefficient budget and the absence of any $L_J x^\delta$ from the mean. For the last term the factors are instead V_f, J_f, $q^{3/2+\eta}$, Λ/w_2, and $\mathcal{B}_c$; they give $6P_{\rm sm}^4 L_Y L_J B_{\rm osc}$. Factors f^{-1}, w_2^{-1} and the additional powers of g^{-1} may be discarded since they are favorable.

Substituting (8.9) and then the upper bound for H gives

$$B_{\rm osc} \le \frac{L^{16} Z^{5/2} x^{3/2+40\omega+16\delta-5\gamma}}{q_0^3 v_0^{5/2}}, \tag{8.21}$$

$$B_2 \le \frac{L^6 Z^{1/2} x^{1/2+16\omega+5\delta-2\gamma}}{q_0^2 v_0^{3/2}}, \tag{8.22}$$

$$B_1 \le \frac{L^6 Z^{1/2} x^{1/2+16\omega+6\delta-2\gamma}}{q_0^3 v_0^{3/2}}. \tag{8.23}$$

For clarity, before inserting the bound for H, the first right-hand side comes from

$$B_{\rm osc} \le \frac{L^4 q_0^7 x^{6\delta} Z^{5/2} M^{3/2} H^{10}}{v_0^{5/2} N^{7/2}}.$$

The factors H^{10} and $M^{3/2} \le (Lx/N)^{3/2}$ then produce 40ω, 16δ, and -5γ. The exact intermediate power of L is $31/2$, rounded up to 16. The corresponding power for each row error is $11/2$, rounded up to 6.

The two source allowances have product q_0^3: q_0^2 for the coefficient budget and q_0 for the incidence density. Both have been incorporated in the outside normalization and (8.17). After all substitutions the remaining denominators are $q_0^2, q_0^3, q_0^2, q_0^3$ for the mean, oscillation, second row, and first row. They remain favorable; no additional condition on ω, δ, γ is needed.

8.5 A uniform saving and completion of the transfer

Here are explicit budgets that avoid an unspecified $O(\varepsilon)$ in the final saving. First choose $\varepsilon > 0$ so small that

$$\begin{aligned} 600\varepsilon &\le 5\gamma - \frac{3}{2} - 40\omega - 16\delta, \\ 250\varepsilon &\le 2\gamma - \frac{1}{2} - 16\omega - 6\delta, \\ 100\varepsilon &\le \min\left\{\gamma - 12\omega - 6\delta,\ \frac{1}{2} - 2\omega - \gamma, 1 - 2\gamma - 4\omega - 2\delta,\ \gamma - 16\omega - 7\delta\right\}. \end{aligned} \tag{8.24}$$

Also impose any smaller fixed retreat required in the initial source reductions. These choices are possible throughout (2.3); the second margin is positive because

$$2\gamma - \frac{1}{2} - 16\omega - 6\delta = \frac{2}{5}\left(5\gamma - \frac{3}{2} - 40\omega - 16\delta\right) + \frac{1}{10} + \frac{2}{5}\delta.$$

The coarse lower bound for the divided row lengths can cost nearly 200ε through powers of L and Z. This is also covered by the first budget: under $\gamma < 1/2 - 2\omega$,

$$\gamma - 16\omega - 7\delta - \frac{1}{2}\left(5\gamma - \frac{3}{2} - 40\omega - 16\delta\right) = \frac{3}{4} - \frac{3}{2}\gamma + 4\omega + \delta > 0.$$

Thus the row margin is in fact greater than 300ε; the separate 100ε line in (8.24) does not limit the available reserve to 100ε. For a closed interval of possible γ, use its lower endpoint in lower bounds and its upper endpoint in upper bounds, before choosing ε.

Next choose the finitely many divisor-loss exponents η sufficiently small, and take x large, to arrange

$$1 \le P_{\mathrm{sm}} \le x^{\varepsilon/100}, \qquad L \le x^{20\varepsilon}, \qquad L_Y, L_J \le x^{25\varepsilon}, \qquad Z = x^{54\varepsilon}. \tag{8.25}$$

All actual indices and moduli have height at most x^{100}, so these bounds hold uniformly for their divisor functions, finite harmonic sums, and q^η. Fixed constants and the logarithmic profile bounds are chosen before the threshold for x. The parameter P_{sm} does not include q_0, κ_0, or either of the coarser factors L_Y, L_J. The density inequality (8.12) follows from its 100ε margin and $L^4 \le x^{80\varepsilon}$, with room for the fixed geometry factors.

Substituting (8.25) into (8.18) and (8.21)–(8.23) gives the following upper exponents, in units of ε, for the four summands of (8.17) before the numerical multipliers $2, 1, 2, 6$:

Summand	Exponent at most
$P_{\mathrm{sm}}^5 B_0$	$5/100 - 54 = -53.95$
$P_{\mathrm{sm}}^5 L_J B_1$	$5/100 + 25 + 120 + 27 - 250 = -77.95$
$P_{\mathrm{sm}}^5 B_2$	$5/100 + 120 + 27 - 250 = -102.95$
$P_{\mathrm{sm}}^4 L_Y L_J B_{\mathrm{osc}}$	$4/100 + 50 + 320 + 135 - 600 = -94.96$

Every term is therefore at most $x^{-50\varepsilon}$. The numerical multipliers total 11, which is at most x^ε for sufficiently large x. Consequently

$$\frac{V_f \mathcal{T}_f}{v_0^2 N^2} \le x^{-49\varepsilon}. \tag{8.26}$$

Finally collect the outside factors of (8.16). Each of $\tau(w_1), \tau_e, W, K, L_{\mathrm{row}}, C_{\mathrm{card}}, \tau_a$ is bounded by P_{sm}. There are at most P_{sm} divisors $f \mid w_1$, and $C_{\mathrm{card}} + 8\tau_a \le 9P_{\mathrm{sm}}$. Thus

$$\mathcal{U} \le (q_0\kappa_0 v_0 N)^2 81 P_{\mathrm{sm}}^7 x^{-49\varepsilon}. \tag{8.27}$$

This is a bound for the full fixed-coefficient masked energy, not merely for a normalized progression sector.

Return to (6.13). The finite number of Taylor terms, their row factors bounded by C_D^{2j}, and the constant $162J^2$ can also be included in P_{sm} by enlarging the threshold for x. The main terms are then at most $(q_0\kappa_0 v_0 N)^2 x^{-48\varepsilon}$, since $P_{\mathrm{sm}}^8 x^{-49\varepsilon} \le x^{-48\varepsilon}$. The remainder (6.14) is at most $x^{-48\varepsilon}$; it is absorbed into the same scale because $q_0\kappa_0 v_0 N \ge 1$. Increasing x once more proves (6.21). This proves the required secondary source estimate through (6.20), with its literal outside normalization and a uniform positive saving.

For completeness, the preliminary low-range dispersion conditions do not add a hidden restriction to (2.3). Its last wall and upper bound imply $\omega < 1/50 < 1/16$ and $\delta < 1/16 < 1/8$; they also imply $\gamma > 1/4 + 4\omega + \delta$ and $\gamma > 8\omega + 2\delta$. These are the preliminary size conditions in the source reductions, with strict room for their fixed ε retreats. Apply those reductions on a small closed interval about the fixed γ in the theorem. The secondary estimate is uniform in every modulus factor, coefficient family and compatible residue. The preceding diagonal and small-modulus estimates, the remaining dyadic sum, and the passage from the dispersed expression to the convolution discrepancy are then those of [9, Lemmas 3–4, proof of Lemma 6, and Appendix A] and [7, Sections 5.1–5.3]. They use the Siegel–Walfisz hypothesis on β stated in Section 2. This proves Theorem 2.1 for its fixed scale N. Covering the entire band $x^a \le N \le x^{1/2}$ is a separate step; its additional handoff conditions are stated below.

Remark 8.1 (The upper range used in the transfer). The source condition $2\gamma + 8\omega + 4\delta < 1$ is not an additional hypothesis of Theorem 2.1. In [9, Lemma 17, equations (3.30) and (3.33)] it controls the complete-sum mean; the same role is explicit in [11, equation (6.4)]. Here that mean is bounded by (8.18). The extraction and the reduction to (6.19) use the preliminary range retained in (2.3).

9 The precise dense-divisor extension

The complete-mode, window, and fraction-energy lemmas require only squarefreeness of the final $\mathfrak{m}$. Smoothness of the original modulus is used earlier to choose three divisors. Keeping those requirements explicit gives a useful extension without assuming smoothness of $\mathfrak{m}$ or an additional level of dense divisibility after the extraction.

Definition 9.1. For fixed $x, \omega, \gamma, \delta, \varepsilon$, let $\mathcal{D}_{\mathrm{inc}}(x; \omega, \gamma, \delta; \varepsilon)$ consist of squarefree $q_{\mathrm{orig}} \ll x^{1/2+2\omega}$ for which there are divisors $r \mid q_{\mathrm{orig}}$, $u \mid q_{\mathrm{orig}}/r$, and $d_1 \mid r$ satisfying

$$\begin{gathered} x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon}, \\ \frac{x^{1-\gamma-6\varepsilon-\delta}}{q_{\mathrm{orig}}} < u < \frac{x^{1-\gamma-6\varepsilon}}{q_{\mathrm{orig}}}, \\ r^2 \frac{x^{2-\gamma-52\varepsilon-\delta}}{q_{\mathrm{orig}}^4} < d_1 < r^2 \frac{x^{2-\gamma-52\varepsilon}}{q_{\mathrm{orig}}^4}. \end{gathered} \tag{9.1}$$

As usual, fixed dyadic comparison constants may be absorbed by slightly smaller positive retreats.

These are the Type IIc factorization witnesses introduced in [10, Lemma 6], in the literal class $\mathcal{D}_{\mathrm{IIc}}$ of [11, Lemma 5.6 and Section 6.4]. Their three choices occur at different stages. The divisor r makes the initial qr factorization; u produces the u_1, v_i splitting; d_1 gives the short divisor window. Substitution of $q_{\mathrm{orig}} \asymp Q_s R_s$ and $r \asymp R_s$ into the last interval gives exactly (6.6). Removing the part of u belonging to q_0 gives (6.4), which is included in the common envelope (6.5).

Corollary 9.2. *The conclusion of Theorem 2.1 holds with the smooth-modulus sum replaced by a sum over $\mathcal{D}_{\mathrm{inc}}(x; \omega, \gamma, \delta; \varepsilon)$, for sufficiently small fixed ε, under the same coefficient hypotheses and coherent primitive residue convention. Small-modulus terms are included by the bilinear Bombieri–Vinogradov theorem.*

Proof. Use the factorization-preserving reduction of [11, Section 6.4], following [10, Section 3.3.6], through the form (6.19). The preliminary removal of small prime parts and the dyadic choices there introduce only the retreats already displayed in (6.1) and (6.4). The three witnesses supply every factorization used before that form. The replacement argument from (6.19) onward is valid for every resulting squarefree $\mathfrak{m}$; it asks for no further divisor. All its normalizations used the larger, factorization-preserving envelopes. The small-modulus range is covered by [7, Theorem 2.9]. This proves the corollary. $\square$

9.1 Strong triple dense divisibility supplies the witnesses

We recall the recursive convention, because weaker notions of "three divisors at convenient sizes" do not suffice here. Let $D^{(0)}(y)$ be all positive integers. For $i \geq 1$, an integer n belongs to $D^{(i)}(y)$ if, for every $j, k \geq 0$ with $j + k = i - 1$ and every $1 \leq R \leq yn$, it has a factorization $n = ab$ such that

$$R/y \leq b \leq R, \qquad a \in D^{(j)}(y), \qquad b \in D^{(k)}(y).$$

This is the multiple dense divisibility of [7, Definition 2.1], also called strong multiple dense divisibility in the fragment-sieve formulation. For $i = 1$, requiring targets only up to n is equivalent, since the remaining targets may use $b = n$.

Corollary 9.3. *Fix $0 < \delta' < \delta$ and parameters satisfying (2.3). The estimate of Theorem 2.1 holds over squarefree moduli $q_{\mathrm{orig}} \ll x^{1/2+2\omega}$ that belong to $D^{(3)}(x^{\delta'})$, uniformly in coherent primitive residues.*

Proof. It is enough to consider $q_{\mathrm{orig}} \geq x^{1/2-\varepsilon}$. Use the orientation $(j, k) = (1, 1)$ in the definition of $D^{(3)}$ to choose r near $x^{\gamma-3\varepsilon}$, so that both r and q_{orig}/r are singly $x^{\delta'}$-densely divisible. Thus

$$x^{\gamma-3\varepsilon-\delta'} \leq r \leq x^{\gamma-3\varepsilon}.$$

In the two singly dense factors choose divisors with upper targets

$$U_* = x^{1-\gamma-6\varepsilon}/q_{\mathrm{orig}}, \qquad D_* = r^2 x^{2-\gamma-52\varepsilon}/q_{\mathrm{orig}}^4. \tag{9.2}$$

The targets do not exceed their available factors, since

$$\frac{U_*}{q_{\mathrm{orig}}/r} \leq x^{-7\varepsilon}, \qquad \frac{D_*}{r} \leq x^{-51\varepsilon}$$

when $q_{\mathrm{orig}} \geq x^{1/2-\varepsilon}$. Their lower bounds have exponents at least

$$\frac{1}{2} - \gamma - 2\omega - 6\varepsilon, \qquad \gamma - 8\omega - 2\delta' - 58\varepsilon,$$

respectively. These are positive for sufficiently small ε by (2.3). Single dense divisibility therefore supplies u, d_1 in intervals of width $x^{\delta'}$.

The strict gap $\delta - \delta'$ allows the targets to be retreated slightly, so their intervals lie strictly inside (9.1). It also absorbs fixed comparison constants and the small prime part transferred in the source reduction. More explicitly, deleting or adjoining a factor $s = x^{o(1)}$ changes a dense-divisor width by at most s, by [7, Lemma 2.10(i)]; choose an intermediate exponent between δ' and δ before taking x large. Corollary 9.2 now applies. □

The corollary is the dense-modulus interface needed when this Type II estimate is used in a fragment sieve. It imposes no identity between fragment predicates arising from different distribution inputs: each predicate supplies a sufficient construction of the dense divisibility required by the arithmetic input. The numerical many-primes application below needs only the smooth special case.

10 The convolution inputs for Harman's sieve

We next explain how the bilinear estimate enters a prime minorant. The terminology in this section refers to the shape of the convolution used in the sieve. In particular a long *smooth* factor is called Type I here, whereas the lower part of the bilinear band in [7, Definition 2.6(i)] is also called a Type I estimate. We specify the factors each time to avoid relying on that terminology.

Let I, a, c be fixed cutoffs and put

$$L = 1 - a, \qquad \lambda = 1 - 2a, \qquad \zeta = 1 - I - a. \tag{10.1}$$

The combinatorial region we use is

$$\begin{aligned}&\frac{2}{5}<a<\frac{5}{12}, \qquad \frac{1}{10}<\lambda<\zeta<I<a<c<\frac{1}{2},\\ &2I+3a<2, \qquad 3I+7a<4, \qquad 2I+a>1.\end{aligned} \tag{10.2}$$

Every inequality is strict; this allows small fixed perturbations when smooth scale partitions are introduced.

10.1 Covering the full bilinear band

For the Harman application both factors in the direct bilinear steps satisfy Siegel–Walfisz. Thus it suffices to cover the shorter factor between x^a and $x^{1/2}$; symmetry then gives $[a, 1-a]$ for a chosen factor. Theorem 2.1 covers the lower interval through a handoff $h < 1/2$ if its decreasing lower errors are negative at a and its upper conditions hold at h.

The existing upper bilinear range is supplied by [7, Theorem 2.8(ii),(iv)]. A convenient sufficient handoff is

$$h > \frac{1}{4} + 14\omega + 4\delta, \qquad 68\omega + 14\delta < 1. \tag{10.3}$$

The first bound is the substitution $\sigma = 1/2 - h$ into $56\omega + 16\delta + 4\sigma < 1$. The published Type II result applies above $x^{1/2-2\omega-c_{\text{near}}}$ for a sufficiently small fixed $c_{\text{near}} > 0$; the Type I result covers up to that point. It is important that c_{near} can be chosen from the strict margin in (10.3). For example, the decisive near-range exponent in [7, equation (5.36), p. 53] is, at that endpoint and before the arbitrarily small retreats,

$$1 + 12\omega + \frac{7}{2}\delta - \frac{5}{2}\left(\frac{1}{2} - 2\omega - c_{\text{near}}\right) = -\frac{1 - 68\omega - 14\delta}{4} + \frac{5}{2}c_{\text{near}}.$$

It is negative if c_{near} is small enough. Fixing $c_{\text{near}} = \delta/2$ instead would require a stronger inequality; that is a different sufficient implementation of the source proof, not the published hypothesis used here. For the strict point in Section 12 we take $h = 9/20$. The additional conservative overlap conditions

$$64\omega + 16\delta < 1, \qquad 112\omega + 41\delta < 2 \tag{10.4}$$

hold there as well. We keep them in the numerical interface rather than infer a larger parameter region from a single boundary inequality.

Proposition 10.1 (Other two convolution inputs)**.** *The following are the inputs used together with this full bilinear band.*

(i) If $I > 1/4 + 7\omega + 2\delta$, a convolution $\alpha * \psi$ with smooth factor at scale $N \geq x^I$, $MN \asymp x$, has the required distribution, provided the complementary coefficient α has the Siegel–Walfisz property whenever the elementary long-variable estimate does not apply.

(ii) Set $\sigma = 1/2 - c$. Suppose $0 < \sigma < 1/6$ and

$$\sigma > \frac{1}{30} + \frac{56}{15}\omega + \frac{4}{15}\delta, \qquad 108\omega + 2\delta > 1. \tag{10.5}$$

A convolution $\alpha * \psi_1 * \psi_2 * \psi_3$ has the required distribution when α is divisor bounded, the three ψ_i are smooth, and

$$MN_1N_2N_3 \asymp x, \qquad x^{1-2c} \ll N_i \ll x^c, \qquad N_iN_j \gg x^{1-c} \quad (i \neq j). \tag{10.6}$$

This Type III input holds on squarefree singly x^δ-densely divisible moduli up to $x^{1/2+2\omega}$.

All statements retain the fixed divisor bounds, coherent primitive residues, and uniformity under smooth subdivisions.

Part (i) is the long-smooth-factor argument of [1, Lemma 3, equations (2.1)–(2.2), and proof of Lemma 5], with the available middle bilinear band substituted in its proof. In its short range, Lemma 3 there requires $4\sigma + 28\omega + 8\delta < 1$ for $N \geq x^{1/2-\sigma}$; this is exactly $\log_x N > 1/4 + 7\omega + 2\delta$. The middle range is covered by the bilinear band. In the final range the smooth variable exceeds the modulus by a fixed power, so direct summation in that variable applies. The complementary Siegel–Walfisz hypothesis in Baker–Irving's Lemma 5 is retained here.

Part (ii) is [15, Theorem 1.1, equations (1.2)–(1.3)]. Its predecessor is the Type III estimate [7, Theorem 2.8(v) and Definition 2.6(iii)], whose parameter condition is $\sigma > 1/18 + 28\omega/9 + 2\delta/9$; the improved condition (10.5) uses the companion theorem. The three individual lower bounds, three upper bounds, and three pair product bounds in (10.6) are part of its hypotheses. A bound on the product M alone does not replace them. The residual factor α need not be smooth.

11 A pointwise minorant and its five-prime losses

We say a fixed arithmetic family is *evaluable* if it satisfies the discrepancy estimate under discussion, with every requested logarithmic saving. This is shorthand only within the proof: it refers to a particular linear combination of coefficient convolutions, with the hypotheses above checked. Linearity and the triangle inequality permit addition and subtraction of such families.

Theorem 11.1 (Parametric Harman minorant)**.** *Assume* (10.2), $\theta + a < 1$, *and the three convolution inputs of Section 10, with small fixed slack in their scale ranges. On* $[x, 2x]$ *there is a fixed family* $\rho(n; x)$ *such that*

$$-24 \leq \rho(n; x) \leq \mathbf{1}_{\mathbb{P}}(n), \qquad \rho(n; x) \neq 0 \Longrightarrow P^-(n) > x^{\lambda}, \tag{11.1}$$

and ρ *has distribution exponent* θ *to squarefree smooth moduli. Its mean is*

$$\sum_{x \leq n \leq 2x} \rho(n; x) = (1 - \kappa(a) + o(1)) \frac{x}{\log x}, \tag{11.2}$$

where $\kappa(a)$ *is the sum of the two labeled prime-density integrals specified below, and*

$$0 \leq \kappa(a) \leq \overline{\kappa}(a) := \frac{9625}{1728} \frac{(a - 2/5)^4}{(1 - 2a)^4 (8a - 3)}. \tag{11.3}$$

The same construction is valid for any other fixed modulus class on which all three required convolution inputs and their stated coefficient uniformity are available.

We give the construction and the signs in detail. The starting identities are those of Baker–Irving [1, Section 3, especially Lemmas 7, 10, 11 and 13], as developed by Stadlmann [9, Sections 4.3–4.8] and in the parametric constructions [10, Proposition 2, Section 4.1] and [11, Proposition 7.2]. The last two propositions include the restrictions $I + 9a < 4$ and $17a < 7$. The region (10.2) extends beyond those restrictions, so the arguments below establish the construction in the required region and keep the complementary Siegel–Walfisz condition explicit in the smooth Type I steps.

11.1 Sifted sums and their actual coefficient families

Define

$$\psi(n,z)=\mathbf{1}_{P^-(n)\ge z},\qquad P^-(1)=+\infty.$$

We write $P(z)=\prod_{p<z}p$ for the product of the primes below z. The exact integer Buchstab identity, used in [2, proof of Lemma 14, equation (4.5)] and [9, Section 4.4], is

$$\psi(n,z_2)=\psi(n,z_1)-\sum_{\substack{z_1\le p<z_2\\ p\mid n}}\psi(n/p,p)\qquad (z_1<z_2). \tag{11.4}$$

We begin with $\mathbf{1}_{\mathbb{P}}(n)=\psi(n,\sqrt{3x})$ for $n\in[x,2x]$ and x large. The exact upper cutoff will be retained until its boundary terms are treated.

In the constructive sieve argument of [2, proof of Lemma 14, pp. 178–181], take its parameters to be

$$\alpha_{\mathrm{BW}}=a,\qquad \beta_{\mathrm{BW}}=\lambda,\qquad M_{\mathrm{BW}}=x^{1-I}.$$

Its two explicit prime-product groups must then have scales $R<x^a$ and $S<x^\zeta$. Write $p_i=x^{\alpha_i}$. The six sifted families follow the pattern in [1, Lemma 7] and [9, Lemma 19]; each line means the sum over the displayed primes multiplying a cofactor, with that cofactor weighted by $\psi(\cdot,x^\lambda)$.

Family	Restrictions on explicit prime exponents	R group	S group
1	No explicit primes	1	1
2	$\lambda\le\alpha_1<a$	p_1	1
3	$\lambda\le\alpha_2<\alpha_1<a$, $\alpha_1+\alpha_2<a$	p_1	p_2
4	$\lambda\le\alpha_2<\alpha_1<a$, $\alpha_1+\alpha_2>L$, $\alpha_2<\zeta$	p_1	p_2
5	$\lambda\le\alpha_3<\alpha_2<\alpha_1<a$, $\alpha_1+\alpha_2<a$, $\alpha_3<\zeta$	p_1p_2	p_3
6	$\lambda\le\alpha_3<\alpha_2<a$, $\alpha_3+\alpha_4<a$, $\alpha_2<\zeta$, $\alpha_4\ge\alpha_3$	p_3p_4	p_2

The required scale inequalities hold directly. For family 3, $\alpha_2<a/2<\zeta$, the last inequality being $2I+3a<2$. Family 6 is the reversed family used in the proof of [9, Lemma 22]. Its strict-order version is [1, Lemma 7(6)]; the equal-prime boundary is treated in Section 11.6.

For these six families we use the coefficient construction in the proof of Baker–Weingartner's lemma. Its stated Type I hypothesis [2, equation (4.1)] allows arbitrary bounded coefficients; here we instead verify the needed properties for the specific coefficients produced by the construction. For example, its actual Type I coefficients have the form

$$a_m=\sum_{\substack{rsd=m\\ d\mid P(x^\lambda)}}u_rv_s\mu(d), \tag{11.5}$$

where u_r,v_s count the prescribed rough-prime products [2, proof of Lemma 14, p. 179]. The Type II stopping step [2, pp. 180–181] groups a product of the form

$$m=rp_1\cdots p_{t+1},\qquad n=p_{t+2}sdk,\qquad d\mid P(p_{t+2}),$$

with the remaining order and product conditions separated as in that proof. These are fixed prime–Möbius convolution families with fixed divisor bounds.

Here is the Siegel–Walfisz check needed for their use. At any coefficient scale $X_0\ge x^{\varepsilon_0}$, with fixed $\varepsilon_0>0$, put $Y_0=\exp(\sqrt{\log x})$. The elementary Rankin bound, with the fixed divisor weights, removes the part supported on Y_0-smooth integers with an arbitrary logarithmic saving. Indeed, dominate the fixed divisor weight by τ_K for a fixed K, and use $s=1-1/\log Y_0$ in

$\sum_{n\le CX_0,\, P^+(n)\le Y_0}\tau_K(n) \le (CX_0)^s\prod_{p\le Y_0}(1-p^{-s})^{-K}$. The Euler product is $(\log x)^{O_K(1)}$, whereas $X_0^{s-1}\le \exp(-\varepsilon_0\sqrt{\log x})$. The part with a repeated prime above Y_0 has the same saving by summing $p^2\mid n$ and the fixed divisor bounds. In every remaining term either an unrestricted integer variable is at least Y_0, or a prime above Y_0 occurs to the first power. Sum that variable last. Once the others are fixed, the actual product and order restrictions give intervals with polylogarithmic total variation. In the stopping step, removing the largest prime from the squarefree stopping product preserves this interval property.

For an unrestricted integer use the elementary progression estimate. For a prime $p\ge Y_0$, a modulus at most $(\log x)^B$ is at most $(\log p)^{2B}$, so classical prime Siegel–Walfisz and partial summation give any requested logarithmic saving. The prime estimate here also follows by retaining an individual modulus in [8, Claim 2.2 and Theorem 2.3], with the prime scale as the large parameter. Sum the other variables using the harmonic divisor bound. Extra coprimality restrictions yield the divisor factor allowed in (2.2). The number of stopping indices and logarithmic subdivisions is polynomial in $\log x$. This proves the required assertion for the coefficient families just described; it is not a claim for arbitrarily thinned prime products.

If the complementary coefficient in a smooth Type I term has scale below x^{ε_0}, choose $\varepsilon_0<1-\theta$. Its smooth partner then exceeds every modulus by a fixed power and the elementary long-variable, or Type 0, estimate [7, Section 3, Type 0 case, pp. 21–22] applies. There is no need to assert Siegel–Walfisz for a point mass at scale one. These observations justify all six sifted families with precisely the coefficient inputs of Section 10.

11.2 The three remainders after Buchstab

Apply (11.4), partitioning pair sums at a, L and the next prime exponent at ζ, as in [9, Section 4.4] and [11, Proposition 7.2, equation (7.2)]. The six families above and the direct Type II terms give an evaluable function Θ_0, with

$$\mathbf{1}_{\mathbb{P}} = \Theta_0+\Gamma_3+\Gamma_5-\Gamma_4. \tag{11.6}$$

At the level of limiting exponents the remaining sums are

$$\Gamma_3(n) = \sum_{\substack{n=p_1p_2n_3\\ \lambda\le\alpha_2<\alpha_1<a\\ \alpha_1+\alpha_2>L,\alpha_2\ge\zeta}} \psi(n_3,p_2), \tag{11.7}$$

$$\Gamma_5(n) = \sum_{\substack{n=p_1p_2p_3p_4n_5\\ \lambda\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<a\\ \alpha_1+\alpha_2<a,\alpha_3<\zeta}} \psi(n_5,p_4), \tag{11.8}$$

$$\Gamma_4(n) = \sum_{\substack{n=p_1p_2p_3n_4\\ \lambda\le\alpha_3<\alpha_2<\alpha_1<a\\ \alpha_1+\alpha_2>L,\alpha_2<\zeta}} \psi(n_4,p_3). \tag{11.9}$$

The additional negative triple term has $\alpha_1+\alpha_2<a$ and $\alpha_3\ge\zeta$, and is empty because $2\zeta>a$. Equation (11.6) comes from literal integer identities. The small discrepancies caused by $\log_x n\ne 1$ are treated in Section 11.6; until then they may be included in Θ_0 with their specified Type II treatment.

11.3 The three-prime remainder and the color argument

In (11.7), the cofactor exceeds one since $p_1p_2<x^{2a}<x$. If it were composite, its prime factors would both be at least $p_2\ge x^{\zeta}$, giving total exponent above $L+2\zeta>1$. Thus it is prime. All three prime factors have exponents in $[\zeta,a]$, up to the endpoint retreats.

Apply Heath–Brown's identity [4, Lemma 1, equation (6)], with the fine smooth decomposition and coefficient properties of [7, Lemmas 3.3–3.4], in the three-prime form used in [9, proof of Lemma 20]. The resulting convolutions have at most forty factors β_i at scales x^{ν_i}. Every factor with exponent at least 1/10 is smooth; every subproduct of positive fixed polynomial scale has the Siegel–Walfisz property. There is also a partition into three colors recording which original prime each factor came from. Each color has total exponent at least ζ. The colors are bookkeeping for the factorization scales; the terms of the identity need not themselves be prime products.

The passage from prime indicators to the von Mangoldt identity keeps the factors $1/\log t$ on the corresponding smooth intervals. Their rescaled derivatives have the required logarithmic bounds. Prime-power corrections are negligible also in progression norm: for a color in $[x^{\zeta+o(1)}, x^{a+o(1)}]$, CRT summed over p^r with $r \geq 2$ and $q \leq x^\theta$ gives $x^{1-\zeta/2+o(1)} + x^{\theta+a/2+o(1)}$. The reciprocal sum of these prime powers is $x^{-\zeta/2+o(1)}$, and their number is $x^{a/2+o(1)}$; primes dividing q contribute nothing to a primitive progression or its coprime mean. This is a power saving under $\theta + a < 1$. Thus logarithmic weights and prime powers are accounted for before the grouping argument.

The following grouping argument develops the parameter geometry of [1, Lemmas 8–10] and [9, proof of Lemma 20] in the present region. Group the small factors so that every group except possibly one has exponent at least 1/20, and a group with exponent at least 1/10 is an original singleton. This can be done by successively combining factors smaller than 1/20 until their sum reaches 1/20; such a new group has exponent below 1/10. Order the grouped exponents as $\mu_1 \geq \cdots \geq \mu_j$.

If a smooth singleton has exponent at least I, Type I applies. Its complement contains the other two color groups and has positive polynomial scale; subsetwise Siegel–Walfisz supplies the required complementary coefficient property. If a subset sum belongs to $[a, L]$, Type II applies with Siegel–Walfisz on both sides. Assume that neither occurs. Then every $\mu_i < I$.

If $\mu_2 + \mu_3 > L$, then

$$\mu_3 > L - I = \zeta > \lambda > 1 - 2c, \qquad \mu_1, \mu_2, \mu_3 < I < c.$$

The first three groups are smooth singletons, and every pair has sum greater than $L > 1 - c$. Thus all the individual conditions (10.6) hold, and Type III applies.

Otherwise $\mu_2 + \mu_3 < a$. Since $\mu_2 + \cdots + \mu_j > 1 - I > a$, a first-crossing argument forces $\mu_4 > \lambda = L - a$: if every subsequent increment were at most λ, partial sums starting from $\mu_2 + \mu_3 < a$ would enter $[a, L]$ at their first crossing of a. Also $\mu_4 \leq \mu_3 < a/2$, and

$$\mu_2 + \mu_3 + \mu_5 + \cdots + \mu_j > 1 - I - a/2 > a.$$

The same argument with μ_4 omitted forces $\mu_5 > \lambda$. The final inequality is precisely $2I + 3a < 2$.

The first five groups are now original smooth singletons. Since $3\lambda > a$, the sum $T = \mu_3 + \mu_4 + \mu_5$ is greater than a, hence greater than L by the absence of Type II. Ordering gives

$$\mu_2 + T \geq \frac{4}{3}T > \frac{4}{3}L, \qquad \mu_1 + \sum_{i \geq 6} \mu_i < \frac{4a-1}{3} < \zeta. \tag{11.10}$$

The last inequality is equivalent to $3I + 7a < 4$. Even a color containing one of the five large singletons and *all* the remaining small groups would have exponent below ζ. Each of the three colors therefore needs at least two of the five singletons, which is impossible. Small groups may mix colors: their total mass was already included in (11.10). This proves that Γ_3 is evaluable.

11.4 The first five-prime error

This is the five-prime remainder treated in [1, Lemma 11] and [9, Lemma 21]. We record the inequalities for the current cutoffs. In (11.8), $\alpha_3 < a/2 < \zeta$ automatically. Its cofactor is greater than one: the four explicit primes have product below x^{2a}. It is prime, because a composite cofactor would give at least six prime factors, each at least x^{λ}, while $6\lambda > 1$. Now $\alpha_2 + \alpha_3 + \alpha_4 \geq 3\lambda > a$. The portion where this sum is at most L is Type II. The remaining positive error is the literal labeled sum

$$E_1(n;x) = \sum_{\substack{n=p_1p_2p_3p_4p_5 \\ \lambda<\alpha_4<\alpha_3<\alpha_2<\alpha_1<a \\ \alpha_1+\alpha_2<a, \quad \alpha_2+\alpha_3+\alpha_4>L \\ \alpha_5>\alpha_4}} 1. \tag{11.11}$$

Thus Γ_5 equals an evaluable family plus E_1, with endpoints interpreted as below.

11.5 Reversal and the sign of the second error

The reversal follows [1, Lemma 13] and [9, Lemma 22]; its signs and the enlarged parameter range are verified here. In (11.9), the cofactor is not one, because $a + 2\zeta < 1$, and cannot be composite, because $L + 3\lambda > 1$. Call it p_4. The ordering $\alpha_1 > \alpha_2$ follows from $\alpha_1 + \alpha_2 > L > 2\zeta$. Also $\alpha_1 \leq 1 - 3\lambda = 6a - 2 < L$. Extending its original upper limit a to L introduces only a Type II family. The extended four-prime sum is

$$\Upsilon_1(n) = \sum_{\substack{n=p_1p_2p_3p_4 \\ \lambda\leq\alpha_3<\alpha_2<a, \quad \alpha_3+\alpha_4<a \\ \alpha_2<\zeta, \quad \alpha_4\geq\alpha_3}} 1. \tag{11.12}$$

The pair-complement replacement is valid off a Type II boundary strip, retained in Section 11.6.

Reverse the role of p_1 and apply Buchstab to its position. The first sifted term is family 6 above. The subtracted term has a new smallest prime p_5 and a cofactor at least p_5; the latter is prime because $6\lambda > 1$. This gives

$$\Upsilon_1 = \text{an evaluable family} - \Upsilon_3, \tag{11.13}$$

where

$$\Upsilon_3(n) = \sum_{\substack{n=p_2p_3p_4p_5p_6 \\ \lambda\leq\alpha_3<\alpha_2<a, \quad \alpha_3+\alpha_4<a \\ \alpha_2<\zeta, \quad \alpha_4\geq\alpha_3, \quad \lambda\leq\alpha_5\leq\alpha_6}} 1. \tag{11.14}$$

We need an equality modulo evaluable families when replacing this remainder, not merely a pointwise upper bound. Consider five exponents at least λ, of total one, with no subset sum in $[a, L]$. Every pair is at most $1 - 3\lambda = 6a - 2 < L$, hence below a. Every triple is at least $3\lambda > a$, hence above L. Ordering the five exponents and applying the bottom-triple argument from (11.10) gives

$$\max_i \alpha_i < \frac{4a-1}{3} < \zeta. \tag{11.15}$$

Each exponent is also at most $b = 1 - 4\lambda = 8a - 3 < a$.

On this no-Type-II set, the conditions in (11.14) are equivalent, apart from equality loci, to

$$\begin{gathered} \lambda < \alpha_i < b \quad (2 \leq i \leq 6), \qquad \alpha_2, \alpha_4 > \alpha_3, \\ \alpha_2 + \alpha_4 < a, \qquad \alpha_3 + \alpha_4 + \alpha_5 > L, \qquad \alpha_6 > \alpha_5. \end{gathered} \tag{11.16}$$

In the forward direction, use the pair and triple conclusions and (11.15). In the reverse direction, $\alpha_3 + \alpha_4 < \alpha_2 + \alpha_4 < a$, and (11.15) restores the omitted $\alpha_2 < \zeta$. Let $E_2(n;x)$ be the nonnegative labeled count of $n = p_2p_3p_4p_5p_6$ satisfying (11.16).

The symmetric difference between Υ_3 and E_2 is supported on tuples having a Type II subset. Fix an ordering of subsets and assign each such tuple to its first eligible subset. On every assigned region the two groups are prime products with Siegel–Walfisz; the remaining order conditions are separated by the source's fine partitions. Both signed pieces of the symmetric difference are therefore evaluable. Hence

$$\Upsilon_3 = \text{an evaluable family} + E_2, \qquad \Gamma_4 = \text{an evaluable family} - E_2.$$

Together with (11.6), this proves

$$\mathbf{1}_{\mathbb{P}} = \text{an evaluable family} + E_1 + E_2. \tag{11.17}$$

In particular the required minorant is $\rho = \mathbf{1}_{\mathbb{P}} - E_1 - E_2$ with the positive signs of both errors fixed by the literal reversal.

11.6 Literal endpoints and prime squares

For $n \in [x, 2x]$, the exponent sum is $\log_x n = 1 + O(1/\log x)$ rather than exactly one. All primeness and grouping arguments with fixed strict slack remain valid for large x. A pair-complement comparison that changes across this perturbation lies in an arbitrarily small fixed enlargement of the Type II band. Such a strip is treated by that distribution estimate; it is not discarded only because its unweighted mean is small. The replacement of the initial $\sqrt{3x}$ cutoff by $x^{1/2}$ is handled in the same way.

The artificial individual cap $b = 8a - 3$ creates no unaccounted strip off Type II, since (11.15) holds with a vanishing endpoint error and

$$b - \frac{4a-1}{3} = \frac{20}{3}(a - 2/5) > 0.$$

Weak versus strict prime ordering differs only on a repeated prime. In the relevant tuples such a prime satisfies $x^\lambda \le p \le x^{a+o(1)}$. If $(p, q) = 1$, CRT gives at most $O(x/(p^2 q) + 1)$ integers in a given primitive progression also divisible by p^2. Thus, even summing over all integers $q \le Q = x^\theta$,

$$\sum_{q \le Q} \#\{n \in [x, 2x] : n \equiv a(q) \pmod{q}, p^2 \mid n \text{ for such a } p\} \ll x^{1-\lambda+o(1)} + Qx^{a+o(1)}. \tag{11.18}$$

When $p \mid q$, a primitive progression has no such integer, and its coprime mean has no such term either. Thus the same bound is valid for the dense classes as for smooth moduli. The principal averages obey the corresponding bound. The assumptions $\lambda > 0$ and $\theta + a < 1$ make (11.18) a power saving. Equality at $p = x^\lambda$ fixes at most one prime and costs $x^{1-\lambda+o(1)} + Q$.

These estimates justify the literal strict versions (11.11) and (11.16). Since both errors are nonnegative and supported on products whose prime factors exceed x^λ, the pointwise minorant and rough-support assertions follow. The uniform lower bound -24 follows from labeling: E_1 has at most four roles for its distinguished p_5, and E_2 has at most $\binom{5}{3} \cdot 2 = 20$ choices for the roles of p_2, p_3, p_4. The other roles are fixed by order. Repeated values do not increase these bounds. No positivity of ρ is required.

11.7 The exact mass and a rational upper bound

For a labeled five-prime region bounded away from the coordinate hyperplanes, use the prime-counting calculation in [1, proofs of Lemmas 11 and 13] and [9, proofs of Lemmas 21–22]: fix its first four primes and sum the last prime in its interval of length comparable to $x/(p_1p_2p_3p_4)$. The prime number theorem and partial summation then give

$$\frac{x}{\log x} \int \frac{d\alpha_1\, d\alpha_2\, d\alpha_3\, d\alpha_4}{\alpha_1\alpha_2\alpha_3\alpha_4(1 - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4)} + o(x/\log x).$$

The integration regions are the limiting regions of (11.11) and (11.16), with the latter relabeled to five coordinates. Their boundaries have measure zero and their exponents have a fixed positive lower bound, so the limiting operation is harmless for the mean. The labeling restrictions are already in these integrals; there is no further factor 5!.

Their sum defines $\kappa(a)$ in (11.2). Put $u = a - 2/5$ and write

$$\alpha_i = 1/5 + uz_i, \qquad \sum_{i=1}^{5} z_i = 0.$$

The two regions become fixed four-dimensional polytopes P_1, P_2, whose volumes are computed in Appendix A:

$$\mathrm{vol}(P_1) = \frac{125}{864}, \qquad \mathrm{vol}(P_2) = \frac{3125}{576}, \qquad \mathrm{vol}(P_1) + \mathrm{vol}(P_2) = \frac{9625}{1728}. \tag{11.19}$$

The coordinate measure is $d\alpha_1 \cdots d\alpha_4 = u^4 dz_1 \cdots dz_4$. There is no Euclidean hyperplane-area factor.

On the simplex $\alpha_i \geq \lambda$, $\sum \alpha_i = 1$,

$$\prod_{i=1}^{5} \alpha_i \geq \lambda^4(1 - 4\lambda). \tag{11.20}$$

One elementary proof replaces two entries $v, w \geq \lambda$ by $\lambda, v + w - \lambda$: their product decreases by $(v - \lambda)(w - \lambda)$. Repeating leaves four entries equal to λ. Multiplying the total volume (11.19) by u^4 and the reciprocal of (11.20) proves (11.3). This bound is a convenient upper bound for the true mass loss, not a replacement for $\kappa(a)$ in the mean identity. It completes the proof of Theorem 11.1.

12 A strict parameter point and the many-primes theorem

The two relevant factor cutoffs must satisfy $a < c$. Combining the principal incidence wall with the Type III wall suggests, before any mass loss,

$$\frac{3}{10} + 8\omega + \frac{16}{5}\delta < a < c < \frac{7}{15} - \frac{56}{15}\omega - \frac{4}{15}\delta.$$

Consequently their limiting common region has

$$\theta = \frac{1}{2} + 2\omega < \frac{93 - 104\delta}{176}. \tag{12.1}$$

This is only a guide to choosing strict parameters. The full source, handoff, and Harman conditions must still be checked, as must the mass loss of the minorant.

We use the rational choices

$$\omega = \frac{71}{5000}, \quad a = \frac{41361}{100000}, \quad I = \frac{34941}{100000}, \quad c = \frac{10341}{25000}, \quad \delta_0 = \frac{1}{10^8}. \tag{12.2}$$

Thus

$$\theta = \frac{1321}{2500} = 0.5284, \quad \lambda = \frac{8639}{50000} = 0.17278, \quad \zeta = \frac{11849}{50000} = 0.23698.$$

For every required distribution retreat choose any sufficiently small $0 < \delta \leq \delta_0$. The cutoffs a, I, c and the literal minorant are fixed independently of that choice.

The following table gives exact positive margins at $\delta = \delta_0$. It includes the two possible bottlenecks and the handoff $h = 9/20$.

Margin	Exact value
$I - 1/4 - 7\omega - 2\delta_0$	499/50000000
$c - a$	3/100000
$5a - 3/2 - 40\omega - 16\delta_0$	623/12500000
$8 - 5(3c/2 - 1/2) - 14\theta - 2\delta_0$	4999/50000000
$h - (1/4 + 14\omega + 4\delta_0)$	29999/25000000
$4 - 3I - 7a$	113/2000

All decreasing lower errors in Theorem 2.1 are checked at $\gamma = a$; all its upper conditions are checked at $\gamma = h$. In particular, $h < 1/2 - 2\omega$ and $2h + 4\omega + 2\delta_0 < 1$. Equations (10.3)–(10.4) hold, so the full interval $[a, 1-a]$ is covered. The elementary Harman inequalities and $\theta + a < 1$ are strict as well. Some additional exact margins are listed in Appendix B.

For clarity, the Type III parameters in the companion are

$$\sigma = \frac{1}{2} - c, \qquad \mu = \frac{1}{4} - \frac{3}{2}\sigma = \frac{6023}{50000}, \qquad s = \frac{4 - \mu - 4\theta + 2\delta_0}{6} = \frac{88297001}{300000000}.$$

Its two largest normalized error exponents are both $-4999/600000000$; the other three are negative. The substitution checks all the individual scale conditions used in the color argument, with the positive $c - a$ reserve.

The mass estimate (11.3) becomes

$$\kappa(a) \le \frac{9381895159877734375}{13513403415750760334592} < \frac{139}{200000} = 0.000695. \tag{12.3}$$

12.1 The precise minorant-to-many-primes transfer

We use the transfer in [9, Lemma 2], equivalently the argument of [1, Lemma 1 and Section 4]. Its hypotheses are a pointwise minorant, fixed rough support, distribution exponent ϑ to squarefree smooth moduli, and mean $(1 - \kappa + o(1))x/\log x$ with $\kappa < 1$. For every fixed $\varepsilon_{\mathrm{BI}} > 0$ it gives

$$H_m \ll_{\varepsilon_{\mathrm{BI}}} \exp\left(\frac{2(1 + \varepsilon_{\mathrm{BI}})}{\vartheta(1-\kappa)} m\right). \tag{12.4}$$

To recall the mechanism, the truncated Maynard variational quantity obeys $M_k^{[\alpha]} \ge \log k - O_\alpha(1)$ for each fixed positive truncation [8, Section 6, proof of Theorem 3.9(xi)]. That proof uses coordinate cutoff $T = \beta/\log k$, with fixed $\beta > 0$, so its test functions eventually fit inside every fixed positive cutoff α. The minorant changes the prime-detection condition to

$$M_k^{[\alpha]} > \frac{2m}{\vartheta(1-\kappa)}.$$

Thus k can be chosen exponentially in m with any slightly larger exponent [1, Section 4]. An admissible tuple of diameter $O(k \log k)$ is obtained by taking k primes larger than k, as in [14, proof of Theorem 1]: smaller primes cannot divide any entry, and a prime larger than k cannot have all its residue classes occupied. Its polynomial factor is absorbed by a further strict increase of the exponential constant. This explains why the denominator in (12.4) contains both the distribution level and the surviving mass. The threshold is the one proved in Section 4 of Baker–Irving, with the statement given in Stadlmann's Lemma 2.

The distribution hypothesis here has a quantifier worth making explicit. For every fixed retreat $\epsilon > 0$, there must be some positive smoothness exponent δ such that the same minorant obeys, for every allowed squarefree P and every primitive $a \pmod{P}$,

$$\sum_{\substack{q \mid P \\ q \le x^{\vartheta - \epsilon}}} |\mathcal{E}_q(\rho; a)| \ll_A x(\log x)^{-A}.$$

Our strict margins persist as δ decreases, and the literal error sets depend only on the fixed cutoffs. This supplies the quantifier in [1, Definition 1]. The fixed roughness choice $\xi = 1/10 < \lambda$ is valid.

Take the explicit retreats

$$\vartheta = \frac{5283999}{10^7}, \qquad \kappa' = \frac{139}{200000}, \qquad \varepsilon_{\mathrm{BI}} = \frac{1}{10^7}.$$

Replacing the true $\kappa(a)$ by its upper bound κ' only weakens the conclusion. The exponent in (12.4) is at most

$$\frac{2(1+\varepsilon_{\mathrm{BI}})}{\vartheta(1-\kappa')} = \frac{4000000400000}{1056065324139} = 3.7876448630\ldots < 3.7877. \tag{12.5}$$

The remaining strict margin absorbs the polynomial admissible-tuple prefactor. This proves Theorem 1.1.

A The two scaled polytopes and their volumes

This appendix makes the geometric part of (11.3) an explicit finite calculation. Use free coordinates z_1, z_2, z_3, z_4 and put $z_5 = -z_1 - z_2 - z_3 - z_4$. Both polytopes have $z_i \ge -2$ for all five indices. Their remaining defining inequalities are

$$\begin{aligned} P_1: \quad & z_1 \ge z_2 \ge z_3 \ge z_4, \\ & z_1 + z_2 + z_3 + 2z_4 \le 0, \quad z_1 + z_2 \le 1, \quad z_2 + z_3 + z_4 \ge -1; \\ P_2: \quad & z_1 \ge z_2, \quad z_3 \ge z_2, \\ & z_1 + z_2 + z_3 + 2z_4 \le 0, \quad z_1 + z_3 \le 1, \quad z_1 + z_2 + z_4 \ge -1. \end{aligned}$$

For P_1 the coordinates correspond to $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. For P_2 they correspond to $(\alpha_4, \alpha_3, \alpha_2, \alpha_5)$, with z_5 corresponding to α_6. This relabeling gives the literal triple inequality in (11.16).

For $2/5 < a < 5/12$, the original upper bound $\alpha_1 < a$ on P_1 is redundant; it can be checked at the vertices below. On P_2, every $z_i \le 8$ already follows from the lower bounds on the other four coordinates, so the individual cap $\alpha_i < 8a - 3$ is redundant apart from boundary sets. Thus these fixed polytopes give the stated volumes throughout the parameter range used above.

Solving four active defining equations and retaining the points satisfying every inequality gives these vertex lists:

Index	P_1	P_2
0	$(-1/3, -1/3, -1/3, -1/3)$	$(-2, -2, -2, 3)$
1	$(0, 0, 0, 0)$	$(-1/3, -1/3, 4/3, -1/3)$
2	$(1/2, 1/2, -3/4, -3/4)$	$(1/2, -2, 1/2, 1/2)$
3	$(1/2, 1/2, -1/3, -1/3)$	$(1/2, 1/2, 1/2, -2)$
4	$(1/2, 1/2, 1/2, -2)$	$(1/2, 1/2, 1/2, -3/4)$
5	$(1/2, 1/2, 1/2, -3/4)$	$(3, -2, -2, -2)$
6	$(4/3, -1/3, -1/3, -1/3)$	$(3, -2, -2, 1/2)$

Pulling triangulations consist of the simplices

$$\begin{array}{llll} P_1: & (0,1,3,5,6), & (0,2,3,5,6), & (0,2,4,5,6), \\ P_2: & (0,1,2,5,6), & (0,1,3,4,6), & (0,1,3,5,6). \end{array}$$

For vertices $v_0, \ldots, v_4$, the simplex volume is

$$\frac{1}{4!} \left| \det(v_1 - v_0, v_2 - v_0, v_3 - v_0, v_4 - v_0) \right|.$$

The sums of these three volumes in the two cases are 125/864 and 3125/576, respectively. More explicitly, the simplex volumes in the displayed orders are

$$P_1: \quad \frac{125}{5184}, \frac{625}{20736}, \frac{625}{6912}; \qquad P_2: \quad \frac{625}{288}, \frac{625}{576}, \frac{625}{288}.$$

All the entries, defining inequalities, and determinants are rational; no numerical integration is used for the upper bound in Theorem 11.1.

B Additional parameter checks and dependency map

At (12.2), several useful decimal margins are terminating rational numbers:

Constraint	Positive margin at δ_0
$a - 12\omega - 6\delta_0$	0.24320994
$a - 16\omega - 7\delta_0$	0.18640993
$1/2 - 2\omega - 9/20$	0.0216
$1 - 2(9/20) - 4\omega - 2\delta_0$	0.04319998
$1 - 68\omega - 14\delta_0$	0.03439986
$1 - 64\omega - 16\delta_0$	0.09119984
$2 - 112\omega - 41\delta_0$	0.40959959
$2 - 2I - 3a$	0.06035
$2I + a - 1$	0.11243
$1 - \theta - a$	0.05799

Each inequality improves when δ is decreased. The remaining ordering conditions follow directly from the displayed cutoffs and $2/5 < a < 5/12$.

The proof separates into the following mathematical inputs. The complete-mode estimate, its smooth-window completion, and the weighted fraction-energy bound are proved in Sections 3–5. The Type II transfer uses the source reduction only up to its positive Gram form; the new estimates of that form are proved in Sections 6–8. Section 9 supplies the additional dense-modulus scope. The only new analytic input proved outside this article is the Type III theorem [15, Theorem 1.1], with the exact hypotheses restated in (10.5)–(10.6). The Harman combinatorics, including coefficient hypotheses, signs and literal boundaries, are proved in Section 11; the standard many-primes transfer is spelled out in Section 12.

The rational parameter and volume calculations provide checkable arithmetic for these statements. They do not replace the analytic estimates, the Siegel–Walfisz arguments, or the pointwise Buchstab identities.

Scope of the formal verification

The public source is `romyar1/prime-gaps-lean` [17]. The verification and source links here refer to commit `9c1ab6648bcd`. The main incidence assembly is `formal/IncidenceRawBlockBound.lean`; its application to the fixed 182 trial is in `formal/SourceIncidenceSecondary182.lean` and `formal/SourceIncidenceBilinear182.lean`.

The Lean development verifies the common-source coefficient estimates, modulus envelope, positive Taylor expansion and remainder, four-term normalization, and raw-to-secondary Cauchy transfer for the literal source expressions. Its scope does not include every analytic conclusion of this article. The raw family permits all five strict low-range guards in (2.3); its finite-field input is the scalar rank-four Kloosterman bound used in (3.11), supplied as an explicit premise. The weaker lower V_s range in that source interface belongs to the public formalization [13]; it is not the literal lower range printed in Stadlmann's Definition 3.

The formal construction of a full bilinear band also records its high-range joining hypotheses. One implemented version requires, in addition, $\omega < 3/200$, $a \le 41361/100000$, $0 < 1/2 - a < 1/4$, the five low guards uniformly on $[a, h]$, and $h > 1/4 + 14\omega + 4\delta$. These imply its convenient high-range condition $68\omega + 20\delta < 1$. They hold at the strict point used here, and on the selected source rows of the fixed-gap application. These additional hypotheses are separate from the five low guards in general. The ordinary handoff in (10.3) instead uses the published $68\omega + 14\delta$ theorem with its freely chosen small near-range cutoff, as explained there.

The formal terminal prime-gap statement has the Type III local Fourier hypothesis and the specified physical integral inequalities as explicit premises. The geometric proof of the Fourier statement is given in [15, Proposition 3.1] as an ordinary analytic argument. The Harman and Baker–Irving argument yielding the all-m statement of Theorem 1.1 is proved here; this all-m conclusion is outside the formalized endpoint.

The public numerical release records a fresh external recomputation of all 262 comparisons for that fixed trial, with 14 implementation audits and four deliberate rejection tests; see `research/numerical_182/README.md`. These computations support the numerical premises of the prime-gap application. They do not change the hypotheses of the Lean theorem.

Reproduction

The repository pins Lean `4.34.0-rc2` and its Mathlib revision. With Elan, Git and Python 3.11 or later installed, obtain the repository revision cited in [17] and run from its root:

```
lake exe cache get
./verify.sh
```

This builds the formal development and checks source integrity, theorem types and axiom dependencies. Use `./verify.sh -fresh` to rebuild the project sources and vendored baseline. `docs/verification.md` gives the detailed build instructions; `docs/assumptions.md` gives the endpoint premises. The numerical release and its separate integrity and regeneration commands are linked from `research/numerical_182/README.md`.

References

[1] R. C. Baker and A. J. Irving, *Bounded intervals containing many primes*, Math. Z. **286** (2017), 821–841. arXiv:1505.01815.

[2] R. C. Baker and A. Weingartner, *A ternary Diophantine inequality over primes*, Acta Arith. **162** (2014), 159–196. Full text; see Lemma 14 and its proof.

[3] D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples I*, Ann. of Math. (2) **170** (2009), 819–862. doi:10.4007/annals.2009.170.819.

[4] D. R. Heath-Brown, *Prime numbers in short intervals and a generalized Vaughan identity*, Canad. J. Math. **34** (1982), 1365–1377. doi:10.4153/CJM-1982-095-9.

[5] N. M. Katz, *Gauss Sums, Kloosterman Sums, and Monodromy Groups*, Annals of Mathematics Studies, vol. 116, Princeton University Press, 1988. Author's copy.

[6] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), 383–413. Published article; arXiv:1311.4600.

[7] D. H. J. Polymath, *New equidistribution estimates of Zhang type*, Algebra Number Theory **8** (2014), 2067–2199. arXiv:1402.0811v3.

[8] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), Art. 12; erratum, **2** (2015), Art. 15. arXiv:1407.4897v4; published erratum.

[9] J. Stadlmann, *On primes in arithmetic progressions and bounded gaps between many primes*, Adv. Math. **468** (2025), 110190. doi:10.1016/j.aim.2025.110190; arXiv:2309.00425v3.

[10] J. Stadlmann, *Bounded gaps between primes*, preprint, 31 August 2026. arXiv:2608.31126v1.

[11] F. Charton, L. Hong, K. Lau, K. Ono, G. Remy, H. C. Siu, A. A. Swaminathan, J. Thorner and Y. Xie, *A new bound for small gaps between primes*, preprint, 3 September 2026. Full text.

[12] OpenAI, *Improved short gaps between primes*, preprint, 30 August 2026. Full text.

[13] OpenAI, *PrimeGaps186*, Lean source and numerical certification for prime gaps at most 186, 2026. Repository.

[14] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), 1121–1174. doi:10.4007/annals.2014.179.3.7.

[15] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Kloosterman correlations with finitely many exceptional Fourier frequencies*, accompanying article, 2026. Article.

[16] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *A fixed-profile proof for prime gaps of at most 182*, accompanying article, 2026. Article.

[17] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Prime gaps in Lean*, source code, numerical certificates and reproduction instructions, 2026. Public repository, commit `9c1ab6648bcd` (accessed 11 September 2026). External numerical release `external-numerics-20260909`.

Kloosterman correlations with finitely many exceptional Fourier frequencies

GPT-6 Astra (OpenAI), prompted by Romyar Sharifi

11 September 2026*

Abstract

We prove a uniform Fourier estimate for a four-cycle of rank-three Kloosterman correlations. When the two row indices and the two column indices are distinct modulo a prime, the exceptional Fourier frequencies form a set of bounded cardinality. The main geometric step excludes curve-supported Fourier constituents: radial tameness rules out nonlinear curves and offset lines, and the explicit local Fourier phases rule out lines through the origin. We give the trace normalization, the rank-six pure correction, the local ramification calculation, and the passage from weights to the stated exponential-sum bound.

We then prove the squarefree completion and averaged matrix estimates needed in the smooth Type III method. The subsequent reduction retains the central frequency, all nonunit completion frequencies, the three Fourier divisors, and the width of dense-divisor extraction. It yields the sufficient condition $\sigma > 1/30 + 56\omega/15 + 4\delta/15$, in the parameter region stated below. This is a convolution-distribution theorem; a numerical bound for prime gaps also requires the separate minorant and sieve arguments.

Contents

*Original version: 4 September 2026. This version expands citations and exposition; the stated results and numerical data are unchanged.

1 Introduction

Yitang Zhang's 2013 breakthrough, published in 2014, established the first finite upper bound for $H_1 = \liminf_{n\to\infty}(p_{n+1} - p_n)$, where p_n is the n-th prime. His proof combined a sieve argument with distribution beyond the Bombieri–Vinogradov range for restricted moduli [1, Theorems 1–2]. The Type III case, treated in [1, Sections 13–14], concerns three substantial smooth factors and uses cancellation in complete exponential sums over finite fields. Polymath8a developed this part of the argument systematically, including its formulation for densely divisible moduli and its reduction to correlations of rank-three Kloosterman sums [2, Definition 2.6(iii), Theorem 2.8(v), and Section 7].

Recent bounds illustrate how distribution estimates and sieve support interact. Stadlmann obtains $H_1 \le 240$ by combining the Bombieri–Vinogradov range with stronger estimates for restricted moduli [4, Theorem 1 and Section 1.1]. Charton, Hong, Lau, Ono, Remy, Siu, Swaminathan, Thorner, and Xie obtain $H_1 \le 212$ in work of Axiom Math [5, Theorem 1.1]. In independent work concurrent with Stadlmann's, OpenAI obtains $H_1 \le 186$, using complementary factorization conditions to enlarge the multidimensional sieve support [6, Theorem 1.1 and Corollary 1.2]. These are applications of distribution estimates and numerical sieve optimization. The present paper addresses a further analytic input: a stronger Type III estimate.

Our new local statement is Proposition 3.1. For a four-cycle of rank-three Kloosterman correlations with distinct row and column indices, it bounds the two-variable Fourier transform by $O(p^3)$ outside a uniformly bounded finite set, with $O(p^{7/2})$ allowed on that set. The improvement depends on excluding every curve-supported Fourier constituent. The rank-six trace construction, the radial phase calculation, and the two curve-exclusion lemmas are proved in this paper. Their inputs are the local Fourier theorems of Laumon and Fu, the purity and support formalism of perverse sheaves, and quantitative sheaf theory; the precise imported statements are recorded in Section 4. None of those general results is itself the four-cycle estimate.

The analytic consequence is Theorem 1.1. For comparison, the Type III estimate in [2, Theorem 2.8(v) and Theorem 7.1] uses

$$\sigma > \frac{1}{18} + \frac{28}{9}\omega + \frac{2}{9}\delta, \qquad \omega < \frac{1}{12}.$$

The finite exceptional set gives a smaller lower threshold for σ in the parameter region of Theorem 1.1. We retain its additional condition $108\omega + 2\delta > 1$ and all its individual factor-length hypotheses. The squarefree completion, averaged matrix estimate, and the optimization producing this range are proved below. A numerical prime-gap bound additionally requires the separate minorant and sieve arguments discussed in Section 18.

1.1 The distribution theorem

The purpose of a Type III estimate is to distribute a convolution with three smooth factors in arithmetic progressions whose moduli can exceed $x^{1/2}$. The term "Type III" specifies this factorization. It does not mean that an arbitrary sequence of the same total length satisfies the conclusion.

For a sequence f and a primitive residue class $a \pmod q$, write

$$\Delta(f; a(q)) = \sum_{n \equiv a \pmod q} f(n) - \frac{1}{\varphi(q)} \sum_{\gcd(n,q)=1} f(n). \tag{1.1}$$

Here and below convolution is Dirichlet convolution. A positive integer q is *y-densely divisible* if, for every real $1 \le R \le q$, it has a divisor in $[R/y, R]$. We denote by $\mathcal{D}(y)$ the squarefree integers with this property, following the one-fold notion in [2, Definition 2.1]. The definition guarantees a factor of approximately a prescribed size, but with a multiplicative uncertainty of y. We will keep that uncertainty in the proof.

A *coherent primitive family* is specified by choosing $a_p \in \mathbf{F}_p^*$ for each relevant prime and requiring $a(q) \equiv a_p \pmod p$ whenever $p \mid q$. Thus two moduli containing the same prime use the same residue at that prime. This condition will ensure that a local matrix parameter remains fixed while we average over two factors of the modulus.

All asymptotic parameters are fixed as $x \to \infty$. A *divisor-bounded sequence at scale M* is supported on $[M, 2M]$ and satisfies $|\alpha(m)| \le \tau_C(m)(\log x)^C$ for a fixed constant C. Here τ_C denotes a fixed divisor function. A *smooth sequence at scale N* has the form $\psi(n) = \Psi(n/N)$, where Ψ is supported in a fixed compact subinterval of $(0, \infty)$ and, for every fixed j, $\left\|\Psi^{(j)}\right\|_\infty \ll_j (\log x)^{C_j}$. Fixed changes to these support intervals cause no change in the argument.

Theorem 1.1. *Let $\omega, \delta > 0$ and $0 < \sigma < 1/6$ satisfy*

$$\sigma > \frac{1}{30} + \frac{56}{15}\omega + \frac{4}{15}\delta, \qquad 108\omega + 2\delta > 1. \tag{1.2}$$

Suppose that

$$\begin{gathered} MN_1N_2N_3 \asymp x, \qquad x^{2\sigma} \ll N_i \ll x^{1/2-\sigma}, \\ N_iN_j \gg x^{1/2+\sigma} \quad (i \neq j). \end{gathered} \tag{1.3}$$

If α is divisor bounded at scale M and each ψ_i is smooth at scale N_i, then, uniformly for coherent primitive families and for every fixed $A > 0$,

$$\sum_{\substack{q \in \mathcal{D}(x^\delta) \\ q \ll x^{1/2+2\omega}}} |\Delta(\alpha * \psi_1 * \psi_2 * \psi_3; a(q))| \ll_A x(\log x)^{-A}. \tag{1.4}$$

The implied constant may depend on the fixed coefficient bounds and on the strict margins in (1.2)*.*

The individual restrictions in (1.3) remain part of the theorem. Multiplying its three pair-product inequalities gives $M \ll x^{1/4-3\sigma/2}$, but this consequence alone is not an adequate hypothesis. The upper and lower bounds on each N_i are also used in the smooth Fourier reduction.

Only ψ_1, ψ_2, ψ_3 are differentiated. The sequence α can be arbitrary subject to its support and divisor bound. We use the Poisson reduction in [2, Sections 7.2–7.3], and check its coefficient hypotheses in Section 11 before inserting a new bound for the second moment. Section 17 explains why the short-modulus step likewise does not require smoothness or a Siegel–Walfisz hypothesis on α.

Theorem 1.1 is proved by the ordinary geometric and analytic argument in this article. Appendix A describes the scope of the Lean formalization, whose local-to-global theorem assumes the finite-field Fourier bounds.

1.2 Why a Fourier exceptional set matters

The eventual second moment contains a matrix of Kloosterman correlations. The fourth power of its operator norm is bounded by a trace, whose expansion consists of four-cycles. Thus the geometric question is about the Fourier transform of a product of four correlations in two modulus-factor variables.

A uniform bound of size p^3 outside a curve, with size $p^{7/2}$ on that curve, already gives cancellation. A finite exceptional set is more useful. For intervals of length $R \le p$, a curve can cost $Rp^{5/2} \log p$

in completion, whereas a bounded set of points costs only $R^2p^{3/2}$. This difference removes a term from the averaged matrix estimate and changes the allowable Type III range. We will derive these powers, rather than treating them as a convention about normalized traces.

There are four main stages of the proof. First we construct a pure rank-six sheaf whose trace differs from the raw correlation by an explicit constant. Next we calculate its radial ramification and show that the Fourier transforms of its relevant constituents have full support. We then pass from the resulting prime estimate to squarefree moduli and to a fourth moment of matrix norms. Finally we insert that estimate into the exact Type III second moment, including all the frequency and divisor terms. The final strict inequalities are proved by explicit identities in Section 16.

2 Finite Fourier conventions and exact correlation identities

Put $e_q(t) = \exp(2\pi i t/q)$. A star on a sum modulo q restricts its variable to the units. All finite Fourier transforms in this paper are *unnormalized* and have a positive sign:

$$\widehat{f}(\xi) = \sum_{x \bmod q} f(x) e_q(\xi x), \qquad f(x) = \frac{1}{q} \sum_{\xi \bmod q} \widehat{f}(\xi) e_q(-\xi x). \tag{2.1}$$

In two variables the inverse factor is q^{-2}. Continuous Fourier transforms have the negative sign, $\widehat{\psi}(\xi) = \int_{\mathbf{R}} \psi(t) e^{-2\pi i \xi t}\, dt$. These choices make the completion formula in Section 12 especially simple.

The normalized rank-three Kloosterman sum and its correlation are

$$\mathrm{Kl}_3(t;q) = \frac{1}{q} \sideset{}{^*}\sum_{u,v \bmod q} e_q\left(u + v + \frac{t}{uv}\right), \tag{2.2}$$

$$C_q(A,B,c) = \sideset{}{^*}\sum_{h \bmod q} \mathrm{Kl}_3(Ah;q)\overline{\mathrm{Kl}_3(Bh;q)} e_q(ch). \tag{2.3}$$

Parameters A, B are units whenever these expressions are used below. We sometimes write $K_3(t) = \mathrm{Kl}_3(t;p)$ at a prime. For comparison,

$$\mathrm{Kl}_2(t;p) = \frac{1}{\sqrt{p}} \sum_{u \in \mathbf{F}_p^*} e_p(u + t/u). \tag{2.4}$$

The factors p^{-1} in Kl_3 and $p^{-1/2}$ in Kl_2 are included in these definitions. There is no additional normalization in C_p or in its Fourier transform. For nonzero arguments, the standard bounds $|\mathrm{Kl}_3(t;p)| \le 3$ and $|\mathrm{Kl}_2(t;p)| \le 2$ follow from the rank and purity statements for Kloosterman sheaves [7, Theorem 4.1.1]. The identities in the next lemma are elementary calculations with the displayed sums.

Lemma 2.1. *Let p be prime, $A, B \in \mathbf{F}_p^*$, and $\kappa_p = 1 + p^{-1} + p^{-2}$. Then*

$$C_p(A,B,0) = p\,\mathbf{1}_{A=B} - \kappa_p. \tag{2.5}$$

If $T_p(A,d) = \sum_{h \in \mathbf{F}_p^*} K_3(Ah) e_p(dh)$, then

$$T_p(A,d) = \begin{cases} \sqrt{p}\,\mathrm{Kl}_2(-A/d;p) - p^{-1}, & d \neq 0, \\ -p^{-1}, & d = 0. \end{cases} \tag{2.6}$$

Proof. Opening the definition of T_p, the sum over h is $p\,\mathbf{1}_{A/(uv)+d=0}-1$. Also $\sum_{u,v\neq 0} e_p(u+v)=1$. For $d=0$ the first term never occurs. For $d\neq 0$ it imposes $uv=-A/d$, giving (2.6).

For (2.5), open both Kloosterman sums and sum over h. With $\beta=B/A$, the condition from the first term is $xy=\beta uv$. Hence

$$C_p(A,B,0)=\frac{1}{p}\sum_{u,v,x\neq 0} e_p\left(u+v-x-\frac{\beta uv}{x}\right)-\frac{1}{p^2}.$$

Summing first over v, the remaining three-variable sum is

$$p\sum_{u\neq 0} e_p((1-\beta)u)-\sum_{u,x\neq 0} e_p(u-x)=p(p\,\mathbf{1}_{\beta=1}-1)-1.$$

Substitution gives the claimed formula. These computations also fix the constant at zero frequency without any appeal to sheaf conventions. □

2.1 The inverse cubes in the Chinese remainder theorem

For coprime q_1,q_2, the additive character splits as

$$e_{q_1q_2}(z)=e_{q_1}(zq_2^{-1})e_{q_2}(zq_1^{-1}).$$

In the q_1-factor of (2.2), replace u,v by q_2u',q_2v'. The linear terms become $u'+v'$, and the last term acquires q_2^{-3}. Therefore

$$\mathrm{Kl}_3(t;q_1q_2)=\mathrm{Kl}_3(tq_2^{-3};q_1)\,\mathrm{Kl}_3(tq_1^{-3};q_2). \tag{2.7}$$

Every inverse is taken modulo the modulus of its factor. The exponent three in this formula is responsible for the equality $mu^3=nv^3$ that will appear at the central frequency. Replacing that equality by $m=n$ would give a different second moment.

3 The prime Fourier statement

Fix $\alpha\in\mathbf{F}_p^*$. For unit indices m,n, define

$$K_{mn}(r_1,r_2)=C_p\left(\frac{\alpha r_2}{mr_1^2},\frac{\alpha r_1}{nr_2^2},1\right),\qquad r_1r_2\neq 0. \tag{3.1}$$

For two row indices m,m' and two column indices n,n', put

$$W(r_1,r_2)=K_{mn}\overline{K_{m'n}}K_{m'n'}\overline{K_{mn'}},\qquad D_0=(m-m')(n-n'), \tag{3.2}$$

and extend W by zero outside $\mathbf{G}_m^2$. In this section the indices are residue classes. Later, when they are integers in intervals, we will distinguish equality of integers from equality modulo a prime.

Proposition 3.1. *There are absolute constants C,p_0 with the following property. For $p>p_0$ and $p\nmid D_0$, a Frobenius-stable geometric set $Z_p\subset\mathbf{A}^2_{\overline{\mathbf{F}}_p}$ of at most C points satisfies*

$$|\widehat{W}(h,k)|\le Cp^3(1+\sqrt{p}\,\mathbf{1}_{Z_p}(h,k)). \tag{3.3}$$

For $p\mid D_0$, a curve E_p of uniformly bounded total degree satisfies

$$|\widehat{W}(h,k)|\le Cp^3(1+\sqrt{p}\,\mathbf{1}_{E_p}(h,k)+p\,\mathbf{1}_{(h,k)=(0,0)}). \tag{3.4}$$

In every case, $|K_{mn}(r_1,r_2)|\ll\sqrt{p}$. All constants are uniform in the unit parameters.

The finite set and the curve can depend on the indices. Their size or degree cannot. This uniformity is required before the Chinese remainder theorem can be applied over arbitrarily many primes.

The large-characteristic condition belongs to the proof of the prime estimate, not to the distribution theorem. For each of the finitely many primes $p \le p_0$, the trivial complete bound can be included by enlarging an absolute constant. Since the product of those primes is fixed, this also causes only an absolute loss at squarefree level.

4 The sheaf facts used in the geometric argument

This section records the conventions and the external inputs needed for Proposition 3.1. Sections 5–8 give the constructions and arguments specific to the four-cycle.

Choose an auxiliary prime different from p and an embedding of its algebraic coefficient field into $\mathbf{C}$. A lisse sheaf is a locally constant family of finite-dimensional representations; its trace function is the trace of geometric Frobenius. A pure sheaf of weight w has Frobenius eigenvalues of absolute value $p^{w/2}$ at $\mathbf{F}_p$-points. A Tate twist (a) multiplies those eigenvalues by p^{-a} and lowers the weight by $2a$.

For a complex P, its trace is the alternating sum of the traces on the ordinary cohomology sheaves $\mathcal{H}^i(P)$. Shifts therefore affect trace signs. If $\mathcal{G}$ is lisse of weight w on a smooth variety of dimension d, then $\mathcal{G}[d]$ is perverse of weight $w+d$. Its intermediate extension is denoted $j_{!*}(\mathcal{G}[d])$, or $\mathrm{IC}(\mathcal{G})$. It agrees with $\mathcal{G}[d]$ on the open set and has no perverse subobject or quotient supported on the complement. Purity of this intermediate extension follows from [12, Corollary 5.3.2]; the open immersions used below are affine. We use the Grothendieck–Lefschetz trace formula in the form [9, Rapport, Theorem 3.2], for constructible sheaves on separated schemes of finite type over a finite field, and its alternating-sum extension to bounded constructible complexes.

4.1 Fourier transform, support, and weights

On $\mathbf{A}^d$, our sheaf Fourier transform is

$$\mathrm{FT}_d(P) = R\pi_{2!}\big(\pi_1^* P \otimes \mathcal{L}_\psi(x \cdot \xi)\big)[d]. \tag{4.1}$$

There is no normalizing Tate twist. It is an equivalence of perverse categories, and

$$\mathrm{FT}_d\,\mathrm{FT}_d(P) \simeq [-1]^* P(-d). \tag{4.2}$$

The inversion formula is [11, Theorem 1.2.2.1 and Corollary 1.2.2.3]; its equivalence on perverse sheaves is [11, Theorem 1.3.2.3]. It takes a pure perverse complex of weight v to one of weight $v+d$, by [13, Corollary 2.1.5 and Theorem 2.2.1]. Compatibility with a linear map is the separate statement [11, Theorem 1.2.2.4]; translation is [11, Proposition 1.2.3.2]. We will use both in Section 7. For $d=2$, the even shift means that its trace is exactly the positive, unnormalized Fourier sum of the input trace.

A simple perverse sheaf has an irreducible support and is the intermediate extension of an irreducible local system on a dense smooth open subset of that support [12, Theorem 4.3.1(ii)]. Pure perverse sheaves are geometrically semisimple [12, Theorem 5.3.8]. Consequently, to understand the possible exceptional Fourier loci, we may study the support of the Fourier transform of each geometric simple constituent.

There is a useful strict-support fact on a smooth surface. For a full-support intermediate extension P,

$$\mathcal{H}^0(P) = 0, \qquad \dim \operatorname{supp} \mathcal{H}^{-1}(P) \le 0, \qquad \mathcal{H}^i(P) = 0 \quad (i \notin \{-2, -1\}). \tag{4.3}$$

This is stronger than ordinary perversity, which would allow a curve in degree -1. It follows from the defining strict boundary condition $i^*P \in {}^pD^{\le -1}$ for intermediate extension [12, Corollaries 1.4.24–1.4.25 and Proposition 2.2.4]. We will explain its numerical consequence in Section 8.

4.2 Local Fourier transform and the meaning of a phase

The inertia representation at a boundary point measures ramification. “Tame” means that wild inertia acts trivially. A rank-one Artin–Schreier sheaf $\mathcal{L}_\psi(a/T)$, with $a \neq 0$, has slope one at $T = 0$. After the cover $t = T^2$, a slope $1/2$ representation in the t-variable can have precisely this form. We call a/T its wild phase; tame and unramified tensor factors do not change the slope or the coefficient a.

We use the following consequences of local stationary phase. Local Fourier contributions are direct sums of inertia representations. A nontrivial tame block at a finite point a contributes at Fourier infinity a nonzero tame block tensored with $\mathcal{L}_\psi(\pm at)$. If $a \neq 0$, this has slope one. The infinity-to-infinity transform receives no contribution from slopes less than one. These are the relevant parts of [11, Proposition 2.3.3.1(iii), Theorem 2.4.3, and Proposition 2.5.3.1(ii)]. Their direct-sum nature will prevent cancellation between distinct points in a projection fiber. At the finite Fourier origin we instead use the vanishing-cycle sequence of [11, Section 2.3.2, Proposition 2.3.2.1(iii), and Lemma 2.4.2.1(ii)]; its application to the correlation family is given in Section 6.

For the explicit rank-three calculation we use [10, Proposition 0.8 and Theorem 0.1(iii)]. The normalized rank-three Kloosterman sheaf has one tame unipotent Jordan block at zero. At infinity its geometric representation, up to tame and unramified factors, is $[3]_*\mathcal{L}_\psi(3x)$, where $[3]$ denotes the cubic cover. Fu’s theorem calculates the infinity-to-zero transform by the derivative coordinates displayed in Section 6. The rank and phase are both part of that local isomorphism. The application has $r = 3$, $s = 1$: thus $s < r$, $s < p$, and the required coprimality of p with $2, r, s, s - r$ holds for $p > 3$. Its six wild phases describe a generic parameter λ; a specialization such as $\lambda = 1$ can also have tame terms. The generic radial calculation below has $\lambda = m/(nz^3)$, so it is in the former case.

4.3 Uniformity in all residue parameters

Bounded rank alone does not bound the number of Fourier exceptional points or the degree of a possible support curve. We use quantitative sheaf theory for that additional purpose. In the fixed ambient dimensions used here, the six operations are controlled by [14, Theorem 6.8], simple constituents and intermediate extensions by [14, Theorem 6.15 and Corollary 6.16], ordinary cohomology by [14, Proposition 6.24], and Fourier transform by [14, Proposition 7.18]. Degrees of nonlisse loci and strata are controlled by [14, Theorem 6.23 and Lemma 6.26]. The quantitative generic base-change theorem [14, Theorem 6.27] provides a dense open with a bounded-degree complement; it does not assert base change at every parameter.

In our application the starting Kloosterman and Artin–Schreier sheaves have fixed complexity by [14, Corollary 7.4 and Proposition 7.5] and their fixed rank and ramification. For each rational map, adjoining the inverses of its denominators realizes its domain and graph by a fixed number of equations of fixed degree. Its complexity is therefore bounded by [14, Proposition 6.21]. The coefficients of those maps contain α, m, n, but their degrees do not. Thus these results give constants independent of both the residue parameters and p. This supplies a bound for the degree of a possible Fourier-support curve *before* the curve is excluded. We may then require p to exceed that bound.

5 The rank-six pure correlation

Let $\mathcal{K}$ be the normalized rank-three Kloosterman sheaf on $\mathbf{G}_m$, pure of weight zero and with trace K_3 [7, Theorem 4.1.1]. The normalization is the integer Tate twist of the usual weight-two rank-three sheaf. For $\lambda \neq 0$, put

$$\mathcal{F}_\lambda = \mathcal{K}(x) \otimes \mathcal{K}(\lambda x)^\vee. \tag{5.1}$$

Because $\mathcal{K}$ has weight zero, duality gives complex conjugation of its trace. Changing variables $x = Ah$ in (2.3) yields

$$C_p(A, \lambda A, 1) = \sum_{x \neq 0} t_{\mathcal{F}_\lambda}(x) e_p(x/A). \tag{5.2}$$

Thus the correlation is a one-variable Fourier transform, except that the point $x = 0$ has been omitted.

5.1 The omitted stalk at zero

Let $j : \mathbf{G}_m \hookrightarrow \mathbf{A}^1$. The stalk of the ordinary sheaf $j_*\mathcal{F}_\lambda$ at zero is the inertia-invariant space. Both Kloosterman factors have one unipotent block of length three. The monodromy and the Frobenius action on its invariant line are given by [7, Theorem 7.4.3]; the geometric description is also given by [10, Proposition 0.8]. After identifying those blocks, the invariant endomorphisms are the centralizer of one regular nilpotent matrix N:

$$\langle I, N, N^2 \rangle.$$

In particular the invariant dimension is three, not nine. Geometric Frobenius conjugates N to $p^{-1}N$. To track the two scaled factors separately, choose a common Jordan basis and set $D = \mathrm{diag}(p^{-1}, 1, p)$. Write their Frobenius operators as $F_i = U_i D$. Each U_i commutes with N, hence is a polynomial in N. Katz's invariant-line normalization, after the Tate twist (1), gives the scalar p^{-1} for both factors; consequently each U_i has constant coefficient one. On the invariant intertwiner space, Frobenius therefore acts by

$$X \longmapsto F_1 X F_2^{-1} = (DXD^{-1})(1 + aN + bN^2)$$

for some coefficient-field scalars a, b. In the ordered basis I, N, N^2 this map is triangular with diagonal $1, p^{-1}, p^{-2}$. Thus coordinate scaling leaves its trace unchanged, even though the two Frobenius operators need not coincide. Therefore

$$t_{(j_*\mathcal{F}_\lambda)_0} = 1 + p^{-1} + p^{-2} = \kappa_p. \tag{5.3}$$

The ordinary-sheaf exact sequence

$$0 \longrightarrow j_!\mathcal{F}_\lambda \longrightarrow j_*\mathcal{F}_\lambda \longrightarrow i_*\mathcal{F}_\lambda^{I_0} \longrightarrow 0$$

now shows exactly what has to be added to (5.2). The complete Fourier sum of the middle-extension trace is $C_p(A, \lambda A, 1) + \kappa_p$. Formula (2.5) gives an independent elementary check on the same constant.

5.2 Why the Fourier fiber has rank six and weight one

Fix a nonzero Fourier coordinate ξ. The rank-nine sheaf $\mathcal{F}_\lambda \otimes \mathcal{L}_\psi(\xi x)$ has Swan conductor nine at infinity. Indeed, the slopes of $\mathcal{F}_\lambda$ there are at most $1/3$, whereas the additive twist has slope one. Every resulting slope is one, and the invariant space at infinity is zero. At zero the sheaf is tame with the three invariants just computed. The Grothendieck–Ogg–Shafarevich formula [11, Theorem 2.2.1.2], applied on $\mathbf{P}^1$ after extension by zero at infinity, therefore gives

$$\chi_c(\mathbf{A}^1, j_*\mathcal{F}_\lambda \otimes \mathcal{L}_\psi(\xi x)) = -9 + 3 = -6. \tag{5.4}$$

There is no compactly supported H^0, since an ordinary middle-extension sheaf has no nonzero section supported at finitely many points. There is no H^2_c, since such a contribution would require a geometrically constant quotient, incompatible with the slope-one inertia at infinity, by Poincaré duality [9, Sommes trig., Remark 1.18(d)]. Thus the only cohomology is six-dimensional H^1_c.

The vanishing of the invariant stalk at infinity also identifies this space with the first cohomology of the ordinary middle extension on $\mathbf{P}^1$. Deligne's purity theorem for a smooth projective curve, in the form [11, Theorem 4.1.3], makes it pure of weight one. The trace formula [9, Rapport, Theorem 3.2] and (5.2) consequently give

$$\mathrm{tr}(\mathrm{Fr} \mid H^1_c) = -(C_p(A, B, 1) + \kappa_p).$$

An unramified twist with Frobenius eigenvalue -1 removes the minus sign at $\mathbf{F}_p$-points and changes no geometric inertia. We use

$$\mathcal{C}_p(A, B) = C_p(A, B, 1) + \kappa_p \tag{5.5}$$

as the trace of that pure rank-six family. The calculation is valid for every pair of unit parameters, including $A = B$, and gives

$$|C_p(A, B, 1)| \leq 6\sqrt{p} + \kappa_p. \tag{5.6}$$

Purity of the fibers and bounded complexity yield a lisse rank-six family on a dense open of the parameter torus. No assertion that this family is lisse on the entire torus is needed.

5.3 The common open set for the four entries

For completeness, the family can be constructed from fixed sheaves by the maps $(x, \lambda) \mapsto x, \lambda x$, ordinary j_*, tensoring by $\mathcal{L}_\psi(\xi x)$, compactly supported pushforward in x, and ordinary H^1. Effective generic base change [14, Theorem 6.27] identifies its fibers with those just calculated. Finally substitute $(\lambda, \xi) = (B/A, 1/A)$. The results cited in Section 4 make every exceptional locus in this construction uniformly bounded in degree.

For the entry (m, n), the parameter map is

$$(r_1, r_2) \longmapsto \left(\frac{\alpha r_2}{m r_1^2}, \frac{\alpha r_1}{n r_2^2}\right). \tag{5.7}$$

Its exponent matrix has determinant three. For $p > 3$ it is dominant and generically separable, so the preimage of the generic lisse locus remains dense. Intersect the four preimages and, if necessary, remove a further bounded-degree hypersurface. This gives one principal affine open $U \subset \mathbf{A}^2$ on which all four core sheaves are lisse. The complement of U has bounded degree.

For a conjugated entry use $\mathcal{H}^\vee(-1)$, where $\mathcal{H}$ is its weight-one core sheaf. A Frobenius eigenvalue z then becomes $p/z = \overline{z}$. Thus a conjugated entry is again represented by a weight-one sheaf. Every nonempty subproduct of the four core factors is pure of weight equal to its number of factors, between one and four. This observation will allow us to retain the constant correction κ_p exactly.

6 The complete radial phase calculation

We now calculate the ramification needed to exclude proper Fourier supports. Work over the algebraically closed constant field $\overline{\mathbf{F}}_p$ and then over an algebraic closure of $\overline{\mathbf{F}}_p(z)$, where z is transcendental. The generic ray is

$$r_1 = t, \qquad r_2 = zt. \tag{6.1}$$

Using a transcendental direction avoids every fixed exceptional direction at once. A geometric constituent over $\overline{\mathbf{F}}_p$ restricts to this ray, and its local representations are subquotients of the corresponding restrictions of the full sheaf.

6.1 One entry: three rank-two local transforms

Up to tame and unramified factors, the local description of $\mathcal{K}$ at infinity is $[3]_*\mathcal{L}_\psi(3x)$. The tensor in (5.1), decomposed over the cubic cover, has three rank-three blocks

$$[3]_*\mathcal{L}_\psi(3d_jx), \qquad d_j = 1 - q_j, \qquad q_j^3 = \lambda. \tag{6.2}$$

There are three choices of q_j. For generic λ, all d_j are nonzero.

Apply Fu's infinity-to-zero formula with $\gamma(x) = x^3$ and $a(x) = 3d_jx$. Its derivative coordinate and transformed phase are

$$\xi = -\frac{a'(x)}{\gamma'(x)} = -\frac{d_j}{x^2}, \qquad a(x) + \gamma(x)\xi = 2d_jx. \tag{6.3}$$

Thus each block produces two phases

$$\pm 2\sqrt{-d_j^3/\xi} = \pm 2\sqrt{A(q_j - 1)^3}, \qquad \xi = 1/A. \tag{6.4}$$

The tame quadratic factors in Fu's formula do not affect these wild coefficients. The theorem applies because the relevant degrees are three and one, and $p > 3$.

To identify these local transforms with the actual correlation cohomology, put

$$K_\lambda = j_*\mathcal{F}_\lambda[1], \qquad K'_\lambda = \mathrm{FT}_\psi(K_\lambda),$$

where j_* is the ordinary direct image. On the punctured Fourier line write $K'_\lambda = H_\lambda[1]$; before the constant sign twist, H_λ is the rank-six cohomology family of Section 5. Let η_0 be a geometric generic point of the strict local line at $\xi = 0$. Laumon's finite-point vanishing-cycle sequence [11, Section 2.3.2] is

$$0 \longrightarrow \mathcal{H}^{-1}(K'_{\lambda,0}) \longrightarrow H_{\lambda,\eta_0} \longrightarrow R^{-1}\Phi_0(K'_\lambda) \longrightarrow \mathcal{H}^0(K'_{\lambda,0}) \longrightarrow 0.$$

The end terms have trivial geometric inertia. By [11, Proposition 2.3.2.1(iii) and Lemma 2.4.2.1(ii)], the vanishing-cycle term is the infinity-to-zero local transform of $\mathcal{F}_\lambda$. For $\lambda \neq 0, 1$, the three rank-two transforms above make it entirely wild of dimension six. It has no nonzero map to the last, trivial-inertia term. Exactness therefore makes the middle map surjective, and (5.4) makes it an isomorphism. Equivalently,

$$H_\lambda|_{I_{\xi=0}} \simeq \mathcal{F}_\psi^{(\infty,0)}(\mathcal{F}_\lambda|_{I_{x=\infty}}) \qquad (\lambda \neq 0, 1).$$

This proves that the six phases exhaust the actual local representation, with no additional tame summand. The constant sign twist changes no geometric inertia. The restriction $\lambda \neq 1$ is essential to this all-wild argument; at $\lambda = 1$ one Mackey block is tame. The generic direction below avoids that value.

On the ray (6.1),

$$A = \frac{\alpha z}{mt}, \qquad B = \frac{\alpha}{nz^2 t}, \qquad \lambda = \frac{m}{nz^3}.$$

Write $q_j = \gamma/z$, with $\gamma^3 = m/n$, and put $t = T^2$. Formula (6.4) becomes

$$\boxed{\pm\frac{2\sqrt{\alpha/m}}{zT}(\gamma - z)^{3/2}, \qquad \gamma^3 = m/n.} \tag{6.5}$$

This is the full six-character wild list, up to tame factors. In particular, the coefficient depends on the direction through a cubic power of $\gamma - z$, not through the reciprocal of a linear function. That distinction will exclude origin-line Fourier supports.

6.2 Why no nonempty subproduct has a zero phase

Assume now $m \neq m'$ and $n \neq n'$ in $\mathbf{F}_p$. A phase of a tensor product is a sum of one chosen phase from each factor. Conjugation changes a sign, which is already allowed by (6.5). Collect terms having the same γ. Every phase coefficient has the form

$$\beta(z) = \frac{1}{z}\sum_{\gamma\in\Gamma} c_\gamma(\gamma - z)^{3/2}. \tag{6.6}$$

We claim that $\Gamma \neq \varnothing$ and every grouped coefficient c_γ is nonzero, for every one of the fifteen nonempty subsets of the four entries.

Two entries sharing a row cannot share γ, since equality of their cubes would force their columns to be equal. The same applies to entries sharing a column. Thus a common γ can occur only in two opposite entries. It cannot occur in three entries. The squared amplitudes in such an opposite pair are proportional to $1/m$ and $1/m'$. If the two signed amplitudes canceled, their squares would agree and force $m = m'$, contrary to the assumption. This proves the claim even when the two opposite pairs have coinciding cubic branch sets.

Finally, the squareclasses of the distinct linear functions $\gamma - z$ are independent in $\overline{\mathbf{F}}_p(z)^*/\overline{\mathbf{F}}_p(z)^{*2}$: the valuation at $z = \gamma$ detects that generator and none of the others. In the resulting multiquadratic extension, the terms in (6.6) belong to distinct character spaces. They cannot sum to zero. Every wild character in every nonempty tensor therefore has a nonzero pole-one phase $\beta(z)/T$.

Corollary 6.1. *For locally distinct row and column indices, no generic lisse constituent of any nonempty core subproduct is a linear Artin–Schreier sheaf $\mathcal{L}_\psi(ar_1 + br_2)$, including the constant sheaf.*

Proof. On the generic ray, a linear Artin–Schreier sheaf is unramified at $t = 0$. Our exhaustive list has no tame character there. Restriction and passage to a constituent cannot create a new wild-inertia character absent from the list. □

6.3 Radial infinity, including repeated indices

There is a second local fact, valid even if row or column indices repeat. Along (6.1), the ratio λ is fixed and the Fourier coordinate $\xi = 1/A$ is a nonzero constant times t. The input $\mathcal{F}_\lambda$ is tame at zero, lisse at every other finite nonzero point, and has slopes at most $1/3$ at infinity. By stationary phase, its Fourier transform has no infinity-to-infinity contribution. The contribution from the finite point zero is tame at Fourier infinity, and there is no other finite singularity to contribute a nonzero linear phase. Thus the core is tame at radial $t = \infty$.

Tensor products, conjugate duals, and subquotients preserve this tameness. A nonconstant linear sheaf $\mathcal{L}_\psi(ar_1 + br_2)$ would restrict to $\mathcal{L}_\psi((a + bz)t)$, which has slope one at infinity because z is transcendental and $(a, b) \neq (0, 0)$. Such a constituent is excluded for all indices. For repeated indices the constant constituent is not excluded. This is exactly why (3.4) permits a larger contribution at the Fourier origin and nowhere else.

7 Why a Fourier-support curve is impossible

Absence of linear Artin–Schreier constituents excludes punctual Fourier supports. We now prove the additional curve-exclusion statement. The following lemma combines an elementary projection argument with the cited local Fourier theorems; it is a result of this paper. We then apply the exact angular phases to the remaining line case.

Lemma 7.1. *Let Q be a geometrically simple, full-support perverse sheaf on $\mathbf{A}^2_{\overline{\mathbf{F}}_p}$. Suppose its generic lisse sheaf is tame at infinity on the ray $(r_1, r_2) = (t, zt)$. If $\mathrm{FT}_2(Q)$ is supported on an irreducible curve Z of degree less than p, then Z is a line through the origin. On a dense open, the underlying lisse sheaf of Q is a pullback $J(ar_1 + br_2)$, up to an unramified twist, for a nonzero one-variable sheaf J over $\overline{\mathbf{F}}_p$ and $(a, b) \neq (0, 0)$.*

Proof. Since Fourier preserves simplicity, write $\mathrm{FT}_2(Q) = i_{Z*}\,\mathrm{IC}_Z(L)$, where L is nonzero and lisse on a dense smooth open of Z. We first show that Z cannot be nonlinear.

Let $f(h, k) = 0$ be an irreducible equation of Z, of degree $D < p$, and consider the projection

$$\pi_z : Z \longrightarrow \mathbf{A}^1, \qquad (h, k) \longmapsto h + zk. \tag{7.1}$$

It is finite of degree D for transcendental z. Indeed, after substituting $h = y - zk$, the coefficient of k^D is the nonzero constant $f_D(-z, 1)$, where f_D is the highest homogeneous part. Dividing by that constant makes k integral over $\overline{\mathbf{F}}_p(z)[y]$. The normalization of Z is finite over Z, so its projection is finite as well. Its degree is below p; hence the induced function-field extension is separable and all its ramification indices are prime to p.

We need a ramification point in the smooth lisse locus whose projected value is not zero. Here are the two geometric reasons it exists. If the Gauss map of a nonlinear Z were constant, a fixed nonzero directional derivative of f would vanish on Z. Its degree is less than D, so it would vanish identically. Since $D < p$, this would make f a polynomial in one linear form. Absolute irreducibility would force a line, a contradiction. The Gauss map is therefore nonconstant and hence dominant onto the line of directions.

Next, the Euler derivative

$$H(h, k) = hf_h(h, k) + kf_k(h, k) \tag{7.2}$$

cannot vanish identically on a nonlinear Z. Otherwise $f \mid H$, and comparison of degrees gives $H = cf$. The homogeneous components of f have distinct degrees modulo p, because $D < p$; consequently only one homogeneous degree can occur. An irreducible homogeneous binary polynomial over $\overline{\mathbf{F}}_p$ is linear. Again this contradicts nonlinearity.

Remove the finitely many singular points of Z, the nonlisse points of L, and the zeros of H and f_h on the remaining curve. The dominant Gauss map still has a point of generic direction in the resulting open set. At such a point,

$$f_k = zf_h, \qquad \pi_z(h, k) = h + zk = H/f_h \neq 0.$$

It is a ramification point of π_z, of tame index $e \geq 2$, with a nonzero critical value, say a. No simple-tangency assumption is being made; e can be larger than two.

Because L is lisse at this point, its local pushforward contains the degree-e tame permutation representation, repeated rank L times. Over the algebraic coefficient field this representation contains every nontrivial character of the tame cyclic group of order e. Other points above a give direct summands. Thus these nontrivial tame blocks cannot be canceled by contributions from the rest of the fiber. Each chosen nontrivial character has zero inertia invariants. Its ordinary middle extension therefore has zero stalk at a, exactly as its zero extension does. This is why its nonzero local Fourier contribution survives; the trivial character in the permutation representation would not justify the same step.

We now relate this local statement to Q. For the ray inclusion $i_z(t) = (t, zt)$, linear Fourier compatibility [11, Theorem 1.2.2.4] gives the exact shift

$$i_z^* \operatorname{FT}_2(i_{Z*} \operatorname{IC}_Z L) \simeq \operatorname{FT}_1(R\pi_{z!} \operatorname{IC}_Z L)[1]. \tag{7.3}$$

This can also be read directly from (4.1): on the ray the phase is $t(h + zk)$; the plane transform has shift two and the one-variable transform shift one. To recover Q from its transform, (4.2), namely [11, Theorem 1.2.2.1 and Corollary 1.2.2.3], additionally gives a reflection and a Tate twist. Neither changes a ramification slope or the fact that a critical value is nonzero.

The finite map π_z has perverse-exact pushforward by [12, Corollary 4.1.3], since it is quasi-finite and affine. Its nontrivial tame blocks are therefore actual local blocks of the middle-extension part, not terms in an alternating cohomology sum. Stationary phase sends a nontrivial tame block at $a \neq 0$ to a nonzero block at Fourier infinity with the factor $\mathcal{L}_\psi(\pm at)$. Its slope is one. The direct-sum stationary-phase decomposition precludes cancellation by other singularities. By (7.3), this is a wild block in the radial restriction of Q, contradicting its tameness at infinity. The Fourier and stationary-phase statements used here are the ones cited in Section 4; translation is [11, Proposition 1.2.3.2].

It remains to consider an affine line. Write it as $Z = u + \overline{\mathbf{F}}_p v$, with $v \neq 0$. Fourier compatibility with a linear inclusion, followed by translation, gives on a dense open

$$Q \simeq (\mathcal{L}_\psi(\pm u \cdot (r_1, r_2)) \otimes J(v \cdot (r_1, r_2)))[2], \tag{7.4}$$

up to a Tate twist. The one-variable sheaf J is defined over the constant field $\overline{\mathbf{F}}_p$. Its generic rank is nonzero: otherwise the one-variable inverse Fourier transform would be punctual, and the pullback in (7.4) would not have full support.

On the generic ray, put $X = (v_1 + zv_2)t$. Radial tameness says that

$$J(X) \otimes \mathcal{L}_\psi(c(z)X) \quad \text{is tame at infinity}, \qquad c(z) = \pm \frac{u_1 + zu_2}{v_1 + zv_2}. \tag{7.5}$$

A fixed nonzero local representation cannot be made tame by two different linear twists: after the first twist it is nonzero and tame, and tensoring by their nontrivial quotient gives slope one. Since J is defined over $\overline{\mathbf{F}}_p$, apply the constant-field automorphism $z \mapsto z + 1$ to (7.5). Uniqueness implies $c(z + 1) = c(z)$. For this ratio of linear functions, that equality gives $u_1v_2 - u_2v_1 = 0$. Hence u is parallel to v, and the line passes through the origin. Its translation factor can be absorbed into J, proving the lemma. $\square$

7.1 Excluding a line through the origin

The preceding lemma uses radial infinity. The angular dependence at radial zero provides the remaining obstruction.

Lemma 7.2. *Assume $m \neq m'$ and $n \neq n'$ modulo p. No geometric full-support simple constituent of an intermediate extension of a nonempty core subproduct can have its Fourier transform supported on a line through the origin.*

Proof. If such a constituent existed, Lemma 7.1 would identify its generic lisse sheaf with $J(ar_1+br_2)$, for $(a,b) \neq (0,0)$. The wild-inertia representation of a constituent is a subquotient of the full representation. After $t = T^2$, the full representation has the exhaustive rank-one list (6.6), with no tame character. Our constituent therefore contains a nonzero character from that list.

We make the passage to a constant coefficient explicit. Choose a constant direction $z_0 \in \overline{\mathbf{F}}_p$ with $a + bz_0 \neq 0$, outside the following finite set of bad directions. Exclude $z_0 = 0$, the cubic branch directions occurring in the four entries, and every direction whose line is contained in the complement of the open set where the constituent and pullback are identified. Also exclude the zeros and poles of all the phase functions in the exhaustive list. These last sets are finite: the list is finite, its members are nonzero algebraic functions on a finite cover of the direction line, and zeros and poles have finite images on that line. The algebraically closed field $\overline{\mathbf{F}}_p$ is infinite, so such a z_0 exists.

On this constant ray, restrict the actual sheaf inclusion giving the constituent. Its nonzero rank is unchanged on a dense open of the ray. The same local calculation as in Section 6 applies directly at $z = z_0$: the excluded directions ensure that its denominators and phases do not degenerate. After $t = T^2$, the full representation has only the characters $\mathcal{L}_\psi(\beta_j(z_0)/T)$ on wild inertia, with $\beta_j(z_0) \in \overline{\mathbf{F}}_p^*$. A subquotient has a nonempty sublist. Choose $u_0 \in \overline{\mathbf{F}}_p^*$ with $u_0^2 = a + bz_0$, and set $Y = u_0 T$. The pullback is now $J(Y^2)$, and its wild characters have coefficients

$$c_j = u_0 \beta_j(z_0) \in \overline{\mathbf{F}}_p^* \quad \text{in the phases} \quad \mathcal{L}_\psi(c_j/Y).$$

This identifies the characters of the fixed one-variable sheaf $J(Y^2)$ over the constant field. It is not an inference from the slope alone.

Return to the generic direction and choose $u(z) \in \overline{\overline{\mathbf{F}}_p(z)}^*$ with $u(z)^2 = a + bz$. The same constant-field representation is base changed, and $Y = u(z)T$. Its wild characters are consequently $\mathcal{L}_\psi(c_j/(u(z)T))$, with the same constants c_j. Matching any one of them to the constituent's exhaustive generic list gives $\beta(z) = c_j/u(z)$. In particular, for a nonzero constant $c \in \overline{\mathbf{F}}_p$,

$$\beta(z)^2 = \frac{c^2}{a+bz}. \tag{7.6}$$

There is no ambiguity from an Artin–Schreier coboundary here. A nonzero Laurent series of the form $g^p - g$ with a pole has leading pole order divisible by p; it cannot change a coefficient of T^{-1} and leave no higher pole.

Equation (7.6) is impossible. If (6.6) has at least two different γ's, its square is not in $\overline{\mathbf{F}}_p(z)$. Indeed, the cross term from an unordered pair γ, γ' is a nonzero vector in the multiquadratic character space corresponding to $(\gamma - z)(\gamma' - z)$. Independent squareclasses make these spaces distinct for distinct unordered pairs, so no cross term can cancel another. The right side of (7.6) is rational, a contradiction.

If there is just one γ, then

$$\beta(z)^2 = c_\gamma^2 \frac{(\gamma - z)^3}{z^2}.$$

It has a cubic zero at the nonzero point $z = \gamma$. A nonzero constant divided by a linear function has no finite zero. This is again incompatible with (7.6). $\square$

7.2 What the two geometric hypotheses exclude

The large-characteristic and angular-phase qualifications have distinct roles. A general radial-tameness statement without the degree restriction would be false: the degree-p graph (u, u^p) has constant Gauss direction. With a nontrivial Kummer local system on its parameter, its inverse plane transform is, up to an unramified factor, Kummer in $r_1^p + r_2$; this is tame on a generic ray at infinity. Our bounded-degree construction excludes this phenomenon for sufficiently large p.

Nor does nonzero wild inertia at radial zero, together with tameness at radial infinity, suffice to exclude an origin line. The sheaf $\mathrm{Kl}_2(1/(r_1 + r_2))$ has those properties, but is a pullback from one linear form and hence has Fourier support on a line. Its local coefficients are proportional to $(1+z)^{-1/2}$. The actual coefficients (6.6) are excluded by the more specific rational-square calculation in Lemma 7.2. These examples explain why both parts of the argument are necessary.

8 From geometric support to the prime Fourier bound

Fix a nonempty subset of the four core factors, of size w, where $1 \le w \le 4$, and let $\mathcal{G}$ be the resulting pure lisse sheaf on the common open set U. Set

$$P = j_{!*}(\mathcal{G}[2]), \qquad j : U \hookrightarrow \mathbf{A}^2. \tag{8.1}$$

Then P is pure perverse of weight $w + 2$, and $\mathrm{FT}_2(P)$ is pure perverse of weight $w + 4$, by [12, Corollary 5.3.2] and [13, Theorem 2.2.1]. All simple constituents of P have full support.

Suppose first that the row and column indices are locally distinct. Corollary 6.1 excludes a punctual Fourier constituent, because its inverse Fourier transform would be a linear Artin–Schreier sheaf. Lemmas 7.1 and 7.2 exclude a curve-supported constituent. The uniform support-degree bound from quantitative sheaf theory permits the first lemma for all $p > p_0$, with one absolute p_0. Thus every simple constituent of $\mathrm{FT}_2(P)$ has full support. Geometric semisimplicity makes the whole transform a sum of full-support intermediate extensions.

8.1 Why the exceptional set is finite

Apply (4.3) to these full-support intermediate extensions. The only ordinary cohomology degrees are -2 and -1; the latter is supported on a finite set. A complex of weights at most v has ordinary degree-i stalks of weights at most $v + i$; this is the weight convention in [12, Section 5.1], also used explicitly in the proof of [14, Proposition 7.11]. Therefore the relevant bounds are

Complex	degree -2	degree -1
P, weight $w + 2$	$O(p^{w/2})$	$O(p^{(w+1)/2})$, on a finite set
$\mathrm{FT}_2(P)$, weight $w + 4$	$O(p^{(w+2)/2})$	$O(p^{(w+3)/2})$, on a finite set

Stalk dimensions are bounded by [14, Theorem 6.8]. For the finite support of $\mathcal{H}^{-1}$, first use [14, Proposition 6.24] to bound its complexity and then [14, Theorem 6.23] to bound the degree of its support. A reduced zero-dimensional set over $\overline{\mathbf{F}}_p$ has degree equal to its cardinality. Thus the finite-set cardinality is uniformly bounded, and we obtain

$$|t_{\mathrm{FT}_2(P)}(h, k)| \ll p^{(w+2)/2} + p^{(w+3)/2}\mathbf{1}_Z(h, k) \tag{8.2}$$

for a uniformly bounded finite geometric set Z.

The strict-support condition is the point that improves a curve to a finite set. At a generic boundary divisor, a nonzero degree--1 sheaf would constitute perverse degree zero on the boundary, forbidden by intermediate extension. At a boundary point, degree zero is forbidden for the same reason. Ordinary perversity alone would not justify (8.2) with finite Z.

If indices repeat locally, Section 6.3 still excludes every nonconstant linear Artin–Schreier constituent. The only permitted punctual Fourier constituent is therefore at the origin. Curve constituents may remain. A degree--1 contribution then has bounded-degree curve support and size $O(p^{(w+3)/2})$; a degree-zero contribution at the origin has size $O(p^{(w+4)/2})$. This gives precisely the alternative type of envelope needed in (3.4).

8.2 The physical boundary is a different issue

The trace of P on the complement of U need not equal zero. Our arithmetic product was extended by zero there, so we must compare the two before applying Fourier inversion. The complement has $O(p)$ rational points, and the first row of the preceding table gives the aggregate boundary estimate

$$\sum_{x\in(\mathbf{A}^2\setminus U)(\mathbf{F}_p)} |t_P(x)| \ll p\,p^{w/2} + p^{(w+1)/2} \ll p^{(w+2)/2}. \tag{8.3}$$

The larger degree--1 stalks occur at only $O(1)$ points. In particular, for $w=4$ it would be incorrect to assert that every physical boundary stalk has size $O(p^2)$; finitely many may have the weaker bound $O(p^{5/2})$. The aggregate $O(p^3)$ bound is sufficient and is the one actually used. Taking a Fourier transform cannot increase the absolute sum in (8.3).

8.3 Returning to the uncorrected correlation

On U, expand each raw factor as $C_p = \mathcal{C}_p - \kappa_p$. The product of four factors has fifteen nonempty pure subproducts and one empty product. Every coefficient in this expansion is bounded absolutely, since κ_p is bounded. Apply the appropriate pure estimate and (8.3) to each nonempty subproduct.

The empty product is $\mathbf{1}_U$, whose complete transform is

$$\widehat{\mathbf{1}_U}(h,k) = p^2\mathbf{1}_{(h,k)=(0,0)} + O(p). \tag{8.4}$$

It is multiplied by the bounded scalar κ_p^4 and fits within the generic $O(p^3)$ bound. Finally, the original raw four-product has size $O(p^2)$, by (5.6), at each of the $O(p)$ omitted torus points. Reinstating those points costs another $O(p^3)$ uniformly in frequency.

Taking the union of the fifteen bounded exceptional sets gives (3.3); retaining curves and the possible origin term gives (3.4). This proves Proposition 3.1, including its raw trace normalization and all boundary corrections.

8.4 A prime-level completion calculation

For an interval I of at most $R \le p$ integers, $\left\|\widehat{\mathbf{1}_I}\right\|_1 \ll p\log(2p)$, and each Fourier coefficient has absolute value at most R. Inverting the two-variable transform in (3.3) gives

$$\left|\sum_{r_1\in I, r_2\in J} W(r_1,r_2)\right| \ll p^3\log^2(2p) + R^2p^{3/2}. \tag{8.5}$$

The finite exceptional frequencies need not avoid either axis or the origin. Their contribution is bounded individually by R^2, and the inverse factor is p^{-2}. By contrast, a bounded-degree curve has weighted Fourier mass $O(pR\log(2p))$, leading to $Rp^{5/2}\log(2p)$. The next section carries out the corresponding calculation when different primes have different kinds of exception.

9 Squarefree completion with finite and curve exceptions

Let s be squarefree. At each $p \mid s$, choose a unit parameter α_p, fixed throughout the two-variable average, and form the product of the prime kernels (3.1). We denote it by $K_s(r_1, r_2)$. Nonunit row, column, and modulus-factor residues are set to zero. The indices m, m', n, n' are now integers, and each is in an interval containing $O(M)$ integers. The modulus-factor variables are in intervals containing $O(R)$ integers. Here $M, R \geq 1$; these intervals are allowed to contain repeated residues modulo s.

Constants of the form $C^{\nu(s)}$, with $\nu(s)$ the number of prime divisors, and fixed powers of divisor functions are written $s^{o(1)}$. Precisely, they are $O_\eta(s^\eta)$ for every fixed $\eta > 0$. Only finitely many such losses occur, so their exponents can be allocated in advance.

Lemma 9.1. *For an interval I containing at most R integers, a divisor $q \mid s$, and any residue a $\pmod q$,*

$$\sum_{\substack{h \bmod s \\ h \equiv a \pmod q}} \left| \sum_{r \in I} e_s(hr) \right| \ll R + \frac{s}{q} \log(2s). \tag{9.1}$$

The constant is uniform in the interval and the residue class.

Proof. For an interval shorter than a period, use the geometric-sum bound $\min(R, O(s/|h|))$, where $|h|$ is the least absolute residue. Frequencies in the indicated coset are spaced by q. The nearest to zero contributes at most R, and the remaining terms are bounded by a harmonic sum of size $O((s/q)\log(2s))$. If $R > s$, remove complete periods of the interval. They contribute only at zero frequency, with total at most R; the remainder is treated as before. $\square$

9.1 The local choices in the CRT expansion

Fix a four-cycle with $D_0 = (m - m')(n - n')$ and let $G = \gcd(s, D_0)$. At a prime not dividing D_0, its Fourier envelope has a generic part and a finite exceptional part. At a prime dividing D_0, it has a generic part, a curve part, and an origin part. Expand the product of these positive envelopes. A term is specified by three pairwise coprime divisors:

divisor	selected local condition	extra multiplier
f	finite set at $p \nmid D_0$	$\sqrt{f}$
d	curve at $p \mid D_0$	$\sqrt{d}$
e	origin at $p \mid D_0$	$e.$

They satisfy

$$fde \mid s, \qquad de \mid G. \tag{9.2}$$

Choices among finitely many exceptional points or curve components cost only a factor $s^{o(1)}$.

Split $d = d_v d_n$. At a prime in d_v, the selected component is a vertical line; at a prime in d_n, it has uniformly bounded vertical fibers. Such a split covers every component of a bounded-degree plane curve. The finite conditions at f and the origin conditions at e fix both coordinate residues, up to bounded choices. The conditions at d_v fix the first coordinate. For each permitted first coordinate, those at d_n give a bounded number of classes for the second coordinate. These second-coordinate classes may depend on the first; uniformity in Lemma 9.1 is precisely what permits that dependence.

Applying the coset bound first in the second coordinate and then in the first, the weighted Fourier mass for this term is at most

$$s^{o(1)} \left(R + \frac{sL}{fed_v} \right) \left(R + \frac{sL}{fed_n} \right), \qquad L \ll \log(2s). \tag{9.3}$$

There is no extra factor for all possible first-coordinate residues: they have already been summed by the first coset bound.

9.2 The four completion terms

The prime generic factors multiply to s^3, the inverse Fourier factor is s^{-2}, and the selected exceptions contribute $\sqrt{fd}\,e$. Thus multiply (9.3) by $s\sqrt{fd}\,e$. Ignoring logarithms, which enter $s^{o(1)}$, its four terms are

$$R^2 s\sqrt{fd}\,e,$$

$$Rs^2\sqrt{\frac{d_n}{fd_v}}, \qquad Rs^2\sqrt{\frac{d_v}{fd_n}},$$

$$\frac{s^3}{f^{3/2}\sqrt{d}\,e}.$$

The first is at most $R^2 s^{3/2}\sqrt{G}$, since $fde^2 \le sG$. Each mixed term is at most $Rs^2\sqrt{G}$, and the last is at most s^3. After summing the divisor choices we have proved

$$\left|\sum_{r_1\in I, r_2\in J} W_s(r_1,r_2)\right| \ll s^{o(1)}\left(s^3 + Rs^2\sqrt{\gcd(s,D_0)} + R^2 s^{3/2}\sqrt{\gcd(s,D_0)}\right). \tag{9.4}$$

This proof did not require either exceptional curves or exceptional points to avoid the coordinate axes. The finitely many small primes are absorbed into the absolute constant as explained after Proposition 3.1.

10 The averaged matrix norm

The relevant operator norm is the norm between the counting-measure ℓ^2 spaces on the row and column integer intervals. Setting nonunit indices to zero is equivalent to inserting orthogonal row and column projections; it causes no increase in this norm.

Proposition 10.1. *For the squarefree product kernel just defined,*

$$\left(R^{-2}\sum_{r_1,r_2}\|K_s(r_1,r_2)\|_{\mathrm{op}}^4\right)^{1/4} \ll s^{o(1)}\left(M^{3/4}s^{1/2} + Ms^{3/4}R^{-1/2} + Ms^{3/8}\right). \tag{10.1}$$

The result is uniform in the fixed local parameters, all interval positions, repeated residues, and $M, R \ge 1$. The case $s = 1$ is included.

Proof. For each matrix K, $\|K\|_{\mathrm{op}}^4 \le \mathrm{tr}((KK^*)^2)$. Expanding the trace gives exactly

$$\sum_{m,m',n,n'} K_{mn}\overline{K_{m'n}}K_{m'n'}\overline{K_{mn'}}.$$

First separate the cycles for which $m = m'$ or $n = n'$ as integers. There are $O(M^3)$ of them. The prime entry bound and CRT give $|K_{mn}| \ll s^{1/2+o(1)}$, so these cycles contribute at most $R^2M^3s^{2+o(1)}$.

For the remaining cycles both integer differences are nonzero, but a prime factor of s can still divide a difference. Such locally repeated indices are measured by the gcd in (9.4). For one interval,

$$\sum_{m\neq m'}\sqrt{\gcd(s,m-m')}\leq\sum_{d|s}\sqrt{d}\,\#\{m\neq m':d\mid m-m'\}$$
$$\ll M^2\sum_{d|s}d^{-1/2}\ll M^2s^{o(1)}.\tag{10.2}$$

To justify the middle bound without a spurious M error, a nonzero difference divisible by d requires d to be no larger than the interval's diameter. In that range the count is $O(M^2/d)$; outside it the count is zero. This also proves the bound when the interval contains several complete residue periods.

The elementary inequality

$$\sqrt{\gcd(s,(m-m')(n-n'))}\leq\sqrt{\gcd(s,m-m')}\sqrt{\gcd(s,n-n')}$$

allows us to use (10.2) independently in rows and columns. Summing (9.4) now yields

$$\sum_{r_1,r_2}\|K_s(r_1,r_2)\|_{\mathrm{op}}^4\ll s^{o(1)}\left(R^2M^3s^2+M^4s^3+RM^4s^2+R^2M^4s^{3/2}\right).\tag{10.3}$$

For all $R,s\geq 1$, $Rs^2\ll s^3+R^2s^{3/2}$: for example, compare R to $s^{3/4}$ and use the arithmetic–geometric mean inequality. Absorb this mixed term, divide by R^2, and take fourth roots. For $s=1$ the kernel is the all-ones matrix, whose norm is $O(M)$, already covered by the last term. $\square$

10.1 Where positivity permits enlargement

In the application the coefficient vectors can depend on both r_1 and r_2. This is allowed. For any family of matrices and vectors, Hölder gives

$$\sum_r|a_r^TK_r\overline{b_r}|\leq\left(\sum_r\|K_r\|_{\mathrm{op}}^4\right)^{1/4}\left(\sum_r(\|a_r\|_2\,\|b_r\|_2)^{4/3}\right)^{3/4}.\tag{10.4}$$

Here r can stand for a pair of modulus factors. One may enlarge a restricted sum of $\|K_r\|_{\mathrm{op}}^4$, since its terms are nonnegative, and then prove its bound by expanding the trace. One may not first expand a signed four-cycle and then discard restrictions in that signed sum. This order is what allows the original coprimality conditions and the variable coefficient vectors to coexist with Proposition 10.1.

11 The smooth Type III reduction and the three Fourier divisors

We turn to the convolution in Theorem 1.1. The reduction in this section is the smooth Poisson reduction of [2, Section 7.2]. We give its relevant details to identify the coefficients, the normalization, and the precise divisor weights that the new matrix estimate must handle.

Put $N=N_1N_2N_3$. Fix a dyadic modulus scale Q and then a narrower interval of relative length $O(x^{-\varepsilon})$, where $\varepsilon>0$ will be chosen smaller than the final strict savings. Write

$$H_i=Q/N_i,\qquad H=H_1H_2H_3=Q^3/N,\qquad H_+=x^{3\varepsilon/2}H.\tag{11.1}$$

The original absolute values in the discrepancy sum are represented by bounded complex coefficients η_q. They retain the restriction to the chosen modulus interval and to the permitted densely divisible moduli. We will keep them inside the signed sums.

11.1 Separating the smooth weights from the modulus

Let $\Psi_i(y) = \sum_n \psi_i(n)e(-ny)$. Smoothness and Poisson summation imply, on a fundamental period,

$$|\Psi_i^{(j)}(y)| \ll_{j,C} N_i^{1+j}(1 + N_i|y|)^{-C}.$$

Consequently frequencies $|h_i| > x^{\varepsilon/2}H_i$ have an arbitrarily small total contribution after choosing C large. For the remaining frequencies, Taylor expansion about h_i/Q separates the q-dependence:

$$\frac{h_i}{q} - \frac{h_i}{Q} \ll \frac{x^{-\varepsilon/2}}{N_i}.$$

Each derivative costs N_i, while each additional Taylor remainder factor saves $x^{-\varepsilon/2}$. A fixed sufficiently long expansion therefore has negligible remainder. Its terms have bounded weights in the individual frequencies and bounded weights in q. This is the separation in [2, Equation (7.13)]. No derivative of $\alpha(m)$ is taken.

For one such separated term, the complete finite sum is

$$\mathfrak{F}(\boldsymbol{h}, a/m; q) = \frac{1}{q} \sideset{}{^*}\sum_{n_1,n_2,n_3 \bmod q} e_q(h_1n_1 + h_2n_2 + h_3n_3)\mathbf{1}_{n_1n_2n_3=a/m \pmod q}. \tag{11.2}$$

The product of the three Fourier normalizations is H^{-1}; the complete sum in (11.2) was divided by q. Thus the remaining expression has prefactor Q/H, with the bounded ratio q/Q included in η_q. This is the normalization in [2, Equations (7.14)–(7.15)].

If an integer Fourier coordinate is zero, the complete sum is independent of the primitive class. For example, solving for n_3 when $h_3 = 0$ gives exactly

$$\mathfrak{F}((h_1, h_2, 0), a/m; q) = q^{-1}c_q(h_1)c_q(h_2), \tag{11.3}$$

where c_q is the Ramanujan sum. Such terms cancel when one subtracts the average over primitive classes in (1.1). This cancellation concerns a zero coordinate in the initial three-variable Fourier expansion; it is different from the central completion frequency studied in Section 14.

11.2 Removing primes that divide a nonzero Fourier coordinate

For $h_1h_2h_3 \neq 0$, factor uniquely

$$h_i = b_i\ell_i, \qquad b_i \mid q^\infty, \qquad \gcd(\ell_i, q) = 1. \tag{11.4}$$

The b_i are positive and can have arbitrary prime-power valuations; signs are assigned to ℓ_i. Define three different products:

$$B = b_1b_2b_3, \qquad b = \mathrm{rad}(B), \qquad A_b = \prod_{i=1}^{3} \mathrm{rad}(b_i). \tag{11.5}$$

Since q is squarefree, $q = bq_0$ with $\gcd(b, q_0) = 1$. Removing the factor b changes dense divisibility from x^δ to

$$y = bx^\delta; \tag{11.6}$$

this is [2, Lemma 2.10(i)]. It is not legitimate to keep the old width x^δ for the quotient.

CRT splits (11.2) into a factor modulo b and a Kloosterman sum modulo q_0. In the latter, all three Fourier coordinates are units, and the change of variables gives

$$\mathrm{Kl}_3\left(\frac{aB\ell_1\ell_2\ell_3}{b^3m}; q_0\right).$$

At a prime $p \mid b$, let $j \in \{1,2,3\}$ be the number of coordinates that vanish modulo p. The other local factor is exactly

$$\frac{(-1)^{3-j}(p-1)^{j-1}}{p}. \tag{11.7}$$

This follows by solving the product constraint for one of the zero coordinates and summing the remaining unit additive characters. It depends only on the zero pattern, not on m, the remaining modulus, or the primitive residue. Its absolute value is at most p^{j-2}. Multiplying over $p \mid b$ gives

$$|\text{local factor modulo } b| \le A_b/b^2. \tag{11.8}$$

Independence, not just this upper bound, permits us to take the factor outside the signed modulus and coefficient sum. This is the content of [2, Lemma 7.3(1)–(2)] needed here.

Group the remaining frequencies by $\ell = \ell_1\ell_2\ell_3$. For fixed ℓ, the number of signed factorizations is bounded by a constant times $\tau_3(|\ell|)$. Set

$$Q_0 = Q/b, \qquad H_b = H_+/B. \tag{11.9}$$

For a fixed triple $\boldsymbol{b} = (b_1, b_2, b_3)$, the quantity to estimate is

$$T(\boldsymbol{b}) = \sum_{0<|\ell|\le H_b} \tau_3(|\ell|) \left| \sum_{\substack{q_0 \asymp Q_0 \\ \gcd(b\ell, q_0)=1}} \eta_{bq_0} \sum_{\gcd(m, bq_0)=1} \alpha(m)\, \mathrm{Kl}_3\left(\frac{aB\ell}{b^3m}; q_0\right) \right|. \tag{11.10}$$

The coefficients still record the original modulus restrictions. If this expression is nonempty, $Q_0 \gg 1$ and $H_b \ge 1$. The units aB/b^3 modulo the remaining primes constitute a coherent family for fixed $\boldsymbol{b}$.

The resulting error bound is

$$|\Sigma_{\text{error}}(Q)| \ll x^{o(1)} \frac{Q}{H} \sum_{\boldsymbol{b}} \frac{A_b}{b^2} T(\boldsymbol{b}), \tag{11.11}$$

apart from the arbitrarily small smooth Fourier and Taylor tails. It is the reduction of [2, Equation (7.19)]. Its proof works with the original complex η's retained in (11.10); replacing a signed inner sum by an unweighted one is neither required nor used.

12 Dense-divisor extraction and the exact second moment

Choose one upper extraction target S, common to every Fourier triple, with

$$1 \le S \le x^{\delta} Q/2 = yQ_0/2. \tag{12.1}$$

Dense divisibility of q_0 permits a chosen factorization

$$q_0 = rs, \qquad S/y \le s \le S, \qquad \gcd(r, s) = 1. \tag{12.2}$$

If $q_0 < S \le yq_0$, take $s = q_0$, $r = 1$. The upper bound in (12.1) ensures that this is allowed for the entire modulus interval. A lower endpoint $S/y < 1$ does not create a difficulty: actual divisors are integers at least one. When we later replace a decreasing function of s by its value at S/y, using an endpoint below one merely overestimates it.

Choose one factorization for each original modulus and record it in $\eta_{r,s}$. There is no multiplicity loss from taking every possible factorization. Apply the triangle inequality over s in (11.10), followed by Cauchy's inequality with weights s^{-1} and s. The divisor second moment $\sum_{0<|\ell|\le H_b} \tau_3(|\ell|)^2 \ll H_b x^{o(1)}$ and $\sum_s s^{-1} \ll \log x$ give

$$T(\boldsymbol{b})^2 \ll x^{o(1)} H_b T_{2,\boldsymbol{b}}, \tag{12.3}$$

where

$$T_{2,\boldsymbol{b}} = \sum_{\gcd(b,s)=1} s \sum_{\substack{\ell\in\mathbf{Z}\\ \gcd(\ell,s)=1}} \psi_{H_b}(\ell) \left| \sum_{\substack{r\\ \gcd(b\ell s,r)=1}} \eta_{r,s} \sum_{\gcd(m,brs)=1} \alpha(m)\,\mathrm{Kl}_3(a^\dagger \ell/m; rs) \right|^2, \qquad a^\dagger = \frac{aB}{b^3}. \tag{12.4}$$

The s-sum retains the interval in (12.2). The smooth function ψ_{H_b} is nonnegative, equals one on $[-H_b, H_b]$, and is supported on a fixed dilation of that interval. The expression $a^\dagger$ means a unit residue at every remaining prime, not division at a prime dividing b.

This square is the point at which cancellation must be preserved. We will complete its ℓ-sum exactly, absorb separate CRT factors into coefficient vectors, and only then use operator norms. An entrywise triangle inequality before that step would lose the matrix average proved in Section 10.

12.1 The shared modulus and the completed period

Split the actual s-range into dyadic blocks $s \asymp V$. In a block put $R = Q_0/V$, so $r_1, r_2 \asymp R$, and write

$$g = \gcd(r_1, r_2), \qquad r_1 = gu, \quad r_2 = gv, \qquad w = sg, \qquad C = uvw. \tag{12.5}$$

On the original squarefree support, u, v, g, s are pairwise coprime. The two original moduli are uw and vw, and their least common multiple is C. Also $u, v \asymp R/g$ and $C \asymp R^2 s/g$.

For fixed m, n, u, v, s, g, the periodic product in the ℓ-sum is

$$G_{m,n}(\ell) = \mathbf{1}_{\gcd(\ell,C)=1}\,\mathrm{Kl}_3(a^\dagger \ell/m; uw)\overline{\mathrm{Kl}_3(a^\dagger \ell/n; vw)}. \tag{12.6}$$

With the sign conventions of (2.1), completion is the exact identity

$$\sum_{\ell} \psi_{H_b}(\ell) G_{m,n}(\ell) = \frac{1}{C} \sum_{c\in\mathbf{Z}} \widehat{\psi}_{H_b}(c/C) \widehat{G}_{m,n}(c). \tag{12.7}$$

The integer c is fixed in the ensuing (u,v)-average; it is not a separately selected representative for each changing period.

For any fixed $A > 1$, rapid decay gives a common positive envelope

$$C^{-1} |\widehat{\psi}_{H_b}(c/C)| \ll_A x^{o(1)} K^{-1} (1 + |c|/K)^{-A}, \qquad K \asymp \frac{R^2 s}{g H_b} > 0. \tag{12.8}$$

The constants in $\asymp$ are fixed across each dyadic block. The envelope is scalar in the matrix indices m, n. It is valid even if $K < 1$, so no fictitious frequency of size one is added.

12.2 Exact CRT factors and their coefficient dependence

Applying (2.7) and splitting the additive character, we obtain

$$\begin{aligned}\widehat{G}_{m,n}(c) = {} & T_u(a^\dagger m^{-1}w^{-3}, c(vw)^{-1}) \\ & \cdot \overline{T_v(a^\dagger n^{-1}w^{-3}, -c(uw)^{-1})} \\ & \cdot C_w(a^\dagger m^{-1}u^{-3}, a^\dagger n^{-1}v^{-3}, c(uv)^{-1}). \end{aligned} \tag{12.9}$$

All inverses are local to the modulus of the corresponding factor. The minus sign in the conjugated T_v makes its conjugate additive character have the required positive sign.

For fixed outer variables, the first factor depends on m, and the second depends on n. They can therefore be included in the row and column coefficient vectors. Formula (2.6) and the prime bound for Kl_2 imply $|T_u| \ll u^{1/2+o(1)}$, including at nonunit frequencies; the zero-frequency local value $-p^{-1}$ is smaller still. The resulting product of vector norms is

$$\|a_{u,v}\|_2 \, \|b_{u,v}\|_2 \ll x^{o(1)} M\sqrt{uv} \ll x^{o(1)} \frac{MR}{g}. \tag{12.10}$$

These vectors can depend on u, v, s, g, c. The estimate (10.4) explicitly allows that dependence. Unit conditions on their indices are retained as orthogonal projections.

13 Every nonzero completion frequency

A nonzero integer c need not be a unit modulo w. Put

$$e = \gcd(c, w), \qquad P = w/e. \tag{13.1}$$

We first handle the primes in e, where the local additive frequency is zero, and then apply the matrix theorem to the remaining primes in P.

13.1 The zero local frequencies are Schur multipliers

At a prime $p \mid e$, (2.5) turns the shared factor of (12.9) into

$$p\,\mathbf{1}_{mu^3 = nv^3 \pmod p} - \kappa_p. \tag{13.2}$$

The indicator is a block-matching mask on the matrix indices. To see its operator-norm bound, partition the row space by the residue mu^3 and the column space by nv^3, with projections P_a, Q_a. Entrywise multiplication by the mask sends a matrix X to $\sum_a P_a X Q_a$. The output blocks have orthogonal row and column spaces, so

$$\left\| \sum_a P_a X Q_a \right\|_{\mathrm{op}} = \max_a \|P_a X Q_a\|_{\mathrm{op}} \le \|X\|_{\mathrm{op}}.$$

Repeated residues cause larger blocks, not a larger norm. Consequently the Schur multiplier in (13.2) has norm at most $p + \kappa_p$. Multiplying over $p \mid e$ costs at most

$$\prod_{p\mid e}(p + \kappa_p) = e^{1+o(1)}. \tag{13.3}$$

This remains valid with all the original unit projections and with the matching pattern depending on u, v.

13.2 The unit local frequencies give a fixed-parameter matrix

At $p \mid P$, normalize the additive frequency of the shared correlation to one. Its entry becomes

$$C_p\left(\frac{\alpha_p v}{mu^2}, \frac{\alpha_p u}{nv^2}, 1\right), \qquad \alpha_p = \frac{a_p^\dagger}{c}(w/p)^{-2} \pmod{p}. \tag{13.4}$$

There are two sources of the cofactor in this expression. The local Kloosterman arguments contain an inverse cube by (2.7); rescaling the shared summation variable to normalize its additive frequency removes one inverse cofactor. The remaining inverse square is $(w/p)^{-2}$.

For fixed $\boldsymbol{b}, s, g, c$, the number α_p is independent of u, v. Coherence of the primitive classes is used here. The operator bound can therefore be applied at modulus P and outer length R/g, with arbitrary fixed unit parameters at its primes. No equidistribution assumption on u, v is needed. The case $P = 1$ is the all-ones matrix already included in Proposition 10.1.

Combining (13.3) with the three terms of (10.1), the residual powers of e, relative to putting $e = 1$, are respectively

$$e^{1/2}, \qquad e^{1/4}, \qquad e^{5/8}. \tag{13.5}$$

For example, the first term contains $eP^{1/2} = w^{1/2}e^{1/2}$; the other two are computed in the same way.

13.3 Summing nonunit frequencies without a short-scale loss

For $A > 1$, $K > 0$, and a positive integer e, monotonicity and integration give

$$\frac{1}{K}\sum_{\substack{c\neq 0\\ e\mid c}}(1+|c|/K)^{-A} = \frac{2}{K}\sum_{j\geq 1}(1+ej/K)^{-A} \leq \frac{2}{e(A-1)}. \tag{13.6}$$

There is no extra constant term on the right. The sum starts at the first nonzero multiple, so the decreasing integral bound remains valid even when $K < 1$. After bounding on each original class $\gcd(c, w) = e$, we may enlarge that class to $e \mid c$ in this positive majorant. For every fixed $\beta < 1$,

$$\frac{1}{K}\sum_{c\neq 0}(1+|c|/K)^{-A}\gcd(c,w)^{\beta} \ll_A \sum_{e\mid w} e^{\beta-1} \ll w^{o(1)}. \tag{13.7}$$

All three powers in (13.5) satisfy this condition. This controls every nonzero integer frequency, including those divisible by the whole shared modulus.

13.4 The three dyadic second-moment terms

We spell out the size calculation. In a block $s \asymp V$, there are $O(V)$ choices of s, each with the outer weight $s \ll V$. For fixed g, there are $O((R/g)^2)$ modulus-factor pairs. The vector-norm product is $O(x^{o(1)}MR/g)$. Before the averaged matrix term, their total contribution is therefore

$$x^{o(1)}\frac{MR^3V^2}{g^3}. \tag{13.8}$$

Apply (10.4), then the matrix theorem at $P = sg/e$, and use (13.7). The three resulting monomials are recorded below; $R = Q_0/V$.

Matrix term	residual e-power	g-power	after replacing R
$M^{3/4}P^{1/2}$	$1/2$	$-5/2$	$M^{7/4}Q_0^3V^{-1/2}$
$MP^{3/4}(R/g)^{-1/2}$	$1/4$	$-7/4$	$M^2Q_0^{5/2}V^{1/4}$
$MP^{3/8}$	$5/8$	$-21/8$	$M^2Q_0^3V^{-5/8}$

All three sums over g converge after choosing the auxiliary divisor exponents sufficiently small. Thus the nonzero-frequency part in this dyadic block satisfies

$$T_{2,\neq 0}(V) \ll x^{o(1)} \left(M^{7/4}Q_0^3V^{-1/2} + M^2Q_0^{5/2}V^{1/4} + M^2Q_0^3V^{-5/8}\right). \tag{13.9}$$

The use of Proposition 10.1 enlarges any necessary restrictions in a nonnegative norm fourth moment, before its four-cycle expansion. At no stage is the original signed cycle treated as monotone under enlargement. Also, no integer diagonal in m, n, u, v was removed from this nonzero-frequency matrix.

14 The central completion frequency

The frequency $c = 0$ is best evaluated directly. Substituting the zero values from (2.6) into (12.9) gives

$$\widehat{G}_{m,n}(0) = \frac{\mu(u)\mu(v)}{uv} C_w\left(\frac{a^\dagger}{mu^3}, \frac{a^\dagger}{nv^3}, 0\right). \tag{14.1}$$

Here $\mu(u)$ denotes the Möbius function. Let

$$D = mu^3 - nv^3. \tag{14.2}$$

At every $p \mid w$, the two correlation parameters agree exactly when $p \mid D$. Therefore (2.5) implies

$$|\widehat{G}_{m,n}(0)| \ll x^{o(1)} \frac{\gcd(w, D)}{uv}.$$

The central Poisson coefficient is $O(H_b/C)$. Multiplying by the outside factor s in (12.4), a single term is bounded by

$$x^{o(1)} \frac{H_b}{g} \frac{\gcd(sg, D)}{u^2v^2}. \tag{14.3}$$

This is a nonnegative bound, so the ensuing divisor sums can be enlarged after the dyadic ranges have been fixed.

14.1 Nonzero integer differences

For $D \neq 0$, use $\gcd(sg, D) \leq g \gcd(s, D)$. For fixed u, v, m, n, the sum over the actual s-block is

$$\sum_{s \asymp V} \gcd(s, D) \leq \sum_{d \| D|} d\, \#\{s \asymp V : d \mid s\} \ll V\tau(|D|). \tag{14.4}$$

Only $d \ll V$ can divide an integer in this block. Thus $\#\{s \asymp V : d \mid s\} \ll V/d$ for every nonzero term, and there is no separate large-divisor error. The integer D is polynomially bounded in x, so its divisor count is part of $x^{o(1)}$.

Let $L = R/g$. The fixed dyadic u, v ranges have $\sum (uv)^{-2} \ll L^{-2}$, and there are $O(M^2)$ coefficient-index pairs. Summing (14.3) and (14.4), the contribution for one g is $O(x^{o(1)}H_bM^2V(g/R)^2)$. Summing $g \ll R$ gives

$$O(x^{o(1)}H_bM^2VR) = O(x^{o(1)}H_bM^2Q_0). \tag{14.5}$$

14.2 The exact integer diagonal

Now coprimality of u, v gives

$$m = v^3 k, \qquad n = u^3 k \tag{14.6}$$

for a positive integer k. For fixed $u, v \asymp L$, there are at most $O(M/L^3)$ possibilities. This bound does not need an added one: $k \geq 1$ and its upper bound is $O(M/L^3)$, so if that upper bound is below one there is no term.

For $D = 0$, (14.3) becomes $O(x^{o(1)} H_b s/(u^2 v^2))$. Summing the $O(L^2)$ pairs u, v, the at most $O(M/L^3)$ choices of k, and then $s \asymp V$, gives $O(x^{o(1)} H_b M V^2 L^{-5})$. Finally $\sum_{g \ll R} (g/R)^5 \ll R$, so this part is

$$O(x^{o(1)} H_b M V^2 R) = O(x^{o(1)} H_b M Q_0 V). \tag{14.7}$$

This is an integer equality, distinct from both the local repeated indices in Proposition 3.1 and the local matching masks at a nonzero integer completion frequency.

Combining the two cases and summing dyadic blocks gives

$$T_{2,0} \ll x^{o(1)} (H_b M Q_0 S + H_b M^2 Q_0). \tag{14.8}$$

The proof uses the exact inverse cubes in CRT.

15 Restoring the extraction width and summing Fourier divisors

For the dyadic s-blocks in (12.2), $S/y \ll V \ll S$. In (13.9), the first and third terms decrease with V; the second increases. Replacing by their appropriate endpoints gives

$$\begin{aligned} T_{2,\neq 0}(\boldsymbol{b}) \ll x^{o(1)} (&M^{7/4} Q_0^3 y^{1/2} S^{-1/2} + M^2 Q_0^{5/2} S^{1/4} \\ &+ M^2 Q_0^3 y^{5/8} S^{-5/8}). \end{aligned} \tag{15.1}$$

The two central terms in (14.8) have no width loss: one is independent of V, and the other is maximized at S. All dyadic counting factors are logarithmic.

Insert $Q_0 = Q/b$, $y = bx^\delta$, and $H_b = x^{3\varepsilon/2} H/B$. Take the square root in (12.3) and multiply by A_b/b^2 from (11.11). The exact divisor weights are

Second-moment term	width before square root	remaining triple weight
$M^{7/4} Q_0^3 S^{-1/2}$	$y^{1/2}$	$A_b/(B^{1/2} b^{13/4})$
$M^2 Q_0^{5/2} S^{1/4}$	1	$A_b/(B^{1/2} b^{13/4})$
$M^2 Q_0^3 S^{-5/8}$	$y^{5/8}$	$A_b/(B^{1/2} b^{51/16})$
Either central term	1	$A_b/(B b^{5/2})$

For example, the first noncentral term has $b^{-3} b^{1/2} = b^{-5/2}$ before its square root, giving $b^{-2} b^{-5/4} = b^{-13/4}$ afterwards. The factor $B^{-1/2}$ comes from $H_b^{1/2}$; the central terms contain one further $H_b^{1/2}$, giving B^{-1}.

15.1 An Euler-product check with arbitrary valuations

For general positive exponents u, v, the local Euler factor of $\sum_{b_1, b_2, b_3} A_b/(B^u b^v)$ is

$$1 + p^{-v} \left[\left(1 + \frac{p^{1-u}}{1 - p^{-u}} \right)^3 - 1 \right]. \tag{15.2}$$

Indeed, a nonzero valuation $a \geq 1$ in one b_i contributes p^{1-ua}. Sum its geometric series, take the product over the three coordinates, subtract the case in which all valuations are zero, and include p^{-v} once for the radical b. For $u = 1/2$, the bracket is $O(p^{3/2})$; convergence holds for every $v > 5/2$. For $u = 1$, the bracket is bounded, and $v > 1$ suffices. All four weights in the table lie strictly inside these ranges.

This check also explains the uniform treatment of divisor losses. Every original nonempty triple has $b \ll Q$ and $B \ll H_+$. Its arithmetic quantities are polynomially bounded in x, so all auxiliary divisor losses are bounded by one predetermined x^η before extending the positive series to infinite triples. The strict convergence margins absorb sufficiently small additional exponents. No infinite sum with a triple-dependent unspecified constant is being used.

15.2 The five discrepancy errors

Combining (11.11), (12.3), (14.8), and (15.1), then summing the positive triple weights, gives

$$\begin{aligned} |\Sigma_{\mathrm{error}}(Q)| \ll x^{o(1)} \Big\{ & x^{3\varepsilon/2} \left(M^{1/2} Q^{3/2} S^{1/2} + M Q^{3/2} \right) \\ & + x^{3\varepsilon/4} H^{-1/2} (x^{\delta/4} M^{7/8} Q^{5/2} S^{-1/4} + M Q^{9/4} S^{1/8} \\ & + x^{5\delta/16} M Q^{5/2} S^{-5/16}) \Big\}. \end{aligned} \tag{15.3}$$

The first line is the central frequency. The other three terms are the three terms of the averaged matrix norm. In particular there is no fourth noncentral term arising from a Fourier exceptional curve at every prime: curves remain only at primes dividing an index difference, and their gcd average was already included in (10.3).

Using $MN \asymp x$ and $H = Q^3/N$, divide by x:

$$\begin{aligned} \frac{|\Sigma_{\mathrm{error}}(Q)|}{x} \ll x^{o(1)} \Big\{ & \frac{x^{3\varepsilon/2}}{x} \left(M^{1/2} Q^{3/2} S^{1/2} + M Q^{3/2} \right) \\ & + \frac{x^{3\varepsilon/4}}{x^{1/2}} (x^{\delta/4} M^{3/8} Q S^{-1/4} + M^{1/2} Q^{3/4} S^{1/8} \\ & + x^{5\delta/16} M^{1/2} Q S^{-5/16}) \Big\}. \end{aligned} \tag{15.4}$$

This explicit bound is the analytic output needed for the theorem. The remaining step is to make every displayed power of x strictly negative with one admissible extraction target.

16 Choosing the extraction target and proving the parameter range

Set the fixed upper exponents

$$\theta = \frac{1}{2} + 2\omega, \qquad \mu = \frac{1}{4} - \frac{3}{2}\sigma. \tag{16.1}$$

The three pair-product hypotheses imply

$$(N_1 N_2 N_3)^2 \gg x^{3/2+3\sigma}, \qquad M \ll x^\mu.$$

Also $\mu > 0$ because $\sigma < 1/6$. We will choose $S = x^s$ using these upper exponents, rather than reoptimizing S at each smaller modulus block.

At $Q = x^\theta$, $M = x^\mu$, and temporarily omitting the small ε and divisor losses, the five exponents in (15.4) are

$$\begin{aligned} d &= \frac{\mu}{2} + \frac{3\theta}{2} + \frac{s}{2} - 1, & j &= \mu + \frac{3\theta}{2} - 1, \\ o_1 &= \frac{3\mu}{8} + \theta - \frac{s}{4} - \frac{1}{2} + \frac{\delta}{4}, & o_2 &= \frac{\mu}{2} + \frac{3\theta}{4} + \frac{s}{8} - \frac{1}{2}, \\ o_3 &= \frac{\mu}{2} + \theta - \frac{5s}{16} - \frac{1}{2} + \frac{5\delta}{16}. \end{aligned} \tag{16.2}$$

Here d, j denote the two central terms, and o_1, o_2, o_3 the three noncentral terms in their displayed order. The explicit H_+ cutoff adds $3\varepsilon/2$ to d, j and $3\varepsilon/4$ to each o_i.

The central diagonal exponent d increases with s, while o_1 decreases. Balance them by taking

$$s = \frac{4 - \mu - 4\theta + 2\delta}{6}. \tag{16.3}$$

Then

$$d = o_1 = \frac{5\mu + 14\theta + 2\delta - 8}{12}. \tag{16.4}$$

It remains to check the other three errors and the validity of this choice of S; a balance of two terms alone is not sufficient.

Define the two strict slacks

$$A_* = 8 - 5\mu - 14\theta - 2\delta, \qquad B_* = 27\theta + \delta - 14. \tag{16.5}$$

Direct substitution of (16.3) gives the exact identities

$$\begin{aligned} d = o_1 &= -\frac{A_*}{12}, \\ j &= -\frac{A_*}{5} - \frac{20 + 13B_* + 95\delta}{270}, \\ o_2 &= -\frac{23A_*}{240} - \frac{B_*}{40} - \frac{\delta}{8}, \\ o_3 &= -\frac{53A_*}{480} - \frac{B_*}{80}. \end{aligned} \tag{16.6}$$

Thus $A_* > 0$, $B_* > 0$, and $\delta \geq 0$ make all five errors strictly negative. These are identities throughout the parameter region, not evaluations at a boundary point.

16.1 The target is positive and available at every needed scale

The same variables give

$$\begin{aligned} \theta - \frac{1}{2} &= \frac{\mu}{8} + \frac{A_* + 2B_*}{40}, \\ \frac{1}{2} - s &= \frac{11\mu}{8} + \frac{29A_*}{120} + \frac{3B_*}{20}. \end{aligned} \tag{16.7}$$

In particular $s < 1/2$. Since $A_* > 0$ implies $\theta < 4/7$, substitution of its upper bound for μ in (16.3) also gives

$$s > \frac{2 - \theta + 2\delta}{5} > 0. \tag{16.8}$$

Hence $1 < S < x^{1/2}$ by a fixed positive power margin.

Keep this same S as the actual Q decreases. With S fixed, every monomial in (15.4) is increasing in both M and Q. The savings proved at the upper values persist at all smaller values. Also, for any fixed $B_0 > 0$,

$$Q \gg x^{1/2}(\log x)^{-B_0} \quad \Longrightarrow \quad S \le x^{\delta}Q/2$$

for large x, by $s < 1/2$. Thus the same extraction target is valid throughout the range where the Type III argument is needed. There is no requirement that $B_* > 0$ hold anew when θ is replaced by the logarithm of each smaller Q; it was used once to choose a target and prove a uniform worst-case saving.

Finally, substituting (16.1) into $A_* > 0$ and $B_* > 0$ gives respectively

$$\sigma > \frac{1}{30} + \frac{56}{15}\omega + \frac{4}{15}\delta, \qquad 108\omega + 2\delta > 1.$$

These are exactly (1.2). They also imply $\omega < 1/28 < 1/12$, so the original upper parameter restriction in the smooth Type III reduction is satisfied.

17 Small moduli, class averaging, and completion of the theorem

The individual scale assumptions finish the source reduction uniformly. The upper bounds $N_i \ll x^{1/2-\sigma}$ imply $H_i = Q/N_i \gg x^{\sigma}$, up to logarithms at the bottom of the large-modulus range. The lower bounds $N_i \gg x^{2\sigma}$ ensure $x^{\varepsilon/2}H_i < q/2$ for sufficiently small ε and large x. Thus the truncated Fourier coordinates used in Section 11 lie in a single period, with the needed positive-power separation of scales.

For Q below $x^{1/2}$ by a suitable logarithmic factor, apply the convolution Bombieri–Vinogradov theorem [3, Theorem 0(b)], in the coefficient-sequence formulation of [2, Theorem 2.9], with

$$\beta = \psi_1, \qquad \widetilde{\alpha} = \alpha * \psi_2 * \psi_3.$$

The latter is divisor bounded at its product scale. The former is smooth of polynomially large length and has the required Siegel–Walfisz property, by the smooth-sequence argument in the proof of [2, Lemma 3.4(iii)]. Both factor lengths are bounded below by a fixed power of x, as follows from (1.3). This supplies the short-modulus estimate without imposing a separate Siegel–Walfisz condition on the arbitrary sequence α.

To recover the discrepancy rather than an uncentered progression sum, fix the modulus coefficients η_q and average the coherent primitive family over the choices of its a_p's. On every fixed modulus this gives the uniform average over primitive classes. All contributions in (11.3), and hence the class-independent main term of the smooth reduction, are unchanged by the average. They cancel in the difference. The error estimate is uniform in the coherent family and in the bounded coefficients, so the same bound, up to an absolute factor, applies to the discrepancy sum. The coefficients can then be those chosen to represent its absolute values.

Let $\kappa > 0$ be a common lower bound for the five strict savings in (16.6). Choose ε, the sum of all auxiliary divisor exponents, and the Fourier-tail budgets sufficiently small relative to κ. The factors $x^{3\varepsilon/2}$, $x^{3\varepsilon/4}$, and the $O(x^{\varepsilon}\log x)$ narrow modulus intervals then consume only a proper fraction of that saving. We obtain a fixed power saving in the large-modulus range. Such a saving is stronger than every fixed negative power of $\log x$. Together with the short-modulus estimate, this proves Theorem 1.1.

18 How to use the theorem in a prime-gap argument

The analytic input is often described by an upper cutoff c for each of three long factors. Put $\sigma = 1/2 - c$. Then the individual Type III assumptions read

$$x^{1-2c} \ll N_i \ll x^c, \qquad N_i N_j \gg x^{1-c},$$

and Theorem 1.1 allows the strict cutoff

$$c < \frac{7}{15} - \frac{56}{15}\omega - \frac{4}{15}\delta, \tag{18.1}$$

in its stated parameter region. In particular an argument invoking this cutoff must still exhibit the three individual factors and their pair-product bounds. A statement only about the remaining short factor is insufficient.

The separate incidence Type II estimate in the companion *Incidence operators and intervals containing many primes* [15, Theorem 2.1 and Corollaries 9.2–9.3] has the lower cutoff

$$a > \frac{3}{10} + 8\omega + \frac{16}{5}\delta.$$

These two cutoff walls have overlap $a < c$ when

$$\theta = \frac{1}{2} + 2\omega < \frac{93 - 104\delta}{176}. \tag{18.2}$$

This calculation identifies the limiting overlap of these two particular estimates; the minorant construction also checks its own combinatorial and coefficient hypotheses. With those additional arguments, the strict point used in [15, Theorem 1.1] gives $H_m \ll \exp(3.7877m)$, with an unevaluated implied constant. Here p_n is the n-th prime and $H_m = \liminf_{n\to\infty}(p_{n+m} - p_n)$.

Neither the distribution theorem proved here nor that asymptotic many-prime statement alone gives a numerical value for H_1. The explicit bounded-gap application additionally uses the prime minorant, the fragment-sieve support and restoration inequalities, and the numerical integral certificate in the main article. The geometric and analytic proofs above are ordinary mathematical proofs; finite computational checks may test their formulas, but no such check is substituted for a sheaf theorem or for a uniform estimate.

A Scope of the Lean formalization

The public development is `romyar1/prime-gaps-lean` [16], at commit `9c1ab6648bcd` for the source links in this appendix. The target is defined in `formal/TypeIIILocal.lean`; the supporting entry is `formal/TypeIIILocalProgress.lean`.

The Lean formalization separates the two Fourier inequalities in Proposition 3.1 from their analytic consequences. In its notation, `TypeIII.LocalFourierHypothesis C D p0` asserts the following uniform statement for the actual finite sums: one fixed C, one fixed exceptional-set or curve-degree bound D, and one fixed small-prime cutoff work for all primes above that cutoff and all unit parameters. Its distinct-row-and-column branch has a finite exceptional set; its repeated-index branch has a curve and the possible origin term. These are the quantifiers and branches of (3.3)–(3.4). The downstream Lean theorems also require the separate assumption $0 \le C$.

The Lean local-to-global Type III argument proves the distribution estimate from that premise and the stated standard local bounds. The final prime-gap theorem is likewise an implication: in addition to this local premise, it requires the explicit numerical source bounds, the two public rank-three/correlation inputs [7, Theorem 4.1.1], [8, Proposition 2], in the form recorded in [6, Lemma A.4], and the scalar rank-four Kloosterman input [7, Theorem 4.1.1] used by the incidence companion. No global distribution theorem is substituted for the local premise. However, proving an implication in Lean, even with only its usual logical axioms, does not prove its explicit hypotheses.

The formalized algebraic components include the nonvanishing and rational-square obstructions for every nonempty rectangle subproduct, with all independent root choices; injectivity of the pole-one Artin–Schreier coefficient in the actual Laurent-series quotient; and constant-field descent under the actual field automorphism group. For the last statement, let $K = \overline{\overline{\mathbf{F}}_p(z)}$. An element of K has finite orbit under all $\overline{\mathbf{F}}_p$-automorphisms precisely when it belongs to $\overline{\mathbf{F}}_p$. Indeed, a nonconstant element is transcendental, its translates over $\overline{\mathbf{F}}_p$ give infinitely many conjugates by extension of rational-field isomorphisms to algebraic closures, and an element algebraic over the algebraically closed field $\overline{\mathbf{F}}_p$ is constant. This gives a second formulation of the scalar descent used in Lemma 7.2: once the actual one-variable sheaf gives a finite invariant list of pole-one characters, their coefficients must be constant.

The public development also contains conditional geometric deductions beyond this algebra. Their source map and precise imported laws are given in `research/type_iii/published_inputs/README.md`. The combined theorem in `formal/TypeIIIPublishedTypeIII.lean` derives the original uniform Fourier statement from general geometric rules and data realizing the actual family. The application bridge is `formal/TypeIIIPublishedApplicationBridge.lean`. Within this framework, coefficient covariance reduces proper Fourier supports to lines through the origin and the origin itself; the phase calculation excludes these supports in the distinct-index case. Strict support and weight bounds, followed by the corrected-product expansion, give the two Fourier envelopes.

The physical construction conditionally derives the rank-six image, purity and corrected trace from cohomological laws and the arithmetic local model. The local construction conditionally derives the phase list from Mackey laws, Fu's formula and finite-origin rank exhaustion. The source recipe and its uniform complexity, the identification of the local representations with the same physical family, and the compatibility of the geometric and arithmetic interfaces remain explicit application data. Thus the posted conditional derivation does not discharge the local Fourier premise in the prime-gap theorem. Sections 4–8 give the ordinary mathematical proof used in this paper.

Public sources and reproduction

Obtain the public revision cited in [16]. With Elan, Git and Python 3.11 or later installed, run from the repository root:

```
lake exe cache get
./verify.sh
```

The pinned toolchain is Lean `4.34.0-rc2`. The verifier builds every formal module, including the supporting Type III entry, and checks source integrity, theorem types and axiom dependencies. `docs/verification.md` gives the exact dependency pins and the fresh-build option `./verify.sh -fresh`; `docs/type-iii.md` gives the detailed code map.

The separate numerical release linked in `research/numerical_182/README.md` reports a fresh external run of all 262 integral-bound comparisons for the fixed 182 trial, with 14 implementation audits and four deliberate rejection tests. Its integrity and regeneration commands concern the numerical inputs to the prime-gap application. The Type III premise is a separate mathematical input, as specified in `docs/assumptions.md`.

References

[1] Y. Zhang, *Bounded gaps between primes*, Annals of Mathematics **179** (2014), 1121–1174. doi:10.4007/annals.2014.179.3.7.

[2] D. H. J. Polymath, *New equidistribution estimates of Zhang type*, Algebra & Number Theory **8** (2014), 2067–2199. arXiv:1402.0811v3.

[3] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Mathematica **156** (1986), 203–251. doi:10.1007/BF02399204.

[4] J. Stadlmann, *Bounded gaps between primes*, preprint, 31 August 2026. arXiv:2608.31126v1.

[5] F. Charton, L. Hong, K. Lau, K. Ono, G. Remy, H. C. Siu, A. A. Swaminathan, J. Thorner, and Y. Xie, *A new bound for small gaps between primes*, preprint, 3 September 2026. Axiom Math text.

[6] OpenAI, *Improved short gaps between primes*, preprint, 30 August 2026. Article; supporting repository.

[7] N. M. Katz, *Gauss Sums, Kloosterman Sums, and Monodromy Groups*, Annals of Mathematics Studies **116**, Princeton University Press, 1988. Author-hosted text.

[8] É. Fouvry, E. Kowalski, and Ph. Michel, *The Friedlander–Iwaniec character sum*, preprint, 14 June 2013. Author-hosted text.

[9] P. Deligne, *Cohomologie étale (SGA $4\frac{1}{2}$)*, Lecture Notes in Mathematics **569**, Springer, 1977. Author-hosted text.

[10] L. Fu, *Calculation of ℓ-adic local Fourier transformations*, Manuscripta Mathematica **133** (2010), 409–464. arXiv:math/0702436v5.

[11] G. Laumon, *Transformation de Fourier, constantes d'équations fonctionnelles et conjecture de Weil*, Publications Mathématiques de l'IHÉS **65** (1987), 131–210. Numdam text.

[12] A. A. Beilinson, J. Bernstein, and P. Deligne, *Faisceaux pervers*, Astérisque **100** (1982), 5–171. Author-hosted text.

[13] N. M. Katz and G. Laumon, *Transformation de Fourier et majoration de sommes exponentielles*, Publications Mathématiques de l'IHÉS **62** (1985), 145–202. Author-hosted text.

[14] W. Sawin, mis en forme par A. Forey, J. Fresán, and E. Kowalski, *Quantitative sheaf theory*, Journal of the American Mathematical Society **36** (2023), 653–726. arXiv:2101.00635v4.

[15] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Incidence operators and intervals containing many primes*, accompanying article, 2026.

[16] GPT-6 Astra (OpenAI), prompted by Romyar Sharifi, *Prime gaps in Lean*, source code, numerical certificates and reproduction instructions, 2026. Public repository, commit `9c1ab6648bcd` (accessed 11 September 2026). External numerical release `external-numerics-20260909`.