A NEW BOUND FOR SMALL GAPS BETWEEN PRIMES

FRANÇOIS CHARTON, LETONG HONG, KENNY LAU, KEN ONO, GUILLAUME REMY, HO CHUNG SIU, ASHVIN A. SWAMINATHAN, JESSE THORNER, AND YUNZHOU XIE

ABSTRACT. Stadlmann recently proved that infinitely many pairs of distinct primes are apart by at most 240, improving upon the longstanding bound 246 established by Polymath8b. Building on her work, we prove that

$$\liminf_{n\to\infty}(p_{n+1}-p_n) \le 212.$$

1. INTRODUCTION

The *twin prime conjecture* asserts that there are infinitely many pairs of prime numbers that differ by 2. In mathematical terms, this can be reformulated as follows: if p_n denotes the n-th prime and

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

then $H_1 = 2$. As an approximation to the twin prime conjecture, one might hope to simply prove that H_1 is finite. All current approaches to bounding H_1 originate with the ideas of Goldston, Pintz, and Yıldırım (GPY) [GPY09], who proved that

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

Their method relates the study of gaps between primes to the distribution of primes in arithmetic progressions. Recall that the primes are said to have *level of distribution* θ if, for every $A > 0$, we have

$$\sum_{q\le x^\theta}\max_{(a,q)=1}\left|\pi(x;q,a)-\frac{\pi(x)}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A}. \tag{1.1}$$

The Bombieri–Vinogradov theorem states that (1.1) holds for every $\theta < 1/2$. If (1.1) were to hold for any $\theta > 1/2$, then the GPY method would show that H_1 is finite.

Zhang [Zha14] proved a variant of (1.1) with $\theta > 1/2$ that, when combined with a modification of the GPY method, established the finiteness of H_1 for the first time. In fact, his method yields

$$H_1 \le 70{,}000{,}000.$$

The Polymath8a project [Pol14a] sharpened Zhang's work, which resulted in $H_1 \le 4680$. Maynard[1] [May15] introduced a multidimensional refinement of the GPY sieve which only requires (1.1) with $\theta > 0$. Therefore, using the Bombieri–Vinogradov theorem, Maynard obtained $H_1 \le 600$. The Polymath8b project [Pol14b] optimized this method to obtain

$$H_1 \le 246.$$

The bound 246 stood for over a decade, and there is a structural reason for this stemming from (1.1). In the multidimensional sieve, one chooses a divisor weight supported on a region $T \subset [0,1]^k$. One notes that squaring that weight produces moduli whose logarithmic sizes are governed by pairs of points of T. The Polymath8b argument arranges T so that every such modulus stays below $x^{1/2-c}$ for a fixed $c > 0$, where Bombieri–Vinogradov is available. Enlarging T improves the variational problem attached to the sieve, but it also produces larger moduli, for which no unconditional distribution estimate is known. Every improvement in the bound for H_1 amounts to a more efficient negotiation of this exchange.

Stadlmann [Sta26] broke the barrier by changing the shape of T rather than simply enlarging it. Her support admits a thin region corresponding to moduli beyond the Bombieri–Vinogradov range, but it controls the total mass carried by coordinates larger than a fixed threshold δ. The effect on the moduli is summarized as follows: The large coordinates record a bounded number of *rough* factors of size at least x^δ, while the

Date: Preliminary Draft: September 3, 2026.

[1]Independently, and at around the same time, Tao arrived at slightly weaker results.

remaining coordinates contribute a factor that is x^δ-smooth. The rough factors can then be assigned to the divisors demanded by the Type I, Type II, and Type III estimates, and the smooth factor can be subdivided to bring each of those divisors to the required scale. The enlargement is therefore confined to those parts of T whose corresponding moduli, although larger than $x^{1/2}$, still factor in one of the ways these estimates require. Combining this with variants of the Polymath8a, Baker–Irving, and her own equidistribution estimates [Pol14a, BI17, Sta25], Stadlmann proved $H_1 \leq 240$ using an admissible 49-tuple.

Stadlmann's results depend on complicated numerical calculations involving delicate choices of parameter. On the basis of extensive numerical experimentation run on a distributed computing cluster, we optimized Stadlmann's methodology and proved the following theorem.

Theorem 1.1. *We have $H_1 \leq 212$.*

The analytic ideas behind Theorem 1.1 originated from Stadlmann. The mixed support, the relaxed Type I/II/III estimates, and the rough-smooth factor-extraction argument all originate in her work, which in turn rests on Maynard, Polymath8a, Polymath8b, Baker–Irving, and her earlier estimates for primes in arithmetic progressions [May15, Pol14a, Pol14b, BI17, Sta25]. To prove Theorem 1.1, we sharpen three parts of her argument.

First, the analytic estimates are stated in terms of the divisors that their proofs actually use, rather than in terms of a uniform sufficient condition on the modulus. This reformulation is Stadlmann's, and we carry it through for each of her five estimates [Sta26, Lemmas 3–7]. For example, the Type IIc case in her Lemma 6 (which relies on Theorem 1 of [Sta25]) becomes our Lemma 5.6, in which the three divisor intervals for r, u, and d_1 remain unsimplified instead of weakening to a convenient, uniform condition. Second, the resulting factorization conditions are verified over the whole continuous range of rough-factor configurations and of modulus sizes, instead of at a worst-case configuration. Third, the argument for moduli near $x^{1/2}$ is made explicit. We divide the moduli by size into four ranges: those well below $x^{1/2}$, those below it by only a power of $\log x$, those at the level $x^{1/2}$ itself, and those above it. The first range is treated by the Bombieri–Vinogradov theorem and the last by Stadlmann's key estimates. For the third we use the factorization that the support provides exactly at $x^{1/2}$, and for the second we show that this same factorization survives a small move inward, so that no range is left uncovered. In Appendix A, we record, statement by statement, which parts of [Sta26] are used unchanged, which are specialized, and which are strengthened.

The proof separates into an arithmetic statement and a variational statement, which are proved independently. On the arithmetic side, a support datum $(\delta, \underline{A}, \underline{B}, \varepsilon)$ determines the family of moduli that can occur when the sieve square is expanded, and the factor-extraction lemmas reduce the treatment of that family to finitely many inequalities among the rough factors. All parameters are rational, and the inequalities are verified exactly. For the reader's convenience, they are listed in Table 3 and Appendix B. On the variational side, the chosen support leads to two rational quadratic forms. After a dilation of the support by a factor of four, the sieve criterion becomes the critical requirement that the associated Rayleigh quotient exceed 4. For the polynomial at which the quadratic forms are evaluated, exact rational arithmetic gives

$$\frac{v^\mathsf{T} J_T v}{v^\mathsf{T} I_T v} = 4.00438409833460131937\ldots > 4.$$

Finally, we note that the sieve criterion proved in Section 4 requires just two results: an equidistribution theorem for the moduli generated by the support, and a strict variational inequality on that support. These results are independent of each other.

Section 3 outlines the proof. Section 4 proves the sieve criterion. Section 5 states the five equidistribution estimates and Section 6 proves them; Section 7 shows that the support supplies the factorizations they require. Section 8 treats the passage through $x^{1/2}$. The exact parameters and the variational problem are given in Section 9, and Sections 10 and 11 evaluate the variational integrals and verify the required inequality. In Section 12, we finish the proof of Theorem 1.1.

2. Notation

Throughout, x denotes a large real parameter, and all asymptotic notation refers to the limit $x \to \infty$. The letter p always denotes a prime, $\mathbb{P}$ the set of primes $1_\mathbb{P}$ the characteristic function of $\mathbb{P}$, and $\pi(x; q, a)$ the count of the primes $p \le x$ with $p \equiv a \pmod q$. We write φ for the Euler totient function, μ for the Möbius function, τ for the divisor function, $\omega(n)$ for the number of distinct prime factors of n, and

$$P(z) = \prod_{p \le z} p, \qquad \log_x d = \frac{\log d}{\log x}.$$

An integer is called x^δ*-smooth* if all of its prime factors are smaller than x^δ. For integers a and b, (a, b) denotes their greatest common divisor and $[a, b]$ their least common multiple. The bracket notation $[a, b]$ is used for a real interval only where the entries are real numbers. We write $\#\mathcal{A}$ for the cardinality of a finite set $\mathcal{A}$.

A finite set $\mathcal{H} = \{h_1, \dots, h_k\}$ of distinct integers is called *admissible* if, for every prime p, the elements of $\mathcal{H}$ omit at least one residue class modulo p. We write

$$\mathrm{DHL}[k, m]$$

for the assertion that for every admissible k-tuple $\mathcal{H}$, there are infinitely many integers n such that at least m of $n + h_1, \dots, n + h_k$ are prime. Thus, if $H(k)$ denotes the diameter of the shortest admissible k-tuple, then $\mathrm{DHL}[k, 2]$ implies $H_1 \le H(k)$.

The integer k denotes a number of shifts, and the letter n is used both for a generic integer variable and, inside the description of a support, for the number of bands of that support. No confusion should arise, as the two never occur in the same expression.

All parameters attached to a support are fixed rational numbers, while ε, ε_0, ε', and ϵ denote small positive quantities that are chosen after those parameters. Whenever such a quantity is required to be sufficiently small in terms of the fixed data, we say so and do not record the dependence again for convenience.

3. Outline of the proof

We consider the Goldston–Pintz–Yıldırım sum

$$N(x; \rho, \nu) = \sum_{\substack{x \le n \le 2x \\ n \equiv b \ (W)}} \nu(n) \left(\sum_{i=1}^{k} \rho(n + h_i; x) - 1 \right), \tag{3.1}$$

where $\nu \ge 0$ is a sieve weight, W is a product of small primes used to pre-sieve, and ρ is a lower-bound weight for the prime indicator. If $N(x; \rho, \nu) > 0$ for arbitrarily large x, then there are infinitely many n for which at least two of the $n + h_i$ are prime, and hence $H_1 \le \max_{i \ne j} |h_i - h_j|$. In the multidimensional method, one takes ν to be the square of a k-dimensional divisor weight built from a function F on a region $T \subset [0, 1]^k$; the two sums appearing in $N(x; \rho, \nu)$ then have main terms proportional to two quadratic forms $I_T(F)$ and $J_T(F)$, and positivity of N reduces to the inequality $kJ_T(F) > I_T(F)$.

Making this rigorous requires equidistribution of ρ in the arithmetic progressions to moduli arising from expanding the square of the divisor weight. Those moduli are products of two divisors, one from each factor of the square, and their possible logarithmic sizes are determined by the geometry of T. The difficulty is therefore to choose T large enough that $kJ_T(F) > I_T(F)$ is attainable for k as small as possible, while keeping the moduli it generates within reach of the hypotheses of the available equidistribution results.

Stadlmann's support (4.1) achieves this by imposing two kinds of constraint on a point $t \in [0, 1]^k$: a constraint on the total mass $t_1 + \cdots + t_k$, which is allowed to exceed $1/2$ inside a band structure indexed by $\underline{A}$, and a constraint on the mass carried by those coordinates exceeding δ, governed by the matrix $\underline{B}$. The second constraint is what makes the enlargement a feasible strategy. A coordinate larger than δ corresponds to a divisor of size at least x^δ, and bounding the number and total size of such coordinates bounds the number and total size of the rough factors of any generated modulus. The remaining coordinates produce an x^δ-smooth factor, which acts as small change. Inserting its prime factors one at a time moves a partial product by at most x^δ, so any target interval of logarithmic width at least δ can be hit exactly. This is the mechanism behind all three factor-extraction lemmas of Section 7.

The proof then has four components, which we now list to help frame the remainder of this paper.

The sieve criterion. Theorem 4.10 states that if $1_{\mathbb{P}}$ is equidistributed for every modulus generated by a support T, and if some nonzero symmetric $F \in L^2(T)$ satisfies

$$\frac{kJ_T(F)}{I_T(F)} > 1,$$

then DHL$[k,2]$ holds. This is proved in Section 4 by approximating F by finite sums of products of one-variable functions, evaluating the resulting divisor sums, and comparing the main term with the error term.

The arithmetic input. Theorem 7.11 states that every modulus generated by the support used here is covered either by the Bombieri–Vinogradov theorem below the half-level, or by the transition argument of Section 8 at the intervening blocks, or by Stadlmann's positive-level Type I/II/III estimates. The support parameters are the rational numbers of Table 3.

The variational input. Theorem 11.1 exhibits a symmetric rational polynomial $F_\star$ with

$$J_T(F_\star) - 4I_T(F_\star) > 0,$$

the factor 4 being the rescaled form of the threshold 1 above.

The admissible tuple. Lemma 12.1 verifies that the 45-tuple $\mathcal{H}_{45}$ displayed in Section 12 is admissible and has diameter 212.

The first three combine to give DHL$[45,2]$, then the fourth gives us that $H_1 \leq 212$.

4. The sieve criterion

This section isolates the only sieve-theoretic implication used later: For a given support, equidistribution for the moduli that support generates with one strict variational inequality implies DHL$[k,2]$. The construction is Maynard's multidimensional sieve [May15], with the support enlargement of Polymath8b [Pol14b], and Stadlmann's mixed support [Sta26].

We proceed in five steps. We first define the support and identify the moduli produced when the sieve square is expanded. We then introduce the sieve weights and the quadratic functionals governing their main terms. After stating the criterion we prove it, by approximating the variational function by finite sums of products of one-variable functions, evaluating the resulting divisor sums, and comparing the main term with the error committed by a signed prime minorant.

4.1. **The support and its generated moduli.** The support determines both the region on which the variational function may live and, after the sieve weight is squared, the moduli for which equidistribution is required. We define it first and then record the modulus family it generates.

Definition 4.1 (Stadlmann support)**.** *Let $k,n \in \mathbb{N}$, let $\delta,\varepsilon > 0$, and let $\underline{A} = (A_0, A_1, \ldots, A_n)$ satisfy*

$$-\varepsilon = A_0 < \cdots < A_n < \frac{1}{2} - \varepsilon.$$

Let $\underline{B} = (B_{j,m})$ be an $n \times \lfloor 1/\delta \rfloor$ matrix satisfying, whenever both sides are defined,

$$\delta < B_{j,m} \leq B_{j,m+1} \leq B_{j,m} + \delta,$$

and set $B_{j,0} = 0$ for every j. For $t = (t_1, \ldots, t_k)$, put

$$S(t) = \sum_{i=1}^{k} t_i, \qquad R_\delta(t) = \{i : t_i > \delta\}, \qquad M_\delta(t) = \sum_{i \in R_\delta(t)} t_i.$$

Then

$$T_k(\delta, \underline{A}, \underline{B}, \varepsilon) = \bigcup_{j=1}^{n} \left\{ t \in [0,1]^k : {\scriptstyle A_{j-1}+\varepsilon \leq S(t) < A_j + \varepsilon, \atop M_\delta(t) \leq B_{j,|R_\delta(t)|}} \right\}.$$

Thus $\underline{A}$ governs the total mass, band by band, and $\underline{B}$ governs the mass carried by the coordinates exceeding δ. The one-step upper bound $B_{j,m+1} \leq B_{j,m} + \delta$ makes the rough-mass constraint hereditary and will be used repeatedly.

Lemma 4.2 (Hereditary rough-mass bound)**.** *Suppose t belongs to the j-th band, and put $I = \{i : t_i > \delta\}$. For every $I' \subseteq I$,*

$$\sum_{i \in I'} t_i \leq B_{j,|I'|}.$$

Proof. Write $|I| = m$, $|I'| = r$. Every removed coordinate is larger than δ, and therefore

$$\sum_{i\in I'} t_i < \sum_{i\in I} t_i - (m-r)\delta \le B_{j,m} - (m-r)\delta.$$

Iterating $B_{j,q+1} \le B_{j,q} + \delta$ gives $B_{j,m} \le B_{j,r} + (m-r)\delta$, which proves the claim. □

The numerical values of $\delta, \underline{A}, \underline{B}$ are fixed only in Section 9, after the analytic estimates have been translated into conditions on the factorizations of the moduli. Expanding the square of a divisor weight supported on $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$ produces the following family.

Definition 4.3 (Generated modulus family)**.** *Let $x > 1$ and $\varepsilon_0 > 0$. For an ordered pair of bands (j, j') and integers $m, m' \le \lfloor 1/\delta \rfloor$, let $Q(j, j', m, m')$ consist of the integers*

$$q = ee' \prod_{a=1}^{m} f_a \prod_{b=1}^{m'} f'_b$$

with the following properties:

(i) every f_a, f'_b is at least x^δ;

(ii) ee' is x^δ-smooth;

(iii)

$$\log_x \prod_a f_a \le (1-\varepsilon_0)B_{j,m}, \qquad \log_x \prod_b f'_b \le (1-\varepsilon_0)B_{j',m'};$$

(iv)

$$\log_x \left(e \prod_a f_a \right) \le (1-\varepsilon_0)(A_j - \varepsilon), \qquad \log_x \left(e' \prod_b f'_b \right) \le (1-\varepsilon_0)(A_{j'} + \varepsilon).$$

The generated family is

$$Q^*(x; T, \varepsilon_0) = \bigcup_{\substack{1 \le j, j' \le n \\ 0 \le m, m' \le \lfloor 1/\delta \rfloor}} Q(j, j', m, m').$$

Here $f_1, \ldots, f_m$ and $f'_1, \ldots, f'_{m'}$ are the rough factors coming from the two sides of the square, and ee' is the smooth part. Conditions (iii) and (iv) are precisely the images of the two constraints in Definition 4.1.

The argument permits the prime indicator to be replaced by a lower-bound weight $\rho \le 1_{\mathbb{P}}$; the distribution property we shall require of ρ is the following.

Definition 4.4 (Equidistribution norm)**.** *For $(a, q) = 1$, set*

$$\Delta(f; x; q, a) = \sum_{\substack{x \le n \le 2x \\ n \equiv a \pmod q}} f(n; x) - \frac{1}{\varphi(q)} \sum_{\substack{x \le n \le 2x \\ (n,q)=1}} f(n; x).$$

We say that f is equidistributed for the moduli generated by T if, for every $A > 0$ and $\varepsilon_0 > 0$,

$$\sum_{\substack{q \in Q^*(x;T,\varepsilon_0) \\ q \text{ squarefree}}} |\Delta(f; x; q, a)| \ll_{A,\varepsilon_0} \frac{x}{(\log x)^A}$$

uniformly in primitive residue classes $a \bmod q$.

4.2. **The weight and the variational functionals.** The divisor coefficients below are the arithmetic realization of a test function on the support. Their main terms are measured by three functionals: I is the unweighted mass of the sieve, J measures the prime contribution, and K measures the loss that a signed prime minorant can cause.

Definition 4.5 (GPY sieve)**.** *Let $(h_1, \ldots, h_k)$ be an admissible k-tuple.*

Let $W = \prod_{p \le \log\log\log x} p$ and choose $b \in \mathbb{N}$ with $(b + h_i, W) = 1$ for all $i \in \{1, \ldots, k\}$.

Let $\rho : \mathbb{N} \times (1, \infty) \to \mathbb{R}$ with $\rho(n; x) \le 1_{\mathbb{P}}(n)$ for all large x and $n \in [x, 2x]$.

For given $L \in \mathbb{N}$, $c_1, \dots, c_L \in \mathbb{R}$ and smooth compactly supported functions $F_{l,i} : [0, \infty) \to \mathbb{R}$, we define the following sieve weights:

$$\nu(n) = \left(\sum_{l=1}^{L} c_l \lambda_{F_{l,1}}(n + h_1) \dots \lambda_{F_{l,k}}(n + h_k) \right)^2 \qquad \text{with} \qquad \lambda_F(n) = \sum_{d|n} \mu(d) F(\log_x(d)).$$

For the chosen weight function $\nu : \mathbb{N} \to [0, \infty)$, we then define the GPY sieve function

$$N(x; \rho, \nu) = \sum_{\substack{x \le n \le 2x \\ n \equiv b(W)}} \nu(n) \left(\sum_{i=1}^{k} \rho(n + h_i; x) - 1 \right).$$

Because $\nu \geq 0$, positivity of $N(x; \rho, \nu)$ forces some $n_0 \in [x, 2x]$ with $\sum_i \rho(n_0 + h_i; x) > 1$. Moreover, since $\rho \le 1_{\mathbb{P}}$, at least two of the $n_0 + h_i$ are then prime. Therefore, positivity for arbitrarily large x implies $H_1 \le h_k - h_1$. Constructing a weight with this positivity property gives rise to the following three integrals.

Definition 4.6 (Key integrals)**.** *Let $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$ be as given in Definition 4.1. Let F be a symmetric, square-integrable function supported on $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$. Then define*

$$I(F; \delta, \underline{A}, \underline{B}, \varepsilon) = \int \cdots \int_{T_k(\delta, \underline{A}, \underline{B}, \varepsilon)} F(t_1, \dots, t_k)^2 \, dt_1 \dots dt_k,$$

$$J(F; \delta, \underline{A}, \underline{B}, \varepsilon) = \sum_{m=1}^{n} \sum_{m'=1}^{n} \int \cdots \int_{\substack{t_1 + \cdots + t_{k-1} \le \max\{A_m - \varepsilon, A_{m'} - \varepsilon\} \\ t_1 + \cdots + t_{k-1} + t_k \in [A_{m-1} + \varepsilon, A_m + \varepsilon] \\ t_1 + \cdots + t_{k-1} + t'_k \in [A_{m'-1} + \varepsilon, A_{m'} + \varepsilon]}} F(t_1, \dots, t_{k-1}, t_k) F(t_1, \dots, t_{k-1}, t'_k) \, dt_1 \dots dt_{k-1} \, dt_k \, dt'_k,$$

$$K(F; \delta, \underline{A}, \underline{B}, \varepsilon) = \sum_{m=1}^{n} \sum_{m'=1}^{n} \int \cdots \int_{\substack{t_1 + \cdots + t_{k-1} > \max\{A_m - \varepsilon, A_{m'} - \varepsilon\} \\ t_1 + \cdots + t_{k-1} + t_k \in [A_{m-1} + \varepsilon, A_m + \varepsilon] \\ t_1 + \cdots + t_{k-1} + t'_k \in [A_{m'-1} + \varepsilon, A_{m'} + \varepsilon]}} F(t_1, \dots, t_k)^2 \, dt_1 \dots dt_k.$$

4.3. **Statement of the criterion.** The hypotheses of the following proposition have two roles: the four conditions on ρ supply arithmetic information, while the strict quadratic inequality supplies the positivity needed to force two primes.

Proposition 4.7. *Let $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$ be as in Definition 4.1. Suppose that $\rho : \mathbb{N} \times (1, \infty) \to \mathbb{R}$ and the constants c_1, c_2 have the following properties:*

1. *Prime minorant: $-c_2 \le \rho(n; x) \le 1_{\mathbb{P}}(n)$ for all large x and $n \in [x, 2x]$.*
2. *ρ has equidistribution properties for moduli corresponding to $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$; see Definition 4.4.*
3. *If $\rho(n; x) \neq 0$, then all prime factors of n exceed x^β, for some $\beta > \max\{B_{j,1} : 1 \le j \le n\}$.*
4. *The sum of $\rho(n; x)$ over $[x, 2x]$ satisfies $\sum_{n \in [x,2x]} \rho(n; x) = (1 - c_1 + o(1)) \frac{x}{\log(x)}$.*

Further let $I(F; \delta, \underline{A}, \underline{B}, \varepsilon)$, $J(F; \delta, \underline{A}, \underline{B}, \varepsilon)$ and $K(F; \delta, \underline{A}, \underline{B}, \varepsilon)$ be as given in Definition 4.6. Suppose there exists a symmetric, square-integrable function F supported on $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$ with

$$\frac{k(1 - c_1) J(F; \delta, \underline{A}, \underline{B}, \varepsilon) - k c_2 K(F; \delta, \underline{A}, \underline{B}, \varepsilon)}{I(F; \delta, \underline{A}, \underline{B}, \varepsilon)} > 1. \tag{4.1}$$

Then $\mathrm{DHL}[k, 2]$ holds. In particular $H_1 = \liminf_{n \to \infty} (p_{n+1} - p_n) \le H(k)$, where $H(k)$ is the diameter of the shortest admissible k-tuple.

4.4. **From an L^2 function to finite tensor sums.** The arithmetic evaluation applies to products of one-variable weights, whereas the variational problem is stated for an arbitrary symmetric function of k variables. To pass between them, we first retreat strictly from every face of the support, and then mollify and approximate by finite sums of products. The retreat guarantees that no smoothing from mollification crosses a band face or a rough-mass face. The argument follows Section 5 of [Pol14b], with the support geometry and the minorant penalty retained explicitly.

Proposition 4.8 (Sieve realization)**.** *Let $T_k(\delta,\underline{A},\underline{B},\varepsilon)$ be as described in Definition 4.1. Let $\varepsilon_0>0$ and define region*

$$R_k^+(\delta,\underline{A},\underline{B},\varepsilon,j,\varepsilon_0)=\left\{(t_1,\dots,t_k)\in[0,1]^k:\ \begin{array}{l}\sum_{i=1}^k t_i<(1-\varepsilon_0)(A_j+\varepsilon),\\ \sum_{i\in I}t_i<(1-\varepsilon_0)B_{j,|I|}\ \text{for } I=\{i:t_i>\delta\}\end{array}\right\}.$$

Suppose there exists a symmetric, square-integrable function F supported on $T_k(\delta,\underline{A},\underline{B},\varepsilon)$ with

$$\frac{k(1-c_1)J(F;\delta,\underline{A},\underline{B},\varepsilon)-kc_2K(F;\delta,\underline{A},\underline{B},\varepsilon)}{I(F;\delta,\underline{A},\underline{B},\varepsilon)}>1.$$

If $\varepsilon_0>0$ is chosen sufficiently small, we can then find integers $L_1,\dots,L_n$, constants $c_{j,l}$ and smooth compactly supported functions $f_{j,l,i}:[0,\infty)\to\mathbb{R}$ with the following two properties:

1. *For every $j\in\{1,\dots,n\}$ and $l\in\{1,\dots,L_j\}$, the support of the product $f_{j,l,1}(t_1)\dots f_{j,l,k}(t_k)$ is contained in $R_k^+(\delta,\underline{A},\underline{B},\varepsilon,j,\varepsilon_0)$.*
2. *Denote by $\mathcal{L}(j,i)$ the set of $l\in\{1,\dots,L_j\}$ for which the support of $\prod_{s\neq i}f_{j,l,s}(t_s)$ is contained in*

$$\left\{(t_s)_{s\neq i}:\sum_{s\neq i}t_s<(1-\varepsilon_0)(A_j-\varepsilon)\right\},$$

and let $\mathcal{U}(j,i)=\{1,\dots,L_j\}\setminus\mathcal{L}(j,i)$. Then the quantities

$$\begin{aligned}\mathcal{I}&=\sum_{j=1}^n\sum_{l=1}^{L_j}\sum_{j'=1}^n\sum_{l'=1}^{L_{j'}}c_{j,l}c_{j',l'}\prod_{s=1}^k\int_0^\infty f'_{j,l,s}(t_s)f'_{j',l',s}(t_s)\,dt_s,\\ \mathcal{J}_i&=\sum_{j=1}^n\sum_{l\in\mathcal{L}(j,i)}\sum_{j'=1}^n\sum_{l'\in\mathcal{L}(j',i)}c_{j,l}c_{j',l'}f_{j,l,i}(0)f_{j',l',i}(0)\prod_{s\neq i}\int_0^\infty f'_{j,l,s}(t_s)f'_{j',l',s}(t_s)\,dt_s\\ &\quad+2\sum_{j=1}^n\sum_{l\in\mathcal{U}(j,i)}\sum_{j'=1}^n\sum_{l'\in\mathcal{L}(j',i)}c_{j,l}c_{j',l'}f_{j,l,i}(0)f_{j',l',i}(0)\prod_{s\neq i}\int_0^\infty f'_{j,l,s}(t_s)f'_{j',l',s}(t_s)\,dt_s\\ \mathcal{K}_i&=\sum_{j=1}^n\sum_{l\in\mathcal{U}(j,i)}\sum_{j'=1}^n\sum_{l'\in\mathcal{U}(j',i)}c_{j,l}c_{j',l'}\prod_{s=1}^k\int_0^\infty f'_{j,l,s}(t_s)f'_{j',l',s}(t_s)\,dt_s\end{aligned}$$

satisfy the inequality

$$\frac{(1-c_1)\sum_{i=1}^k\mathcal{J}_i-c_2\sum_{i=1}^k\mathcal{K}_i}{\mathcal{I}}>1.$$

The sieve weights used in the GPY method will then be

$$\nu(n)=\left(\sum_{j=1}^n\sum_{l=1}^{L_j}c_{j,l}\lambda_{f_{j,l,1}}(n+h_1)\dots\lambda_{f_{j,l,k}}(n+h_k)\right)^2.$$

Proof. Let $F:[0,\infty)^k\to\mathbb{R}$ be a symmetric square-integrable function, supported on $T_k(\delta,\underline{A},\underline{B},\varepsilon)$, which satisfies the inequality

$$\frac{k(1-c_1)J(F;\delta,\underline{A},\underline{B},\varepsilon)-kc_2K(F;\delta,\underline{A},\underline{B},\varepsilon)}{I(F;\delta,\underline{A},\underline{B},\varepsilon)}>1. \tag{4.2}$$

Let $S_j=\{(t_1,\dots,t_k)\in[0,1]^k:\sum_{i=1}^k t_i\in[A_{j-1}+\varepsilon,A_j+\varepsilon)$, and let $1_{S_j}(n)$ be the indicator function of S_j. For $I=\{i:t_i>\delta\}\}$, let $\sum_{i\in I}t_i\le B_{j,|I|}$. Since $\{S_1,\dots,S_n\}$ is a partition of $T_k(\delta,\underline{A},\underline{B},\varepsilon)$, we then have $F=\sum_{j=1}^n F\cdot 1_{S_j}$.

Next, we choose small $\varepsilon_1,\varepsilon_2>0$ and define $F_{2,j}:[0,\infty)^k\to\mathbb{R}$ as follows:

$$F_{2,j}(t_1,\dots,t_k)=\begin{cases}F\left(\frac{t_1-\varepsilon_2}{1-\varepsilon_1},\dots,\frac{t_k-\varepsilon_2}{1-\varepsilon_1}\right)\cdot 1_{S_j}\left(\frac{t_1-\varepsilon_2}{1-\varepsilon_1},\dots,\frac{t_k-\varepsilon_2}{1-\varepsilon_1}\right) & \text{if } t_1,\dots,t_k\ge\varepsilon_2,\\ 0 & \text{otherwise.}\end{cases}$$

We also take $\varepsilon_0 = \frac{\varepsilon_1}{100}$ and define the integrals

$$I(F,G) = \int_0^\infty \cdots \int_0^\infty F(t_1,\dots,t_k)G(t_1,\dots,t_k)\ \mathrm{d}t_1\dots\mathrm{d}t_k,$$

$$J_i(F,G;B) = \underset{\sum_{s\neq i} t_s<B}{\int_0^\infty \cdots \int_0^\infty} F(t_1,\dots,t_{i-1},t_i,t_{i+1},\dots,t_k)G(t_1,\dots,t_{i-1},t_i',t_{i+1},\dots,t_k)\,\mathrm{d}t_i'\mathrm{d}t_1\dots\mathrm{d}t_k,$$

$$K_i(F,G;B) = \underset{\sum_{s\neq i} t_s>B}{\int_0^\infty \cdots \int_0^\infty} F(t_1,\dots,t_k)G(t_1,\dots,t_k)\,\mathrm{d}t_1\dots\mathrm{d}t_k,$$

Because F is symmetric, square-integrable, and satisfies the strict inequality (4.2), continuity under dilation and translation allows us to choose $\varepsilon_1 > 0$ and then $0 < \varepsilon_2 \ll \varepsilon_1$ so that the shifted functions $F_{2,j}$ still satisfy

$$\frac{\sum_{i=1}^k \sum_{j=1}^n \sum_{j'=1}^n ((1-c_1)J_i(F_{2,j},F_{2,j'};B^*(j,j')) - c_2K_i(F_{2,j},F_{2,j'};B^*(j,j')))}{\sum_{j=1}^n \sum_{j'=1}^n I(F_{2,j},F_{2,j'})} > 1 \tag{4.3}$$

where $B^*(j,j') = (1-\varepsilon_0)\max\{A_j - \varepsilon, A_{j'} - \varepsilon\}$.

We next check that this deformation has created room for mollification. Let $t \in \operatorname{supp} F_{2,j}$ and set $L_t = \{i : t_i > \delta\}$. For $i \in L_t$, the preimage coordinate $(t_i - \varepsilon_2)/(1-\varepsilon_1)$ is still larger than δ, provided ε_2 is sufficiently small relative to ε_1. Therefore, the rough-mass bound for the preimage, together with Lemma 4.2, gives

$$\sum_{i\in L_t} \frac{t_i - \varepsilon_2}{1-\varepsilon_1} \le B_{j,|L_t|}.$$

Therefore, the dilation and translation place $\operatorname{supp} F_{2,j}$ strictly inside every total-mass, rough-mass, and coordinate face of $R_k^+(\delta, \underline{A}, \underline{B}, \varepsilon, j, 2\varepsilon_0)$, and we may convolve with an approximate identity of sufficiently small radius. This produces smooth functions $F_{3,j}$, supported in the same open region, for which continuity of the three quadratic forms preserves the strict inequality

$$\frac{\sum_{i=1}^k \sum_{j=1}^n \sum_{j'=1}^n ((1-c_1)J_i(F_{3,j},F_{3,j'};B^*(j,j')) - c_2K_i(F_{3,j},F_{3,j'};B^*(j,j')))}{\sum_{j=1}^n \sum_{j'=1}^n I(F_{3,j},F_{3,j'})} > 1. \tag{4.4}$$

Following [Pol14b], define $f_{3,j} : [0,\infty)^k \to \mathbb{R}$ by setting

$$f_{3,j}(t_1,\dots,t_k) = \int_{t_1}^\infty \cdots \int_{t_k}^\infty F_{3,j}(s_1,\dots,s_k)\,\mathrm{d}s_1\dots\mathrm{d}s_k.$$

The condition $B_{j,m} \le B_{j,m+1} \le B_{j,m} + \delta$ again ensures that the support of $f_{3,j}$ is contained in $R_k^+(\delta, \underline{A}, \underline{B}, \varepsilon, j, 2\varepsilon_0)$. Write $D = \partial_{t_1}\cdots\partial_{t_k}$ and $D_{\widehat{i}} = \prod_{s\neq i}\partial_{t_s}$. Then

$$I(F_{3,j},F_{3,j'}) = \int (Df_{3,j})(t)(Df_{3,j'})(t)\,dt,$$

$$J_i(F_{3,j},F_{3,j'};B) = \int_{\sum_{s\neq i} t_s<B} (D_{\widehat{i}}f_{3,j})(t_{\widehat{i}},0)(D_{\widehat{i}}f_{3,j'})(t_{\widehat{i}},0)\,dt_{\widehat{i}},$$

$$K_i(F_{3,j},F_{3,j'};B) = \int_{\sum_{s\neq i} t_s>B} (Df_{3,j})(t)(Df_{3,j'})(t)\,dt.$$

Write $I^*(f_{3,j},f_{3,j'}) = I(F_{3,j},F_{3,j'})$, $J_i^*(f_{3,j},f_{3,j'};B) = J_i(F_{3,j},F_{3,j'};B)$, and $K_i^*(f_{3,j},f_{3,j'};B) = K_i(F_{3,j},F_{3,j'};B)$. Hence

$$\frac{\sum_{i=1}^k \sum_{j=1}^n \sum_{j'=1}^n ((1-c_1)J_i^*(f_{3,j},f_{3,j'};B^*(j,j')) - c_2K_i^*(f_{3,j},f_{3,j'};B^*(j,j')))}{\sum_{j=1}^n \sum_{j'=1}^n I^*(f_{3,j},f_{3,j'})} > 1 \tag{4.5}$$

where $B^*(j,j') = (1-\varepsilon_0)\max\{A_j - \varepsilon, A_{j'} - \varepsilon\}$.

With this in hand, we next pass from smooth multivariable functions to finite tensor sums, as in Section 5.3 of [Pol14b]. We cover the compact support of each $f_{3,j}$ by coordinate boxes of side at most ε_3, choose a smooth partition of unity subordinate to this cover, and approximate on each box by finite sums of products of one-variable smooth functions. Tensor-product density in the required Sobolev norm makes the three quadratic forms converge simultaneously. Because the support of $f_{3,j}$ lies strictly inside the admissible region, ε_3 and the approximation error may be chosen so that every tensor retains the same support conditions. We thereby obtain integers $L_1, \dots, L_n$, constants $c_{j,l}$, and smooth functions $f_{j,l,i} : [0, \infty) \to \mathbb{R}$ that satisfy the following key properties:

(i) Each $f_{j,l,i}$ is supported on an interval of length at most ε_3.
(ii) For every j and l, the support of $\prod_{i=1}^{k} f_{j,l,i}(t_i)$ is contained in $R_k^+(\delta, \underline{A}, \underline{B}, \varepsilon, j, \varepsilon_0)$.
(iii) For each j, the tensor sum

$$h_j(t_1, \dots, t_k) = \sum_{l=1}^{L_j} c_{j,l} \prod_{i=1}^{k} f_{j,l,i}(t_i)$$

approximates $f_{3,j}$ closely enough that

$$\frac{\sum_{i=1}^{k} \sum_{j=1}^{n} \sum_{j'=1}^{n} ((1 - c_1) J_i^*(h_j, h_{j'}; B^*(j, j')) - c_2 K_i^*(h_j, h_{j'}; B^*(j, j')))}{\sum_{j=1}^{n} \sum_{j'=1}^{n} I^*(h_j, h_{j'})} > 1. \tag{4.6}$$

With our goal in mind, it remains to identify the three tensor expressions with the quadratic forms in Proposition 4.8. Expanding $\sum_{j,j'} I^*(h_j, h_{j'})$ gives

$$\sum_{j=1}^{n} \sum_{l=1}^{L_j} \sum_{j'=1}^{n} \sum_{l'=1}^{L_{j'}} c_{j,l} c_{j',l'} \int_0^\infty \cdots \int_0^\infty f'_{j,l,1}(t_1) \dots f'_{j,l,k}(t_k) f'_{j',l',1}(t_1) \dots f'_{j',l',k}(t_k) \, \mathrm{d}t_1 \dots \mathrm{d}t_k.$$

This is exactly $\mathcal{I}$. We next compare the numerator terms. For a fixed i, expansion of $\sum_{j,j'} J_i^*(h_j, h_{j'}; B^*(j, j'))$ gives

$$\mathcal{J}_i^* = \sum_{j=1}^{n} \sum_{l=1}^{L_j} \sum_{j'=1}^{n} \sum_{l'=1}^{L_{j'}} c_{j,l} c_{j',l'} f_{j,l,i}(0) f_{j',l',i}(0) \underset{\sum_{s \neq i} t_s < (1-\varepsilon_0) \max\{A_j - \varepsilon, A_{j'} - \varepsilon\}}{\int_0^\infty \cdots \int_0^\infty} \prod_{s \neq i} f'_{j,l,s}(t_s) \prod_{s \neq i} f'_{j',l',s}(t_s) \prod_{s \neq i} \mathrm{d}t_s.$$

Partition $\{1, \dots, L_j\}$ according to whether the marginal support lies inside the retained region. To this end, we let $\mathcal{L}(j, i)$ consist of those l for which the support of $\prod_{s \neq i} f_{j,l,s}(t_s)$ is contained in

$$\left\{ (t_s)_{s \neq i} : \sum_{s \neq i} t_s < (1 - \varepsilon_0)(A_j - \varepsilon) \right\},$$

and put $\mathcal{U}(j, i) = \{1, \dots, L_j\} \setminus \mathcal{L}(j, i)$.

The union of the supports of $\prod_{s \neq i} f_{j,l,s}(t_s) \prod_{s \neq i} f_{j',l',s}(t_s)$, with $l \in \mathcal{U}(j, i)$ and $l' \in \mathcal{U}(j', i)$, meets $\{(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_k) : \sum_{s \neq i} t_s < (1 - \varepsilon_0) \max\{A_j - \varepsilon, A_{j'} - \varepsilon\}\}$ on a set of measure $O(\varepsilon_3)$. Therefore, for any given $\varepsilon_4 > 0$, we can choose the value of ε_3 sufficiently small that the contribution of pairs $(l, l') \in \mathcal{U}(j, i) \times \mathcal{U}(j', i)$ to $\mathcal{J}_i^*$ is bounded by ε_4 and

$$\begin{aligned} \mathcal{J}_i &= \sum_{j=1}^{n} \sum_{l \in \mathcal{L}(j,i)} \sum_{j'=1}^{n} \sum_{l' \in \mathcal{L}(j',i)} c_{j,l} c_{j',l'} f_{j,l,i}(0) f_{j',l',i}(0) \int_0^\infty \cdots \int_0^\infty \prod_{s \neq i} f'_{j,l,s}(t_s) \prod_{s \neq i} f'_{j',l',s}(t_s) \prod_{s \neq i} \mathrm{d}t_s \\ &\quad + 2 \sum_{j=1}^{n} \sum_{l \in \mathcal{U}(j,i)} \sum_{j'=1}^{n} \sum_{l' \in \mathcal{L}(j',i)} c_{j,l} c_{j',l'} f_{j,l,i}(0) f_{j',l',i}(0) \int_0^\infty \cdots \int_0^\infty \prod_{s \neq i} f'_{j,l,s}(t_s) \prod_{s \neq i} f'_{j',l',s}(t_s) \prod_{s \neq i} \mathrm{d}t_s \\ &> \sum_{j=1}^{n} \sum_{j'=1}^{n} J_i^*(h_j, h_{j'}; (1 - \varepsilon_0) \max\{A_j - \varepsilon, A_{j'} - \varepsilon\}) - \varepsilon_4. \end{aligned}$$

Finally, a pair (l, l') for which $l \in \mathcal{L}(j, i)$ or $l' \in \mathcal{L}(j', i)$ makes no contribution to $\sum_{j,j'} K_i^*(h_j, h_{j'}; B^*(j, j'))$. Indeed, its support is disjoint from the region of integration. The entire K_i^*-sum therefore equals

$$\mathcal{K}_i^* = \sum_{j=1}^{n} \sum_{l \in \mathcal{U}(j,i)} \sum_{j'=1}^{n} \sum_{l' \in \mathcal{U}(j',i)} c_{j,l} c_{j',l'} \underset{\sum_{s \neq i} t_s > (1-\varepsilon_0) \max\{A_j - \varepsilon, A_{j'} - \varepsilon\}}{\int_0^\infty \cdots \int_0^\infty} \prod_{s=1}^{k} f'_{j,l,s}(t_s) \prod_{s=1}^{k} f'_{j',l',s}(t_s) \, \mathrm{d}t_1 \dots \mathrm{d}t_k.$$

Since the support of functions $\prod_{s\neq i} f_{j,l,s}(t_s)\prod_{s\neq i} f_{j',l',s}(t_s)$ with $l\in\mathcal{U}(j,i)$ and $l'\in\mathcal{U}(j',i)$ only overlaps with $\{(t_1,\dots,t_{i-1},t_{i+1},\dots,t_k):\sum_{s\neq i}t_s<(1-\varepsilon_0)\max\{A_j-\varepsilon,A_{j'}-\varepsilon\}\}$ on a set of measure $O(\varepsilon_3)$, we can drop the condition $\sum_{s\neq i}t_s>(1-\varepsilon_0)\max\{A_j-\varepsilon,A_{j'}-\varepsilon\}$ with a loss of at most ε_4, provided that ε_3 was chosen small enough. Overall we then have

$$\begin{aligned}\mathcal{K}_i&=\sum_{j=1}^{n}\sum_{l\in\mathcal{U}(j,i)}\sum_{j'=1}^{n}\sum_{l'\in\mathcal{U}(j',i)}c_{j,l}c_{j',l'}\int_0^\infty\cdots\int_0^\infty\prod_{s=1}^{k}f'_{j,l,s}(t_s)\prod_{s=1}^{k}f'_{j',l',s}(t_s)\,\mathrm{d}t_1\dots\mathrm{d}t_k\\&<\sum_{j=1}^{n}\sum_{j'=1}^{n}K_i^*(h_j,h_{j'};B^*(j,j'))+\varepsilon_4.\end{aligned}$$

Looking back at inequality (4.6) we therefore have

$$\frac{(1-c_1)\sum_{i=1}^k\mathcal{J}_i-c_2\sum_{i=1}^k\mathcal{K}_i+((1-c_1)k+c_2k)\varepsilon_4}{\mathcal{I}}>1.$$

Hence, taking ε_4 sufficiently small (and the corresponding ε_3 even smaller), we arrive at the desired inequality $\frac{(1-c_1)\sum_{i=1}^k\mathcal{J}_i-c_2\sum_{i=1}^k\mathcal{K}_i}{\mathcal{I}}>1$. $\square$

4.5. **Asymptotic evaluation of the divisor sums.** Having reduced to products of one-variable weights, we evaluate the corresponding divisor sums, of which there are two terms. The principal residue-class average produces the variational functionals, and the discrepancy from that average is summed over the modulus family of Definition 4.3. The coprimality conditions introduced by the W-trick are recorded explicitly, since they are part of the uniformity demanded of the equidistribution estimates in Section 5.

Both sums appearing in $N(x;\rho,\nu)$ reduce to the mixed term

$$\sum_{x\leq n\leq 2x}\rho(n+h_i;x)\prod_{s=1}^{k}\lambda_{F_s}(n+h_s)\lambda_{G_s}(n+h_s),$$

which is estimated as follows. The calculation follows Section 4.2 of [Pol14b], with the minorant modification of Baker and Irving [BI17] and with the generated modulus family tracked explicitly.

Lemma 4.9. *Suppose that $\rho:\mathbb{N}\times(1,\infty)\to\mathbb{R}$ has the following properties:*

1. *Prime minorant: $-c_2\leq\rho(n;x)\leq 1_{\mathbb{P}}(n)$ for all large x and $n\in[x,2x]$.*
2. *ρ has equidistribution properties for moduli corresponding to $T_k(\delta,\underline{A},\underline{B},\varepsilon)$; see Definition 4.4.*
3. *If $\rho(n;x)\neq 0$, then all prime factors of n exceed x^β, for some $\beta>\max\{B_{j,1}:1\leq j\leq n\}$.*
4. *The sum of $\rho(n;x)$ over $[x,2x]$ satisfies $\sum_{n\in[x,2x]}\rho(n;x)=(1-c_1+o(1))\frac{x}{\log(x)}$.*

Fix $i_0\in\{1,\dots,k\}$, and let $F_i,G_i:[0,\infty)\to\mathbb{R}$ be smooth and compactly supported. Suppose

$$\operatorname{supp}\prod_{i=1}^{k}F_i(t_i)\subset R_k^+(\delta,\underline{A},\underline{B},\varepsilon,j,\varepsilon_0)$$

and the analogous inclusion for $\prod_i G_i(t_i)$ holds with j' in place of j. Finally, assume that the marginal support satisfies

$$\operatorname{supp}\prod_{i\neq i_0}F_i(t_i)\subset\left\{(t_i)_{i\neq i_0}:\sum_{i\neq i_0}t_i<(1-\varepsilon_0)(A_j-\varepsilon)\right\}.$$

Let $W=\prod_{p\leq\log\log\log x}p$ and $b\in\mathbb{N}$ with $(b+h_i,W)=1$ for all $i\in\{1,\dots,k\}$. Then

$$\begin{aligned}&\sum_{\substack{x\leq n\leq 2x\\ n\equiv b(W)}}\rho(n+h_{i_0};x)\prod_{i=1}^{k}\lambda_{F_i}(n+h_i)\lambda_{G_i}(n+h_i)\\&=\left((1-c_1)F_{i_0}(0)G_{i_0}(0)\prod_{i\neq i_0}\int_0^\infty F_i'(t_i)G_i'(t_i)dt_i+o(1)\right)\left(\frac{xW^{k-1}}{\varphi(W)^k\log(x)^k}\right).\end{aligned}$$

Proof. By relabeling the coordinates, assume $i_0 = k$. Recall that $\lambda_F(n) = \sum_{d|n} \mu(d)F(\log_x(d))$. We call the sum we are interested in S. Expanding, we have

$$S := \sum_{\substack{x\le n\le 2x\\ n\equiv b(W)}} \rho(n+h_k;x)\prod_{i=1}^{k}\lambda_{F_i}(n+h_i)\lambda_{G_i}(n+h_i)$$

$$= \sum_{d_1,\ldots,d_k,d_1',\ldots,d_k'}\prod_{i=1}^{k}\mu(d_i)\mu(d_i')F_i(\log_x(d_i))G_i(\log_x(d_i'))\sum_{\substack{x\le n\le 2x\\ n\equiv b(W)\\ n\equiv -h_i[d_i,d_i']\,\forall i}}\rho(n+h_k;x).$$

If $\rho(n+h_k;x)\neq 0$, every nontrivial divisor $d\mid n+h_k$ satisfies $\log_x d > B_{j,1}$. Since $B_{j,m}\le B_{j,1}+(m-1)\delta$, adjoining any further $m-1$ rough coordinates would give total rough mass greater than $B_{j,m}$. Such a point cannot lie in the support of $\prod_i F_i(t_i)$. Consequently only $d_k = 1$ contributes. Moreover, same argument gives $d_k' = 1$. Hence

$$S = F_k(0)G_k(0)\sum_{d_1,\ldots,d_k,d_1',\ldots,d_k'}\prod_{i=1}^{k-1}\mu(d_i)\mu(d_i')F_i(\log_x(d_i))G_i(\log_x(d_i'))\sum_{\substack{x\le n\le 2x\\ n\equiv b(W)\\ n\equiv -h_i[d_i,d_i']\,\forall i\neq k}}\rho(n+h_k;x).$$

We now use the W-trick in the form of [Pol14b]. The residue class b was chosen so that $(b+h_i,W)=1$ for every i. Once x is large enough that

$$|h_i - h_{i'}| < \log\log\log x \qquad (i\neq i'),$$

the preceding congruences can be compatible only if the least common multiples $[d_i,d_i']$ are pairwise coprime and are all coprime to $W = \prod_{p\le \log\log\log x} p$.

The conditions ($n\equiv b(W)$ and $n\equiv -h_i[d_i,d_i']$ for $i\neq k$) can thus be replaced by a single congruence $n+h_k\equiv a_{W,d_1,\ldots,d_{k-1}'} \bmod q_{W,d_1,\ldots,d_{k-1}'}$. Here $a_{W,d_1,\ldots,d_{k-1}'}$ can be chosen to be coprime to all primes $p\le x$, and $q_{W,d_1,\ldots,d_{k-1}'}$ is defined to equal

$$q_{W,d_1,\ldots,d_{k-1}'} = W[d_1,d_1']\ldots[d_{k-1},d_{k-1}'].$$

Hence we split S into a main term S_M and an error term S_E as follows:

$$S_M = F_k(0)G_k(0)\sum_{d_1,\ldots,d_k,d_1',\ldots,d_k'}\frac{\prod_{i=1}^{k-1}\mu(d_i)\mu(d_i')F_i(\log_x(d_i))G_i(\log_x(d_i'))}{\varphi(W[d_1,d_1']\ldots[d_{k-1},d_{k-1}'])}\sum_{x+h_k\le m\le 2x+h_k}\rho(m;x),$$

$$S_E = |F_k(0)G_k(0)|\sum_{\substack{d_1,\ldots,d_k,d_1',\ldots,d_k'\\ q_{W,d_1,\ldots,d_{k-1}'}\ \text{squarefree}}}\prod_{i=1}^{k-1}|F_i(\log_x(d_i))G_i(\log_x(d_i'))|\Delta(\rho;a_{W,d_1,\ldots,d_{k-1}'}(q_{W,d_1,\ldots,d_{k-1}'})),$$

$$\text{where } \Delta(\rho;a(q)) = \left|\sum_{\substack{x+h_k\le m\le 2x+h_k\\ m\equiv a(q)}}\rho(m;x) - \frac{1}{\varphi(q)}\sum_{\substack{x+h_k\le m\le 2x+h_k\\ (m,q)=1}}\rho(m;x)\right|.$$

For the main term S_M, use the assumption $\sum_{n\in[x,2x]}\rho(n;x) = (1-c_1+o(1))\frac{x}{\log(x)}$ and apply Lemma 4.1 of [Pol14b], with $[d_i,d_i']$ replaced by $\varphi([d_i,d_i'])$. This gives

$$S_M = \left(F_k(0)G_k(0)\left(\prod_{i=1}^{k-1}\int_0^\infty F_i'(t_i)G_i'(t_i)\mathrm{d}t_i\right)(1-c_1)+o(1)\right)\left(\frac{xW^{k-1}}{\varphi(W)^k\log(x)^k}\right).$$

For the error term, consider a contributing modulus

$$q = W[d_1,d_1']\cdots[d_{k-1},d_{k-1}'].$$

Among the factors d_i, denote those at least x^δ by $f_1,\dots,f_m$, and absorb the remaining factors, together with W, into e. Apply the same procedure to the coprime quotients $d_i'/(d_i,d_i')$, obtaining rough factors $f_1',\dots,f_{m'}'$ and a residual factor e'. Then

$$q = ee'\prod_{i=1}^{m} f_i \prod_{i=1}^{m'} f_i',$$

where e and e' are x^δ-smooth. The two tensor-support conditions give

$$\log_x\left(e\prod_i f_i\right) \le \left(1-\frac{\varepsilon_0}{2}\right)(A_j-\varepsilon), \qquad \log_x\left(e'\prod_i f_i'\right) \le \left(1-\frac{\varepsilon_0}{2}\right)(A_{j'}+\varepsilon).$$

The one-step inequalities for the B-rows similarly give

$$\log_x \prod_i f_i \le (1-\varepsilon_0)B_{j,m}, \qquad \log_x \prod_i f_i' \le (1-\varepsilon_0)B_{j',m'}.$$

Thus Definition 4.3 places every modulus contributing to S_E in $Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\varepsilon_0/2)$.

By hypothesis (2), ρ has equidistribution properties for the moduli corresponding to $T_k(\delta,\underline{A},\underline{B},\varepsilon)$: if $a\in\mathbb{Z}$ with $(a,p)=1$ for all $p\le x$, and if $C>0$, then

$$\sum_{\substack{q\in Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\frac{\varepsilon_0}{2})\\ q \text{ is squarefree}}} \left|\sum_{\substack{n\in[x,2x]\\ n\equiv a(q)}} \rho(n;x) - \frac{1}{\varphi(q)}\sum_{\substack{n\in[x,2x]\\ (n,q)=1}} \rho(n;x)\right| \ll_{C,\varepsilon_0} \frac{x}{\log(x)^C}. \tag{4.7}$$

The congruence class a above is fixed, whereas in S_E the class $a_{W,d_1,\dots,d_{k-1}'}$ depends on the divisors. The averaging device from the proof of Theorem 3.6 of [Pol14b] uses the set

$$\mathcal{A} = \left\{a\in\left\{1,\dots,\prod_{p\le x} p\right\} : a\equiv b+h_k \bmod W,\ \forall p\in(\log\log\log x, x]\ \exists i\in\{1,\dots,k-1\} : a\equiv h_k-h_i \bmod p\right\}.$$

Fix $d_1,d_1',\dots,d_{k-1},d_{k-1}'$, and put

$$q = q_{W,d_1,\dots,d_{k-1}'} = W[d_1,d_1']\cdots[d_{k-1},d_{k-1}'].$$

The set $\mathcal{A}$ was chosen so that its reduction modulo q contains the residue class $a_{W,d_1,\dots,d_{k-1}'}$. The Chinese remainder theorem gives its multiplicity: among the elements of $\mathcal{A}$, this class occurs

$$\frac{|\mathcal{A}|}{(k-1)^{\omega(q/W)}}$$

times, where $\omega(q/W)$ denotes the number of prime factors of q/W. Since q/W is squarefree, $(k-1)^{\omega(q/W)} \ll \tau(q)^{O(1)}$. Thus averaging over $\widetilde{a}\in\mathcal{A}$, with the factor $\tau(q)^{O(1)}/|\mathcal{A}|$, dominates the discrepancy in the required class. Conversely, a fixed $q\in Q^*$ arises from at most $\tau(q)^{O(1)}$ tuples of divisors. These two multiplicity bounds give the following estimate for S_E:

$$S_E \ll \sum_{\substack{q\in Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\frac{\varepsilon_0}{2})\\ q \text{ squarefree}}} \frac{\tau(q)^{O(1)}}{|\mathcal{A}|}\sum_{a\in\mathcal{A}}\left|\sum_{\substack{n\in[x+h_k,2x+h_k]\\ n\equiv a(q)}} \rho(n;x) - \frac{1}{\varphi(q)}\sum_{\substack{n\in[x+h_k,2x+h_k]\\ (n,q)=1}} \rho(n;x)\right|.$$

To finish, apply Cauchy–Schwarz and (4.7). Choosing C sufficiently large gives

$$\begin{aligned} S_E &\ll \frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}}\left(\sum_{\substack{q\in Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\frac{\varepsilon_0}{2})\\ q \text{ squarefree}}} \tau(q)^{O(1)}\Delta(\rho,a(q))\right)^{1/2}\left(\sum_{\substack{q\in Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\frac{\varepsilon_0}{2})\\ q \text{ squarefree}}} \Delta(\rho,a(q))\right)^{1/2} \\ &\ll_C \frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}} (x\log(x)^{O(1)})^{1/2}(x\log(x)^{-C})^{1/2} \ll_C \frac{x}{\log(x)^{C/2}}. \end{aligned}$$

Since $W\ll(\log\log x)^{O(1)}$, summing the asymptotic for S_M and the bound on S_E for a large C completes the proof. $\square$

4.6. **Proof of the criterion.** We now assemble the approximation and the asymptotic formula. One point requires care: the prime detector may be a signed minorant, and the tensor approximation was made on a slightly enlarged support. We therefore separate the retained marginal terms from the discarded ones, and show that the latter are charged precisely to the functionals K_i.

Proof of Proposition 4.7. Fix a sufficiently small $\varepsilon_0 > 0$. Lemma 4.8 then supplies integers $L_1, \dots, L_n$, coefficients $c_{j,l}$, and smooth compactly supported functions $f_{j,l,i} : [0,\infty) \to \mathbb{R}$, each tensor $\prod_i f_{j,l,i}(t_i)$ being supported in $R_k^+(\delta, \underline{A}, \underline{B}, \varepsilon, j, \varepsilon_0)$. Let $\mathcal{L}(j,i)$ be the set of $l \in \{1, \dots, L_j\}$ for which the support of $\prod_{s \neq i} f_{j,l,s}(t_s)$ is contained in

$$\left\{ (t_s)_{s\neq i} : \sum_{s \neq i} t_s < (1-\varepsilon_0)(A_j - \varepsilon) \right\},$$

and let $\mathcal{U}(j,i) = \{1, \dots, L_j\} \setminus \mathcal{L}(j,i)$. Then the quantities

$$\begin{aligned}
\mathcal{I} &= \sum_{j=1}^{n} \sum_{l=1}^{L_j} \sum_{j'=1}^{n} \sum_{l'=1}^{L_{j'}} c_{j,l} c_{j',l'} \prod_{s=1}^{k} \int_0^\infty f'_{j,l,s}(t_s) f'_{j',l',s}(t_s)\, \mathrm{d}t_s, \\
\mathcal{J}_i &= \sum_{j=1}^{n} \sum_{l \in \mathcal{L}(j,i)} \sum_{j'=1}^{n} \sum_{l' \in \mathcal{L}(j',i)} c_{j,l} c_{j',l'} f_{j,l,i}(0) f_{j',l',i}(0) \prod_{s \neq i} \int_0^\infty f'_{j,l,s}(t_s) f'_{j',l',s}(t_s)\, \mathrm{d}t_s \\
&\quad + 2 \sum_{j=1}^{n} \sum_{l \in \mathcal{U}(j,i)} \sum_{j'=1}^{n} \sum_{l' \in \mathcal{L}(j',i)} c_{j,l} c_{j',l'} f_{j,l,i}(0) f_{j',l',i}(0) \prod_{s \neq i} \int_0^\infty f'_{j,l,s}(t_s) f'_{j',l',s}(t_s)\, \mathrm{d}t_s \\
\mathcal{K}_i &= \sum_{j=1}^{n} \sum_{l \in \mathcal{U}(j,i)} \sum_{j'=1}^{n} \sum_{l' \in \mathcal{U}(j',i)} c_{j,l} c_{j',l'} \prod_{s=1}^{k} \int_0^\infty f'_{j,l,s}(t_s) f'_{j',l',s}(t_s)\, \mathrm{d}t_s
\end{aligned}$$

satisfy the inequality

$$\frac{(1-c_1)\sum_{i=1}^k \mathcal{J}_i - c_2 \sum_{i=1}^k \mathcal{K}_i}{\mathcal{I}} > 1.$$

Choose the sieve weights

$$\nu(n) = \left(\sum_{j=1}^{n} \sum_{l=1}^{L_j} c_{j,l} \lambda_{f_{j,l,1}}(n+h_1) \dots \lambda_{f_{j,l,k}}(n+h_k) \right)^2.$$

Next we consider the ratio

$$\widetilde{N}(x;\rho,\nu) = \frac{\sum_{i=1}^k \sum_{x \le n \le 2x, n \equiv b(W)} \nu(n) \rho(n+h_i;x)}{\sum_{x \le n \le 2x, n \equiv b(W)} \nu(n)}.$$

On the one hand, Theorem 3.6 of [Pol14b] gives

$$\sum_{\substack{x \le n \le 2x \\ n \equiv b(W)}} \prod_{s=1}^{k} \lambda_{f_{j,l,s}}(n+h_s) \lambda_{f_{j',l',s}}(n+h_s) = \left(\prod_{s=1}^{k} \int_0^\infty f'_{j,l,s}(t_s) f'_{j',l',s}(t_s) \mathrm{d}t_s + o(1) \right) \left(\frac{x W^{k-1}}{\varphi(W)^k \log(x)^k} \right).$$

Expanding the square in $\nu(n)$, this gives

$$\sum_{x \le n \le 2x, n \equiv b(W)} \nu(n) = (\mathcal{I} + o(1)) \left(\frac{x W^{k-1}}{\varphi(W)^k \log(x)^k} \right). \tag{4.8}$$

On the other hand, a lower bound for the numerator follows from $\rho(n+h_i;x)+c_2 \geq 0$ and the identity below.

$$\begin{aligned}
\nu(n)\rho(n+h_i;x) \geq{}& \left(\sum_{j=1}^{n}\sum_{l\in\mathcal{L}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)^2 (\rho(n+h_i;x)+c_2) \\
&+ 2\left(\sum_{j=1}^{n}\sum_{l\in\mathcal{L}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)\left(\sum_{j=1}^{n}\sum_{l\in\mathcal{U}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)(\rho(n+h_i;x)+c_2) \\
&- c_2\left(\sum_{j=1}^{n}\sum_{l\in\mathcal{L}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s) + \sum_{j=1}^{n}\sum_{l\in\mathcal{U}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)^2 \\
={}& \left(\sum_{j=1}^{n}\sum_{l\in\mathcal{L}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)^2 \rho(n+h_i;x) \\
&+ 2\left(\sum_{j=1}^{n}\sum_{l\in\mathcal{L}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)\left(\sum_{j=1}^{n}\sum_{l\in\mathcal{U}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)\rho(n+h_i;x) \\
&- c_2\left(\sum_{j=1}^{n}\sum_{l\in\mathcal{U}(j,i)} c_{j,l}\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\right)^2 .
\end{aligned}$$

Lemma 4.9 gives the asymptotic formula

$$\begin{aligned}
&\sum_{\substack{x\leq n\leq 2x\\ n\equiv b(W)}} \rho(n+h_i;x)\prod_{s=1}^{k}\lambda_{f_{j,l,s}}(n+h_s)\lambda_{f_{j',l',s}}(n+h_s) \\
&= \left((1-c_1)f_{j,l,i}(0)f_{j',l',i}(0)\prod_{s\neq i}\int_0^\infty f'_{j,l,s}(t_s)f'_{j',l',s}(t_s)\mathrm{d}t_s + o(1)\right)\left(\frac{xW^{k-1}}{\varphi(W)^k\log(x)^k}\right).
\end{aligned}$$

Summing the lower bound for $\nu(n)\rho(n+h_i;x)$ over $[x,2x]$ and substituting the asymptotic from Lemma 4.9 for terms involving $\rho(n+h_i;x)$, and substituting the asymptotic from Theorem 3.6 of [Pol14b] for terms involving c_2, we then get

$$\sum_{i=1}^{k}\sum_{x\leq n\leq 2x, n\equiv b(W)}\nu(n)\rho(n+h_i;x) = \left((1-c_1)\sum_{i=1}^{k}\mathcal{J}_i - c_2\sum_{i=1}^{k}\mathcal{K}_i + o(1)\right)\left(\frac{xW^{k-1}}{\varphi(W)^k\log(x)^k}\right). \tag{4.9}$$

Combining asymptotics (4.8) and (4.9), we finally get

$$\widetilde{N}(x;\rho,\nu) = \frac{(1-c_1)\sum_{i=1}^{k}\mathcal{J}_i - c_2\sum_{i=1}^{k}\mathcal{K}_i + o(1)}{\mathcal{I}+o(1)} > 1+o(1).$$

Hence, for sufficiently large x, we have the desired inequality $N(x;\rho,\nu) > 0$. ☐

Specializing to $\rho = 1_{\mathbb{P}}$ gives the form of the criterion that we shall use. Here and below we abbreviate $I_T(F) = I(F;\delta,\underline{A},\underline{B},\varepsilon)$ and $J_T(F) = J(F;\delta,\underline{A},\underline{B},\varepsilon)$ for a support T with data $(\delta,\underline{A},\underline{B},\varepsilon)$.

Theorem 4.10 (Direct-prime sieve criterion)**.** *Let T be a Stadlmann support, and suppose that $1_{\mathbb{P}}$ is equidistributed for every modulus generated by T. If a nonzero symmetric $F \in L^2(T)$ satisfies*

$$kJ_T(F) > I_T(F),$$

then DHL$[k,2]$ *holds.*

Proof. Apply Proposition 4.7 with

$$\rho = 1_{\mathbb{P}}, \qquad c_1 = c_2 = 0.$$

Its variational inequality becomes $kJ_T(F) > I_T(F)$. The admissible k-tuple in Definition 4.5 is arbitrary, so the conclusion of that proposition is exactly DHL$[k,2]$. ☐

Theorem 4.10 leaves two things to prove: equidistribution for the moduli generated by the chosen support, and the existence of an F satisfying the strict inequality. The first occupies Sections 5–8 and the second Sections 9–11.

5. The equidistribution estimates

The sieve criterion reduces the arithmetic input to a discrepancy estimate over $Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\varepsilon_0)$, and this section states the estimates from which that discrepancy bound will be assembled. A modulus lying below $x^{1/2}$ is covered by the Bombieri–Vinogradov theorem, whereas a modulus above that level must exhibit divisors of prescribed sizes in order to meet one of the Type I, Type II, or Type III hypotheses below.

The analytic inputs are the Type I/II/III estimates for multiply x^δ-densely divisible moduli of Polymath8a [Pol14a], Stadlmann's estimate for smooth moduli [Sta25], and the Baker–Irving Type I estimate [BI17]. A large smooth divisor does not by itself imply multiple dense divisibility, and this is where Stadlmann's reformulation enters: one retains the particular factorizations that the proofs of those estimates actually use, and states the conclusions directly in terms of the divisors r, u, d_1 that the combinatorial partition of Section 7 can supply. Every such modification is recorded, and the proofs are given in Section 6. The labels I, IIa, IIb, IIc, and III refer to the shape and relative lengths of the convolution variables.

5.1. Coefficient sequences and discrepancy.

We first fix the analytic language. All implicit constants in the divisor bound below are independent of the integer variable and may depend on the fixed sequence.

Definition 5.1 (Coefficient sequences)**.** *A coefficient sequence is a function $\alpha : \mathbb{N}\times(1,\infty)\to\mathbb{R}$ which satisfies $|\alpha(n;x)| \ll \tau(n)^{O(1)}\log(x)^{O(1)}$.*

(i) *α is said to be located at scale $N(x)$ if there are some constants $1\ll c\ll C\ll 1$ such that $\alpha(\,\cdot\,;x)$ is supported on $[cN(x), CN(x)]$ for every $x>1$.*

(ii) *If α is located at scale $N(x)$, it is said to have the Siegel-Walfisz property if*

$$\left|\sum_{\substack{n=a(q)\\(n,r)=1}}\alpha(n;x)-\frac{1}{\varphi(q)}\sum_{(n,qr)=1}\alpha(n;x)\right|\ll_A \tau(qr)^{O(1)}\frac{N(x)}{\log(x)^A}$$

for any $x>1$, any $q,r\geq 1$, any $A>1$, and any residue class a mod q with $(a,q)=1$.

(iii) *α is said to be smooth at scale $N(x)$ if there are some constants $1\ll c\ll C\ll 1$ so that for every $x>1$, there exists a smooth function $\psi:\mathbb{R}\to\mathbb{C}$ supported on $[c,C]$, with $|\psi^{(j)}(t)|\ll_j \log(x)^{O_j(1)}$ for all $t\in\mathbb{R}$, such that $\alpha(n;x)=\psi\left(\frac{n}{N}\right)$.*

For sequences $\alpha,\beta:\mathbb{N}\to\mathbb{C}$, their Dirichlet convolution is

$$(\alpha\star\beta)(n)=\sum_{d|n}\alpha(d)\beta\left(\frac{n}{d}\right).$$

The estimates below use the following discrepancy convention.

Definition 5.2 (Equidistribution)**.** *Consider $f:\mathbb{N}\times(1,\infty)\to\mathbb{C}$. Let $D(x;\omega,\gamma,\delta;\varepsilon)$ be a set of integers dependent on parameters ω, γ and δ, and dependent on a small parameter $\varepsilon>0$.*

For a given set of (ω,γ,δ), we say f has equidistribution properties for moduli in $D(x;\omega,\gamma,\delta;\varepsilon)$, provided the following holds: For any sufficiently small $\varepsilon>0$ and for all $x>1$ and $A>0$, and for all $a\in\mathbb{Z}$ with $(a,p)=1$ for $p\leq x$,

$$\sum_{\substack{d\in D(x;\omega,\gamma,\delta;\varepsilon)\\ d\text{ is squarefree}}}\left|\sum_{\substack{n\in[x,2x]\\ n\equiv a(d)}}f(n;x)-\frac{1}{\varphi(d)}\sum_{\substack{n\in[x,2x]\\(n,d)=1}}f(n;x)\right|\ll_{A,\varepsilon}\frac{x}{\log(x)^A}. \tag{5.1}$$

Note that the implied constant is only allowed to depend on ω, γ and δ insofar that they determine which ε are sufficiently small for (5.1) to hold.

5.2. The five estimates.

Harman's decomposition, which we shall use in Section 7 to construct the prime minorant, sorts its convolution terms by the exponent γ of one distinguished variable. Different parts of the γ-range are accessible to different exponential-sum estimates, and those estimates in turn require different divisors of the modulus. Table 1 maps this correspondence. A row records the convolution shape, the factorization that must be extracted from the modulus, the range in which the estimate is used, and its source. The formal statements that follow quantify all coefficient and residue-class uniformity.

Table 1. The five equidistribution estimates.

estimate	convolution shape	factorization of q	range	source
Type I	$\alpha \star \beta$, β smooth	$r \mid q$	all γ	Baker–Irving
Type IIa	$\alpha \star \beta$	$r \mid q$	lower Type II	Polymath8a
Type IIb	$\alpha \star \beta$	$ru \mid q$	middle Type II	Polymath8a
Type IIc	$\alpha \star \beta$	$q = ruv,\ d_1 \mid r$	upper Type II	Stadlmann
Type III	$\alpha \star \psi_1 \star \psi_2 \star \psi_3$	$r \mid q$	Type III	Polymath8a

In the statements below, γ denotes the logarithmic size of the indicated convolution factor, ω measures how far above $x^{1/2}$ the moduli are allowed to reach, δ is the roughness threshold, and ε is the local retreat. The first estimate is the form of Lemma 5 of Baker and Irving [BI17] needed for the present modulus family.

Lemma 5.3 (Type I). *Suppose $f(n;x) = (\alpha \star \beta)(n;x)$, where α is a coefficient sequence at scale M and β is a smooth coefficient sequence at scale N with $M(x)N(x) \asymp x$. Suppose also that α has the Siegel–Walfisz property; this is used only in the range $\gamma \in (\frac{1}{2}, \frac{1}{2} + 2\omega + \varepsilon]$. Write $N(x) = x^\gamma$. Define the set of moduli*

$$D_I(x;\omega,\gamma,\delta;\varepsilon) = \begin{cases} \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \exists r \mid d \text{ with } x^{\gamma-\delta-3\varepsilon} < r < x^{\gamma-3\varepsilon} \right\} & \text{if } \gamma \le \frac{1}{2}, \\ \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \exists r \mid d \text{ with } x^{1-\gamma-\delta-3\varepsilon} < r < x^{1-\gamma-3\varepsilon} \right\} & \text{if } \gamma \in (\frac{1}{2}, \frac{1}{2} + 2\omega + \varepsilon], \\ [1, x^{1/2+2\omega}] \cap \mathbb{N} & \text{if } \gamma > \frac{1}{2} + 2\omega + \varepsilon. \end{cases}$$

Consider triples (ω, γ, δ) which satisfy the following inequalities:

$$\begin{aligned} &3\gamma - 12\omega - 3\delta > 1 && \text{if} \quad \gamma \le \tfrac{1}{2}, \\ &68\omega + 14\delta < 1 && \text{if} \quad \gamma \in (\tfrac{1}{2}, \tfrac{1}{2} + 2\omega + \varepsilon]. \end{aligned}$$

For such triples (ω, γ, δ), convolution f has equidistribution properties for moduli in $D_I(x;\omega,\gamma,\delta;\varepsilon)$.

The next estimate comes from Theorem 2.8(iv) of [Pol14a]; in the lower Type-II range it takes the place of Theorem 2.8(i), which was used at this point in the original formulation.

Lemma 5.4 (Type IIa). *Suppose $f(n;x) = (\alpha \star \beta)(n;x)$, where α is a coefficient sequence at scale M and β is a coefficient sequence at scale N with $M(x)N(x) \asymp x$ and $N(x) \le x^{1/2}$. Suppose also that β has the Siegel–Walfisz property. Write $N(x) = x^\gamma$. Define the set of moduli*

$$D_{IIa}(x;\omega,\gamma,\delta;\varepsilon) = \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \text{There exists a divisor } r \mid d \text{ with } x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon} \right\}.$$

Consider triples (ω, γ, δ) which satisfy the following inequalities:

$$\begin{aligned} 24\omega + 7\delta - 5\gamma &< -2, \\ 8\omega + 3\delta - \gamma &< 0. \end{aligned}$$

For such triples (ω, γ, δ), convolution f has equidistribution properties for moduli in $D_{IIa}(x;\omega,\gamma,\delta;\varepsilon)$.

In Lemma 5.4, and in each of the lemmas that follow, the permissible triples (ω, γ, δ) are defined by strict inequalities. If one restricts attention to triples satisfying these inequalities with room to spare—say, for a fixed $\varepsilon_1 > 0$, to triples with $24\omega + 7\delta - 5\gamma + \varepsilon_1 < -2$ and $8\omega + 3\delta - \gamma + \varepsilon_1 < 0$—then ε may be chosen uniformly in (ω, γ, δ), and the implied constant in the equidistribution estimate does not depend on the triple. This is how the lemmas will be applied.

The middle Type-II estimate is the corresponding specialization of Theorem 2.8(ii) of [Pol14a].

Lemma 5.5 (Type IIb). *Suppose $f(n;x) = (\alpha \star \beta)(n;x)$, where α is a coefficient sequence at scale M and β is a coefficient sequence at scale N with $M(x)N(x) \asymp x$ and $N(x) \le x^{1/2}$. Suppose also that β has the Siegel–Walfisz property. Write $N(x) = x^\gamma$. Define the set of moduli*

$$D_{IIb}(x;\omega,\gamma,\delta;\varepsilon) = \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \begin{array}{l} \exists r \mid d \ \exists u \mid \frac{d}{r} \text{ with } x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon} \\ \text{and } x^{1/2-\gamma-2\omega-6\varepsilon-\delta} < u < x^{1/2-\gamma-2\omega-6\varepsilon} \end{array} \right\}.$$

Consider triples (ω, γ, δ) which satisfy the following inequalities:

$$24\omega + 7\delta - 3\gamma < -1,$$
$$8\omega + 3\delta - \gamma < 0.$$

For such triples (ω, γ, δ), convolution f has equidistribution properties for moduli in $D_{IIb}(x; \omega, \gamma, \delta; \varepsilon)$.

The upper Type-II estimate is obtained by retaining the factorization requirements in the proof of Theorem 1 of [Sta25].

Lemma 5.6 (Type IIc)**.** *Suppose $f(n; x) = (\alpha \star \beta)(n; x)$, where α is a coefficient sequence at scale M and β is a coefficient sequence at scale N with $M(x)N(x) \asymp x$ and $N(x) \leq x^{1/2}$. Suppose also that β has the Siegel–Walfisz property. Write $N(x) = x^{\gamma}$. Define the set of moduli*

$$D_{IIc}(x; \omega, \gamma, \delta; \varepsilon) = \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \begin{array}{l} \exists r \mid d\ \exists u \mid \frac{d}{r}\ \exists d_1 \mid r \text{ with } x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon} \\ \text{and } x^{1-\gamma-6\varepsilon-\delta} d^{-1} < u < x^{1-\gamma-6\varepsilon} d^{-1} \\ \text{and } r^2 x^{2-\gamma-52\varepsilon-\delta} d^{-4} < d_1 < r^2 x^{2-\gamma-52\varepsilon} d^{-4} \end{array} \right\}.$$

Consider triples (ω, γ, δ) which satisfy the following inequalities:

$$8\omega + 4\delta + 2\gamma < 1,$$
$$32\omega + 10\delta - \gamma < 0,$$
$$48\omega + 16\delta - 4\gamma < -1.$$

For such triples (ω, γ, δ), convolution f has equidistribution properties for moduli in $D_{IIc}(x; \omega, \gamma, \delta; \varepsilon)$.

Finally, the Type-III estimate is the fixed-divisor form of Theorem 2.8(v) of [Pol14a].

Lemma 5.7 (Type III)**.** *Suppose $f(n; x) = (\alpha \star \psi_1 \star \psi_2 \star \psi_3)(n; x)$, where α is a coefficient sequence at scale M and ψ_1, ψ_2 and ψ_3 are smooth coefficient sequences at scales N_1, N_2 and N_3 with*

$$M(x)N_1(x)N_2(x)N_3(x) \asymp x,$$
$$N_1(x)N_2(x),\ N_1(x)N_3(x),\ N_2(x)N_3(x) \gg x^{1-\gamma},$$
$$x^{1-2\gamma} \ll N_1(x), N_2(x), N_3(x) \ll x^{\gamma}.$$

Define the set of moduli

$$D_{III}(x; \omega, \gamma, \delta; \varepsilon) = \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \exists r \mid d \text{ with } x^{1/3+4\delta/3-4\omega/3-\delta} < r < x^{1/3+4\delta/3-4\omega/3} \right\}.$$

Consider triples (ω, γ, δ) which satisfy the following inequality:

$$28\omega + 9\gamma + 8\delta < 4.$$

For such triples (ω, γ, δ), convolution f has equidistribution properties for moduli in $D_{III}(x; \omega, \gamma, \delta; \varepsilon)$.

6. Proofs of the equidistribution estimates

The five lemmas of Section 5 are variants of the results cited there, not restatements of them. This section details the differences. Their common analytic core is Theorem 5.8 of [Pol14a]. The Type I argument uses in addition Lemma 5 of [BI17], and the Type IIc argument Theorem 1 of [Sta25]. In each case we first record which source estimate is being invoked, then match its variables to ours, and only then simplify the resulting exponent inequalities. A reader can therefore see which part of each cited proof is unchanged and which part is supplied here.

The proofs are ordered by increasing complexity of the required factorization. Type I and Type IIa need a single divisor in a prescribed interval. Type IIb needs two divisors at compatible scales. Type IIc is the delicate case: it needs a nested divisor $d_1 \mid r$, and it is responsible for the active parameter wall in the final optimization. Type III again needs one divisor interval, but for a three-variable convolution.

The common reduction. With the exception of Type III, the estimates pass through the reduction in Section 5 of [Pol14a]. That paper is formulated for i-tuply x^δ-densely divisible integers, but its reduction uses only a factorization $m = qr$ with $Nx^{-3\varepsilon-\delta_0} \ll r \ll Nx^{-3\varepsilon}$. We therefore restate its Theorem 5.8 with precisely the relaxed modulus hypothesis needed here. Since the statement involves exponential sums, we first fix notation.

For $q \in \mathbb{N}$ and $n \in \mathbb{Z}$, we define $e_q(n) = \exp(2\pi i n/q)$. For $a, b \in \mathbb{Z}$, we set

$$e_q\left(\frac{a}{b}\right) = \begin{cases} e_q\big(\overline{a/b}\big) & \text{if } a/b \text{ is well-defined mod } q, \\ 0 & \text{otherwise,} \end{cases}$$

where $\overline{a/b}$ represents an integer n which satisfies $n \equiv a/b \bmod q$.

The reduction produces the following family of exponential sums.

Definition 6.1. *For a given triple (ω, γ, δ) with $\gamma - 4\omega - \delta > 0$ and $\gamma \le \frac{1}{2}$ and for $x > 0$, let $Z(\omega, \gamma, \delta; x)$ denote the set of $(Q, R, q_0, b_1, b_2, \ell)$ with the following properties:*

(i) *$R \in (0, \infty)$ with $x^{\gamma-\delta-3\varepsilon} \le R \le x^{\gamma-2\varepsilon}$.*
(ii) *$Q \in (0, \infty)$ with $x^{1/2-\varepsilon} \ll QR \ll x^{1/2+2\omega}$.*
(iii) *$q_0 \in \mathbb{N}$ with $q_0 \ll Q$.*
(iv) *$b_1, b_2 \in \mathbb{Z}$ with $(b_1 b_2, p) = 1$ for all $p \le x$.*
(v) *$\ell \in \mathbb{Z}$ with $1 \le |\ell| \ll \frac{N}{R}$.*

Consider some choice of sets $\mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \subseteq [Q, 2Q] \cap \mathbb{N}$ and $\mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon) \subseteq [R, 2R] \cap \mathbb{N}$.

For a given $z = (Q, R, q_0, b_1, b_2, \ell) \in Z(\omega, \gamma, \delta; x)$, we then set $C(n) = 1_{\frac{b_1}{n} \equiv \frac{b_2}{n+\ell r} \bmod q_0}$ and define

$$\Sigma_{\mathcal{Q},\mathcal{R}}(z) = \sum_{r \in \mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon)} \sum_{\substack{q_1, q_2 \\ q_0 q_1, q_0 q_2 \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \\ (q_1, q_2) = 1 \\ q_0 q_1 r,\, q_0 q_2 r \text{ squarefree}}} \sum \sum_{0 < |h| \le \frac{x^\varepsilon R Q^2}{q_0 M}} \Bigg| \sum_{\substack{n \\ (n, r q_0 q_1) = 1 \\ (n + \ell r, q_0 q_2) = 1}} C(n) \beta(n) \overline{\beta(n + \ell r)} \Phi_\ell(h, n, r, q_0, q_1, q_2) \Bigg|,$$

where $\Phi_\ell(h, n, r, q_0, q_1, q_2) = e_r\left(\frac{ah}{n q_0 q_1 q_2}\right) e_{q_0 q_1}\left(\frac{b_1 h}{n r q_2}\right) e_{q_2}\left(\frac{b_2 h}{(n + \ell r) r q_0 q_1}\right)$.

Lemma 6.2 (Variant of Theorem 5.8 of [Pol14a])**.** *Suppose $f(n; x) = (\alpha \star \beta)(n; x)$, where α is a coefficient sequence at scale M and β is a coefficient sequence at scale N with $M(x)N(x) \asymp x$. Suppose β has the Siegel–Walfisz property. Write $N(x) = x^\gamma$.*

Consider some sets $\mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \subseteq [Q, 2Q] \cap \mathbb{N}$ and $\mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon) \subseteq [R, 2R] \cap \mathbb{N}$.

Let $Z(\omega, \gamma, \delta; x)$ and $\Sigma_{\mathcal{Q},\mathcal{R}}(Q, R, q_0, b_1, b_2, \ell)$ be as defined in Definition 6.1.

Let $D_0 = \exp(\log(x)^{1/3})$ and suppose that $q \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon)$ implies that $(q, p) = 1$ for $p \le D_0$.

Suppose $D(x; \omega, \gamma, \delta; \varepsilon)$ is a set of moduli with

$$D(x; \omega, \gamma, \delta; \varepsilon) \subseteq \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \begin{array}{l} \exists R \exists Q \text{ with } \log_2(Q) \in \mathbb{Z} \text{ and } \log_2(R) \in \mathbb{Z} \\ \text{and } x^{\gamma-\delta-3\varepsilon} \le R \le x^{\gamma-2\varepsilon} \text{ so that} \\ d = qr \text{ for some } q \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \text{ and } r \in \mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon), \\ \text{or } \prod_{p | d, p \le D_0} p > \exp(\log^{2/3} x) \end{array} \right\}.$$

Consider triples (ω, γ, δ) with $\gamma - 4\omega - \delta > 0$ and $\gamma \le \frac{1}{2}$. Suppose a collection of such triples satisfies

$$\sup_{(Q, R, q_0, b_1, b_2, \ell) \in Z(\omega, \gamma, \delta; x)} \Sigma_{\mathcal{Q},\mathcal{R}}(Q, R, q_0, b_1, b_2, \ell) \ll_\varepsilon x^{-2\varepsilon} Q^2 R N (q_0, \ell) q_0^{-2}.$$

Then convolution f has equidistribution properties for moduli in $D(x; \omega, \gamma, \delta; \varepsilon)$.

Proof. Only one preliminary change is required in the proof of Theorem 5.8 of [Pol14a]. There the congruence class a obeys $(a, p) = 1$ for $p \in I$, with I a bounded set depending on x. We take $I = [1, x]$ and replace P_I by $\prod_{p \le x} p$.

The contribution of d with $\prod_{p | d,\, p \le D_0} p > \exp(\log^{2/3} x)$ is bounded trivially. It remains to consider $d = qr$, with $q \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon)$ and $r \in \mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon)$.

The source factors every i-tuply x^δ-densely divisible integer as $d = qr$ with $x^{\gamma-\delta-3\varepsilon} \le r \le x^{\gamma-3\varepsilon}$, then removes from q its prime factors at most D_0. We replace the resulting densely divisible sums by sums over $q \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon)$ and $r \in \mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon)$, where $x^{\gamma-\delta-3\varepsilon} \le R \le x^{\gamma-2\varepsilon}$. Elements of $\mathcal{Q}$ already have no prime factor at most D_0. The conditions $\log_2 Q, \log_2 R \in \mathbb{Z}$ leave only $O(\log^2 x)$ dyadic pairs, and qr remains squarefree.

The required bound $RQ^2 \ll x$ follows from $RQ \ll x^{1/2+2\omega}$ and $RQ^2 \ll x^{1+4\omega}/R \ll x^{1+4\omega-\gamma+\delta+3\varepsilon}$ when $\gamma - 4\omega - \delta > 0$. Here and below ε is chosen sufficiently small in terms of (ω, γ, δ) and is suppressed in the recorded limiting inequalities.

Cauchy–Schwarz introduces two copies q_1, q_2 and their gcd q_0. Relabel them as q_0q_1, q_0q_2, where $(q_1, q_2) = 1$. The resulting sums satisfy $q_0q_1, q_0q_2 \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon)$, and (5.31) in the proof of Theorem 5.8 becomes

$$\Sigma_{\mathcal{Q},\mathcal{R}}(Q, R, q_0, b_1, b_2, \ell) \ll_\varepsilon x^{-2\varepsilon} Q^2 RN(q_0, \ell) q_0^{-2}.$$

□

6.1. **The Type I estimate.** This proof specializes Lemma 5 of [BI17], using the second estimate in Corollary 4.16 of [Pol14a]. The task is to convert the divisor r supplied by the definition of D_I into the factor R required by the source theorem. Once r has been localized dyadically, the complementary factor is forced, and the source inequalities reduce to the two exponent walls in Lemma 5.3.

Proof of Lemma 5.3. We first consider $\gamma \le \frac{1}{2}$. Recall that we work with moduli

$$D_I(x; \omega, \gamma, \delta; \varepsilon) = \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \exists r \mid d \text{ with } x^{\gamma-\delta-3\varepsilon} < r < x^{\gamma-3\varepsilon} \right\}.$$

Accordingly, choose

$$\begin{aligned} \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) &= \{q \in [Q, 2Q] \cap \mathbb{N} : (q, p) = 1 \text{ for } p \le D_0\}, \\ \mathcal{R}(R; x, \omega, \gamma, \delta, \varepsilon) &= [R, 2R] \cap \mathbb{N}. \end{aligned}$$

Applying Lemma 6.2 with this choice of $\mathcal{Q}$ and $\mathcal{R}$, we must now show that

$$\sup_{(Q,R,q_0,b_1,b_2,\ell) \in Z(\omega,\gamma,\delta;x)} \Sigma_{\mathcal{Q},\mathcal{R}}(Q, R, q_0, b_1, b_2, \ell) \ll_\varepsilon x^{-2\varepsilon} Q^2 RN(q_0, \ell) q_0^{-2}.$$

As in the proof of Lemma 3 of [BI17], we observe that by fixing $r \in [R, 2R]$, this problem can be reduced to showing that

$$\begin{aligned} &\sum_{\substack{q_1, q_2 \\ q_0q_1, q_0q_2 \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \\ (q_1, q_2) = 1 \\ q_0q_1r, q_0q_2r \text{ squarefree}}} \sum_{0 < |h| \le \frac{x^\varepsilon RQ^2}{q_0 M}} (q_0, \ell) \sup_{t(q_0)} \left| \sum_{n \equiv t(q_0)} \beta(n) \overline{\beta(n+\ell r)} e_{rq_0q_1}\left(\frac{zh}{nq_2}\right) e_{q_2}\left(\frac{b_2 h}{(n+\ell r) r q_0 q_1}\right) \right| \\ &\ll_\varepsilon x^{-2\varepsilon} Q^2 N(q_0, \ell) q_0^{-2}, \end{aligned}$$

for a suitable choice of $z = z(a, b_1, r, q_0, q_1, q_2)$ with $(z, rq_0q_1) = 1$.

We apply Lemma 6.3 with $\beta(n)\overline{\beta(n+\ell r)}$, t, q_0, rq_0q_1, and q_2 in place of $\psi_N(n)$, a, d, d_1, and d_2, respectively. We also take zh/q_2, $b_2h/(rq_0q_1)$, 0, and ℓr in place of c_1, c_2, l_1, l_2. These are the same substitutions as in [BI17]. Consequently $[d_1, d_2] = rq_0q_1q_2$, $\delta'_1 = rq_1$, and $\delta'_2 = q_2$, while $(c_1, \delta'_1) = (h, rq_1)$ and $(c_2, \delta'_2) = (h, q_2)$. It follows that

$$\begin{aligned} &\sum_{\substack{q_1, q_2 \\ q_0q_1, q_0q_2 \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \\ (q_1, q_2) = 1 \\ q_0q_1r, q_0q_2r \text{ squarefree}}} \sum_{0 < |h| \le \frac{x^\varepsilon RQ^2}{q_0 M}} (q_0, \ell) \sup_{t(q_0)} \left| \sum_{n \equiv t(q_0)} \beta(n) \overline{\beta(n+\ell r)} e_{rq_0q_1}\left(\frac{zh}{nq_2}\right) e_{q_2}\left(\frac{b_2 h}{(n+\ell r) r q_0 q_1}\right) \right| \\ &\ll_\varepsilon \sum_{\substack{q_1, q_2 \\ q_0q_1, q_0q_2 \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \\ (q_1, q_2) = 1 \\ q_0q_1r, q_0q_2r \text{ squarefree}}} \sum_{0 < |h| \le \frac{x^\varepsilon RQ^2}{q_0 M}} (q_0, \ell)(rq_0q_1q_2)^\varepsilon \left((rq_1q_2)^{1/2} + \frac{(h, rq_1q_2)}{rq_0q_1q_2} N \right) \\ &\ll_\varepsilon x^{3\varepsilon} \sum_{\substack{q_1, q_2 \\ q_0q_1, q_0q_2 \in \mathcal{Q}(Q; x, \omega, \gamma, \delta, \varepsilon) \\ (q_1, q_2) = 1 \\ q_0q_1r, q_0q_2r \text{ squarefree}}} (q_0, \ell) \left(\frac{R^{1/2}Q}{q_0} + \frac{q_0 N}{RQ^2} \right) \left(\frac{RQ^2}{q_0 M} \right) \ll_\varepsilon x^{3\varepsilon} \frac{(q_0, \ell)}{q_0^2} \left(\frac{R^{3/2}Q^5}{M} + \frac{Q^2 N}{M} \right). \end{aligned}$$

To obtain the desired upper bound $\ll_\varepsilon x^{-2\varepsilon}Q^2N(q_0,\ell)q_0^{-2}$, it is enough to verify

$$\left(\frac{R^{3/2}Q^3}{MN}+\frac{1}{M}\right)\ll_\varepsilon x^{-5\varepsilon}.$$

Recalling $NM\asymp x$ and $RQ\ll x^{1/2+2\omega}$ and $R\gg x^{\gamma-\delta-3\varepsilon}$, and choosing ε sufficiently small (depending on (ω,γ,δ)), the inequality holds provided that

$$3(1/2+2\omega)-3/2(\gamma-\delta)<1.$$

This rearranges to condition $3\gamma-12\omega-3\delta>1$ and concludes the treatment of $\gamma\le\frac12$.

Consider next $\gamma\in[\frac12,\frac12+2\omega+\varepsilon]$. Interchanging α and β replaces γ by $1-\gamma$ and exchanges the roles of the two sequences; since $1-\gamma\le\frac12$ and α has the Siegel–Walfisz property, Lemma 5.4 applies when

$$24\omega+7\delta-5(1-\gamma)<-2,\qquad 8\omega+3\delta-(1-\gamma)<0.$$

Since $\gamma\le\frac12+2\omega+\varepsilon$, it is enough that

$$\min\{\tfrac35-\tfrac75\delta-\tfrac{24}{5}\omega,1-8\omega-3\delta\}>\tfrac12+2\omega.$$

Rearranging, the condition becomes

$$68\omega+14\delta<1.$$

Finally, if $\gamma>\frac12+2\omega+\varepsilon$, the terminal case of the proof of Theorem 2.8(ii) of [Pol14a], as used by Baker and Irving, gives the desired estimate for every $d\le x^{1/2+2\omega}$ without an additional factorization condition.

Overall we thus have the desired equidistribution properties if $68\omega+14\delta<1$ and $3\gamma-12\omega-3\delta>1$. □

6.2. **The Type IIa estimate.** This is the required extension of Theorem 2.8(iv) of [Pol14a] into the lower Type-II range. As in Type I, one divisor of a suitable size suffices. The only additional point is that the very small prime factors removed in the preliminary decomposition must be absorbed into that divisor; their total size is $x^{o(1)}$, so the reserved ε-width of the interval accommodates them.

Proof of Lemma 5.4. We first apply Lemma 6.2 with a specific choice of $\mathcal{Q}$ and $\mathcal{R}$: Recall first that in Lemma 5.4 we work with moduli

$$D_{IIa}(x;\omega,\gamma,\delta;\varepsilon)=\left\{d\in[1,x^{1/2+2\omega}]\cap\mathbb{N}:\exists r\mid d\text{ with }x^{\gamma-\delta-3\varepsilon}<r<x^{\gamma-3\varepsilon}\right\}.$$

So $d\in D_{IIa}(x;\omega,\gamma,\delta;\varepsilon)$ implies $d=qr$ with $x^{\gamma-\delta-3\varepsilon}<r<x^{\gamma-3\varepsilon}$. Suppose now that $\prod_{p\mid d,p\le D_0}p\le\exp(\log^{2/3}x)$. Then $q=q^*s$ for some q^* with $(q^*,p)=1$ for $p\le D_0$ and some s with $s\ll x^\varepsilon$. Thus we can factorise d as $d=q^*(rs)$ where $x^{\gamma-\delta-3\varepsilon}<rs<x^{\gamma-2\varepsilon}$. Accordingly, choose

$$\begin{aligned}\mathcal{Q}(Q;x,\omega,\gamma,\delta,\varepsilon)&=\{q\in[Q,2Q]\cap\mathbb{N}:(q,p)=1\text{ for }p\le D_0\},\\ \mathcal{R}(R;x,\omega,\gamma,\delta,\varepsilon)&=[R,2R]\cap\mathbb{N}.\end{aligned}$$

Applying Lemma 6.2, we must now show that

$$\sup_{(Q,R,q_0,b_1,b_2,\ell)\in Z(\omega,\gamma,\delta;x)}\Sigma_{\mathcal{Q},\mathcal{R}}(Q,R,q_0,b_1,b_2,\ell)\ll_\varepsilon x^{-2\varepsilon}Q^2RN(q_0,\ell)q_0^{-2}.\tag{6.1}$$

Fixing r, enlarging the conditions $q_0q_1,q_0q_2\in\mathcal{Q}(Q;x,\omega,\gamma,\delta,\varepsilon)$ to $q_0q_1,q_0q_2\in[Q,2Q]$, and linearizing the absolute values by constants c_{h,q_1,q_2} gives condition (5.33) of [Pol14a]. The remainder of the proof of its Theorem 2.8(iv) uses no further modulus factorization and proves (6.1) provided that

$$\begin{aligned}&\frac{x^{1+12\omega+7\delta/2}}{N^{5/2}}\ll_\varepsilon x^{-100\varepsilon},\\ &\frac{x^{8\omega+3\delta}}{N}\ll_\varepsilon x^{-100\varepsilon}.\end{aligned}$$

These are the two inequalities immediately following (5.36) of [Pol14a]. There one imposes $\gamma>\frac12-2\omega-c$; here we instead substitute $N=x^\gamma$ directly. For sufficiently small ε, the preceding inequalities follow from

$$\begin{aligned}&24\omega+7\delta-5\gamma<-2,\\ &8\omega+3\delta-\gamma<0.\end{aligned}$$

Lemma 6.2 also assumes $4\omega+\delta-\gamma<0$. This follows from $8\omega+3\delta-\gamma<0$ in the present parameter range, and therefore imposes no additional restriction. □

6.3. **The Type IIb estimate.** After the reduction to Lemma 6.2, the usual next step [Pol14a, BI17] is the first inequality in Corollary 4.16 of [Pol14a], which uses the q-van der Corput method and requires one further factor of the modulus in a prescribed range. For the present support it is more efficient to use the second inequality of that corollary instead. The resulting exponential-sum bound is weaker, but it requires no additional modulus factorization, and this exchange of analytic strength for a simpler geometric condition is favourable here. We record the inequality first, and then show that the two divisors defining D_{IIb} place us within its range.

Lemma 6.3 (Corollary 4.16 of [Pol14a], inequality 2)**.** *Let $N\geq 1$ and suppose $\psi_N(n)$ is a smooth coefficient sequence at scale N. Let $d_1,d_2\in\mathbb{Z}$ be squarefree, and let $a,c_1,c_2,l_1,l_2\in\mathbb{Z}$. Suppose $d\mid[d_1,d_2]$. Write $\delta_i=d_i/(d_1,d_2)$ and $\delta_i'=\delta_i/(d,\delta_i)$. Then for any $\varepsilon>0$,*

$$\left|\sum_{n\equiv a(d)}\psi_N(n)e_{d_1}\left(\frac{c_1}{n+l_1}\right)e_{d_2}\left(\frac{c_2}{n+l_2}\right)\right|\ll_\varepsilon [d_1,d_2]^\varepsilon\left(\left(\frac{[d_1,d_2]}{d}\right)^{1/2}+\frac{1}{d}\frac{(c_1,\delta_1')}{\delta_1'}\frac{(c_2,\delta_2')}{\delta_2'}N\right).$$

Since this bound involves greatest common divisors, we shall also use the following elementary estimate.

Lemma 6.4 (Lemma 1.4 of [Pol14a])**.** *Let $q\in\mathbb{N}$ and $K\geq 1$. Then*

$$\sum_{1\leq k\leq K}(k,q)\leq K\tau(q).$$

Proof of Lemma 5.5. Recall first that in Lemma 5.5 we work with moduli

$$D_{IIb}(x;\omega,\gamma,\delta;\varepsilon)=\left\{d\in[1,x^{1/2+2\omega}]\cap\mathbb{N}:\begin{array}{l}\exists r\mid d\ \exists u\mid \frac{d}{r}\text{ with } x^{\gamma-3\varepsilon-\delta}<r<x^{\gamma-3\varepsilon}\\ \text{and } x^{1/2-\gamma-2\omega-6\varepsilon-\delta}<u<x^{1/2-\gamma-2\omega-6\varepsilon}\end{array}\right\}.$$

Hence we can take

$$\begin{aligned}\mathcal{Q}(Q;x,\omega,\gamma,\delta,\varepsilon)&=\left\{q\in[Q,2Q]\cap\mathbb{N}:\begin{array}{l}(q,p)=1\text{ for } p\leq D_0\\ \text{and } \exists u\mid q\text{ with } x^{1/2-\gamma-2\omega-7\varepsilon-\delta}<u<x^{1/2-\gamma-2\omega-6\varepsilon}\end{array}\right\},\\ \mathcal{R}(R;x,\omega,\gamma,\delta,\varepsilon)&=[R,2R]\cap\mathbb{N}.\end{aligned}$$

Applying Lemma 6.2, we must now show that

$$\sup_{(Q,R,q_0,b_1,b_2,\ell)\in Z(\omega,\gamma,\delta;x)}\Sigma_{\mathcal{Q},\mathcal{R}}(Q,R,q_0,b_1,b_2,\ell)\ll_\varepsilon x^{-2\varepsilon}Q^2RN(q_0,\ell)q_0^{-2}. \tag{6.2}$$

In the proof of Theorem 2.8(ii) of [Pol14a], the $x^{\delta+o(1)}$-dense divisibility of q_0q_1 is used to obtain a factor u_1 of q_1 in the interval from $q_0^{-1}x^{-\delta-2\varepsilon}Q/H$ to $x^{-2\varepsilon}Q/H$. Our hypothesis $q_0q_1\in\mathcal{Q}(Q;x,\omega,\gamma,\delta,\varepsilon)$ instead gives $q_0q_1=uv$ with

$$x^{1/2-\gamma-2\omega-7\varepsilon-\delta}<u<x^{1/2-\gamma-2\omega-6\varepsilon}.$$

Recalling that q_0q_1r is squarefree and writing $u_1=u/(u,q_0)$ and $v_1=v/(v,q_0)$, we then have $q_1=u_1v_1$ with $q_0^{-1}x^{1/2-\gamma-2\omega-7\varepsilon-\delta}<u_1<x^{1/2-\gamma-2\omega-6\varepsilon}$. Thus (5.39) of [Pol14a] is replaced by

$$q_0^{-1}x^{1/2-\gamma-2\omega-7\varepsilon-\delta}\ll U\ll x^{1/2-\gamma-2\omega-6\varepsilon}.$$

The complementary factor satisfies $UV\asymp Q/q_0$. The condition $\gamma<1/2-2\omega$ ensures, for sufficiently small ε, that the displayed upper bound for U exceeds 1; thus the factorization range is nonempty.

The unchanged Cauchy–Schwarz and completion steps leading to (5.42) reduce the argument to the following estimate:

$$\begin{aligned}\Upsilon_2&:=\sum_{u_1\asymp U}\sum_{q_2\asymp Q/q_0}\sum_{\substack{v_1,v_2\asymp V\\(u_1v_1v_2,q_0q_2)=1}}\sum_{1\leq|h_1|,|h_2|\leq\frac{x^\varepsilon RQ^2}{q_0M}}\sup_{t(q_0)}\left|\sum_{n\equiv t(q_0)}\psi_N(n)\Phi_\ell(h_1,n,r,q_0,u_1v_1,q_2)\overline{\Phi_\ell(h_2,n,r,q_0,u_1v_2,q_2)}\right|\\&\ll_\varepsilon x^{-10\varepsilon}q_0^{-1}Q^2VN.\end{aligned}\tag{6.3}$$

At this stage [Pol14a] applies the first inequality of Corollary 4.16. We instead apply its second inequality, recorded as Lemma 6.3. As there, substitute q_0, $rq_0u_1[v_1, v_2]$, and q_2 for d, d_1, d_2, respectively. The accompanying gcd factors satisfy

$$\frac{(c_2, \delta_2')}{\delta_2'} \leq 1, \qquad \frac{(c_1, \delta_1')}{\delta_1'} \leq \frac{(h_1v_2 - h_2v_1, r)}{r}.$$

Therefore

$$\begin{aligned}
\Upsilon_2 &\ll_\varepsilon x^\varepsilon \sum_{u_1 \asymp U} \sum_{q_2 \asymp Q/q_0} \sum_{\substack{v_1, v_2 \asymp V \\ (u_1v_1v_2, q_0q_2)=1}} \sum_{1 \leq |h_1|, |h_2| \leq \frac{x^\varepsilon RQ^2}{q_0M}} \left((ru_1[v_1, v_2]q_2)^{1/2} + \frac{N}{q_0} \frac{(h_1v_2 - h_2v_1, r)}{r} \right) \\
&\ll_\varepsilon x^{2\varepsilon} q_0^{-3} U^{3/2} V^3 R^{5/2} Q^{11/2} M^{-2} + x^{4\varepsilon} \sum_{u_1 \asymp U} \sum_{q_2 \asymp Q/q_0} \sum_{|\Delta| \ll \frac{x^\varepsilon RQ^2V}{q_0M}} \frac{N(\Delta, r)}{q_0 r} \sum_{v \asymp V} \sum_{1 \leq |h| \leq \frac{x^\varepsilon RQ^2}{q_0M}} \tau(\Delta - hv) \\
&\ll_\varepsilon x^{2\varepsilon} q_0^{-3} U^{3/2} V^3 R^{5/2} Q^{11/2} M^{-2} + x^{10\varepsilon} q_0^{-4} UV^2 RQ^5 NM^{-2} + x^{10\varepsilon} q_0^{-3} UVRQ^3 NM^{-1}.
\end{aligned}$$

The bound (6.3) is therefore satisfied if the following inequalities are true:

$$\begin{aligned}
&q_0^{-2} U^{3/2} V^2 R^{5/2} Q^{7/2} M^{-2} N^{-1} \ll_\varepsilon x^{-100\varepsilon}, \\
&q_0^{-3} UVRQ^3 M^{-2} \ll_\varepsilon x^{-100\varepsilon}, \\
&q_0^{-2} URQM^{-1} \ll_\varepsilon x^{-100\varepsilon}.
\end{aligned}$$

Recall now that $RQ \ll x^{1/2+2\omega}$ and $R \gg x^{\gamma-3\varepsilon-\delta}$ and $q_0^{-1} x^{1/2-\gamma-2\omega-7\varepsilon-\delta} \ll U \ll x^{1/2-\gamma-2\omega-6\varepsilon}$ and $UV \asymp Q/q_0$. Hence it suffices to have (for sufficiently small ε):

$$\begin{aligned}
&-\tfrac{1}{2}(\tfrac{1}{2} - \gamma - 2\omega - \delta) + \tfrac{11}{2}(\tfrac{1}{2} + 2\omega) - 3(\gamma - \delta) + (\gamma - 2) < 0, \\
&4(\tfrac{1}{2} + 2\omega) - 3(\gamma - \delta) - 2(1 - \gamma) < 0, \\
&(\tfrac{1}{2} - \gamma - 2\omega - 6\varepsilon) + (\tfrac{1}{2} + 2\omega) - (1 - \gamma) < 0.
\end{aligned}$$

Rearranging, we obtain our proposed conditions

$$\begin{aligned}
&24\omega + 7\delta - 3\gamma < -1, \\
&8\omega + 3\delta - \gamma < 0.
\end{aligned}$$

The condition $4\omega + \delta - \gamma < 0$ of Lemma 6.2 is again already implied by $8\omega + 3\delta - \gamma < 0$. □

6.4. **The Type IIc estimate.** This is the most constrained of the five estimates. The outer divisor r, its complementary divisor u, and the nested divisor $d_1 \mid r$ are used at three different stages of the exponential-sum argument, and we keep all three intervals unsimplified: no interval endpoint is replaced by a worst-case value during the translation into Stadlmann's variables. We first make that translation, and then check that the three displayed exponent inequalities are exactly what is needed to make every error term smaller than the target discrepancy.

Proof of Lemma 5.6. Recall that we work with moduli

$$D_{IIc}(x; \omega, \gamma, \delta; \varepsilon) = \left\{ d \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \begin{array}{l} \exists r \mid d \ \exists u \mid \frac{d}{r} \ \exists d_1 \mid r \text{ with } x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon} \\ \text{and } x^{1-\gamma-6\varepsilon-\delta} d^{-1} < u < x^{1-\gamma-6\varepsilon} d^{-1} \\ \text{and } r^2 x^{2-\gamma-52\varepsilon-\delta} d^{-4} < d_1 < r^2 x^{2-\gamma-52\varepsilon} d^{-4} \end{array} \right\}.$$

We thus choose the following sets $\mathcal{Q}(Q, R; x, \omega, \gamma, \delta, \varepsilon)$ and $\mathcal{R}(Q, R; x, \omega, \gamma, \delta, \varepsilon)$:

$$\mathcal{Q}(Q, R; x, \omega, \gamma, \delta, \varepsilon) = \left\{ q \in [Q, 2Q] \cap \mathbb{N} : \begin{array}{l} (q, p) = 1 \text{ for } p \leq D_0 \\ \text{and } \exists u \mid q \text{ with } x^{1-\gamma-6\varepsilon-\delta}(QR)^{-1} \ll u \ll x^{1-\gamma-6\varepsilon}(QR)^{-1} \end{array} \right\},$$

$$\mathcal{R}(Q, R; x, \omega, \gamma, \delta, \varepsilon) = \left\{ r \in [R, 2R] \cap \mathbb{N} : \ \exists d_1 \mid r \text{ with } x^{2-\gamma-52\varepsilon-\delta}(Q^4R^2)^{-1} \ll d_1 \ll x^{2-\gamma-52\varepsilon}(Q^4R^2)^{-1} \right\}.$$

For any $d \in D_{IIc}(x; \omega, \gamma, \delta; \varepsilon)$ with $\prod_{p|d, p \leq D_0} p \leq \exp(\log^{2/3} x)$ there exist Q and R with $\log_2(Q) \in \mathbb{Z}$ and $\log_2(R) \in \mathbb{Z}$ and $x^{\gamma-3\varepsilon-\delta} \ll R \ll x^{\gamma-2\varepsilon}$ so that $d = qr$ for some $q \in \mathcal{Q}(Q, R; x, \omega, \gamma, \delta, \varepsilon)$ and some $r \in \mathcal{R}(Q, R; x, \omega, \gamma, \delta, \varepsilon)$.

We modify the proof of Theorem 1 of [Sta25]. Its initial reduction follows the proof of Theorem 2.8(ii) of [Pol14a], so the preceding transformation applies. The source uses $d \mid P(x^\delta)$ to write $d = qr$ with $r \in [R, 2R]$ and $x^{-4\varepsilon-\delta}N \ll R \ll x^{-2\varepsilon}N$. The definition of $D(x;\omega,\gamma,\delta;\varepsilon)$ supplies the same factorization here. Accordingly, replace the conditions $q_1 \asymp Q$, $r \asymp R$, and $q_1 r \mid P(x^\delta)$ by $q_1 \in \mathcal{Q}(Q,R;x,\omega,\gamma,\delta,\varepsilon)$, $r \in \mathcal{R}(Q,R;x,\omega,\gamma,\delta,\varepsilon)$, and $q_1 r$ squarefree. Replace $(ab_1b_2, P(x^\delta)) = 1$ by $(ab_1b_2, P(x)) = 1$.

The source next uses $q_1 \mid P(x^\delta)$ to extract a factor of size U. Here $q_0q_1 \in \mathcal{Q}(Q,R;x,\omega,\gamma,\delta,\varepsilon)$ gives $q_0q_1 = uv$ for some u with $x^{1-\gamma-6\varepsilon-\delta}(QR)^{-1} \ll u \ll x^{1-\gamma-6\varepsilon}(QR)^{-1}$. Writing $u_1 = u/(u,q_0)$ and $v_1 = v/(v,q_0)$ gives $q_1 = u_1v_1$ for some $u_1 \asymp U$ with $q_0^{-1}x^{1-\gamma-6\varepsilon-\delta}(QR)^{-1} \ll U \ll x^{1-\gamma-6\varepsilon}(QR)^{-1}$. Since $H = x^\varepsilon RQ^2/(q_0M)$, this is

$$q_0^{-2}x^{-\delta-5\varepsilon}Q/H \ll U \ll q_0^{-1}x^{-5\varepsilon}Q/H$$

Hence we also get $x^{5\varepsilon}H \ll V \ll q_0x^{\delta+5\varepsilon}H$. Thus the ranges of U, V, and R in the summary lemma and its associated definition are replaced by the displayed ranges.

The proof then uses $r \mid P(x^\delta)$ to extract from r a factor d_1 of size $D = N/(x^{50\varepsilon}H^2)$. Instead, $r \in \mathcal{R}(R;x,\omega,\gamma,\delta,\varepsilon)$ gives $r = d_1r_1$ for some $d_1 \asymp \Delta$ with $x^{2-\gamma-52\varepsilon-\delta}(Q^4R^2)^{-1} \ll \Delta \ll x^{2-\gamma-52\varepsilon}(Q^4R^2)^{-1}$. Equivalently,

$$\frac{x^{-\delta}N}{q_0^2x^{50\varepsilon}H^2} \ll \Delta \ll \frac{N}{q_0^2x^{50\varepsilon}H^2}.$$

In particular, we have replaced $D \approx N/H^2$ by $D \approx q_0^{-2}N/H^2$, so that the value of D no longer depends on q_0. The ensuing q-van der Corput argument is unchanged after these substitutions except in the diagonal estimate. Since V has increased by a factor q_0, replace $\max\{x^{\delta+10\varepsilon}H^3/(w_2,m), H^4/(w_2,m)\}$ by $\max\{x^{\delta+10\varepsilon}H^3q_0^2/(w_2,m), H^4/(w_2,m)\}$ in the upper bound for Σ_3. This extra factor q_0^2 persists in the subsequent estimates.

Definition 7 and Lemma 8 of [Sta25] summarize the reduction. The extra factors of q_0 require the following amendments: $v_1, v_2 \ll x^{\delta+5\varepsilon}H$ becomes $v_1, v_2 \ll q_0x^{\delta+5\varepsilon}H$, while $N/(x^{\delta+55\varepsilon}H^2) \ll \Delta_1 \ll N/(x^{55\varepsilon}H^2)$ becomes $\frac{N}{q_0^2x^{\delta+55\varepsilon}H^2} \ll \Delta_1 \ll \frac{N}{q_0^2x^{55\varepsilon}H^2}$. Further, $\frac{RQ^2H}{q_0(v_1,v_2)\Delta_1} \ll m \ll \frac{x^\delta RQ^2H}{q_0(v_1,v_2)\Delta_1}$ becomes $\frac{RQ^2H}{q_0(v_1,v_2)\Delta_1} \ll m \ll \frac{x^\delta RQ^2H}{(v_1,v_2)\Delta_1}$, and $|\Lambda| \ll \frac{x^{\delta+5\varepsilon}H^2}{w_1(v_1,v_2)}$ becomes $|\Lambda| \ll \frac{q_0x^{\delta+5\varepsilon}H^2}{w_1(v_1,v_2)}$. The remainder of the argument depends on Δ_1, Λ, and m only through these displayed ranges. In Lemma 10 of [Sta25], $\Delta^* = \min\{N/(\Lambda x^{5\varepsilon}), \Delta_1\}$. For our new choice of Δ_1 and ε sufficiently small compared with δ, we have $\frac{N}{\Lambda x^{5\varepsilon}} \gg \frac{w_1(v_1,v_2)N}{q_0x^{\delta+10\varepsilon}H^2} \gg \frac{N}{q_0x^{55\varepsilon}H^2} \ll \Delta_1$, so we may take $\Delta^* = \Delta_1$.

It remains to carry the extra factors q_0 through equations (3.30)–(3.32) of [Sta25]. They become

$$\frac{N\Delta_1}{m} \ll \min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3q_0^2}, \frac{1}{H^4}\right\}\frac{(v_1,v_2)q_0^2(q_0,\ell)N\Delta_1}{x^{\delta+131\varepsilon}H^2}, \tag{6.4}$$

$$N \ll \min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3q_0^2}, \frac{1}{H^4}\right\}\frac{(v_1,v_2)q_0^2(q_0,\ell)N\Delta_1}{x^{\delta+131\varepsilon}H^2}, \tag{6.5}$$

$$m \ll \min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3q_0^2}, \frac{1}{H^4}\right\}\frac{(v_1,v_2)q_0^2(q_0,\ell)N\Delta_1}{x^{\delta+131\varepsilon}H^2}. \tag{6.6}$$

We begin with (6.4). This condition rearranges to

$$\left(\frac{x^{\delta+131\varepsilon}}{(v_1,v_2)q_0^2(q_0,\ell)}\right)\frac{\max\left\{q_0^2x^{\delta+10\varepsilon}H^5, H^6\right\}}{m} \ll 1. \tag{6.7}$$

Recalling $m \gg \frac{RQ^2H}{q_0(v_1,v_2)\Delta_1} \gg \frac{q_0x^{55\varepsilon}RQ^2H^3}{(v_1,v_2)N} \gg \frac{q_0x^{54\varepsilon}MH^4}{(v_1,v_2)N}$, we get that the LHS of (6.7) is bounded by

$$\left(\frac{x^{\delta+131\varepsilon}}{(v_1,v_2)q_0^2(q_0,\ell)}\right)\frac{\max\left\{q_0^2x^{\delta+10\varepsilon}H^5, H^6\right\}}{m} \ll x^{-1+\delta+77\varepsilon}N^2\max\left\{x^{\delta+10\varepsilon}H, H^2\right\} \ll x^{-1+8\omega+3\delta+100\varepsilon}N^2.$$

This gives us the condition $8\omega + 4\delta + 2\gamma < 1$.

Next we consider (6.5). Rearranging, we find that this condition holds if

$$\frac{x^{\delta+131\varepsilon}}{q_0^2\Delta_1}\max\left\{q_0^2x^{\delta+10\varepsilon}H^5, H^6\right\} \ll 1.$$

Recalling that $\Delta_1 \gg \frac{N}{q_0^2 x^{\delta+55\varepsilon} H^2}$ and $H \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0}$, we get that (6.5) holds if $32\omega + 10\delta - \gamma < 0$.

Finally, we consider (6.6). Recall that $m \ll \frac{x^\delta R Q^2 H}{(v_1,v_2)\Delta_1} \ll \frac{q_0 x^\delta M H^2}{(v_1,v_2)\Delta_1}$. So (6.6) is true if

$$\frac{x^{1+2\delta+131\varepsilon}}{N^2\Delta_1^2} \max\left\{q_0^2 x^{\delta+10\varepsilon} H^7, H^8\right\} \ll 1.$$

Using $H \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0}$, we see that (6.6) holds if the following two inequalities are satisfied:

$$\frac{x^{1+32\omega+10\delta+200\varepsilon}}{N^2}\left(\frac{x^{5\varepsilon}\Lambda}{\Delta_1 N}\right) \ll 1 \quad \text{and} \quad \frac{x^{1+32\omega+10\delta+200\varepsilon}}{N^2}\left(\frac{1}{\Delta_1^2}\right) \ll 1. \tag{6.8}$$

Using $\Lambda \ll x^{\delta+5\varepsilon} H^2$ and $\Delta_1 \gg \frac{N}{x^{\delta+55\varepsilon}H^2}$, it suffices to have $48\omega + 16\delta - 4\gamma < -1$.

Finally, the standing hypothesis $\gamma - 4\omega - \delta > 0$ of Lemma 6.2 holds in this range: the second displayed inequality gives $\delta < \gamma/10 - 32\omega/10$, whence $\gamma - 4\omega - \delta > \frac{9}{10}\gamma - \frac{4}{5}\omega > 0$. □

6.5. **The Type III estimate.** Theorem 2.8(v) of [Pol14a] is stated using factorizations $d = qr$ over a range of values $r \asymp R$, but only one such value is consumed by the final exponential-sum estimate. We therefore isolate that value and formulate the argument with a fixed R, which produces precisely the divisor interval in Lemma 5.7. In translating notation, our γ corresponds to $\frac{1}{2} - \sigma$ in [Pol14a].

Proof of Lemma 5.7. The reduction to exponential sums in the proof of Theorem 2.8(v) of [Pol14a] does not use a modulus factorization; only the final treatment must be modified. Immediately after (7.17), that proof sets $b = \prod_{p \mid b_1b_2b_3} p$ and uses $q = bd$ to infer $d \in \mathcal{D}_I(bx^\delta)$. In our setting we instead have

$$D_{III}(x;\omega,\gamma,\delta;\varepsilon) = \left\{q \in [1, x^{1/2+2\omega}] \cap \mathbb{N} : \exists r \mid q \text{ with } x^{1/3+4\delta/3-4\omega/3-\delta} < r < x^{1/3+4\delta/3-4\omega/3}\right\}.$$

We thus define the set

$$\mathcal{D}'(b;x;\omega,\gamma,\delta) = \left\{d \in \mathbb{N} : \exists r \mid d \text{ with } b^{-1}x^{1/3+4\delta/3-4\omega/3-\delta} < r < x^{1/3+4\delta/3-4\omega/3}\right\}$$

Consider a squarefree $q \in D(x;\omega,\gamma,\delta;\varepsilon)$. We have $q = rs$ for some r with $x^{1/3+4\delta/3-4\omega/3-\delta} < r < x^{1/3+4\delta/3-4\omega/3}$. If also $q = bd$, we take $r' = r/(r,b)$ and $s' = s/(s,b)$, so that $d = r's'$ with $b^{-1}x^{1/3+4\delta/3-4\omega/3-\delta} < r' < x^{1/3+4\delta/3-4\omega/3}$. Hence we replace condition $d \in \mathcal{D}_I(bx^\delta)$ by condition $d \in \mathcal{D}'(b;x;\omega,\gamma,\delta)$.

The preceding reductions leave the estimate

$$\Sigma_2' = \frac{N_1N_2N_3}{Q^2} \sum_{(b_1,b_2,b_3)} \frac{b_1^\flat b_2^\flat b_3^\flat}{b^2} T(b_1,b_2,b_3) \ll_\varepsilon x^{1-3\varepsilon} \tag{6.9}$$

$$\text{where } T(b_1,b_2,b_3) = \sum_{0<|\ell|\le \frac{x^{3\varepsilon/2}Q^3}{b_1b_2b_3N}} \left| \sum_{\substack{d \asymp Q/b \\ d \in \mathcal{D}'(b;x;\omega,\gamma,\delta) \\ d \text{ squarefree} \\ (b\ell,d)=1}} \eta'_{bd} \sum_{(m,bd)=1} \alpha(m) \mathrm{Kl}_3\left(\frac{a\ell b_1b_2b_3}{b^3 m}; d\right)\right|$$

Here $x^{1/2-\varepsilon} \ll Q \ll x^{1/2+2\omega}$, η'_{bd} is some sequence of coefficients supported on short interval $\mathcal{I}(Q)$ and $b = \prod_{p\mid b_1b_2b_3} p$ for some $b_i \in \mathbb{N}$ with $b \ll Q$ and $b_1b_2b_3 \ll x^{3\varepsilon/2}Q^3/N$. Further, $b_i^\flat$ is the largest squarefree divisor of b_i and $b = (b_1b_2b_3)^\flat$.

The source uses $d \in \mathcal{D}_I(bx^\delta)$ to write $d = rs$ with $(bx^\delta)^{-1}S \le s \le S$. Our condition $d \in \mathcal{D}'(b;x;\omega,\gamma,\delta)$ gives the same conclusion with $S = x^{1/3+4\delta/3-4\omega/3}$. The requirement $1 \le S \le x^\delta Q/2$ follows when $4\omega - 4\delta < 1$ and $2\delta - 8\omega < 1$.

The subsequent argument is uniform in S. With $H_b = x^{3\varepsilon/2}H/(b_1b_2b_3) = x^{3\varepsilon/2}Q^3/(b_1b_2b_3N)$ and $N = N_1N_2N_3$, it gives $|T(b_1,b_2,b_3)|^2 \le T_1T_2$, where $T_1 \ll_\varepsilon x^{3\varepsilon}Q^3/(b_1b_2b_3N)$ and

$$T_2 = T_2' + T_2'' \ll_\varepsilon \frac{x^{3\varepsilon}MQ^4S}{bb_1b_2b_3N} + \frac{x^{3\varepsilon}M^2Q^4S^{1/2}}{bb_1b_2b_3N} + \frac{x^{3\varepsilon}x^{\delta/2}M^2Q^3}{b^{5/2}S^{1/2}}.$$

Multiplying T_1 and T_2 together, substituting $S = x^{1/3+4\delta/3-4\omega/3}$ and $Q \ll x^{1/2+2\omega}$ and recalling that $x^{3/2-3\gamma/2} \ll N \ll x^{3\gamma}$, we have

$$\begin{aligned}|T(b_1,b_2,b_3)| &\ll_\varepsilon \frac{x^{3\varepsilon}M^{1/2}Q^{7/2}S^{1/2}}{b^{1/2}b_1b_2b_3N} + \frac{x^{3\varepsilon}MQ^{7/2}S^{1/4}}{b^{1/2}b_1b_2b_3N} + \frac{x^{3\varepsilon}x^{\delta/4}MQ^3}{b^{5/4}(b_1b_2b_3)^{1/2}S^{1/4}N^{1/2}} \\ &\ll_\varepsilon \left(\frac{x^{3\varepsilon}Q^2}{N_1N_2N_3}\right)\left(\frac{x^{1/2+3\omega+3\gamma/4}S^{1/2}}{b^{1/2}b_1b_2b_3} + \frac{x^{1/4+3\omega+3\gamma/2}S^{1/4}}{b^{1/2}b_1b_2b_3} + \frac{x^{\delta/4+3/4+2\omega+3\gamma/4}}{b^{5/4}(b_1b_2b_3)^{1/2}S^{1/4}}\right) \\ &\ll_\varepsilon \left(\frac{x^{3\varepsilon}Q^2}{N_1N_2N_3}\right)\left(\frac{x^{2/3+7\omega/3+3\gamma/4+2\delta/3}}{b(b_1b_2b_3)^{1/2}} + \frac{x^{1/3+8\omega/3+3\gamma/2+\delta/3}}{b(b_1b_2b_3)^{1/2}} + \frac{x^{7\omega/3+3\gamma/4+2/3-\delta/12}}{b(b_1b_2b_3)^{1/2}}\right).\end{aligned}$$

Substituting this upper bound on $|T(b_1,b_2,b_3)|$ back into (6.9), we get

$$\Sigma_2' \ll_\varepsilon x^{3\varepsilon}\left(x^{2/3+7\omega/3+3\gamma/4+2\delta/3} + x^{1/3+8\omega/3+3\gamma/2+\delta/3} + x^{7\omega/3+3\gamma/4-5/6-\delta/12}\right)\sum_{(b_1,b_2,b_3)} \frac{b_1^\flat b_2^\flat b_3^\flat}{b^3(b_1b_2b_3)^{1/2}}.$$

But by Lemma 7.5 of [Pol14a], the inner sum over (b_1,b_2,b_3) is convergent. Therefore we have $\Sigma_2' \ll_\varepsilon x^{1-3\varepsilon}$ provided the following three inequalities hold:

$$\begin{aligned}2/3+7\omega/3+3\gamma/4+2\delta/3 &< 1,\\ 1/3+8\omega/3+3\gamma/2+\delta/3 &< 1,\\ 7\omega/3+3\gamma/4+2/3-\delta/12 &< 1.\end{aligned}$$

The second and third inequality are already implied by the first, and so this leaves us with the condition $28\omega+9\gamma+8\delta<4$. □

7. From the support to the required factorizations

The sieve criterion requires a prime minorant together with equidistribution over the moduli generated by the support. The estimates of Section 5 supply that equidistribution under the assumption that each such modulus is known to admit a prescribed factorization. Now we join the two. We construct the minorant from Harman's identity, then prove the factor-extraction lemmas that turn the geometry of $T_k(\delta,\underline{A},\underline{B},\varepsilon)$ into the divisor configurations the estimates need. The result is Proposition 7.8.

7.1. **Harman's sieve and the prime minorant.** The criterion permits a lower-bound weight ρ in place of the prime indicator, provided ρ decomposes into convolution sequences to which the estimates of Section 5 apply; Harman's identity supplies such a decomposition. We first define the three convolution classes, then state the general minorant, and finally specialize to $\xi_2 = 2/5$, where the exceptional Buchstab sums are empty and the minorant is exactly $1_{\mathbb{P}}$. We keep the general statement because it shows which part of the argument would change at a different parameter point.

Definition 7.1 (Decomposition)**.** *Let $\epsilon = 10^{-10}$ and let $\xi_1,\xi_2,\xi_3 \in (0,1)$. Then $\mathcal{H}(\xi_1,\xi_2,\xi_3)$ denotes the set of functions $f : \mathbb{N}\times(1,\infty) \to \mathbb{C}$ which satisfy one of the following three conditions:*

- (I) *$f(n;x) = (\alpha\star\beta)(n;x)$, where α is a coefficient sequence at scale M with the Siegel–Walfisz property and β is a smooth coefficient sequence at scale N with $M(x)N(x) \asymp x$. Additionally, $N(x) = x^{\gamma(x)}$ with $\gamma(x) \ge \xi_1 - \epsilon$.*
- (II) *$f(n;x) = (\alpha\star\beta)(n;x)$, where α is a coefficient sequence at scale M, β is a coefficient sequence at scale N with $M(x)N(x) \asymp x$, and α and β both have the Siegel–Walfisz property. Additionally, $N = x^{\gamma(x)}$ with $\xi_2 - \epsilon \le \gamma(x) \le 1-\xi_2+\epsilon$.*
- (III) *$f(n;x) = (\alpha\star\psi_1\star\psi_2\star\psi_3)(n;x)$, where α is a coefficient sequence at scale M and ψ_1, ψ_2 and ψ_3 are smooth coefficient sequences at scales N_1, N_2 and N_3 with $M(x)N_1(x)N_2(x)N_3(x) \asymp x$. Additionally, $x^{1-2\xi_3-\epsilon} \le N_1(x), N_2(x), N_3(x) \le x^{\xi_3+\epsilon}$ and $N_i(x)N_j(x) \ge x^{1-\xi_3-\epsilon}$ for $i \ne j$.*

The five scalar inequalities in the hypotheses of the next proposition permit the use of Harman's decomposition. The discrepancy estimate makes valid every convolution term that the decomposition produces.

Proposition 7.2 (General Harman minorant)**.** *Let $T_k(\delta,\underline{A},\underline{B},\varepsilon)$ be as given in Definition 4.1, let $Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\varepsilon_0)$ be as given in Definition 4.3 and let $\mathcal{H}(\xi_1,\xi_2,\xi_3)$ be as given in Definition 7.1.*

Suppose $\xi_1, \xi_2, \xi_3 \in (0,1)$ *satisfy all of the following inequalities:*

$$
\begin{aligned}
&2\xi_1 + 3\xi_2 < 2,\\
&\xi_2 \le \xi_3,\\
&\xi_1 + 9\xi_2 < 4,\\
&2\xi_1 + \xi_2 > 1,\\
&17\xi_2 < 7.
\end{aligned}
$$

Suppose further that for every $f \in \mathcal{H}(\xi_1, \xi_2, \xi_3)$, $x > 1$, $\varepsilon_0 > 0$, $A > 0$ *and* $a \in \mathbb{Z}$ *with* $(a, P(x)) = 1$,

$$
\sum_{\substack{q \in Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\varepsilon_0)\\ q \text{ is squarefree}}} \left| \sum_{\substack{n\in[x,2x]\\ n\equiv a(q)}} f(n;x) - \frac{1}{\varphi(q)} \sum_{\substack{n\in[x,2x]\\ (n,q)=1}} f(n;x) \right| \ll_{A,\varepsilon_0} \frac{x}{\log(x)^A}. \tag{7.1}
$$

We then write $p_j = x^{\alpha_j}$ *and define the minorant*

$$
\rho(n;x) = 1_{\mathbb{P}}(n) - \sum_{\substack{n=p_1p_2p_3p_4p_5\\ 1-2\xi_2<\alpha_4<\alpha_3<\alpha_2<\alpha_1<\xi_2\\ \alpha_1+\alpha_2<\xi_2\\ \alpha_2+\alpha_3+\alpha_4>1-\xi_2\\ \alpha_5>\alpha_4}} 1 - \sum_{\substack{n=p_2p_3p_4p_5p_6\\ 1-2\xi_2<\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6<8\xi_2-3\\ \alpha_2,\alpha_4>\alpha_3\\ \alpha_2+\alpha_4<\xi_2\\ \alpha_3+\alpha_4+\alpha_5>1-\xi_2\\ \alpha_6>\alpha_5}} 1.
$$

Furthermore, we choose the constants

$$
\begin{aligned}
c_1 = &\int_{1-2\xi_2}^{\xi_2} \int_{1-2\xi_2}^{\min\{\alpha_1,\xi_2-\alpha_1\}} \int_{1-2\xi_2}^{\alpha_2} \int_{\max\{1-2\xi_2,1-\xi_2-\alpha_2-\alpha_3\}}^{\min\{\alpha_3,(1-\alpha_1-\alpha_2-\alpha_3)/2\}} \frac{1}{\alpha_1\alpha_2\alpha_3\alpha_4(1-\sum_{i=1}^4 \alpha_i)}\, d\alpha_1\, d\alpha_2\, d\alpha_3\, d\alpha_4\\
&+ \int_{1-2\xi_2}^{8\xi_2-3} \int_{1-2\xi_2}^{\min\{8\xi_2-3,\alpha_2\}} \int_{\max\{\alpha_3,1-2\xi_2\}}^{\min\{8\xi_2-3,\xi_2-\alpha_2\}} \int_{\max\{1-2\xi_2,1-\xi_2-\alpha_3-\alpha_4\}}^{\min\{8\xi_2-3,(1-\sum_{i=2}^4 \alpha_i)/2\}} \frac{1}{\prod_{i=2}^5 \alpha_i(1-\sum_{i=2}^5 \alpha_i)}\, d\alpha_2\, d\alpha_3\, d\alpha_4\, d\alpha_5,
\end{aligned}
$$

$$
c_2 = \begin{cases} 24 & \text{if } \xi_2 > \frac{4}{10},\\ 0 & \text{if } \xi_2 \le \frac{4}{10}. \end{cases}
$$

Then $\rho(n;x)$ *has the following properties:*

(1) *Prime minorant:* $-c_2 \le \rho(n;x) \le 1_{\mathbb{P}}(n)$ *for all large* x *and* $n \in [x, 2x]$.
(2) ρ *has equidistribution properties for moduli corresponding to* $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$*; see Definition 4.4.*
(3) *If* $\rho(n;x) \neq 0$*, then all prime factors of* n *exceed* $x^{1-2\xi_2}$.
(4) *The sum of* $\rho(n;x)$ *over* $[x,2x]$ *satisfies* $\sum_{n\in[x,2x]} \rho(n;x) = (1 - c_1 + o(1))\frac{x}{\log(x)}$.

Proof. The construction of $\rho(n;x)$ follows Baker and Irving [BI17] and Section 4 of [Sta25], with their numerical cutoffs replaced by parameters. Specifically, replace $0.33856, 0.40481, 0.59519, 0.19, 0.405, 0.595$ by $\xi_1, \xi_2, 1-\xi_2, 1-2\xi_3, \xi_3, 1-\xi_3$, respectively. In Lemma 19 of [Sta25] this gives $\alpha = \xi_2$, $\beta = 1 - 2\xi_2 = \lambda$, $M = x^{1-\xi_1}$, and $\zeta = 1 - \xi_1 - \xi_2$. Thus 0.19038 and 0.25663 become $1-2\xi_2$ and $1-\xi_1-\xi_2$. The Buchstab decomposition in that lemma, under the assumption

$$
2\zeta = 2(1-\xi_1-\xi_2) > \xi_2,
$$

we then arrive at

$$
1_{\mathbb{P}}(n) = \theta_0(n;x) + \sum_{\substack{n=p_1p_2n_3\\ 1-2\xi_2\le\alpha_2<\alpha_1<\xi_2\\ \alpha_1+\alpha_2>1-\xi_2\\ \alpha_2>1-\xi_1-\xi_2}} \psi(n_3,p_2) + \sum_{\substack{n=p_1p_2p_3p_4n_5\\ 1-2\xi_2\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<\xi_2\\ \alpha_1+\alpha_2<\xi_2\\ \alpha_3<1-\xi_1-\xi_2}} \psi(n_5,p_4) - \sum_{\substack{n=p_1p_2p_3n_4\\ 1-2\xi_2\le\alpha_3<\alpha_2<\alpha_1<\xi_2\\ \alpha_1+\alpha_2>1-\xi_2\\ \alpha_2<1-\xi_1-\xi_2}} \psi(n_4,p_3), \tag{7.2}
$$

where $\theta_0(n;x)$ has equidistribution properties for moduli corresponding to $T_k(\delta, \underline{A}, \underline{B}, \varepsilon)$.

Lemma 20 of [Sta25] treats the first sum in (7.2). Its proof remains valid for the general parameters provided the following conditions hold. First, n_3 must be prime. Since $\alpha_1 + \alpha_2 > 1 - \xi_2$, a composite n_3 would satisfy $n_3 > x^{2(1-\xi_1-\xi_2)}$. Therefore, it is enough that $2(1-\xi_1-\xi_2) > \xi_2$. The Type-III step further requires

$$
\xi_2 \le \xi_3.
$$

Furthermore, the exponent variables μ_i attached there to the ordered prime factors satisfy $\mu_2 + \mu_3 < \xi_2$. The inequality $1 - \xi_1 - \xi_2/2 > \xi_2$ then gives $\mu_4, \mu_5 > 1 - 2\xi_2$; this is exactly the notation and ordering used in the proof of Lemma 20 of [Sta25]. Next, $\mu_1 + (\mu_6 + \cdots + \mu_j) < 1 - \xi_1 - \xi_2$ follows from $1 - 4(1-2\xi_2) < 1 - \xi_1 - \xi_2$, which rearranges to

$$\xi_1 + 9\xi_2 < 4.$$

Hence, the proof of that lemma requires $2\xi_1 + 3\xi_2 < 2$, $\xi_2 < \xi_3$, and $\xi_1 + 9\xi_2 < 4$.

Lemma 21 of [Sta25] treats the second sum in (7.2). To use its proof for general ξ_1, ξ_2, and ξ_3, we track the inequalities at its two case splits. The first split is according as $\alpha_2 + \alpha_3 + \alpha_4 \le 1 - \xi_2$ or $\alpha_2 + \alpha_3 + \alpha_4 > 1 - \xi_2$. In the first case, $3(1 - 2\xi_2) > \xi_2$ is required for the desired equidistribution estimate; equivalently, we require $7\xi_2 < 3$.

In the second case, we use $\xi_2/2 < 1 - \xi_1 - \xi_2$ to drop the assumption $\alpha_3 < 1 - \xi_1 - \xi_2$, and $\alpha_2 + \alpha_3 + \alpha_4 > 1 - \xi_2$ means that we need $(4/3)(1 - \xi_2) + 2(1 - 2\xi_2) > 1$ to deduce that n_5 is prime. The condition $(4/3)(1 - \xi_2) + 2(1 - 2\xi_2) > 1$ rearranges to $16\xi_2 < 7$, which is already implied by $7\xi_2 < 3$. Therefore, we are left with the sum

$$\Gamma_2(n; x) = \sum_{\substack{n = p_1p_2p_3p_4p_5 \\ 1-2\xi_2 < \alpha_4 < \alpha_3 < \alpha_2 < \alpha_1 < \xi_2 \\ \alpha_1 + \alpha_2 < \xi_2 \\ \alpha_2 + \alpha_3 + \alpha_4 > 1 - \xi_2 \\ \alpha_5 > \alpha_4}} 1, \tag{7.3}$$

This is precisely the residual range for which the available equidistribution estimates do not apply. Therefore, it must be subtracted from $1_{\mathbb{P}}(n)$. Prime-sum partial summation gives

$$\frac{x(1+o(1))}{\log(x)} \int_{1-2\xi_2}^{\xi_2} \int_{1-2\xi_2}^{\min\{\alpha_1, \xi_2 - \alpha_1\}} \int_{1-2\xi_2}^{\alpha_2} \int_{\max\{1-2\xi_2, 1-\xi_2-\alpha_2-\alpha_3\}}^{\min\{\alpha_3, (1-\alpha_1-\alpha_2-\alpha_3)/2\}} \frac{1}{\alpha_1\alpha_2\alpha_3\alpha_4(1 - \sum_{i=1}^4 \alpha_i)} \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \mathrm{d}\alpha_3 \mathrm{d}\alpha_4.$$

Finally, Lemma 22 of [Sta25] treats the third sum in (7.2). To force n_4 to be prime it uses $(1 - \xi_2) + 3(1 - 2\xi_2) > 1$, which follows from $7\xi_2 < 3$. Removing the condition $\alpha_2 < \alpha_1$ requires $(1 - \xi_2)/2 > 1 - \xi_1 - \xi_2$, or equivalently

$$2\xi_1 + \xi_2 > 1.$$

Moreover, $7\xi_2 < 3$ implies $1 - 3(1 - 2\xi_2) < 1 - \xi_2$, which removes the condition $\alpha_1 < 1 - \xi_2$. The proof of Lemma 22 then divides the sum into Υ_1 and Υ_2. The latter already lies in the available equidistribution range; interchanging the relevant variables treats the former. The remaining term is

$$\Upsilon_3(n; x) = \sum_{\substack{n = p_2p_3p_4p_5p_6 \\ 1-2\xi_2 \le \alpha_3 < \alpha_2 < \xi_2 \\ \alpha_3 + \alpha_4 < \xi_2 \\ \alpha_2 < 1 - \xi_1 - \xi_2 \\ \alpha_4 \ge \alpha_3 \\ 1 - 2\xi_2 \le \alpha_5 \le \alpha_6}} 1,$$

where $6(1 - 2\xi_2) > 1$ was needed to get the prime p_6. This rearranges to condition $12\xi_2 < 5$, which implies $7\xi_2 < 3$. Next we observe that $\alpha_i > 1 - 2\xi_2$ for all i implies that also $\alpha_i < 1 - 4(1 - 2\xi_2) = 8\xi_2 - 3$ for all i. Furthermore, $8\xi_2 - 3 < \xi_2$ since $7\xi_2 < 3$. So condition $\alpha_i < \xi_2$ can be replaced by $\alpha_i < 8\xi_2 - 3$. Finally, the proof of Lemma 22 also requires $2(8\xi_2 - 3) < 1 - \xi_2$. We need

$$17\xi_2 < 7$$

to get that Υ_3 is the sum of a function with the desired equidistribution properties and of $\Gamma_2^*(n; x)$, where

$$\Gamma_2^*(n; x) = \sum_{\substack{n = p_2p_3p_4p_5p_6 \\ 1-2\xi_2 \le \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \le 8\xi_2 - 3 \\ \alpha_2, \alpha_4 > \alpha_3 \\ \alpha_2 + \alpha_4 < \xi_2 \\ \alpha_3 + \alpha_4 + \alpha_5 > 1 - \xi_2 \\ \alpha_6 > \alpha_5}} 1. \tag{7.4}$$

Using standard prime-sum partial summation, we find that $\sum_{n \in [x, 2x]} \Gamma_2^*(n; x)$ is asymptotic to

$$\frac{x(1+o(1))}{\log(x)} \int_{1-2\xi_2}^{8\xi_2-3} \int_{1-2\xi_2}^{\min\{8\xi_2-3, \alpha_2\}} \int_{\max\{\alpha_3, 1-2\xi_2\}}^{\min\{8\xi_2-3, \xi_2-\alpha_2\}} \int_{\max\{1-2\xi_2, 1-\xi_2-\alpha_3-\alpha_4\}}^{\min\{8\xi_2-3, (1-\sum_{i=2}^4 \alpha_i)/2\}} \frac{1}{\prod_{i=2}^5 \alpha_i (1 - \sum_{i=2}^5 \alpha_i)} \mathrm{d}\alpha_2 \mathrm{d}\alpha_3 \mathrm{d}\alpha_4 \mathrm{d}\alpha_5.$$

We have therefore proved the required equidistribution for ρ on the moduli associated with $T_k(\delta,\underline{A},\underline{B},\varepsilon)$, together with

$$\sum_{n\in[x,2x]}\rho(n;x)=(1-c_1+o(1))\frac{x}{\log x}.$$

By construction, $\rho\le 1_{\mathbb{P}}$, and $\rho(n;x)\ne 0$ only when every prime factor of n exceeds $x^{1-2\xi_2}$. It remains to obtain the pointwise lower bound encoded by c_2.

If $\xi_2\le 2/5$, then $1-2\xi_2\ge 1/5$. Every term in either Γ_2 or Γ_2^* would require five prime-factor exponents, each strictly larger than $1/5$, which is impossible because their sum is 1. Both exceptional sums are therefore empty, so $\rho=1_{\mathbb{P}}$ and we may take $c_2=0$.

If instead $\xi_2>2/5$, a fixed integer can contribute to Γ_2 only through a factorization into five distinct primes. Once the smallest prime is assigned the role of p_4, there are at most four choices for the remaining exceptional role; the strict ordering determines the other assignments. Hence $\Gamma_2(n;x)\le 4$. For Γ_2^*, there are at most $\binom{5}{3}=10$ choices for the three primes assigned the roles p_2,p_3,p_4, and at most two orders for the remaining pair (p_5,p_6). Thus $\Gamma_2^*(n;x)\le 20$. Since both terms enter ρ with a minus sign, we obtain $\rho(n;x)\ge -4-20=-24$, and may take $c_2=24$. □

The endpoint $\xi_2=2/5$ is the case we shall use, which we explicate with following corollary.

Corollary 7.3 (Direct-prime endpoint)**.** *Assume the parameter inequalities in Proposition 2 of* [Sta26] *and take $\xi_2=2/5$. Then the exceptional Buchstab sums vanish, the minorant is*

$$\rho=1_{\mathbb{P}},\qquad c_1=c_2=0,$$

and it is enough to prove equidistribution for every convolution in $\mathcal{H}(\xi_1,\xi_2,\xi_3)$.

Proof. The exceptional terms in the decomposition are supported on products whose prime-factor exponents all exceed $1-2\xi_2$. At $\xi_2=2/5$, each such exponent is at least $1/5$, while the strict ordering and remaining support inequalities would require five exponents with total strictly greater than 1. Both exceptional sums are therefore empty. The integration intervals defining c_1 are empty at this endpoint, so $c_1=0$, while Proposition 7.2 gives $c_2=0$. □

This is why we choose the endpoint $\xi_2=2/5$: it removes both the density loss c_1 and the pointwise penalty c_2, while retaining the Type-II range required below.

7.2. **The factor-extraction lemmas.** Corollary 7.3 reduces the arithmetic problem to equidistribution for every $f\in\mathcal{H}(\xi_1,\xi_2,\xi_3)$, and the estimates of Section 5 turn this into a demand for divisors of prescribed sizes. We now show when the support forces those divisors to exist.

The mechanism is the same in all three lemmas below. The coordinates exceeding δ account for finitely many rough factors, and the rest of the modulus is x^δ-smooth. One first partitions the rough logarithmic masses, placing each rough factor into a target bin; one then subdivides the smooth part, inserting its prime factors into a bin until its lower endpoint is crossed. Since the last inserted prime is smaller than x^δ, a target interval of logarithmic width at least δ controls the overshoot, and since every factor is assigned whole, no prime is split. Schematically,

$$\underbrace{\{\text{rough factors}\}}_{\text{assigned to prescribed bins}}\quad+\quad\underbrace{\{\text{smooth prime factors}\}}_{\text{inserted until a lower endpoint is crossed}}\quad\longrightarrow\quad\underbrace{q=q_1\cdots q_r}_{\text{required exponent intervals}}\quad.$$

We treat in turn one target divisor, two target divisors, and the nested three-divisor pattern needed for Type IIc. The rough-factor configurations themselves are described by the following sets.

Definition 7.4. *For given $B_1,B_2,\delta\in(0,\infty)$ and given $m_1,m_2\in\mathbb{N}\cup\{0\}$, we set*

$$\Xi(B_1,B_2,m_1,m_2,\delta)=\left\{(y_1,\dots,y_{m_1+m_2})\in[\delta,1]^{m_1+m_2}:\sum_{i=1}^{m_1}y_i\le B_1\ \text{and}\ \sum_{i=m_1+1}^{m_1+m_2}y_i\le B_2\right\}.$$

Thus $\Xi(B_{j,m},B_{j',m'},m,m',\delta)$ is the set of possible rough profiles of a modulus in

$$Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0).$$

7.2.1. *Two-factor extraction.* The first problem is to extract one divisor in a fixed interval $[x^a, x^b]$. The rough factors are assigned to the two sides of the factorization, and the interval width $b - a \geq \delta$ then allows the smooth remainder to move the first side to the desired scale without overshooting.

Lemma 7.5 (Two-factor extraction). *Let $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ be as given in Definition 4.3 and let $\Xi(B_1, B_2, m_1, m_2, \delta)$ be as given in Definition 7.4. Consider $a, b \in (0, \frac{1}{2})$ with $b - a \geq \delta$.*

Suppose that for every $(y_1, \ldots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ there exists a partition $I_1 \cup I_2$ of $\{1, \ldots, m + m'\}$ with

$$\sum_{i \in I_1} y_i \leq b \quad \text{and} \quad \sum_{i \in I_2} y_i \leq \frac{1}{2} - a.$$

Then there exists a constant $\varepsilon_1 > 0$ (dependent only on ε_0 and δ), so that the following is true for large x: If $q \in Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ and $q > x^{1/2 - \varepsilon_1}$, then q has some $r \mid q$ with $r \in [x^a, x^b]$.

Proof. Recall that we defined $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ to be the following set:

$$\left\{ ee' \prod_{i=1}^{m} f_i \prod_{i=1}^{m'} f_i' \in [1, x] : \begin{array}{ll} \log_x(f_1 \ldots f_m) \leq (1 - \varepsilon_0) B_{j,m}, & \log_x(f_1' \ldots f_{m'}') \leq (1 - \varepsilon_0) B_{j',m'}, \\ \log_x(f_1 \ldots f_m e) \leq (1 - \varepsilon_0)(A_j - \varepsilon), & \log_x(f_1' \ldots f_{m'}' e') \leq (1 - \varepsilon_0)(A_{j'} + \varepsilon), \\ ee' \text{ is } x^\delta\text{-smooth}, & \log_x(f_i), \log_x(f_i') \geq \delta \text{ for all } i \end{array} \right\}$$

Fix some $q = ee' \prod_{i=1}^{m} f_i \prod_{i=1}^{m'} f_i' \in Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ with $q > x^{1/2 - \epsilon}$, where ϵ is very small. Let $p_1 \ldots p_s$ be the prime factorization of ee'. We then know that $p_i < x^\delta$ for all i. We set

$$\begin{aligned} y_i &= \frac{1}{(1 - \varepsilon_0)} \log_x(f_i) && \text{for } 1 \leq i \leq m, \\ y_{m+i} &= \frac{1}{(1 - \varepsilon_0)} \log_x(f_i') && \text{for } 1 \leq i \leq m', \\ z_i &= \log_x(p_i) && \text{for } 1 \leq i \leq s. \end{aligned}$$

Since $y_i \geq \delta$ for all i and $\sum_{i=1}^{m} y_i \leq B_{j,m}$ and $\sum_{i=m+1}^{m+m'} y_i \leq B_{j',m'}$, there exists a partition $I_1 \cup I_2$ of $\{1, \ldots, m + m'\}$ with

$$\sum_{i \in I_1} (1 - \varepsilon_0) y_i \leq (1 - \varepsilon_0) b < b \quad \text{and} \quad \sum_{i \in I_2} (1 - \varepsilon_0) y_i \leq (1 - \varepsilon_0) \left(\tfrac{1}{2} - a \right).$$

Since $q > x^{1/2 - \epsilon}$, we also know that

$$\sum_{i \in I_1} (1 - \varepsilon_0) y_i + \sum_{i=1}^{s} z_i \geq \tfrac{1}{2} - \epsilon - (1 - \varepsilon_0) \left(\tfrac{1}{2} - a \right) \geq a,$$

provided ϵ is sufficiently small compared to ε_0. But $z_i < \delta$ and so there exists $J \subseteq \{1, \ldots, s\}$ with

$$\sum_{i \in I_1} (1 - \varepsilon_0) y_i + \sum_{i \in J}^{s} z_i \in [a, b].$$

Multiplying the corresponding f_i, f_i' and p_i together, we then get a factor $r \mid q$ with $r \in [x^a, x^b]$. □

7.2.2. *Three-factor extraction.* The second problem is to construct two controlled divisors at once: $q = uvr$, with $r \in [x^{a_1}, x^{b_1}]$ and $u \in [x^{a_2}, x^{b_2}]$. Three rough bins now reserve mass for r, for u, and for the unconstrained remainder, and the relations among the four endpoints guarantee that the smooth reservoir can fill the first two bins in sequence.

Lemma 7.6 (Three-factor extraction). *Let $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ be as given in Definition 4.3 and let $\Xi(B_1, B_2, m_1, m_2, \delta)$ be as given in Definition 7.4. Consider $a_1, a_2, b_1, b_2 \in (0, \frac{1}{2})$ with $b_1 - b_2 \geq a_1 - a_2$ and $b_1 + b_2 < \frac{1}{2}$ and $b_1 - a_1 \geq \delta$ and $b_2 - a_2 \geq \delta$.*

Suppose that for every $(y_1, \ldots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ there exists a partition $I_1 \cup I_2 \cup I_3$ of $\{1, \ldots, m + m'\}$ with

$$\sum_{i \in I_1} y_i \leq b_1 \quad \text{and} \quad \sum_{i \in I_2} y_i \leq b_2 \quad \text{and} \quad \sum_{i \in I_3} y_i \leq \frac{1}{2} - b_1 - a_2.$$

Then there exists a constant $\varepsilon_1 > 0$ (dependent only on ε_0 and δ), so that the following is true for large x: If $q \in Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ and $q > x^{1/2-\varepsilon_1}$, then q has a factorization $q = uvr$ with $r \in [x^{a_1}, x^{b_1}]$ and $u \in [x^{a_2}, x^{b_2}]$.

Proof. The construction of Lemma 7.5 is applied successively to the two target intervals.

We first fix some $q = ee' \prod_{i=1}^m f_i \prod_{i=1}^{m'} f_i' \in Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ with $q > x^{1/2-\epsilon}$, where ϵ is very small. Denoting by $p_1 \dots p_s$ the prime factorization of ee', we know that $p_i < x^\delta$ for all i and set

$$\begin{aligned} y_i &= \frac{1}{(1-\varepsilon_0)} \log_x(f_i) && \text{for } 1 \le i \le m, \\ y_{m+i} &= \frac{1}{(1-\varepsilon_0)} \log_x(f_i') && \text{for } 1 \le i \le m', \\ z_i &= \log_x(p_i) && \text{for } 1 \le i \le s. \end{aligned}$$

Then there exists a partition $I_1 \cup I_2 \cup I_3$ of $\{1,\dots,m+m'\}$ with

$$\sum_{i\in I_1}(1-\varepsilon_0)y_i < b_1 \quad \text{and} \quad \sum_{i\in I_2}(1-\varepsilon_0)y_i < b_2 \quad \text{and} \quad \sum_{i\in I_3}(1-\varepsilon_0)y_i \le (1-\varepsilon_0)\left(\tfrac12 - b_1 - a_2\right).$$

Since $q > x^{1/2-\epsilon}$, we also know that

$$\sum_{i\in I_2}(1-\varepsilon_0)y_i + \sum_{i=1}^s z_i > \tfrac12 - \epsilon - b_1 - (1-\varepsilon_0)\left(\tfrac12 - b_1 - a_2\right) \ge a_2.$$

But $z_i < \delta$ for all i and so there exists $J_2 \subseteq \{1,\dots,s\}$ with

$$\sum_{i\in I_2}(1-\varepsilon_0)y_i + \sum_{i\in J_2} z_i \in [a_2, b_2].$$

Further, we then also have

$$\sum_{i\in I_1}(1-\varepsilon_0)y_i + \sum_{i\notin J_2} z_i > \tfrac12 - \epsilon - b_2 - (1-\varepsilon_0)\left(\tfrac12 - b_1 - a_2\right) \ge b_1 - b_2 + a_2 \ge a_1.$$

Hence there also exists $J_1 \subseteq \{1,\dots,s\} \setminus J_2$ with

$$\sum_{i\in I_1}(1-\varepsilon_0)y_i + \sum_{i\in J_1}^{s} z_i \in [a_1, b_1].$$

Multiplying the corresponding f_i, f_i' and p_i together, we then get a factorization $q = uvr$ with $r \in [x^{a_1}, x^{b_1}]$ and $u \in [x^{a_2}, x^{b_2}]$. □

7.2.3. *Four-factor extraction.* The Type IIc estimate requires the nested pattern $q = uvr$ together with a divisor $d_1 \mid r$, so that r, u, and d_1 must all lie in prescribed intervals while a fourth bin holds the unused rough mass. Because the available capacities depend on the actual size of q, we work on a dyadic block $q \asymp x^{1/2+2\omega_0}$ rather than replacing ω_0 at the outset by a global maximum. This is the most delicate of the three extraction statements.

Lemma 7.7 (Four-factor extraction)**.** *Let $Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ be as given in Definition 4.3 and let $\Xi(B_1,B_2,m_1,m_2,\delta)$ be as given in Definition 7.4. Consider a_1,a_2,a_3,b_1,b_2,b_3 with $3(b_1-a_1)+(a_3-b_3) \ge 0$ and $b_1 - b_2 \ge a_1 - a_2$ and $b_1 - a_1 \ge \delta$ and $b_2 - a_2 \ge \delta$ and $b_3 - a_3 \ge \delta$.*

Let $\omega_0 \ge -\varepsilon_0$ and suppose that for every $(y_1,\dots,y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ there exists a partition $I_1 \cup I_2 \cup I_3 \cup I_4$ of $\{1,\dots,m+m'\}$ with

$$\sum_{i\in I_1} y_i \le 2a_1 + b_3 \quad \textit{and} \quad \sum_{i\in I_2} y_i \le b_2 \quad \textit{and} \quad \sum_{i\in I_3} y_i \le \frac12 + 2\omega_0 - b_1 - a_2 \quad \textit{and} \quad \sum_{i\in I_4} y_i \le a_1 - 2b_1 - a_3.$$

Then for large x, every $q \in Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ with $q \asymp x^{1/2+2\omega_0}$ has a factorization $q = uvr$ with the following properties: $r \in [x^{a_1}, x^{b_1}]$ and $u \in [x^{a_2}, x^{b_2}]$ and there exists $d_1 \mid r$ with $d_1 \in [r^2 x^{a_3}, r^2 x^{b_3}]$.

Proof. We first fix some $q = ee'\prod_{i=1}^m f_i \prod_{i=1}^{m'} f_i' \in Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ with $q \asymp x^{1/2+2\omega_0}$. Denoting by $p_1\dots p_s$ the prime factorization of ee', we know that $p_i < x^\delta$ and set

$$\begin{aligned} y_i &= \log_x(f_i) && \text{for } 1\le i\le m,\\ y_{m+i} &= \log_x(f_i') && \text{for } 1\le i\le m',\\ z_i &= \log_x(p_i) && \text{for } 1\le i\le s. \end{aligned}$$

Then there exists a partition $I_1\cup I_2\cup I_3\cup I_4$ of $\{1,\dots,m+m'\}$ with

$$\sum_{i\in I_1} y_i \le 2a_1+b_3 \quad\text{and}\quad \sum_{i\in I_2} y_i\le b_2 \quad\text{and}\quad \sum_{i\in I_3} y_i \le \frac12+2\omega_0-b_1-a_2 \quad\text{and}\quad \sum_{i\in I_4} y_i\le a_1-2b_1-a_3.$$

Since $q\asymp x^{1/2+2\omega_0}$, we also know that

$$\begin{aligned}\sum_{i\in I_2} y_i+\sum_{i=1}^s z_i &> \tfrac12+2\omega_0-(2a_1+b_3)-(\tfrac12+2\omega_0-b_1-a_2)-(a_1-2b_1-a_3)\\ &\ge a_2+(3b_1-3a_1+a_3-b_3)\ge a_2.\end{aligned}$$

But $z_i<\delta$ for all i and so there exists $J_2\subseteq\{1,\dots,s\}$ with

$$\sum_{i\in I_2} y_i+\sum_{i\in J_2} z_i\in[a_2,b_2].$$

This sum corresponds to factor u. Next, we observe that

$$\begin{aligned}&\sum_{i\in I_1} y_i+\sum_{i\in I_4} y_i\le(2a_1+b_3)+(a_1-2b_1-a_3)=b_1+(3a_1-3b_1+b_3-a_3)\le b_1,\\ &\sum_{i\in I_1} y_i+\sum_{i\in I_4} y_i+\sum_{i\notin J_2} z_i>\tfrac12+2\omega_0-b_2-\big(\tfrac12+2\omega_0-b_1-a_2\big)\ge b_1-b_2+a_2\ge a_1.\end{aligned}$$

Hence there also exists $J_1\subseteq\{1,\dots,s\}\setminus J_2$ with

$$\sum_{i\in I_1} y_i+\sum_{i\in I_4} y_i+\sum_{i\in J_1} z_i\in[a_1,b_1].$$

This sum corresponds to factor r. Finally, consider $c\in[a_1,b_1]$. Notice that

$$\begin{aligned}&\sum_{i\in I_1} y_i\le 2a_1+b_3\le 2c+b_3,\\ &\sum_{i\in I_1} y_i+\sum_{i\in J_1} z_i\ge a_1-(a_1-2b_1-a_3)\ge 2b_1+a_3\ge 2c+a_3.\end{aligned}$$

Hence for any $c\in[a_1,b_1]$ we can find $J'\subseteq J_1$ with

$$\sum_{i\in I_1} y_i+\sum_{i\in J'} z_i\in[2c+a_3,2c+b_3].$$

This sum corresponds to factor d_1 of r. Overall, multiplying the corresponding f_i, f_i' and p_i together, we then get that $q\asymp x^{1/2+2\omega_0}$ has a factorization $q=uvr$ with $r\in[x^{a_1},x^{b_1}]$ and $u\in[x^{a_2},x^{b_2}]$ so that for every $R\in[x^{a_1},x^{b_1}]$, there exists $d_1\mid r$ with $d_1\in[R^2x^{a_3},R^2x^{b_3}]$. ☐

7.3. **An equidistribution criterion.** We now compose the two preceding steps. The analytic estimates specify which divisors a modulus must possess. And the extraction lemmas specify when the support forces those divisors to exist. Proposition 7.8 lists the resulting sufficient conditions in terms of δ, $\underline{A}$, $\underline{B}$, and the rough-profile sets Ξ.

Proposition 7.8. *Let $T_k(\delta,\underline{A},\underline{B},\varepsilon)$ be as given in Definition 4.1, let $Q^*(x;\delta,\underline{A},\underline{B},\varepsilon,\varepsilon_0)$ be as given in Definition 4.3 and let $\Xi(B_1,B_2,m_1,m_2,\delta)$ be as given in Definition 7.4. Let $\epsilon=10^{-10}$. Suppose $\xi_1,\xi_2,\xi_3\in(0,1)$ satisfy the following conditions:*

(I) *Type I condition:* $\min\left\{\xi_1-4A_n+\frac23,\frac97-\frac{34}{7}A_n\right\}-2\epsilon>\delta$. (*Recall that* $\underline{A}=(A_1,\dots,A_n)$.)

(II) *Type II condition:* $\frac{19}{2}-36A_n-13\delta-9\epsilon\ge 0$ *and*

$$\min\left\{\tfrac{\xi_2}{10}-\tfrac{32A_n}{10}+\tfrac{8}{10},\ \tfrac{\xi_2}{4}+\tfrac{11}{16}-3A_n\right\}-2\epsilon\ge\delta.$$

(III) *Type III condition:* $\frac{11}{8} - \frac{7}{2}A_n - \frac{9}{8}\xi_3 - 2\epsilon > \delta$.

Suppose further that for all $j, j' \in \{1, \dots, n\}$ *and for all* $m, m' \in \{0, 1, \dots, \lfloor \frac{1}{\delta} \rfloor\}$ *with* $m + m' > 0$*, every tuple* $(y_1, \dots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ *has all of the following properties:*

(A) *Type I condition: There exists a partition* $I_1 \cup I_2$ *of* $\{1, \dots, m + m'\}$ *with*

$$\sum_{i \in I_1} y_i \le \xi_1 - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{1}{6} - 4\omega(j, j') - 2\epsilon \quad \text{where} \quad \omega(j, j') = \left(\frac{A_j + A_{j'}}{2} - \frac{1}{4} \right),$$

and there exists a partition $I_1 \cup I_2$ *of* $\{1, \dots, m + m'\}$ *with*

$$\sum_{i \in I_1} y_i \le \frac{1}{2} - 2\omega(j, j') - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{1}{14} - \frac{68}{14}\omega(j, j') - 2\epsilon.$$

(*The first partition serves the range* $\gamma \le \frac{1}{2}$*, the second the range* $\gamma \in (\frac{1}{2}, \frac{1}{2} + 2\omega(j, j') + \varepsilon]$*.*)

(B) *Type IIa condition: There exists a partition* $I_1 \cup I_2$ *of* $\{1, \dots, m + m'\}$ *with*

$$\sum_{i \in I_1} y_i \le \frac{2}{5} + \frac{24\omega(j, j')}{5} + \frac{7\delta}{5} - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{1}{14} - \frac{24\omega(j, j')}{7} - 2\epsilon.$$

(C) *Type IIb condition: There also exists a partition* $I_1 \cup I_2 \cup I_3$ *of* $\{1, \dots, m + m'\}$ *with*

$$\sum_{i \in I_1} y_i \le \frac{1}{3} + \frac{24\omega(j, j')}{3} + \frac{7\delta}{3} - 4\epsilon,$$

$$\sum_{i \in I_2} y_i \le \frac{1}{10} - \frac{34\omega(j, j')}{5} - \frac{7\delta}{5} - 4\epsilon,$$

$$\sum_{i \in I_3} y_i \le 2\omega(j, j') + \delta - 4\epsilon.$$

(D) *Type IIc condition: For any* $\omega_0 \in [0, \omega(j, j')]$ *and any* γ *with* $\xi_2 - \epsilon \le \gamma \le \frac{1}{3} + \frac{24\omega(j,j')}{3} + \frac{7\delta}{3} + 3\epsilon$*, there exists a partition* $I_1 \cup I_2 \cup I_3 \cup I_4$ *with*

$$\sum_{i \in I_1} y_i \le \gamma - 2\delta - 8\omega_0 - \epsilon,$$

$$\sum_{i \in I_2} y_i \le \frac{1}{2} - \gamma - 2\omega_0 - \epsilon,$$

$$\sum_{i \in I_3} y_i \le 4\omega_0 + \delta - \epsilon,$$

$$\sum_{i \in I_4} y_i \le 8\omega_0.$$

(E) *Type III condition: There also exists a partition* $I_1 \cup I_2$ *of* $\{1, \dots, m + m'\}$ *with*

$$\sum_{i \in I_1} y_i \le 1 - 6\omega(j, j') - \frac{3\xi_3}{2} - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{5\omega(j, j')}{2} + \frac{3\xi_3}{8} - 2\epsilon.$$

Then for every $f \in \mathcal{H}(\xi_1, \xi_2, \xi_3)$*,* $x > 1$*,* $\varepsilon_0 > 0$*,* $A > 0$ *and* $a \in \mathbb{Z}$ *with* $(a, P(x)) = 1$*,*

$$\sum_{\substack{q \in Q^*(x; \delta, \underline{A}, \underline{B}, \varepsilon, \varepsilon_0) \\ q \text{ is squarefree}}} \left| \sum_{\substack{n \in [x, 2x] \\ n \equiv a(q)}} f(n; x) - \frac{1}{\varphi(q)} \sum_{\substack{n \in [x, 2x] \\ (n, q) = 1}} f(n; x) \right| \ll_{A, \varepsilon_0} \frac{x}{\log(x)^A}.$$

Proof. The class $\mathcal{H}(\xi_1, \xi_2, \xi_3)$ consists of the three convolution types in Definition 7.1. We prove the asserted estimate separately for each type and each fixed choice of $j, j' \in \{1, \dots, n\}$ and $m, m' \in \{1, \dots, \lfloor 1/\delta \rfloor\}$. It is enough to establish

$$\sum_{\substack{q \in Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0) \\ q \text{ is squarefree}}} \left| \sum_{\substack{n \in [x, 2x] \\ n \equiv a(q)}} f(n; x) - \frac{1}{\varphi(q)} \sum_{\substack{n \in [x, 2x] \\ (n, q) = 1}} f(n; x) \right| \ll_{A, \varepsilon_0} \frac{x}{\log(x)^A}.$$

We fix some choice of $j, j' \in \{1, \dots, n\}$ and $m, m' \in \{0, 1, \dots, \lfloor \frac{1}{\delta} \rfloor\}$ and may assume $m + m' > 0$ (as the $m + m' = 0$ case follows trivially from the other cases). For $q \in Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ we have $q \le x^{(1-\varepsilon_0)(A_j + A_{j'})}$. The largest exponent available in this band pair is therefore encoded by

$$\omega_{\max} = \frac{A_j + A_{j'}}{2} - \frac{1}{4}.$$

Type I. First let $f(n; x) = (\alpha \star \beta)(n; x)$, where α is a coefficient sequence at scale M and β is a smooth coefficient sequence at scale N with $M(x)N(x) \asymp x$ and $N(x) \ge x^{\xi_1 - \epsilon}$.

Initially we assume $\gamma \le \frac{1}{2}$ and define $\delta^*(\gamma) = \gamma - 4\omega_{\max} - \frac{1}{3} - \epsilon$. For $\gamma \ge \xi_1 - \epsilon$ we have

$$3\gamma - 12\omega_{\max} - 3\delta^*(\gamma) > 1.$$

Also, the assumption $\xi_1 - 4A_n + \frac{2}{3} - 2\epsilon > \delta$ ensures that $\delta^*(\gamma) \ge \delta$.

Apply Lemma 5.3 with $(\omega, \gamma, \delta) = (\omega_{\max}, \gamma, \delta^*(\gamma))$. For sufficiently small ε', it gives the required Type-I discrepancy estimate on $D_I(x; \omega_{\max}, \gamma, \delta^*(\gamma); \varepsilon')$. It remains only to prove that the present modulus family is contained in this set. When $\gamma \le \frac{1}{2}$, the definition is

$$D_I(x; \omega, \gamma, \delta; \varepsilon) = \left\{ d \in [1, x^{1/2 + 2\omega}] \cap \mathbb{N} : \exists r \mid d \text{ with } x^{\gamma - \delta - 3\varepsilon} < r < x^{\gamma - 3\varepsilon} \right\}.$$

To show $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0) \subseteq D_I(x; \omega_{\max}, \gamma, \delta^*(\gamma); \varepsilon')$, apply Lemma 7.5. We need the relevant moduli q to have a factor r with $x^{\gamma - \delta^*(\gamma) - 3\varepsilon'} < r < x^{\gamma - 3\varepsilon'}$ when $\xi_1 - \epsilon \le \gamma \le \frac{1}{2}$. But

$$\gamma - \delta^*(\gamma) - 3\varepsilon' = 4\omega_{\max} + \tfrac{1}{3} + \epsilon - 3\varepsilon'.$$

By Lemma 7.5 elements of $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ thus have the desired factors provided that for every $(y_1, \dots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ there exists a partition $I_1 \cup I_2$ of $\{1, \dots, m + m'\}$ with

$$\sum_{i \in I_1} y_i \le \xi_1 - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{1}{6} - 4\omega_{\max} - 2\epsilon.$$

On the other hand, if $\gamma \in (\frac{1}{2}, \frac{1}{2} + 2\omega_{\max} + \varepsilon']$, we may take any $\delta^*(\gamma) \in \left[\delta, \frac{1}{14} - \frac{68}{14}\omega_{\max} - \epsilon\right]$, so that

$$68\omega_{\max} + 14\delta^*(\gamma) < 1. \tag{7.5}$$

Then $f(n; x)$ equidistributes for moduli in $D_I(x; \omega_{\max}, \gamma, \delta^*(\gamma); \varepsilon')$. This time, D_I consists of q which have a factor r with $x^{1 - \gamma - \delta^*(\gamma) - 3\varepsilon'} < r < x^{1 - \gamma - 3\varepsilon'}$. Thus Lemma 7.5 applies to $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0)$ provided that for every $(y_1, \dots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ there exists a partition $I_1 \cup I_2$ with

$$\sum_{i \in I_1} y_i \le \tfrac{1}{2} - 2\omega_{\max} - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \tfrac{1}{14} - \tfrac{68}{14}\omega_{\max} - 2\epsilon. \tag{7.6}$$

Type II. Next let $f(n; x) = (\alpha \star \beta)(n; x)$, where α is a coefficient sequence at scale M, β is a coefficient sequence at scale N with $M(x)N(x) \asymp x$, and $x^{\xi_2 - \epsilon} \le N(x) = x^\gamma \le x^{1 - \xi_2 + \epsilon}$. Assume also that α and β both have the Siegel–Walfisz property. Exchanging the two convolution factors replaces γ by $1 - \gamma$, so it is enough to treat $\gamma \le 1/2$. We cover this interval by three overlapping ranges, handled respectively by Lemmas 5.4, 5.5, and 5.6. For each range we choose an auxiliary divisor-window width $\delta^*(\gamma)$, verify the analytic inequalities of the corresponding lemma, and then invoke a factor-extraction lemma to construct the required divisors of q.

Type IIa. We begin with Lemma 5.4 and set $\delta^*(\gamma) = \frac{5\gamma}{7} - \frac{2}{7} - \frac{24\omega_{\max}}{7} - \epsilon$. This choice ensures that

$$24\omega_{\max} + 7\delta^*(\gamma) - 5\gamma < -2,$$
$$8\omega_{\max} + 3\delta^*(\gamma) - \gamma < 0,$$

so Lemma 5.4 applies with $(\omega, \gamma, \delta) = (\omega_{\max}, \gamma, \delta^*(\gamma))$. For sufficiently small ε', the analytic task is therefore reduced to the containment $Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0) \subseteq D_{IIa}(x; \omega_{\max}, \gamma, \delta^*(\gamma); \varepsilon')$. Lemma 7.5 supplies this containment once $\delta^*(\gamma) = \frac{5\gamma}{7} - \frac{2}{7} - \frac{24\omega_{\max}}{7} - \epsilon \ge \delta$. This means for now we only consider

$$\gamma \ge \tfrac{2}{5} + \tfrac{24\omega_{\max}}{5} + \tfrac{7\delta}{5} + 2\epsilon.$$

We then get that any $q \in Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ has a divisor $r \mid q$ with $x^{\gamma-3\varepsilon-\delta^*(\gamma)} < r < x^{\gamma-3\varepsilon}$ provided that for every $(y_1,\dots,y_{m+m'}) \in \Xi(B_{j,m},B_{j',m'},m,m',\delta)$ there exists a partition $I_1 \cup I_2$ with

$$\sum_{i\in I_1} y_i \le \tfrac{2}{5} + \tfrac{24\omega_{\max}}{5} + \tfrac{7\delta}{5} - 2\epsilon \quad \text{and} \quad \sum_{i\in I_2} y_i \le \tfrac{1}{14} - \tfrac{24\omega_{\max}}{7} - 2\epsilon. \tag{7.7}$$

Type IIb. For the intermediate range $\gamma < \frac{2}{5} + \frac{24}{5}\omega_{\max} + \frac{7}{5}\delta + 2\epsilon$, use Lemma 5.5 and choose $\delta^*(\gamma) = \frac{3\gamma}{7} - \frac{1}{7} - \frac{24\omega_{\max}}{7} - \epsilon$, which ensures that

$$\begin{aligned} &24\omega_{\max} + 7\delta^*(\gamma) - 3\gamma < -1, \\ &8\omega_{\max} + 3\delta^*(\gamma) - \gamma < 0. \end{aligned}$$

Lemma 5.5 now applies with $(\omega,\gamma,\delta) = (\omega_{\max},\gamma,\delta^*(\gamma))$. As in the IIa range, the remaining task is a modulus containment, this time in D_{IIb}; Lemma 7.6 proves it provided $\delta^*(\gamma) = \frac{3\gamma}{7} - \frac{1}{7} - \frac{24\omega_{\max}}{7} - \epsilon \ge \delta$, so have

$$\tfrac{1}{3} + \tfrac{24\omega_{\max}}{3} + \tfrac{7\delta}{3} + 3\epsilon < \gamma < \tfrac{2}{5} + \tfrac{24\omega_{\max}}{5} + \tfrac{7\delta}{5} + 2\epsilon$$

Recall now that

$$D_{IIb}(x;\omega,\gamma,\delta;\varepsilon) = \left\{ d \in [1,x^{1/2+2\omega}] \cap \mathbb{N} : \begin{array}{l} \exists r \mid d\ \exists u \mid \frac{d}{r} \text{ with } x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon} \\ \text{and } x^{1/2-\gamma-2\omega-6\varepsilon-\delta} < u < x^{1/2-\gamma-2\omega-6\varepsilon} \end{array} \right\}.$$

Any $q \in Q(x;\delta,\underline{A},\underline{B},\varepsilon,j,j',m,m',\varepsilon_0)$ has a divisor $r \mid q$ and a divisor $u \mid \frac{q}{r}$ with $x^{\gamma-3\varepsilon-\delta^*(\gamma)} < r < x^{\gamma-3\varepsilon}$ and $x^{1/2-\gamma-2\omega_{\max}-6\varepsilon-\delta^*(\gamma)} < u < x^{1/2-\gamma-2\omega_{\max}-6\varepsilon}$ provided that the following is true: For every $(y_1,\dots,y_{m+m'}) \in \Xi(B_{j,m},B_{j',m'},m,m',\delta)$ there exists a partition $I_1 \cup I_2 \cup I_3$ with

$$\sum_{i\in I_1} y_i \le \gamma - 2\epsilon, \quad \sum_{i\in I_2} y_i \le 1/2 - \gamma - 2\omega_{\max} - 2\epsilon \quad \text{and} \quad \sum_{i\in I_3} y_i \le \frac{3\gamma}{7} - \frac{1}{7} - \frac{10\omega_{\max}}{7} - 2\epsilon.$$

The three cases depend on γ. The first and third increase with γ, so their worst values occur at the lower endpoint of the IIb range, where $\delta^*(\gamma) = \delta$. The second decreases with γ, so its worst value occurs at the upper endpoint. Taking these values gives the following uniform sufficient condition: there exists a partition $I_1 \cup I_2 \cup I_3$ with

$$\begin{aligned} &\sum_{i\in I_1} y_i \le \tfrac{1}{3} + \tfrac{24\omega_{\max}}{3} + \tfrac{7\delta}{3} - 4\epsilon, \\ &\sum_{i\in I_2} y_i \le \tfrac{1}{10} - \tfrac{34\omega_{\max}}{5} - \tfrac{7\delta}{5} - 4\epsilon, \\ &\sum_{i\in I_3} y_i \le 2\omega_{\max} + \delta - 4\epsilon. \end{aligned} \tag{7.8}$$

Type IIc. Finally, we use Lemma 5.6 to treat the case $\xi_2 - \epsilon \le \gamma \le \frac{1}{3} + \frac{24\omega_{\max}}{3} + \frac{7\delta}{3} + 3\epsilon$. Choose $\delta^*(\gamma)$ subject to

$$\delta^*(\gamma) \le \min\left\{ \frac{1}{4} - 2\omega_{\max} - \frac{\gamma}{2}, \frac{\gamma}{10} - \frac{32\omega_{\max}}{10}, \frac{\gamma}{4} - \frac{1}{16} - 3\omega_{\max} \right\} - \epsilon.$$

This ensures that we have the following inequalities:

$$\begin{aligned} &8\omega_{\max} + 4\delta^*(\gamma) + 2\gamma < 1, \\ &32\omega_{\max} + 10\delta^*(\gamma) - \gamma < 0, \\ &48\omega_{\max} + 16\delta^*(\gamma) - 4\gamma < -1. \end{aligned}$$

Therefore, Lemma 5.6 applies with $(\omega,\gamma,\delta) = (\omega_{\max},\gamma,\delta^*(\gamma))$ and gives the Type-IIc discrepancy estimate on $D_{IIc}(x;\omega_{\max},\gamma,\delta^*(\gamma);\varepsilon')$. This modulus class is

$$D_{IIc}(x;\omega,\gamma,\delta;\varepsilon) = \left\{ d \in [1,x^{1/2+2\omega}] \cap \mathbb{N} : \begin{array}{l} \exists r \mid d\ \exists u \mid \frac{d}{r}\ \exists d_1 \mid r \text{ with } x^{\gamma-3\varepsilon-\delta} < r < x^{\gamma-3\varepsilon} \\ \text{and } x^{1-\gamma-6\varepsilon-\delta} d^{-1} < u < x^{1-\gamma-6\varepsilon} d^{-1} \\ \text{and } r^2 x^{2-\gamma-52\varepsilon-\delta} d^{-4} < d_1 < r^2 x^{2-\gamma-52\varepsilon} d^{-4} \end{array} \right\}.$$

The given upper bound on $\delta^*(\gamma)$ must permit $\delta^*(\gamma) \geq \delta$ for $\xi_2 - \epsilon \leq \gamma \leq \frac{1}{3} + \frac{24\omega_{\max}}{3} + \frac{7\delta}{3} + 3\epsilon$. The first required inequality is

$$\tfrac{1}{4} - 2\omega_{\max} - \tfrac{1}{2}\left(\tfrac{1}{3} + \tfrac{24\omega_{\max}}{3} + \tfrac{7\delta}{3} + 3\epsilon\right) \geq \delta,$$

which, since $\omega_{\max} \leq A_n - \frac{1}{4}$, is equivalent to

$$\tfrac{19}{2} - 36A_n - 13\delta - 9\epsilon \geq 0.$$

On the other hand, we also need that

$$\min\left\{\tfrac{\xi_2}{10} - \tfrac{32\omega_{\max}}{10},\ \tfrac{\xi_2}{4} - \tfrac{1}{16} - 3\omega_{\max}\right\} - 2\epsilon \geq \delta.$$

Recalling that $\omega_{\max} \leq A_n - \frac{1}{4}$, it suffices to have

$$\min\left\{\tfrac{\xi_2}{10} - \tfrac{32A_n}{10} + \tfrac{8}{10},\ \tfrac{\xi_2}{4} + \tfrac{11}{16} - 3A_n\right\} - 2\epsilon \geq \delta.$$

These two criteria form condition (II) of Proposition 7.8. Hence we can work with any $\delta^*(\gamma)$ which satisfies

$$\delta \leq \delta^*(\gamma) \leq \min\left\{\frac{1}{4} - 2\omega_{\max} - \frac{\gamma}{2},\ \frac{\gamma}{10} - \frac{32\omega_{\max}}{10},\ \frac{\gamma}{4} - \frac{1}{16} - 3\omega_{\max}\right\} - \epsilon.$$

Lemma 7.7 converts these three nested divisor intervals into the required four-bin condition. In particular, it proves $Q^*(x; \delta, \underline{A}, \underline{B}, \varepsilon, \varepsilon_0) \subseteq D_{IIc}(x; \omega_{\max}, \gamma, \delta^*(\gamma); \varepsilon')$ provided that every rough profile admits the following partition: For every $\omega_0 \in [0, \omega_{\max}]$ and every tuple $(y_1, \ldots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ there exists a partition $I_1 \cup I_2 \cup I_3 \cup I_4$ of $\{1, \ldots, m+m'\}$ which satisfies inequalities

$$\sum_{i \in I_1} y_i \leq \gamma - 2\delta^*(\gamma) - 8\omega_0 - \epsilon,$$
$$\sum_{i \in I_2} y_i \leq \frac{1}{2} - \gamma - 2\omega_0 - \epsilon,$$
$$\sum_{i \in I_3} y_i \leq 4\omega_0 + \delta^*(\gamma) - \epsilon,$$
$$\sum_{i \in I_4} y_i \leq 8\omega_0.$$

Since $\delta^*(\gamma) \geq \delta$, and since a smaller δ makes these conditions easier to satisfy, we may replace $\delta^*(\gamma)$ by δ and settle on the criterion

$$\sum_{i \in I_1} y_i \leq \gamma - 2\delta - 8\omega_0 - \epsilon,$$
$$\sum_{i \in I_2} y_i \leq \frac{1}{2} - \gamma - 2\omega_0 - \epsilon,$$
$$\sum_{i \in I_3} y_i \leq 4\omega_0 + \delta - \epsilon,$$
$$\sum_{i \in I_4} y_i \leq 8\omega_0.$$

Type III. Finally, let $f(n; x) = (\alpha \star \psi_1 \star \psi_2 \star \psi_3)(n; x)$, where α is a coefficient sequence at scale M and ψ_1, ψ_2 and ψ_3 are smooth coefficient sequences at scales N_1, N_2 and N_3 with $M(x)N_1(x)N_2(x)N_3(x) \asymp x$. Additionally, $x^{1-2\xi_3-\epsilon} \leq N_1(x), N_2(x), N_3(x) \leq x^{\xi_3+\epsilon}$ and $N_i(x)N_j(x) \geq x^{1-\xi_3-\epsilon}$ for $i \neq j$.

We define $\delta^* = \frac{1}{2} - \frac{7}{2}\omega_{\max} - \frac{9}{8}\xi_3 - \epsilon$, so that

$$28\omega_{\max} + 9\xi_3 + 8\delta^* < 4.$$

The condition $\frac{11}{8} - \frac{7}{2}A_n - \frac{9}{8}\xi_3 - 2\epsilon > \delta$ ensures $\delta^* > \delta$. Apply Lemma 5.7 with $(\omega, \gamma, \delta) = (\omega_{\max}, \gamma, \delta^*)$. Every Type-III convolution then satisfies the desired discrepancy estimate on $D_{III}(x; \omega_{\max}, \gamma, \delta^*; \varepsilon')$. To obtain

$$Q(x; \delta, \underline{A}, \underline{B}, \varepsilon, j, j', m, m', \varepsilon_0) \subseteq D_{III}(x; \omega_{\max}, \gamma, \delta^*; \varepsilon')$$

we use once more Lemma 7.5. We need to show that elements q of Q have a factor $r \mid q$ with

$$x^{1/3+4\delta^*/3-4\omega_{\max}/3-\delta^*} < r < x^{1/3+4\delta^*/3-4\omega_{\max}/3}.$$

This is guaranteed if every $(y_1, \dots, y_{m+m'}) \in \Xi(B_{j,m}, B_{j',m'}, m, m', \delta)$ has a partition $I_1 \cup I_2$ with

$$\sum_{i \in I_1} y_i \le \frac{1}{3} + \frac{4\delta^*}{3} - \frac{4\omega_{\max}}{3} \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{1}{6} - \frac{\delta^*}{3} + \frac{4\omega_{\max}}{3}.$$

Plugging in $\delta^* = \frac{1}{2} - \frac{7}{2}\omega_{\max} - \frac{9}{8}\xi_3 - \epsilon$, this is implied by

$$\sum_{i \in I_1} y_i \le 1 - 6\omega_{\max} - \frac{3\xi_3}{2} - 2\epsilon \quad \text{and} \quad \sum_{i \in I_2} y_i \le \frac{5\omega_{\max}}{2} + \frac{3\xi_3}{8} - 2\epsilon. \tag{7.9}$$

This concludes the proof. □

Remark 7.9. *Lemmas 7.5 and 7.6 produce the required divisors only for $q > x^{1/2-\varepsilon_1}$, and Lemma 7.7 only for $q \asymp x^{1/2+2\omega_0}$ with $\omega_0 \ge 0$. The containments obtained above concern the moduli of $Q^*(x; \delta, \underline{A}, \underline{B}, \varepsilon, \varepsilon_0)$ in that range. The complementary moduli are smaller than $x^{1/2}(\log x)^{-C_{\mathrm{BV}}}$ for large x and are handled directly by Lemma 8.1, which requires no factorization of the modulus; the two ranges together give the displayed bound over all of Q^*.*

7.4. **The Type-IIc packing condition.** At the parameter point chosen in Section 9, the rough cap depends on the number of rough factors. We therefore verify all five packing conditions simultaneously over their complete continuous domains.

Lemma 7.10 (Exact continuum packing)**.** *Let $1 \le m \le m' \le 13$, and let*

$$(y_1, \dots, y_{m+m'}) \in \Xi(B_{1,m}, B_{1,m'}, m, m', \delta)$$

for the datum of Table 3. Then conditions (A)–(E) *of Proposition 7.8 hold. In condition* (D) *this assertion is uniform for*

$$\frac{2}{5} - \varepsilon_{\mathrm{a}} \le \gamma \le \frac{1}{3} + 8\omega + \frac{7\delta}{3} + 3\varepsilon_{\mathrm{a}}, \qquad 0 \le \omega_0 \le \omega.$$

Proof. The assertion is a finite family of exact rational polyhedral implications. By symmetry we may order the rough coordinates on each side. For each of the 91 pairs (m, m') and each condition (A)–(E), subdivide the corresponding rational polytope whenever a proposed bin inequality changes sign. On every terminal cell an explicit partition assigns every rough coordinate to one of the required bins, and a nonnegative rational dual vector proves each bin inequality throughout that cell. At every internal node the two children are obtained by adjoining an affine inequality and its reverse, so they cover the parent cell. An empty child carries a rational Farkas certificate. Exact verification gives 455 successful root certificates, with no uncovered, failed, or indeterminate cell. Cases with one rough side empty follow by adjoining a part of size δ, applying the corresponding $(1, m)$ certificate, and deleting that part; the case with both sides empty is immediate. □

Together with the half-level argument of the next section, this gives the arithmetic input to the sieve criterion.

Theorem 7.11 (Arithmetic admissibility of the chosen support)**.** *For the rational datum in Table 3, every modulus generated by the support is covered by Bombieri–Vinogradov, the half-level theorem, or one of Stadlmann's positive-level Type I, Type II, and Type III estimates.*

Proof. Lemma 7.10 proves conditions A–E of Proposition 7.8, including Condition D throughout its continuous (γ, ω_0)-range. Theorem 8.3 supplies the sub-half, transition, and endpoint blocks, and the strict positive-level walls license the remaining blocks through the estimates of Section 5. □

8. The transition through the half-level

The factor-extraction argument is naturally formulated at and above the half-level, whereas the dyadic decomposition of the modulus family also produces a logarithmically thin range just below it. This section separates the ranges and treats each one. Throughout we write

$$\log_x q = \frac{1}{2} + 2\omega_0 = \frac{1}{2} - \kappa, \qquad \omega_0 = -\frac{\kappa}{2}.$$

There are three steps. We first divide the modulus family into a Bombieri–Vinogradov range, a shrinking transition range, the endpoint block, and the positive-level range. We then show that every divisor interval available at the endpoint persists across the transition range after a small inward retreat. Finally we sum the discrepancy estimates uniformly over all four ranges. One consequence of this separation is that the formally negative fourth Type-IIc capacity appearing in the printed statement of Proposition 3(D) of [Sta26] never enters the argument: Type IIc is invoked only for $\omega_0 \geq 0$, and sub-half moduli are handled by Bombieri–Vinogradov and endpoint transport.

8.1. **The four modulus ranges.** The distinction between these ranges is analytic. Far enough below $x^{1/2}$, the logarithmic saving in Bombieri–Vinogradov absorbs all of the decompositions. At the endpoint, the factorization has a zero-capacity bin. Between them lies a short range in which neither statement can be invoked verbatim, and above the endpoint one uses Stadlmann's Type estimates. We therefore make the four sets disjoint from the outset.

Fix $C_{\mathrm{BV}} > 0$, after all finite decompositions have been chosen. The dyadic modulus blocks are divided into

$$\begin{aligned} \mathcal{Q}_{\mathrm{BV}} &= \{q : \theta \leq \tfrac{1}{2} - C_{\mathrm{BV}} \tfrac{\log\log x}{\log x}\}, \\ \mathcal{Q}_{\mathrm{tr}} &= \{q : \tfrac{1}{2} - C_{\mathrm{BV}} \tfrac{\log\log x}{\log x} < \theta < \tfrac{1}{2}\}, \\ \mathcal{Q}_0 &= \{q : \theta = \tfrac{1}{2}\}, \\ \mathcal{Q}_+ &= \{q : \theta > \tfrac{1}{2}\}, \end{aligned}$$

where $\theta = \log_x q$. Stratifying by the exponent rather than by dyadic blocks makes the four classes disjoint and exhaustive. The class $\mathcal{Q}_0$ contains at most one integer and is generically empty; it is retained because the endpoint capacities at $\theta = 1/2$ are the ones the factorization argument supplies, and they govern the two neighbouring ranges.

For $\mathcal{Q}_{\mathrm{BV}}$ we use the bilinear form of the Bombieri–Vinogradov theorem given in Theorem 2.9 of [Pol14a]. That statement already applies to divisor-bounded convolution coefficients in primitive residue classes; it is not inferred from a theorem about Λ alone.

Lemma 8.1 (Bilinear Bombieri–Vinogradov)**.** *Let α and β be divisor-bounded coefficient sequences at scales M, N, with $MN \asymp x$, and suppose that at least one sequence has the Siegel–Walfisz property uniformly with the auxiliary coprimality condition in Definition 5.1. Assume also that $\min(M, N) \geq x^{\eta}$ for some fixed $\eta > 0$. For every $A > 0$ there is $B = B(A) > 0$ such that, uniformly in primitive residue classes a,*

$$\sum_{q \leq x^{1/2}(\log x)^{-B}} |\Delta(\alpha \star \beta; x; q, a)| \ll_A \frac{x}{(\log x)^A}.$$

The conclusion is uniform after the fixed smooth dyadic decompositions and coprimality restrictions used in the prime decomposition.

The hypotheses hold in the present application. The coefficient sequences are divisor bounded, and every nontrivial factor has scale at least x^{η} for a fixed $\eta > 0$. The sequence designated as β has the Siegel–Walfisz property with the auxiliary coprimality integer from Definition 5.1, and restricting both variables to be coprime to the fixed pre-sieving modulus is permitted by that definition. The residue classes produced by the shifts are primitive: otherwise a prime divisor of the modulus would divide both the shift and its pre-sieved residue, and the corresponding sieve term would vanish.

Finally, smooth dyadic partitions and partial summation produce $O((\log x)^{C_1})$ admissible pieces, while the finite Heath–Brown decomposition, the residue subclasses, and the coprimality subdivisions contribute another $O((\log x)^{C_2})$ pieces. Applying Theorem 2.9 of [Pol14a] with saving $A + C_1 + C_2 + 2$, and then summing absolute values, leaves the saving A required in Definition 4.4.

8.2. **Transport of the endpoint factorization.** We next quantify how much each endpoint interval changes when the modulus level moves slightly below $x^{1/2}$. The point is that the inward retreat ε_t leaves a positive reserve in every bin that must receive a rough factor, while the fourth bin whose endpoint capacity is zero is kept empty and so is never interpreted as having negative capacity.

For the Type-IIc leaf, write $d = \delta^*(\gamma)$ and fix $\varepsilon_t = \varepsilon_{\mathrm{a}}/100$. The unsimplified factor-extraction intervals are

$$\begin{aligned} a_1 &= \gamma - 3\varepsilon_t - d, & b_1 &= \gamma - 3\varepsilon_t, \\ a_2 &= 1 - \gamma - 6\varepsilon_t - \theta - d, & b_2 &= 1 - \gamma - 6\varepsilon_t - \theta, \\ a_3 &= 2 - \gamma - 52\varepsilon_t - d - 4\theta, & b_3 &= 2 - \gamma - 52\varepsilon_t - 4\theta, \end{aligned}$$

where $\theta = \log_x q = 1/2 - \kappa$. The four bin capacities before the final uniform weakening are

$$2a_1 + b_3, \qquad b_2, \qquad \theta - b_1 - a_2, \qquad a_1 - 2b_1 - a_3.$$

Substitution gives, term by term,

$$\begin{aligned} 2a_1 + b_3 &= 2(\gamma - 3\varepsilon_t - d) + 2 - \gamma - 52\varepsilon_t - 4(\tfrac{1}{2} - \kappa) \\ &= \gamma - 2d + 4\kappa - 58\varepsilon_t, \\ b_2 &= 1 - \gamma - 6\varepsilon_t - (\tfrac{1}{2} - \kappa) = \tfrac{1}{2} - \gamma + \kappa - 6\varepsilon_t, \\ \theta - b_1 - a_2 &= (\tfrac{1}{2} - \kappa) - (\gamma - 3\varepsilon_t) - (1 - \gamma - 6\varepsilon_t - \tfrac{1}{2} + \kappa - d) \\ &= d - 2\kappa + 9\varepsilon_t, \\ a_1 - 2b_1 - a_3 &= \gamma - 3\varepsilon_t - d - 2(\gamma - 3\varepsilon_t) - (2 - \gamma - 52\varepsilon_t - d - 4(\tfrac{1}{2} - \kappa)) \\ &= -4\kappa + 55\varepsilon_t. \end{aligned}$$

The comparison with the endpoint capacities is summarized in Table 2.

Table 2. Transport of the Type-IIc capacities below the half-level.

bin	raw capacity	endpoint capacity	difference
1	$\gamma - 2d + 4\kappa - 58\varepsilon_t$	$\gamma - 2d - \varepsilon_{\mathrm{a}}$	$4\kappa + 42\varepsilon_{\mathrm{a}}/100$
2	$\frac{1}{2} - \gamma + \kappa - 6\varepsilon_t$	$\frac{1}{2} - \gamma - \varepsilon_{\mathrm{a}}$	$\kappa + 94\varepsilon_{\mathrm{a}}/100$
3	$d - 2\kappa + 9\varepsilon_t$	$d - \varepsilon_{\mathrm{a}}$	$109\varepsilon_{\mathrm{a}}/100 - 2\kappa$
4	$-4\kappa + 55\varepsilon_t$	0	$55\varepsilon_{\mathrm{a}}/100 - 4\kappa$

All four differences are nonnegative for

$$0 \le \kappa \le \frac{11}{80}\varepsilon_{\mathrm{a}}.$$

Since $C_{\mathrm{BV}} \log\log x/\log x = o(1)$, every block in $\mathcal{Q}_{\mathrm{tr}}$ lies in this interval once x is sufficiently large. The same phenomenon holds for the other factorization leaves, and the following lemma states it in general.

Lemma 8.2 (Monotone endpoint transport). *Consider any factor-extraction leaf used in the arithmetic certificate. Suppose its bin capacities are affine functions of $\theta = \log_x q$, and an endpoint partition at $\theta = 1/2$ has reserve $\eta > 0$ in every strict bin. If L_{leaf} is the maximum absolute slope of these capacities, then the same partition is valid whenever*

$$\left|\theta - \frac{1}{2}\right| < \frac{\eta}{L_{\mathrm{leaf}}}.$$

Here the minimum defining η is taken over every bin, a bin whose endpoint capacity is zero contributing reserve 0*; such a bin may not simply be ignored, because the extraction argument uses the nonnegativity of its capacity even when the bin is empty.*

Proof. Changing θ by u changes each capacity by at most $L_{\mathrm{leaf}}|u|$. The displayed bound makes this loss smaller than the endpoint reserve. For a bin that is empty at the endpoint the relevant requirement is that its capacity remain nonnegative: in Lemma 7.7 the fourth capacity $a_1 - 2b_1 - a_3$ enters the construction of d_1 through the inequality $\sum_{i\in I_1} y_i + \sum_{i\in J_1} z_i \ge a_1 - (a_1 - 2b_1 - a_3) = 2b_1 + a_3$, which fails once that capacity turns negative, whether or not I_4 is empty. This is why row 4 of Table 2 determines the window. The two-factor, three-factor, and four-factor extraction leaves all have finitely many affine endpoints. In particular, the minimum of their positive radii therefore gives one uniform transition window. □

8.3. **Uniform discrepancy across the boundary.** It remains to combine the local estimates without losing uniformity. The number of convolution, dyadic, and residue-class pieces is only a fixed power of $\log x$, so we may demand a correspondingly larger logarithmic saving from each input estimate and then sum absolute discrepancies. The following theorem is the arithmetic statement passed back to the sieve criterion.

Theorem 8.3 (Half-level arithmetic coverage)**.** *For the convolution sequences and support used in this paper, the absolute sum of discrepancies over $\mathcal{Q}_{\mathrm{BV}} \cup \mathcal{Q}_{\mathrm{tr}} \cup \mathcal{Q}_0 \cup \mathcal{Q}_+$ is $O_A(x(\log x)^{-A})$ for every $A > 0$, uniformly in all primitive residue classes and auxiliary coprimality conditions occurring in the sieve.*

Proof. Lemma 8.1 treats $\mathcal{Q}_{\mathrm{BV}}$. The Heath–Brown decomposition, smooth dyadic partitions, residue subclasses, partial summation, and coprimality restrictions introduce only a fixed power of $\log x$; choosing the saving in that lemma sufficiently large absorbs all of these losses.

The endpoint partitions apply verbatim on $\mathcal{Q}_0$ when it is nonempty. Table 2 transports the four-factor Type-IIc partition throughout $\mathcal{Q}_{\mathrm{tr}}$, and Lemma 8.2 transports every two- and three-factor leaf used there. Stadlmann's positive-level estimates from Section 5 treat $\mathcal{Q}_+$. Their strict parameter inequalities are uniform on each compact positive-level chamber.

There are $O((\log x)^C)$ decomposition and dyadic pieces for a fixed C. Applying each estimate with an additional logarithmic saving and summing absolute discrepancies proves the assertion. ☐

9. The parameters and the variational problem

This section fixes the rational support used in the arithmetic argument and derives the finite-dimensional variational problem attached to it. We first state the parameter datum in both physical and rescaled coordinates, then explain which analytic walls determine it and verify the active margins. A dilation by four converts the physical sieve criterion into a Rayleigh quotient with threshold 4. Finally we describe the rescaled support and compute the moving endpoint of a marginal integral, which is the geometric input for the exact evaluation in Section 10.

9.1. **The parameter datum.** Table 3 is the single parameter point used in both halves of the proof. The "physical" column is in the normalization of the Stadlmann support, while the last column records the dilation used for the finite-dimensional variational problem. Both normalizations are listed so that every subsequent inequality can be checked without an implicit change of scale.

Table 3. Exact parameter datum $\mathscr{P}$.

quantity	physical	rescaled by 4
k	45	45
ω	7/1000	–
(A_0, A_1)	$(-1/125, 257/1000)$	–
ε_{s}	1/125	–
δ	41/2500	$g = 41/625$
$B_{1,m}$, $1 \le m \le 10$	$\frac{1}{5000}(777, 794, 875, 917, 953, 983, 1016, 1042, 1063, 1081)$	$C(m) = 4B_{1,m}$
$B_{1,m}$, $m \ge 11$	1081/5000	$C(m) = 1081/1250$
(ξ_1, ξ_2, ξ_3)	$(19/50, 2/5, 2/5)$	–
ε_{a}	10^{-10}	–

With the notation of Definition 4.1, the support is

$$T = \left\{ t \in [0,\infty)^{45} : \sum_{i=1}^{45} t_i < \frac{53}{200}, \quad \sum_{t_i > \delta} t_i \le B_{1,\#\{i : t_i > \delta\}} \right\}, \qquad \delta = \frac{41}{2500},$$

where the cap sequence is the one in Table 3, with $B_{1,0} = 0$. The marginal cutoff is $A_1 - \varepsilon_{\mathrm{s}} = 249/1000$.

9.2. **Why these parameters.** The choices of ξ_1, ξ_2, ξ_3 are forced by the analytic constraints. The endpoint $\xi_2 = 2/5$ is the largest value for which the Harman decomposition of Section 7 reduces to the prime indicator, so that $\rho = 1_{\mathbb{P}}$ and $c_1 = c_2 = 0$. The Type-III constraint is monotone increasing in ξ_3, so its weakest admissible value is $\xi_3 = \xi_2 = 2/5$.

At these values, the active Type-II inequality in Proposition 7.8 is

$$\delta < \frac{\xi_2}{4} + \frac{11}{16} - 3A_1 - 2\varepsilon_{\mathrm{a}} = \frac{63}{80} - 3A_1 - 2\varepsilon_{\mathrm{a}}.$$

The chosen rational point satisfies

$$3A_1 + \delta = 3 \cdot \frac{513}{2000} + \frac{179}{10000} = \frac{3937}{5000} = \frac{63}{80} - \frac{1}{10000},$$

so it lies 10^{-4} inside the zero-loss frontier, before the fixed analytic retreat $2\varepsilon_{\mathrm{a}}$.

The active zero-loss bounds for ξ_1 give

$$\frac{2833}{7500} < \xi_1 < \frac{2}{5},$$

and we take $\xi_1 = 19/50$.

The remaining quantities $A_1, \varepsilon_{\mathrm{s}}, \delta$, and $(B_{1,m})$ are chosen to balance the available support against the active Type-II wall. Table 4 records the six inequalities that are closest to being tight; the complete list, including every analytic wall and packing inequality, appears in Appendix B.

TABLE 4. The six principal exact reserves.

condition	left	right	slack
Definition 1: $B_3 \le B_2 + \delta$	7/40	438/2500	1/5000
Proposition 2: $\xi_2 \le 2/5$	2/5	2/5	0
Prop. 3(II), cap wall 2	41/2500	82499999/5000000000	499999/5000000000
Dominant Type-II wall	3937000001/5000000000	63/80	499999/5000000000
First-rung cap wall	777/5000	778/5000	1/5000
Transition bin 3	11/400000000000	109/1000000000000	163/2000000000000

9.3. **Rescaling.** Put $x_i = 4t_i$ and define $\widetilde{F}(x) = F(x/4)$. The change of variables in 45 dimensions gives

$$I(F) = \int_T F(t)^2\, dt = 4^{-45} \int_{\widetilde{T}} \widetilde{F}(x)^2\, dx = 4^{-k} I(\widetilde{F}).$$

For a marginal, the outer integral has 44 variables and the two inner integrals each contribute one factor 1/4. Hence

$$J_i(F) = 4^{-44} \cdot 4^{-1} \cdot 4^{-1} J_i(\widetilde{F}) = 4^{-(k+1)} J_i(\widetilde{F}).$$

Consequently

$$\frac{kJ(F)}{I(F)} = \frac{1}{4} \frac{kJ(\widetilde{F})}{I(\widetilde{F})},$$

so the physical threshold 1 is exactly the rescaled threshold 4. The corresponding dictionaries are

	$A_1 + \varepsilon_{\mathrm{s}}$	$A_1 - \varepsilon_{\mathrm{s}}$	δ	B_m
physical	53/200	249/1000	41/2500	$B_{1,m}$
rescaled	$L = 53/50$	$\tau = 249/250$	$g = 41/625$	$C(m) = 4B_{1,m}$.

9.4. **Support and marginal geometry.** The denominator of the Rayleigh quotient integrates over the full support, whereas a numerator term first fixes all but one coordinate and then integrates along the remaining fiber. We describe both regions explicitly. The statistics $s(x)$ and $R_g(x)$ record the number and total mass of the rough coordinates, and C converts that information into the relevant cap; the second lemma determines the exact endpoint of each marginal fiber.

For $x \in [0,\infty)^{45}$, put

$$P_j(x) = \sum_{i=1}^{45} x_i^j, \qquad s(x) = \#\{i : x_i > g\}, \qquad R_g(x) = \sum_{x_i > g} x_i,$$

and extend the cap function by $C(0) = 0$.

Lemma 9.1 (Rescaled support)**.** *The dilation $x = 4t$ maps T onto*

$$\widetilde{T} = \{x \in [0,\infty)^{45} : P_1(x) \le L,\ R_g(x) \le C(s(x))\}.$$

Proof. The total-mass inequality is multiplied by 4. A coordinate satisfies $t_i > \delta$ precisely when $x_i > g = 4\delta$, and multiplication by 4 also sends the two physical rough caps to $C(1) = C(2) = 31/50$ and $C(s) = 17/25$ for $s \ge 3$. ☐

Lemma 9.2 (Moving marginal endpoint)**.** *Fix $\widehat{x} = (x_1, \dots, x_{44})$ and write*

$$S = \sum_{i=1}^{44} x_i, \qquad s = \#\{i : x_i > g\}, \qquad \rho_1 = \sum_{\substack{1 \le i \le 44 \\ x_i > g}} x_i.$$

For

$$\widehat{T} = \{\widehat{x} : S \le \tau,\ \rho_1 \le C(s)\},$$

the last coordinate ranges over $0 \le u \le U(\widehat{x})$, where

$$U(\widehat{x}) = \min\{L - S, \max\{g, C(s+1) - \rho_1\}\}.$$

Proof. The total-radius condition gives $u \le L - S$. Every $0 \le u \le g$ is available because it does not change the rough count or rough mass. If $u > g$, the rough count becomes $s+1$, and the rough cap gives $u \le C(s+1) - \rho_1$. The union of the always available small interval with this possible rough continuation therefore ends at $\max\{g, C(s+1) - \rho_1\}$; intersecting with the total-radius interval gives the formula. The condition $S \le \tau$ is exactly the marginal cutoff on the other 45 coordinates. ☐

For a symmetric polynomial F, the two quadratic forms are therefore

$$I_T(F) = \int_{\widetilde{T}} F(x)^2 \, dx$$

and

$$J_T(F) = 45 \int_{\widehat{T}} \left(\int_0^{U(\widehat{x})} F(\widehat{x}, u) \, du \right)^2 d\widehat{x}.$$

10. Exact evaluation of the variational integrals

We now explain why the variational claim can be reduced to exact rational arithmetic. The support is stratified by its set of rough coordinates; those coordinates are translated by the rough threshold, and the remaining upper bounds are handled by inclusion–exclusion. For the numerator one adds a subdivision according to which affine expression gives the moving fiber endpoint. Every resulting integral is a polynomial moment over a rational simplex, so the whole calculation reduces to the elementary identity

$$\int_{\substack{z_i \ge 0 \\ \sum_{i=1}^m z_i \le R}} \prod_{i=1}^{m} z_i^{e_i} \, dz = R^{m + \sum_i e_i} \frac{\prod_i e_i!}{(m + \sum_i e_i)!}, \qquad e_i \in \mathbb{Z}_{\ge 0}. \tag{10.1}$$

We state the consequence we need separately, in order to keep the geometry apart from the algebra that follows.

Lemma 10.1 (Rational polytope moments). *Let $P \subset \mathbb{R}^m$ be a bounded polytope cut out by finitely many affine inequalities with rational coefficients. If $G \in \mathbb{Q}[z_1, \dots, z_m]$, then*

$$\int_P G(z)\,dz \in \mathbb{Q}.$$

Moreover, the integral is a finite rational linear combination of the simplex moments in (10.1).

Proof. If P has empty interior, its m-dimensional integral is zero, so suppose that P is full-dimensional. Every vertex of P is rational, since it is the unique solution of a full-rank subsystem of the defining rational affine equations. Triangulate P using its vertices. On a simplex with vertices $v_0, \dots, v_m$, the affine substitution

$$z = v_0 + \sum_{i=1}^{m} u_i(v_i - v_0), \qquad u_i \geq 0, \quad \sum_i u_i \leq 1,$$

has rational coefficients and rational Jacobian determinant. Expanding $G(z)$ after this substitution gives a polynomial in the u_i with rational coefficients. Formula (10.1), with $R = 1$, evaluates each monomial moment rationally. Summing over the triangulation proves both assertions. □

Proposition 10.2 (Exactness of the variational integrals). *If F has rational coefficients, then $I_T(F)$ and $J_T(F)$ are rational.*

Proof. For each subset $R \subseteq \{1, \dots, 45\}$, consider the stratum on which $x_i > g$ exactly when $i \in R$. On this stratum the support conditions are

$$x_i \geq g \ (i \in R), \qquad 0 \leq x_i \leq g \ (i \notin R), \qquad \sum_i x_i \leq L, \qquad \sum_{i \in R} x_i \leq C(|R|).$$

Changing open faces to closed faces does not alter the integral. This is a bounded rational polytope. Translating $x_i = g + z_i$ for $i \in R$ makes all lower coordinate bounds zero and preserves rationality of every remaining face. Since F^2 is a rational polynomial, Lemma 10.1 evaluates its integral on the stratum as a finite rational combination of simplex moments. There are only finitely many choices of R, so their sum is $I_T(F)$.

For $J_T(F)$, fix in the same way the rough set among the first 44 coordinates. Lemma 9.2 writes the moving endpoint as

$$U(\widehat{x}) = \min\{L - S, \max\{g, C(s+1) - \rho_1\}\}.$$

The hyperplanes on which two of the three displayed affine functions agree partition the stratum into finitely many bounded rational polytopes. On each cell, U is one specified affine function. If $F(\widehat{x}, u) = \sum_{r=0}^{d} f_r(\widehat{x}) u^r$, then

$$\int_0^{U(\widehat{x})} F(\widehat{x}, u)\,du = \sum_{r=0}^{d} \frac{f_r(\widehat{x})}{r+1} U(\widehat{x})^{r+1},$$

which is a rational polynomial in $\widehat{x}$ on that cell. Its square is again a rational polynomial, and Lemma 10.1 applies. Summing over the finitely many rough sets and ordering cells proves that $J_T(F)$ is rational. □

11. The variational inequality

Sections 5–8 shows that the support T is permissible. It remains to show that T is large enough for the sieve. We use the symmetric space of degree 21 generated by products of even power sums and powers of the remaining simplex radius. In that basis I_T and J_T are rational Gram matrices, by Proposition 10.2, so any explicit rational coefficient vector gives a rigorous lower bound for the largest generalized eigenvalue.

More precisely, if $\lambda = (\lambda_1, \dots, \lambda_r)$ is a partition all of whose parts are positive and even, write $|\lambda| = \lambda_1 + \cdots + \lambda_r$ and $P_\lambda = \prod_{j=1}^{r} P_{\lambda_j}$, with $P_\varnothing = 1$. We use the space

$$\mathcal{B}_{21} = \operatorname{span}_{\mathbb{Q}}\bigl\{(L - P_1)^a P_\lambda : a \geq 0,\ \lambda_j \in 2\mathbb{Z}_{>0},\ a + |\lambda| \leq 21\bigr\}.$$

After repeated monomials are identified, the displayed family has dimension 846. Let $P_\star \in \mathcal{B}_{21}$ be the polynomial specified by the exact rational coefficient vector used in the certificate, and define

$$F_\star(x) = \mathbf{1}_{\widetilde{T}}(x) P_\star(x).$$

Theorem 11.1 (Variational estimate). *For the symmetric rational polynomial $P_\star$ defined above, the supported function $F_\star = \mathbf{1}_{\widetilde{T}} P_\star$ satisfies*

$$\frac{J_T(F_\star)}{I_T(F_\star)} = 4.00438409833460131937\ldots > 4.$$

Proof. Because both functionals integrate only over the support and its marginal fibers, $I_T(F_\star) = I_T(P_\star)$ and $J_T(F_\star) = J_T(P_\star)$. Proposition 10.2 therefore reduces both values to rational combinations of the simplex identity (10.1). Their exact values satisfy

$$I_T(F_\star) > 0, \qquad J_T(F_\star) - 4I_T(F_\star) > 0.$$

Division by the first positive quantity gives the stated inequality. □

12. Completion of the proof

The two substantive inputs are now in place: Theorem 7.11 licenses every modulus produced by the support, and Theorem 11.1 supplies the strict variational inequality. The sieve criterion therefore gives DHL$[45, 2]$, and it remains only to exhibit an admissible 45-tuple of diameter 212.

$$\begin{aligned}\mathcal{H}_{45} = \{&0, 2, 12, 14, 24, 26, 30, 36, 44, 50, 54, 56, 60, 66, 72, 74, 80, 84, 92, 96, 102, 110, 114,\\ &116, 122, 126, 134, 140, 144, 150, 156, 162, 164, 170, 176, 180, 182, 186, 192, 194,\\ &200, 204, 206, 210, 212\}.\end{aligned}$$

Lemma 12.1. *The set $\mathcal{H}_{45}$ is an admissible 45-tuple of diameter 212.*

Proof. The displayed list has 45 distinct elements, with minimum 0 and maximum 212. For

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43,$$

respectively, its reductions omit the residue classes

$$1, 1, 3, 6, 9, 3, 1, 13, 17, 4, 11, 4, 4, 3.$$

Thus the tuple does not cover every residue class for any prime $p \leq 43$. If $p \geq 47$, its 45 elements cannot cover all p residue classes. Hence $\mathcal{H}_{45}$ is admissible. □

Proof of Theorem 1.1. Theorem 7.11 supplies equidistribution for every generated modulus. Theorem 11.1, together with the rescaling in Section 9, supplies the strict physical variational inequality. Theorem 4.10 then gives DHL$[45, 2]$. Apply DHL$[45, 2]$ to $\mathcal{H}_{45}$. Infinitely many translates contain two primes, and their difference is at most 212. □

Acknowledgements

We thank Shubho Sengupta for setting up the compute environment and Angus Joshi and Alistair Tucker for helping with the preparation of the draft. We would also like to thank Rohan Aanegola, Alberto Alfarano, Evan Chen, Ruiz Chen, Constantin Eichenberg, Srihari Ganesh, Tobias Gessler, Erik Gregory, Leopold Haller, Sidharth Hariharan, Vasily Ilin, Albert Jiang, Yunjiang Jiang, Tadeusz Jordan, Andranik Kurghinyan†, Simon Mahns, Rishi Malhotra, Bhavik Mehta, Michał Mogielnicki, Connor Olson, Gaurang Pendharkar, Bartosz Piotrowski, Anish Rajeev, Karun Ram, Aditya Ramabadran, Charlie Ray, Aleksey Tsaplin, Niels Voss, Chao Wang, Chenkai Wang, Jimmy Xin, and Yinglun Zhu for many helpful discussions, comments, and contributions to the broader collaborative effort surrounding this work.

Appendix A. Source correspondences

Some statements used in the proof are taken unaltered from [Sta26] or from the earlier distribution estimates on which it rests. Others are specialized to the direct-prime regime used here. Still more are strengthened by retaining information that was discarded when [Sta26] passed to uniform sufficient conditions. Table 5 provides a reference.

Table 5: Source-to-paper correspondences.

present item	exact source	changed hypothesis	changed conclusion	proof location
Sieve realization	Stadlmann [Sta26, Proposition 1]; Polymath8b [Pol14b, Section 5]	Specialize the mixed support to the direct-prime datum and retain the inward support retreat.	Replace the general minorant quotient by $kJ_T/I_T > 1$.	Section 4
Prime decomposition	Stadlmann [Sta26, Proposition 2]	Set $\xi_2 = 2/5$.	The exceptional Buchstab terms vanish, so $\rho = 1_{\mathbb{P}}$ and $c_1 = c_2 = 0$.	Section 7
Type I	Stadlmann [Sta26, Lemma 5]; Baker–Irving [BI17, Lemma 5]	Require only the divisor interval used by the proof.	Equidistribution for the factorization class D_I.	Sections 5 and 6
Types IIa and IIb	Stadlmann [Sta26, Lemmas 3–4]; Polymath8a [Pol14a, Theorem 2.8(iv),(ii)]	Replace multiple dense divisibility by the displayed divisors r and u.	Equidistribution for D_{IIa} and D_{IIb}.	Sections 5 and 6
Type IIc	Stadlmann [Sta26, Lemma 6]; Stadlmann [Sta25, Theorem 1]	Retain the factors r, u, d_1 and their unsimplified intervals.	Equidistribution for D_{IIc} under the three displayed walls.	Sections 5 and 6
Type III	Stadlmann [Sta26, Lemma 7]; Polymath8a [Pol14a, Theorem 2.8(v)]	Retain the single divisor interval used in the source argument.	Equidistribution for D_{III}.	Sections 5 and 6
Factor extraction	Stadlmann [Sta26, Lemmas 11–13]	Insert the chosen rational caps and keep the endpoint intervals before the uniform weakening.	Two-, three-, and four-factor decompositions for every generated modulus.	Section 7
Half-level transition	Stadlmann [Sta26, proof of Lemma 6]; the Bombieri–Vinogradov estimate in [Pol14a, Theorem 2.9]	Separate the dyadic decomposition into four disjoint ranges.	Uniform coverage below, at, and above $x^{1/2}$.	Section 8
Variational certificate	Stadlmann [Sta26, Section 6]; Polymath8b [Pol14b, Section 4]	Use the 45-shift rational support and a symmetric rational polynomial.	The exact rescaled quotient exceeds 4.	Sections 9–11

Appendix B. Exact verification of the parameter inequalities

This appendix collects the numerical substitutions used when Theorem 7.11 applies the support and factorization results to the parameters of Table 3. The linked description in the first column returns the reader to the definition, proposition, lemma, or theorem in which the inequality is used.

The columns headed "left" and "right" give the two sides after inserting the chosen rational parameters, oriented so that the required comparison is left $\leq$ right, and the slack is right $-$ left. Positive slack indicates a strict inequality; zero slack occurs only where equality is permitted or where an endpoint is handled separately. A decimal column is included to give an idea of scale but the proof uses the exact fractions. The final column identifies the part of the argument for which the comparison is needed. The rows fall into five groups: support conditions, the direct-prime decomposition, analytic range conditions, the factor-packing conditions A–E of Proposition 7.8, and the transition through $x^{1/2}$.

Table 6: Exact analytic and combinatorial slack ledger.

condition	left	right	slack	decimal	use
Support: $\delta > 0$	0	41/2500	41/2500	0.0164	support
Support: $B_1 > \delta$	41/2500	777/5000	139/1000	0.139	support
Support: $B_m \leq B_{m+1}$	0	17/5000	17/5000	0.0034	monotonicity
Support: $B_{m+1} \leq B_m + \delta$	81/5000	82/5000	1/5000	0.0002	heredity
First-rung cap wall	777/5000	778/5000	1/5000	0.0002	arithmetic certificate
Prime decomposition: $\xi_2 \leq 2/5$	2/5	2/5	0	0	direct prime
Prime decomposition: $\xi_3 \geq \xi_2$	2/5	2/5	0	0	Type III split
Prime decomposition: $2\xi_1 + 3\xi_2 < 2$	49/25	2	1/25	0.04	Harman decomposition
Prime decomposition: $\xi_1 + 9\xi_2 < 4$	199/50	4	1/50	0.02	Harman decomposition
Prime decomposition: $2\xi_1 + \xi_2 > 1$	1	29/25	4/25	0.16	Harman decomposition
Prime decomposition: $17\xi_2 < 7$	34/5	7	1/5	0.2	Harman decomposition
Type I admissibility, first wall	41/2500	279999997/15000000000	33999997/15000000000	0.0022666665	Prop. 3(I)
Type I admissibility, second wall	41/2500	1309999993/35000000000	735999993/35000000000	0.02102857123	Prop. 3(I)
Type II admissibility, first wall	0	69599997/2000000000	69599997/2000000000	0.0347999985	Prop. 3(II)
Type II admissibility, first cap wall	41/2500	87999999/5000000000	5999999/5000000000	0.0011999998	Prop. 3(II)
Type II admissibility, second cap wall	41/2500	82499999/5000000000	499999/5000000000	$9.99998e-05$	Prop. 3(II)
Type III admissibility	41/2500	127499999/5000000000	45499999/5000000000	0.0090999998	Prop. 3(III)
Dominant Type II wall	3937000001/5000000000	63/80	499999/5000000000	$9.99998e-05$	arithmetic certificate
Exact packing roots A–E	455	455	0	0	all continuous packing cells
Half-level transition, third bin	11/400000000000	109/1000000000000	163/2000000000000	$8.15e-11$	transition
Half-level transition, fourth bin	11/200000000000	11/200000000000	0	0	transition

Appendix A. Formal certificate by AxiomProver

AxiomProver, an AI system under development by Axiom Math, autonomously generated from natural-language specifications a Lean certificate of the deduction of Theorem 1.1. The formal development takes as hypotheses the five Type I, Type II, and Type III equidistribution estimates stated in Section 5, the bilinear Bombieri–Vinogradov theorem used below the half-level, the Harman decomposition reducing the prime indicator to the relevant convolution classes, and the variational certificate of Theorem 11.1. The mathematical sources for these analytic inputs are [Pol14a, BI17, Sta25, Sta26]. The exact rational verification of the variational certificate is presently performed separately and supplied to the formal development as a hypothesis.

References

[BI17] Roger C Baker and Alastair J Irving. Bounded intervals containing many primes. *Mathematische Zeitschrift*, 286:821–841, 2017.

[GPY09] Daniel A. Goldston, János Pintz, and Cem Y. Yıldırım. Primes in tuples. I. *Annals of Mathematics*, 170(2):819–862, 2009.

[May15] James Maynard. Small gaps between primes. *Annals of mathematics*, 181:383–413, 2015.

[Pol14a] D. H. J. Polymath. New equidistribution estimates of Zhang type. *Algebra & Number Theory*, 8(9):2067–2199, 2014.

[Pol14b] D. H. J. Polymath. Variants of the Selberg sieve, and bounded intervals containing many primes. *Research in the Mathematical sciences*, 1:1–83, 2014.

[Sta25] Julia Stadlmann. On primes in arithmetic progressions and bounded gaps between many primes. *Advances in Mathematics*, 468(110190), 2025.

[Sta26] Julia Stadlmann. Bounded gaps between primes, 2026.

[Zha14] Yitang Zhang. Bounded gaps between primes. *Annals of mathematics*, 179(3):1121–1174, 2014.