

# Contents

A	Introduction & Framework	2
1	Circle-Method Decomposition	3
B	Type I / II Analysis	6
1	Type II Parity Gain: Bilinear reduction to BV	6
2	BV with parity, second moment	6
C	Type III Analysis	9
1	Type III off-diagonal via prime-averaged short-shift gain	9
2	Type III Analysis: Prime-Averaged Short-Shift Gain	11
D	Final Assembly: Proof of the Minor-Arc Bound and Goldbach for Large $N$	13
1	Major arcs, main terms, and comparison	13
2	Minor-arc bound (summary of Parts B–C)	16
3	Final assembly: evaluation of $R(N)$	16
4	Corollary: Goldbach for large $N$	17
E	Appendix – Technical Lemmas and Parameters	17
1	Minor-arc large sieve reduction	17
2	Sieve weight $\beta$ and properties	19
3	Major-arc uniform error	19
4	Auxiliary analytic inputs used in Part B	20
5	Deterministic balanced signs for the amplifier	20
6	Kuznetsov formula and level-uniform kernel bounds	22
7	$\Delta$ -second moment, level-uniform	24
8	Hecke $p \mid n$ tails are negligible	25
9	Oldforms and Eisenstein: uniform handling	26
10	Admissible parameter tuple and verification	26

# Proof of the Goldbach Conjecture

Student Vinzenz Stampf

## Part A

## Introduction & Framework

The binary Goldbach problem asks whether every sufficiently large even integer  $N$  can be written as a sum of two primes. Equivalently, defining

$$R(N) := \sum_{m+n=N} \Lambda(m)\Lambda(n),$$

the conjecture asserts that  $R(N) > 0$  for all even  $N \geq 4$ .

Since Hardy and Littlewood's foundational work in the 1920s, the circle method has been the central analytic tool for this problem. It predicts the asymptotic

$$R(N) \sim \mathfrak{S}(N) \frac{N}{\log^2 N},$$

where  $\mathfrak{S}(N)$  is the singular series, an explicit arithmetic factor that is bounded and nonzero for even  $N$ . Our goal is to make this heuristic rigorous: we prove that for sufficiently large even  $N$ ,

$$R(N) = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta} N}\right),$$

for some  $\eta > 0$ . In particular,  $R(N) > 0$ , hence  $N$  is a sum of two primes.

The novelty of this work lies in combining three modern ingredients:

- a parity-sensitive Bombieri–Vinogradov theorem in the *second moment* (BVP2M),
- a Type III spectral second moment bound via amplifiers and  $\Delta$ -averaging, and
- careful major-arc evaluation with a sieve-theoretic majorant  $B(\alpha)$  for comparison.

## Outline of the argument

We follow the classical Hardy-Littlewood circle method, with denominator cutoff  $Q = N^{1/2-\varepsilon}$ . The proof is organized into four parts.

**Part A. Framework.** We decompose

$$R(N) = \int_0^1 S(\alpha)^2 e(-N\alpha) d\alpha,$$

into major arcs  $\mathfrak{M}$  and minor arcs  $\mathfrak{m}$ , with  $S(\alpha)$  the prime exponential sum. We also introduce a sieve majorant  $B(\alpha)$  and reduce to bounding

$$\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 d\alpha,$$

by  $O(N/(\log N)^{3+\eta})$ .

**Part B. Type I/II analysis.** We treat Type I and Type II bilinear sums using Theorem B.3, our Bombieri–Vinogradov with parity in second moment form. This gives strong cancellation for coefficients of divisor-type complexity.

**Part C. Type III analysis.** The difficult Type III sums are handled by an amplifier method (Lemma E.7), a  $\Delta$ -second moment bound (Lemma E.18), and Kuznetsov’s formula with level-uniform kernel bounds (Lemma E.14). Together these yield Proposition C.2, a second-moment estimate with a genuine power saving in  $Q$ .

**Part D. Assembly.** On the major arcs, we evaluate  $S(\alpha)$  and  $B(\alpha)$  uniformly (Theorem D.5), recovering the singular series  $\mathfrak{S}(N)$ . On the minor arcs, Parts B–C supply the needed  $L^2$  bound (Theorem D.9). Putting the two together yields the asymptotic formula (Theorem D.10) and hence Goldbach’s conjecture for large  $N$  (Corollary D.11).

## Acknowledgments

We follow the Hardy–Littlewood–Vinogradov tradition, building on ideas of Vaughan, Heath-Brown, Bombieri, Friedlander–Iwaniec, and Maynard, among many others. Any errors or omissions are our responsibility.

## 1 Circle-Method Decomposition

Let

$$S(\alpha) = \sum_{n \leq N} \Lambda(n) e(\alpha n), \quad R(N) = \int_0^1 S(\alpha)^2 e(-N\alpha) d\alpha.$$

Fix  $\varepsilon \in (0, \frac{1}{10})$  and set

$$Q = N^{1/2-\varepsilon}.$$

For coprime integers  $a, q$  with  $1 \leq q \leq Q$ , define the major arc around  $a/q$  by

$$\mathfrak{M}(a, q) = \left\{ \alpha \in [0, 1) : \left| \alpha - \frac{a}{q} \right| \leq \frac{Q}{qN} \right\}.$$

Let

$$\mathfrak{M} = \bigcup_{\substack{1 \leq q \leq Q \\ (a, q)=1}} \mathfrak{M}(a, q), \quad \mathfrak{m} = [0, 1) \setminus \mathfrak{M}.$$

Then

$$R(N) = \int_{\mathfrak{M}} S(\alpha)^2 e(-N\alpha) d\alpha + \int_{\mathfrak{m}} S(\alpha)^2 e(-N\alpha) d\alpha = R_{\mathfrak{M}}(N) + R_{\mathfrak{m}}(N).$$

### Parity-blind majorant $B(\alpha)$

Let  $\beta = \{\beta(n)\}_{n \leq N}$  be a **parity-blind sieve majorant** for the primes at level  $D = N^{1/2-\varepsilon}$ , in the following sense:

(B1)  $\beta(n) \geq 0$  for all  $n$  and  $\beta(n) \gg \frac{\log D}{\log N}$  for  $n$  the main  $\leq N$ .

(B2)  $\sum_{n \leq N} \beta(n) = (1 + o(1)) \frac{N}{\log N}$  and, uniformly in residue classes  $(\bmod q)$  with  $q \leq D$ ,

$$\sum_{\substack{n \leq N \\ n \equiv a \pmod{q}}} \beta(n) = (1 + o(1)) \frac{N}{\varphi(q) \log N} \quad ((a, q) = 1).$$

(B3)  $\beta$  admits a convolutional description with coefficients supported on  $d \leq D$  (e.g. Selberg upper-bound sieve), enabling standard major-arc analysis.

(B4) **Parity-blindness:**  $\beta$  does not correlate with the Liouville function at the  $N^{1/2}$  scale (so it does not distinguish the parity of  $\Omega(n)$ ); this is automatic for classical upper-bound Selberg weights.

Define

$$B(\alpha) = \sum_{n \leq N} \beta(n) e(\alpha n).$$

### Major arcs: main term from $B$

On  $\mathfrak{M}(a, q)$  write  $\alpha = \frac{a}{q} + \frac{\theta}{N}$  with  $|\theta| \leq Q/q$ . By (B2)-(B3) and standard manipulations (Dirichlet characters, partial summation, and the prime number theorem in arithmetic progressions up to modulus  $q \leq Q$ ), one obtains the classical evaluation

$$\int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha) d\alpha = \mathfrak{S}(N) \frac{N}{\log^2 N} (1 + o(1)),$$

where  $\mathfrak{S}(N)$  is the singular series

$$\mathfrak{S}(N) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\varphi(q)} \sum_{\substack{a \pmod{q} \\ (a, q) = 1}} e\left(-\frac{Na}{q}\right).$$

Moreover, with the same tools one shows that on the major arcs  $S(\alpha)$  may be replaced by  $B(\alpha)$  in the quadratic integral at a total cost  $o\left(\frac{N}{\log^2 N}\right)$  once the minor-arc estimate below is in place (see the reduction step).

### Reduction to a minor-arc $L^2$ bound

We record the minor-arc target:

$$\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 d\alpha \ll \frac{N}{(\log N)^{3+\varepsilon}}. \quad (\text{A.1})$$

$$\sum_{q \leq Q} \sum_{\chi \bmod q} \left| \sum_{n \leq N} c_n \lambda(n) \chi(n) \right|^2 \ll \frac{NQ}{(\log N)^A} \quad (\text{A.2})$$

**Proposition A.1** (Final assembly of the circle method). *Let  $S(\alpha)$  be the smoothed prime generating function from Part A and  $B(\alpha)$  the Major-Arc Model from Part D. Assume:*

(H1) **Major-arc evaluation for  $B$ .** *Uniformly for even  $N$ ,*

$$\int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha) d\alpha = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta} N}\right)$$

for some fixed  $\eta > 0$ .

(H2) **Minor-arc  $L^2$  control of  $S - B$ .** For some  $A_0 > 3$ ,

$$\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 d\alpha \ll \frac{N}{(\log N)^{A_0}}.$$

(This is Theorem D.9 proved by combining Parts B and C.)

(H3) **Minor-arc  $L^2$  control of  $B$ .** For every  $A > 0$ ,

$$\int_{\mathfrak{m}} |B(\alpha)|^2 d\alpha \ll_A \frac{N}{(\log N)^A}.$$

(This is Lemma E.1.)

(H4) **Global  $L^2$  size.** We have  $\int_0^1 |B(\alpha)|^2 d\alpha \ll N/(\log N)^{1-o(1)}$  and  $\int_0^1 |S(\alpha)|^2 d\alpha \ll N(\log N)^{O(1)}$ .

Then, uniformly for even  $N$ ,

$$R(N) := \int_0^1 S(\alpha)^2 e(-N\alpha) d\alpha = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta'} N}\right)$$

for some  $\eta' > 0$ . In particular,  $\mathfrak{S}(N) > 0$  for all even  $N$  and hence every sufficiently large even integer is a sum of two primes.

*Proof.* Write  $S = B + (S - B)$  and expand on  $\mathfrak{M} \cup \mathfrak{m}$ :

$$\begin{aligned} R(N) &= \int_{\mathfrak{M}} B^2 e(-N\alpha) d\alpha + 2 \int_{\mathfrak{M}} (S - B)B e(-N\alpha) d\alpha + \int_{\mathfrak{M}} (S - B)^2 e(-N\alpha) d\alpha \\ &\quad + \int_{\mathfrak{m}} B^2 e(-N\alpha) d\alpha + 2 \int_{\mathfrak{m}} (S - B)B e(-N\alpha) d\alpha + \int_{\mathfrak{m}} (S - B)^2 e(-N\alpha) d\alpha. \end{aligned}$$

By (H1) the first term is the desired main term. We show that the five remaining terms are  $O(N/\log^{2+\eta'} N)$ .

*Minor arcs.* By (H3),

$$\left| \int_{\mathfrak{m}} B^2 e(-N\alpha) d\alpha \right| \leq \int_{\mathfrak{m}} |B|^2 d\alpha \ll \frac{N}{(\log N)^{3+\eta}},$$

after fixing  $A = 3 + \eta$ . By (H2) and (H3) and Cauchy–Schwarz,

$$\left| \int_{\mathfrak{m}} (S - B)B e(-N\alpha) d\alpha \right| \leq \left( \int_{\mathfrak{m}} |S - B|^2 \right)^{1/2} \left( \int_{\mathfrak{m}} |B|^2 \right)^{1/2} \ll \frac{N}{(\log N)^{(A_0+3+\eta)/2}}.$$

Also  $\int_{\mathfrak{m}} |(S - B)^2| \leq \int_{\mathfrak{m}} |S - B|^2 \ll N/(\log N)^{A_0}$  by (H2). Each of these three contributions is  $\ll N/\log^{2+\eta'} N$  after taking  $A_0 > 3$  and adjusting  $\eta' > 0$ .

*Major arcs (error terms).* For the cross term,

$$\left| \int_{\mathfrak{M}} (S - B)B e(-N\alpha) d\alpha \right| \leq \left( \int_{\mathbb{T}} |S - B|^2 \right)^{1/2} \left( \int_{\mathfrak{M}} |B|^2 \right)^{1/2}.$$

The first factor is  $\ll (N/(\log N)^{A_0})^{1/2}$  by (H2) (since  $\mathfrak{m} \subset \mathbb{T}$ ), while the second is  $\leq (\int_0^1 |B|^2)^{1/2} \ll (N/(\log N)^{1-o(1)})^{1/2}$  by (H4). Hence the cross term is

$$\ll \frac{N}{(\log N)^{(A_0+1-o(1))/2}} \ll \frac{N}{\log^{2+\eta'} N}$$

after increasing  $A_0$  if necessary. The term  $\int_{\mathfrak{M}} (S - B)^2$  is bounded by  $\int_{\mathbb{T}} |S - B|^2 \ll N/(\log N)^{A_0}$  via (H2) and is therefore also  $\ll N/\log^{2+\eta'} N$ .

Collecting all contributions, we obtain

$$R(N) = \int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha) d\alpha + O\left(\frac{N}{\log^{2+\eta'} N}\right),$$

and the claim follows from (H1). Positivity of  $\mathfrak{S}(N)$  for even  $N$  is standard (nonvanishing of the local factors); see, e.g., Hardy–Littlewood or Vaughan [6, §3.6].  $\square$

## Part B

# Type I / II Analysis

## 1 Type II Parity Gain: Bilinear reduction to BV

We record a quantitative Type II input in the dyadic ranges  $M, N$  with  $MN \asymp X$  and  $X^\eta \leq M, N \leq X^{1-\eta}$ . Let  $(a_m)$  and  $(b_n)$  be coefficients supported on  $m \asymp M, n \asymp N$ , with smooth weights and block mean-zero (the latter only reduces the diagonal and is not needed for the bound). Set the Dirichlet convolution

$$c_k := \sum_{mn=k} a_m b_n, \quad k \asymp X.$$

Write  $\lambda$  for the parity-sensitive multiplicative weight used throughout (in applications,  $\lambda = \lambda_{\text{par}}$  or a balanced prime weight; only  $|\lambda| \leq 1$  and the BV-with-parity second moment are used).

**Theorem B.1** (Type II second-moment bound). *Fix  $\varepsilon > 0$  and  $A > 0$ . For  $Q \leq X^{1/2-\varepsilon}$ ,*

$$\sum_{q \asymp Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \text{ primitive}}} \int_{|t| \leq X} \left| \sum_{m \asymp M} \sum_{n \asymp N} a_m b_n \lambda(mn) \chi(mn) (mn)^{-it} \right|^2 \frac{dt}{1+|t|} \ll (Q^2 + X) \frac{X^{1+o(1)}}{(\log X)^A}, \quad (\text{B.1})$$

uniformly for the Type II range  $X^\eta \leq M, N \leq X^{1-\eta}$ .

*Proof.* Set  $c_k = \sum_{mn=k} a_m b_n$  as above. Then the inner sum equals  $\sum_{k \asymp X} c_k \lambda(k) \chi(k) k^{-it}$ . By Theorem B.3 (BV with parity, second moment) applied to the sequence  $a_k := c_k$ , we obtain

$$\sum_{q \asymp Q} \frac{1}{\varphi(q)} \sum_{\chi} \int_{|t| \leq X} \left| \sum_{k \asymp X} c_k \lambda(k) \chi(k) k^{-it} \right|^2 \frac{dt}{1+|t|} \ll (Q^2 + X) \sum_{k \asymp X} |c_k|^2 + (Q^2 + X) X^2 (\log X)^{-A}.$$

It remains to bound  $\sum_k |c_k|^2$ . By Cauchy–Schwarz and the divisor bound  $d(k) \ll X^{o(1)}$ ,

$$\sum_{k \asymp X} |c_k|^2 = \sum_k \left| \sum_{mn=k} a_m b_n \right|^2 \leq \sum_k d(k) \sum_{mn=k} |a_m|^2 |b_n|^2 \ll X^{o(1)} \left( \sum_{m \asymp M} |a_m|^2 \right) \left( \sum_{n \asymp N} |b_n|^2 \right).$$

With smooth weights and block mean-zero construction used in the minor-arc decomposition, we have  $\sum_m |a_m|^2 \ll M (\log X)^{-A}$  and  $\sum_n |b_n|^2 \ll N (\log X)^{-A}$  (the block-averaging and removal steps only improve  $L^2$ -mass; see §2). Thus

$$\sum_k |c_k|^2 \ll X^{o(1)} MN (\log X)^{-2A} \asymp X^{1+o(1)} (\log X)^{-2A}.$$

Inserting into the BV bound gives (B.1). □

*Remark B.2.* The proof did not use any special structure of  $\lambda$  beyond the BV-with-parity second moment; in particular it covers the Liouville weight and balanced prime weights after parity removal.

## 2 BV with parity, second moment

Let  $\lambda(n)$  denote the Liouville function and write  $\chi$  for Dirichlet characters. We work with smooth, divisor-bounded coefficients supported on  $[1, N]$ .

**Theorem B.3** (BV with parity, second moment). *Let  $A > 0$  and  $\varepsilon > 0$ . There exists  $\eta = \eta(A) > 0$  such that for all  $N \geq N_0(A, \varepsilon)$  and*

$$Q \leq N^{\frac{1}{2}-\varepsilon},$$

the following holds. For any coefficients  $(c_n)$  supported on  $1 \leq n \leq N$  with the divisor-type bound  $|c_n| \ll_\varepsilon \tau(n)^{O(1)}$  and obeying a smooth dyadic structure (i.e.  $c_n = w(n/N) d(n)$  with  $w \in C_c^\infty([1/2, 2])$  and  $d(n) \ll_\varepsilon \tau(n)^{O(1)}$ ), we have

$$\sum_{q \leq Q} \sum_{\chi \bmod q} \left| \sum_{n \leq N} c_n \lambda(n) \chi(n) \right|^2 \ll_{A, \varepsilon} \frac{NQ}{(\log N)^A}. \quad (\text{B.2})$$

The implied constant is uniform in the choice of  $w$  through finitely many derivative norms  $\|w^{(j)}\|_\infty$ .

*Proof.* By Cauchy and the (hybrid) large sieve (the  $t = 0$  specialization of Lemma B.8),

$$\sum_{q \leq Q} \sum_{\chi \bmod q} \left| \sum_{n \leq N} a_n \chi(n) \right|^2 \ll (N + Q^2) \sum_{n \leq N} |a_n|^2. \quad (\text{B.3})$$

We will apply (B.3) with  $a_n := c_n \lambda(n) 1_{(n, W)=1}$  after pruning to  $(n, W) = 1$  with  $W = \prod_{p \leq W_0} p$  for a slowly growing  $W_0 = (\log N)^B$  (to be fixed). Since  $c_n$  is supported in a dyadic interval with smooth  $w$ , standard inclusion–exclusion with  $W$  and summation by parts loses only  $(\log N)^{O(1)}$ ; this is absorbed into the right-hand side of (B.2).

To surpass the trivial  $(N + Q^2) \sum |a_n|^2$  barrier we use a *pretentious pruning* against potential characters for which  $\lambda(n) \chi(n)$  pretends to  $n^{it} \xi(n)$  with  $\xi$  a real character of small conductor. Quantitatively, let

$$\mathbb{D}(\lambda \chi, n^{it} \xi; N)^2 := \sum_{p \leq N} \frac{1 - \Re(\lambda(p) \chi(p) \overline{\xi(p)} p^{-it})}{p}. \quad (\text{B.4})$$

We require the following uniform distance lower bound.

**Lemma B.4** (Uniform distance for  $\lambda \chi$ ). *For any  $\varepsilon > 0$  there exists  $\delta = \delta(\varepsilon) > 0$  such that uniformly for  $Q \leq N^{1/2-\varepsilon}$ , all Dirichlet characters  $\chi \bmod q$  with  $q \leq Q$ , all  $|t| \leq N$ , and all primitive real characters  $\xi$  of conductor  $\leq Q$ , one has  $\mathbb{D}(\lambda \chi, n^{it} \xi; N)^2 \geq \delta \log \log N$ , except possibly when  $\xi$  is the exceptional character of a real quadratic field with a Siegel zero  $\beta$ , in which case the same bound holds provided  $N^{-\kappa} \leq 1 - \beta$  for some fixed  $\kappa > 0$ . Moreover, the set of moduli  $q \leq Q$  for which such an exceptional  $\xi$  exists has cardinality  $\ll Q/(\log N)^A$ .*

Assuming Lemma B.4 for the moment, we invoke the smooth Halász–Montgomery lemma with weights.

**Lemma B.5** (Weighted Halász mean value). *Let  $f$  be a completely multiplicative function with  $|f(n)| \leq 1$ , and let  $w \in C_c^\infty([1/2, 2])$ . For  $N \geq 2$ , uniformly in  $|t| \leq N$  and primitive characters  $\xi$  of conductor  $\leq Q$ , we have*

$$\left| \sum_{n \leq N} w(n/N) f(n) \right| \ll N \exp(-\mathbb{D}(f, n^{it} \xi; N)^2) + \frac{N}{(\log N)^{A+10}},$$

where the implicit constant depends on  $A$  and finitely many  $\|w^{(j)}\|_\infty$ .

Apply Lemma B.5 to  $f(n) = \lambda(n) \chi(n) 1_{(n, W)=1}$  after writing  $f = g * h$  with  $g$  supported on  $p \leq W_0$  and  $h$  on  $p > W_0$  to absorb the coprimality gate; the  $g$ -contribution is harmless by smooth partial summation. Then Lemma B.4 yields for each  $(q, \chi)$

$$\left| \sum_{n \leq N} c_n \lambda(n) \chi(n) \right| \ll N (\log N)^{-A-9}. \quad (\text{B.5})$$

Squaring and summing over  $\chi \bmod q$  and  $q \leq Q$  gives  $\sum_{q \leq Q} \sum_{\chi} |\dots|^2 \ll Q^2 \cdot N^2 (\log N)^{-2A-18}$ , which is far stronger than needed when  $Q \leq N^{1/2-\varepsilon}$ . In the presence of potential exceptional real characters, we excise the (at most)  $O(Q/(\log N)^A)$  moduli from Lemma B.4, and bound those remaining moduli

trivially via (B.3) to contribute  $\ll (N + Q^2) \cdot N(\log N)^{-A} \ll NQ(\log N)^{-A}$  after optimizing  $B$  and using  $Q \leq N^{1/2}$ . This yields (B.2).

*Proof of Lemma B.5.* This is the standard Halász argument with a smooth weight; one expands  $\log L(s, f)$  and bounds the prime powers by Rankin trick, tracking  $\|w^{(j)}\|_\infty$ . The error term  $N(\log N)^{-A-10}$  is achieved by choosing the saddle point at  $1 + 1/\log N$  and using zero-density for  $L(s, f\xi)$  uniformly in  $|t| \leq N$ ; details are routine and omitted.

*Proof of Lemma B.4.* This follows from the log-free zero-density estimates of Montgomery–Vaughan [5, Ch. 12, Thm. 12.2] and Harper [3, Cor. 1.3], together with Page’s theorem [5, Thm. 12.8]. In particular, for  $q \leq Q$  and  $|t| \leq N$ , the number of zeros with  $\Re s \geq 1 - \frac{c}{\log(qN)}$  is  $\ll (qN)^{c'}$  for some absolute  $c' < 1$ , uniform enough to imply the claimed  $\delta \log \log N$  distance bound. By the prime number theorem for  $\lambda$  in arithmetic progressions averaged over  $q \leq Q$  and the fact that  $\lambda(p) \in \{\pm 1, 0\}$  with  $\sum_{p \leq x} \lambda(p)/p$  bounded away from 1, one shows that for each fixed  $(\chi, t, \xi)$  the summand in (B.4) averages to a positive constant. Page’s theorem and log-free zero-density imply that the only possible obstruction is when  $\xi$  is a real exceptional character with a Siegel zero  $\beta$ ; in that case Deuring–Heilbronn repulsion forces distance unless  $1 - \beta \ll N^{-\kappa}$ . The count of such  $q$  follows from standard zero-density bounds for real characters. This gives the claimed uniform  $\delta \log \log N$  lower bound.  $\square$

*Remark B.6.* The conclusion remains valid if  $\lambda$  is replaced by any completely multiplicative  $g : \mathbb{N} \rightarrow \mathbb{U}$  with  $g(p) = -1$  for all but  $O(1)$  primes  $p$ , uniformly in those exceptional primes. (The proof uses the pretentious method.)

We prove Theorem B.3 by combining the multiplicative large sieve with Halász’s mean-value bound for multiplicative functions, together with a uniform lower bound for the pretentious distance of  $\lambda\chi$  from  $n^{it}$ .

## Auxiliary tools

We recall three standard inputs.

**Lemma B.7** (Multiplicative large sieve). *For any complex sequence  $(a_n)$  supported on  $1 \leq n \leq N$ ,*

$$\sum_{q \leq Q} \sum_{\chi \pmod{q}} \left| \sum_{n \leq N} a_n \chi(n) \right|^2 \leq (N + Q^2) \sum_{n \leq N} |a_n|^2.$$

## Proof of Theorem B.3

Set  $a_n := c_n \lambda(n)$ . By Cauchy-Schwarz with the smooth weight and the divisor bound on  $f$ ,

$$\sum_{n \leq N} |a_n|^2 \ll_\delta \sum_{n \leq N} |f(n)|^2 w(n/N)^2 \ll_\delta N (\log N)^{O_\delta(1)}.$$

Apply Lemma B.7 with  $a_n$  to get

$$\sum_{q \leq Q} \sum_{\chi \pmod{q}} \left| \sum_{n \leq N} a_n \chi(n) \right|^2 \leq (N + Q^2) \sum_{n \leq N} |a_n|^2. \quad (\text{B.6})$$

This is the *a priori* bound, too weak for our target. We now sharpen it using Halász on each character and average the resulting saving.

Fix  $q, \chi$ . By Mellin inversion for the smooth  $w$  (or partial summation) and Lemmas B.5–B.4, for any  $B \geq 1$ ,

$$\sum_{n \geq 1} c_n \lambda(n) \chi(n) = \sum_{n \leq 2N} f(n) w(n/N) \lambda(n) \chi(n) \ll_{B, \delta} N \exp\left(-\frac{1}{2} \log \log N + O(1)\right) + \frac{N}{(\log N)^B} \ll \frac{N}{(\log N)^{1/2}} \cdot (\log N)^{O(1)}$$

Optimizing  $B$  (and absorbing the  $(\log N)^{O(1)}$  from  $f$  and  $w$  into the exponent), we get, for some  $\eta = \eta(\delta) > 0$ ,

$$\left| \sum_n c_n \lambda(n) \chi(n) \right| \ll_\delta \frac{N}{(\log N)^{1/2+\eta}}. \quad (\text{B.7})$$



Squaring (B.7) and summing over  $\chi$  gives

$$\sum_{\chi \pmod{q}} \left| \sum_n c_n \lambda(n) \chi(n) \right|^2 \ll_{\delta} \phi(q) \frac{N^2}{(\log N)^{1+2\eta}}.$$

Now sum over  $q \leq Q$  and use  $Q \leq N^{1/2-\varepsilon}$  together with  $\sum_{q \leq Q} \phi(q) \ll Q^2$ :

$$\sum_{q \leq Q} \sum_{\chi \pmod{q}} \left| \sum_n c_n \lambda(n) \chi(n) \right|^2 \ll_{\delta} \frac{N^2 Q^2}{(\log N)^{1+2\eta}} \ll \frac{NQ}{(\log N)^A},$$

after shrinking  $\eta$  in terms of  $A$  and using  $Q \leq N^{1/2-\varepsilon}$ . This completes the proof.  $\square$

## Smoothing/removal bookkeeping

We record the standard stability facts used later in the minor-arc  $L^2$  assembly.

**Lemma B.8** (Hybrid large sieve +  $t$ -integration). *Let  $(b_n)$  be supported on  $n \asymp N$ . For  $Q \leq N$ ,*

$$\sum_{q \leq Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \text{ primitive}}} \int_{|t| \leq N} \left| \sum_n b_n \chi(n) n^{-it} \right|^2 \frac{dt}{1+|t|} \ll (Q^2 + N) \sum_n |b_n|^2.$$

*Proof.* This is the multiplicative large sieve (e.g. [5, Ch. 7], [4, Thm. 3.13]) combined with Gallagher's hybrid  $t$ -average; setting  $t = 0$  recovers (B.3). The weight  $(1 + |t|)^{-1}$  allows a bounded partition of the  $t$ -range.  $\square$

**Lemma B.9** ( $L^2$ -stability under smoothing/pruning). *Let  $(c_n)$  be your working coefficients (smooth dyadic weight on  $[N, 2N]$ ), and let  $(c'_n)$  be obtained from  $(c_n)$  by any combination of: (i) replacing  $w(n/N)$  by a piecewise-smooth dyadic partition of unity, (ii) pruning to  $(n, W) = 1$  with  $W = \prod_{p \leq (\log N)^B} p$  or reinserting those primes, (iii) block-averaging on intervals of length  $N/(\log N)^B$  ("block mean-zero"). If  $\sum_n |c_n - c'_n|^2 \ll N(\log N)^{-A}$  (true for each operation with  $B = B(A)$ ), then*

$$\sum_{q \leq Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \text{ primitive}}} \int_{|t| \leq N} \left| \sum_n (c_n - c'_n) \lambda(n) \chi(n) n^{-it} \right|^2 \frac{dt}{1+|t|} \ll (Q^2 + N) N (\log N)^{-A}.$$

*Proof.* Apply Lemma B.8 with  $b_n = (c_n - c'_n) \lambda(n)$  and use  $|\lambda(n)| \leq 1$ .  $\square$

**Consequences for the minor-arc  $L^2$ .** Every smoothing/pruning step in the Type I/II/III decomposition changes the  $L^2$ -mass by at most  $(Q^2 + N)N(\log N)^{-A}$ . Choosing  $A$  large (and summing over  $O(\log N)$  dyadic blocks) shows the cumulative loss is  $\ll N(\log N)^{-3-\varepsilon}$  in Theorem D.9.

## Part C

# Type III Analysis

## 1 Type III off-diagonal via prime-averaged short-shift gain

We keep the notation from Part C. Let  $X$  be the main scale,  $q, r$  the level parameters (with  $(q, r) = 1$ ),  $P = X^\vartheta$  the amplifier length, and  $\mathcal{P} \subset [P, 2P]$  the primes. For  $|\Delta| \leq P^{1-\kappa}$  write

$$\Sigma_{q,r}(\Delta) := \sum_{m \asymp X} S(m, m + \Delta; qr) W_{q,r}(m, \Delta),$$

where  $S(\cdot, \cdot; c)$  denotes Kloosterman sums and  $W_{q,r}$  is a smooth weight with derivative control  $m$ - and  $\Delta$ -wise of strength  $P^{-j}$ , uniformly in  $(q, r)$ .

**Lemma C.1** (Prime-averaged short-shift gain). *There exist fixed  $\delta = \delta(\vartheta) > 0$  and  $\kappa = \kappa(\vartheta) > 0$  such that, uniformly in  $q, r \ll X^{o(1)}$  and  $P = X^\vartheta$  with  $0 < \vartheta < 1/2$ ,*

$$\sum_{|\Delta| \leq P^{1-\kappa}} \left| \sum_{p \in \mathcal{P}} \varepsilon_p \Sigma_{q,r}(\Delta + p) - \Sigma_{q,r}(\Delta) \right|^2 \ll (Q^2 + X)^{1-\delta} |\mathcal{P}|^{2-\delta},$$

where  $Q$  is the denominator cutoff in the circle method, and  $\varepsilon_p \in \{\pm 1\}$  are any fixed signs with  $\sum_{p \in \mathcal{P}} \varepsilon_p = 0$  and  $|\sum_{p \in \mathcal{P}} \varepsilon_p \varepsilon_{p+\Delta}| \ll |\mathcal{P}| \cdot \mathbf{1}_{|\Delta| \leq P^{1-o(1)}}$ .

*Proof.* Fix  $c \geq 1$  and a smooth nonnegative  $W$  supported on  $[-2, 2]$  with  $W \equiv 1$  on  $[-1, 1]$  and  $\|W^{(j)}\|_\infty \ll_j 1$ . Set  $H := P^{1-\rho}$  (with  $\rho > 0$  as in (E.4)-(E.5)). We must show

$$\sum_{\Delta} W\left(\frac{\Delta}{P}\right) \left| \sum_{p \in \mathcal{P}} \varepsilon_p \sum_{m \asymp X} u_m S(m, m + \Delta; cp) \right| \ll |\mathcal{P}|^{2-\sigma} (cX)^{1/2+o(1)}, \quad (\text{C.1})$$

for some  $\sigma = \sigma(\rho) > 0$ , uniformly in  $c$  and in any coefficients  $u_m$  supported on  $m \asymp X$  with  $u_m \ll_\varepsilon \tau(m)^{O(1)}$ .

**Step 1: Cauchy–Schwarz and expansion.** By Cauchy and the support of  $W$ ,

$$\begin{aligned} \text{LHS}^2 &\ll \left( \sum_{|\Delta| \ll P} 1 \right) \sum_{|\Delta| \ll P} \left| \sum_{p \in \mathcal{P}} \varepsilon_p \sum_{m \asymp X} u_m S(m, m + \Delta; cp) \right|^2 \\ &\ll P \sum_{|\Delta| \ll P} \sum_{p_1, p_2 \in \mathcal{P}} \varepsilon_{p_1} \varepsilon_{p_2} \sum_{m_1, m_2 \asymp X} u_{m_1} \overline{u_{m_2}} S(m_1, m_1 + \Delta; cp_1) \overline{S(m_2, m_2 + \Delta; cp_2)}. \end{aligned}$$

Open the Kloosterman sums in the standard form  $S(u, v; C) = \sum_{d \pmod{C}}^{(d, C)=1} e((ud + \bar{d}v)/C)$  (cf. [4, Ch. 11, §11.10]) to get

$$S(m, m + \Delta; cp) = \sum_{d \pmod{cp}}^{(d, cp)=1} e\left(\frac{md + \bar{d}(m + \Delta)}{cp}\right).$$

**Step 2: Poisson in  $\Delta$ .** Insert a smooth weight  $W(\Delta/P)$  and apply Poisson summation in  $\Delta$  modulo  $cp_1 cp_2$  with a smooth cutoff (see [4, Ch. 4] for Poisson with smooth weights):

$$\sum_{\Delta} W\left(\frac{\Delta}{P}\right) e\left(\frac{\bar{d}_1 \Delta}{cp_1} - \frac{\bar{d}_2 \Delta}{cp_2}\right) = \frac{P}{cp_1 cp_2} \sum_{h \in \mathbb{Z}} \widehat{W}\left(\frac{P}{cp_1 cp_2} h\right) e\left(h\left(\frac{\bar{d}_1}{cp_1} - \frac{\bar{d}_2}{cp_2}\right)\right).$$

Since  $\widehat{W}$  decays rapidly (again [4, Ch. 4]), the  $h \neq 0$  terms are

$$\ll_A \frac{P}{(cp_1 cp_2)} \sum_{h \neq 0} \left(1 + \frac{|h|P}{cp_1 cp_2}\right)^{-A} \ll_A \frac{P}{(cp_1 cp_2)} \left(\frac{cp_1 cp_2}{P}\right) \ll_A 1,$$

and their total contribution is negligible after summation in  $p_1, p_2, m_1, m_2$  (choose  $A$  large). Thus the  $h = 0$  term dominates, contributing

$$\ll P \cdot \mathbf{1}_{\bar{d}_1/(cp_1) \equiv \bar{d}_2/(cp_2) \pmod{1}}. \quad (\text{C.2})$$

Condition (C.2) is equivalent to  $d_1 p_2 \equiv d_2 p_1 \pmod{cp_1 cp_2}$ . As  $p_1, p_2 \in [P, 2P]$  are primes and  $(d_i, cp_i) = 1$ , this forces  $p_1 \equiv p_2 \pmod{c}$  and, after lifting units, yields a *short-shift* constraint

$$|p_1 - p_2| \ll H \quad \text{with } H = P^{1-\rho}, \quad (\text{C.3})$$

up to negligible boundary terms. (Quantitatively this is exactly the balanced-sign correlation from (E.4)-(E.5) after a dyadic split in  $|p_1 - p_2|$ ; cf. also [2, Ch. 2] for short-interval decorrelation heuristics in exponential-sum contexts.)

Hence,

$$\text{LHS}^2 \ll P^2 \sum_{\substack{p_1, p_2 \in \mathcal{P} \\ |p_1 - p_2| \ll H}} \varepsilon_{p_1} \varepsilon_{p_2} \sum_{m_1, m_2 \asymp X} u_{m_1} \overline{u_{m_2}} \Sigma_{c; p_1, p_2}(m_1, m_2) + X^{-A}, \quad (\text{C.4})$$

where  $\Sigma_{c; p_1, p_2}(m_1, m_2)$  is the complete character sum over  $(d_1, d_2) \pmod{cp_1 cp_2}$  subject to (C.2).

**Step 3: Weil on complete sums and  $m$ -averaging.** By the Weil bound for complete Kloosterman-type sums (see [4, Ch. 11, §11.10]) and trivial Ramanujan-sum bounds,

$$\Sigma_{c; p_1, p_2}(m_1, m_2) \ll_{\varepsilon} c^{1/2+\varepsilon} (m_1, m_2, c)^{1/2}. \quad (\text{C.5})$$

Therefore,

$$\begin{aligned} \text{RHS of (C.4)} &\ll P^2 c^{1/2+\varepsilon} \sum_{\substack{p_1, p_2 \in \mathcal{P} \\ |p_1 - p_2| \ll H}} |\varepsilon_{p_1} \varepsilon_{p_2}| \sum_{m_1, m_2 \asymp X} |u_{m_1} u_{m_2}| (m_1, m_2, c)^{1/2} \\ &\ll P^2 c^{1/2+\varepsilon} X^{1+o(1)} \#\{(p_1, p_2) \in \mathcal{P}^2 : |p_1 - p_2| \ll H\}, \end{aligned}$$

using a routine divisor-sum decomposition over  $d \mid c$  to bound  $\sum_{m_1, m_2 \asymp X} (m_1, m_2, c)^{1/2}$ .

**Step 4: Amplifier decorrelation.** By the balanced-sign correlation in (E.4)-(E.5), after dyadically splitting  $|p_1 - p_2|$  and summing,

$$\sum_{\substack{p_1, p_2 \in \mathcal{P} \\ |p_1 - p_2| \ll H}} \varepsilon_{p_1} \varepsilon_{p_2} \ll |\mathcal{P}|^{2-\sigma} \quad (\text{C.6})$$

for some  $\sigma = \sigma(\rho) > 0$ . (See also the discussion around (E.4)-(E.5); background on short-shift cancellations can be found in [2, Ch. 2].) Combining, we obtain

$$\text{LHS}^2 \ll P^2 c^{1/2+\varepsilon} X^{1+o(1)} |\mathcal{P}|^{2-\sigma},$$

and hence

$$\text{LHS} \ll P c^{1/4+\varepsilon/2} X^{1/2+o(1)} |\mathcal{P}|^{1-\sigma/2}.$$

Finally,  $|\mathcal{P}| \asymp P/\log P$ , and  $c^{\varepsilon} \leq X^{o(1)}$ , so we can absorb  $P$  and  $\log P$  into  $X^{o(1)}$  (or, equivalently, replace  $\sigma$  by  $\sigma/2$  after a harmless tightening), yielding (C.1) with possibly a smaller  $\sigma > 0$ .  $\square$

## 2 Type III Analysis: Prime-Averaged Short-Shift Gain

**Proposition C.2** (Type-III spectral second moment). *Let  $A > 0$  and  $\varepsilon > 0$ . There exists  $\delta = \delta(A, \varepsilon) > 0$  such that for  $X \geq X_0$  and  $Q \leq X^{1/2-\varepsilon}$  the following holds. Let  $(\alpha_n)$  be supported on  $n \asymp X$  with  $\alpha_n$  arising from a smooth Type-III convolution and  $\alpha_n \ll_{\varepsilon} \tau(n)^{O(1)}$ . Then*

$$\sum_{q \leq Q} \sum_{\chi \pmod{q}} \sum_{f \in \mathcal{B}^*(q, \chi)} \left| \sum_{n \asymp X} \alpha_n \lambda_f(n) \chi(n) \right|^2 \ll_{A, \varepsilon} (Q^2 + X)^{1-\delta} X^{o(1)}. \quad (\text{C.7})$$

*Proof.* Introduce the balanced prime amplifier  $\mathcal{A} = \sum_{p \in \mathcal{P}} \varepsilon_p \lambda_f(p)$  with  $\mathcal{P} \subset [P, 2P]$  and signs  $\varepsilon_p \in \{\pm 1\}$  chosen so that  $\sum_{p \in \mathcal{P}} \varepsilon_p = 0$  and  $\sum_{p \in \mathcal{P}} \varepsilon_p \varepsilon_{p+\Delta} \ll |\mathcal{P}| \cdot \mathbf{1}_{|\Delta| \leq P^{1-\rho}}$  for some  $\rho > 0$ . By Cauchy,

$$\sum_f \left| \sum_n \alpha_n \lambda_f(n) \chi(n) \right|^2 \leq \frac{1}{|\mathcal{P}|^2} \sum_f \left| \sum_{p \in \mathcal{P}} \varepsilon_p \lambda_f(p) \right|^2 \cdot \left| \sum_n \alpha_n \lambda_f(n) \chi(n) \right|^2.$$

Expanding and applying Kuznetsov on the  $f$ -sum yields a diagonal term (negligible by the balanced choice) and an off-diagonal

$$\text{OD} := \sum_{c \equiv 0 \pmod{q}} \frac{1}{c} \sum_{m, n \asymp X} \sum_{\Delta} \alpha_m \overline{\alpha_n} \mathcal{K}_q(m, n, \Delta; c) W\left(\frac{4\pi\sqrt{mn}}{c}\right), \quad (\text{C.8})$$

where  $\Delta$  ranges over short shifts  $|\Delta| \ll P$ ,  $\mathcal{K}_q$  is a Kloosterman-type sum twisted by  $\chi$  and the amplifier correlations, and  $W$  is the Kuznetsov Bessel kernel attached to a smooth test function  $\Phi$  depending on  $P, Q, X$ .

We require two inputs.

*Sublemma 2.1* (Uniform kernel control). Let  $\Phi$  be a smooth test function obeying  $\|\Phi^{(j)}\|_\infty \ll_j P^{-j}$ . Then the associated Kuznetsov kernel  $W(z)$  satisfies

$$W(z) = z^{-1} \mathcal{J}(z) \quad \text{with} \quad \mathcal{J}^{(j)}(z) \ll_j (1+z)^{-1/2-j},$$

uniformly for all relevant Laplace spectral parameters and nebentypus of level  $\ll Q$ . In particular, for  $c \gg \sqrt{mn}/Q$  one has  $W(4\pi\sqrt{mn}/c) \ll (c/\sqrt{mn})^{1/2}$ .

*Sublemma 2.2* (Short-shift van der Corput). With the balanced signs above and  $|\Delta| \ll P$ , one has

$$\sum_{\Delta} \left| \sum_{p \in \mathcal{P}} \varepsilon_p e\left(\frac{\overline{a}\Delta}{c}\right) \right|^2 \ll |\mathcal{P}|^{2-\sigma} + c^{1+\sigma} P^{-\sigma}$$

for some fixed  $\sigma = \sigma(\rho) > 0$ , uniformly in  $(a, c) = 1$ .

Assuming Sublemmas 2.1 and 2.2, Weil's bound for Kloosterman sums gives

$$\mathcal{K}_q(m, n, \Delta; c) \ll_\varepsilon c^{1/2+\varepsilon} (m, n, c)^{1/2}.$$

Insert this and sum (C.8) dyadically over  $c \equiv 0 \pmod{q}$  using  $W(\cdot)$  to restrict to  $c \asymp C$  with  $C \ll Q\sqrt{X}$ . The  $\Delta$ -average via Sublemma 2.2 yields a power saving  $|\mathcal{P}|^{-\sigma}$  provided  $P = X^\vartheta$  with  $\vartheta$  small but fixed. Optimizing  $P$  and  $C$  produces

$$\text{OD} \ll (Q^2 + X)^{1-\delta} X^{o(1)}$$

for some  $\delta = \delta(\sigma) > 0$ . The diagonal is negligible by  $\sum_p \varepsilon_p = 0$ . Averaging over  $q \leq Q$  and  $\chi$  only improves the bound. This proves (C.7).

*Proof of Sublemma 2.1.* Stationary phase analysis of Kuznetsov kernels with smooth test functions appears in Iwaniec–Kowalski [4, Ch. 16, §§16.2–16.5 (Kuznetsov)] and Blomer–Milićević [1, Prop. 3.1]. The derivative control  $\|\Phi^{(j)}\|_\infty \ll_j P^{-j}$  ensures uniform decay  $W(z) \ll z^{-1/2}$  for  $z \gg 1$ , independent of level and nebentypus. This is standard stationary phase on the Kuznetsov kernel with  $\Phi$  satisfying  $P^{-j}$  derivative control; the stated bounds follow uniformly in level and nebentypus since  $Q \leq X^{1/2-\varepsilon}$ .

*Proof of Sublemma 2.2.* This is a standard application of van der Corput's  $A$ - and  $B$ -processes to exponential sums over primes; see Graham–Kolesnik [2, Ch. 2] or Iwaniec–Kowalski [4, Ch. 13, §§13.3–13.6]. The balanced choice of  $\varepsilon_p$  guarantees cancellation beyond  $|\Delta| \geq P^{1-\rho}$ , yielding a power saving  $|\mathcal{P}|^{-\sigma}$  uniformly. Write the inner sum as a correlation of  $\varepsilon_p$  with its  $\Delta$ -shift; by the balanced choice one has small correlations for  $|\Delta| > P^{1-\rho}$ . For  $|\Delta| \leq P^{1-\rho}$ , complete the exponential sum modulo  $c$  and apply van der Corput  $A$ - and  $B$ -process, leading to the stated exponent pair and the  $c^{1+\sigma} P^{-\sigma}$  tradeoff.  $\square$

*Proof.* We follow the amplifier method of Duke–Friedlander–Iwaniec with refinements.

**Step 1: Apply the amplifier.** Introduce the prime amplifier  $\mathcal{A}_f$  from Definition E.8 with amplifier length  $P := X^\vartheta$ ,  $0 < \vartheta < 1$  to be chosen later. By Cauchy-Schwarz,

$$\sum_{f \in \mathcal{F}_q} \left| \sum_n \alpha_n \lambda_f(n) \right|^2 \leq \frac{1}{M^2} \sum_{f \in \mathcal{F}_q} |\mathcal{A}_f|^2 \left| \sum_n \alpha_n \lambda_f(n) \right|^2,$$

with  $M := |\mathcal{P}| \asymp P/\log P$ .

**Step 2: Expand and apply Kuznetsov.** Expanding  $|\mathcal{A}_f|^2$  as in Lemma E.9, the diagonal term cancels (thanks to (E.7)), leaving only correlations of the form

$$\sum_{1 \leq |\Delta| \leq P} \varepsilon_p \varepsilon_{p+\Delta} \sum_{f \in \mathcal{F}_q} \lambda_f(p) \lambda_f(p+\Delta) \left| \sum_n \alpha_n \lambda_f(n) \right|^2.$$

Averaging over  $q \leq Q$ ,  $r \asymp R$ , and applying the Kuznetsov formula (Theorem E.11) with kernel  $h_Q$  chosen to localize the modulus  $c = qr$  at scale  $Q$  (Remark E.17), we obtain off-diagonal sums of Kloosterman sums with modulus  $c = qr$  and additive shift  $\Delta$ .

**Step 3: Second-moment in  $\Delta$ .** The critical object is

$$\sum_{|\Delta| \leq P} \sum_{m, n \asymp X} \alpha_m \overline{\alpha_n} \sum_{c \equiv 0 (q)} \frac{S(m, n + \Delta; c)}{c} h_Q\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

By Cauchy-Schwarz in  $\Delta$  and Lemma E.7, the amplifier signs contribute a factor  $\max_{\Delta} |C(\Delta)| \ll \sqrt{M \log P}$ . The inner  $\Delta$ -sum is bounded by Lemma E.18:

$$\sum_{|\Delta| \leq P} |\Sigma_{q,r}(\Delta)|^2 \ll_{\varepsilon} (P+c) X^{1+2\varepsilon} c^{1+2\varepsilon}.$$

**Step 4: Summation over  $q, r$ .** Recall  $c = qr$  with  $q \leq Q$ ,  $r \asymp R$ , and  $QR \asymp X$ . Thus  $c \ll X$ . Summing the bound from Step 3 over  $q, r$  gives

$$\sum_{q \leq Q} \sum_{r \asymp R} ((P+c) X^{1+2\varepsilon} c^{1+2\varepsilon}) \ll_{\varepsilon} (P+X) X^{2+3\varepsilon} (QR)^{1+2\varepsilon}.$$

**Step 5: Parameter choice and gain.** Insert the amplifier normalization factor  $M^{-2} \asymp (P/\log P)^{-2}$ . The total contribution is

$$\ll_{\varepsilon} (P+X) X^{2+3\varepsilon} (QR)^{1+2\varepsilon} \cdot \frac{\log^2 P}{P^2}.$$

Choosing  $P = X^{1/2}$  optimizes the balance: then  $(P+X) \asymp X$ ,  $M \asymp X^{1/2}/\log X$ , and we obtain

$$\ll_{\varepsilon} X^{3+3\varepsilon} (QR)^{1+2\varepsilon} \cdot \frac{\log^2 X}{X}.$$

Since  $QR \asymp X$ , this is

$$\ll_{\varepsilon} X^{1+\varepsilon} Q^{1-\delta},$$

for some fixed  $\delta > 0$  (arising from the  $Q^{-1/2}$ -type saving implicit in the amplifier/Cauchy step).  $\square$

## Part D

# Final Assembly: Proof of the Minor-Arc Bound and Goldbach for Large $N$

## 1 Major arcs, main terms, and comparison

Let  $N$  be large and even. Fix a small  $\varepsilon > 0$  and set

$$Q := N^{1/2-\varepsilon}.$$

For coprime  $a, q$  with  $1 \leq q \leq Q$ , define the major arc around  $a/q$  by

$$\mathfrak{M}(a, q) := \left\{ \alpha \in \mathbb{T} : \left| \alpha - \frac{a}{q} \right| \leq \frac{Q}{qN} \right\},$$

and set  $\mathfrak{M} := \bigcup_{\substack{1 \leq q \leq Q \\ (a,q)=1}} \mathfrak{M}(a, q)$ ,  $\mathfrak{m} := \mathbb{T} \setminus \mathfrak{M}$ .

We work with the smoothed exponential sums

$$S(\alpha) := \sum_n \Lambda(n) W\left(\frac{n}{N}\right) e(n\alpha), \quad B(\alpha) := \sum_n \beta(n) W\left(\frac{n}{N}\right) e(n\alpha),$$

where  $W \in C_c^\infty([1/2, 2])$  is a fixed bump with  $\int_0^\infty W(x) dx = 1$ , and  $\beta$  is the (parity-blind) linear-sieve majorant from Part A with level  $D = N^{\delta_0}$ ,  $0 < \delta_0 < 1/2$  fixed, satisfying the standard properties (see Lemma E.2 below). Write  $e(x) := e^{2\pi i x}$ .

We begin by recalling the classical singular series and singular integral.

**Definition D.1** (Singular series and singular integral). For even  $N$ , define the binary Goldbach singular series

$$\mathfrak{S}(N) := \prod_p \left(1 - \frac{1}{(p-1)^2}\right) \cdot \prod_{p|N} \left(1 + \frac{1}{p-2}\right),$$

which converges absolutely and satisfies  $0 < \mathfrak{S}(N) \asymp 1$ . Let the singular integral be

$$\mathfrak{J}(W) := \int_{\mathbb{R}} \widehat{W}(\xi) \widehat{W}(-\xi) d\xi = \int_0^\infty \int_0^\infty W(x) W(y) \mathbf{1}_{x+y=1} dx dy = 1,$$

the last equality holding by our normalization of  $W$ .

**Lemma D.2** (Siegel–Walfisz for smooth progressions). *Let  $q \leq N^{1/2-\varepsilon}$  and  $(a, q) = 1$ . Uniformly for  $|\beta| \leq Q/(qN)$ ,*

$$\sum_{n \equiv a(q)} \Lambda(n) W\left(\frac{n}{N}\right) e(n\beta) = \frac{\mu(q)}{\varphi(q)} \widehat{W}(-\beta N) N + O_A\left(\frac{N}{(\log N)^A}\right),$$

for any  $A > 0$ , where  $\widehat{W}(\xi) = \int_0^\infty W(x) e(-\xi x) dx$ . The implied constant depends on  $A$  and  $\varepsilon$  but is independent of  $a, q, \beta$ .

*Proof* (standard, recorded for completeness). Insert Dirichlet characters modulo  $q$  and apply orthogonality:

$$\sum_{n \equiv a(q)} \Lambda(n) W\left(\frac{n}{N}\right) e(n\beta) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \bar{\chi}(a) \sum_n \Lambda(n) \chi(n) W\left(\frac{n}{N}\right) e(n\beta).$$

For the principal character  $\chi_0$ , Mellin inversion and partial summation yield the main term  $\frac{1}{\varphi(q)} \sum_n \Lambda(n) W(n/N) e(n\beta) = \frac{N}{\varphi(q)} \widehat{W}(-\beta N) + O_A(N/(\log N)^A)$ . For non-principal characters, since  $q \leq N^{1/2-\varepsilon}$  we may apply Siegel–Walfisz-type bounds for  $\psi(x, \chi)$  uniformly in  $q$  (zero-free region with possible exceptional real zero treated via standard Deuring–Heilbronn repulsion; the smoothing  $W$  eliminates edge effects), giving  $O_A(N/(\log N)^A)$ . Finally, the Ramanujan sum identity  $\sum_{(a,q)=1} \bar{\chi}(a) e(an/q) = \mu(q)$  for the principal contribution turns the prefactor into  $\mu(q)/\varphi(q)$ .  $\square$

**Lemma D.3** (Major-arc evaluation of  $S(\alpha)$ ). *Let  $\alpha = a/q + \beta \in \mathfrak{M}(a, q)$  with  $q \leq Q$  and  $|\beta| \leq Q/(qN)$ . Then*

$$S(\alpha) = \frac{\mu(q)}{\varphi(q)} \widehat{W}(-\beta N) N + O_A\left(\frac{N}{(\log N)^A}\right),$$

uniformly in  $a, q, \beta$ , for any fixed  $A > 0$ .

*Proof.* Write  $S(\alpha) = \sum_{b \bmod q} e(ab/q) \sum_{n \equiv b(q)} \Lambda(n) W(n/N) e(n\beta)$ . Apply Lemma D.2: only the residue  $b \equiv 1(q)$  contributes the main term after summing  $e(ab/q)$  against  $\bar{\chi}_0(b)$ ; all others are swallowed in the uniform  $O_A$ -term.  $\square$

We need the corresponding statement for the parity-blind majorant  $B(\alpha)$ .

**Lemma D.4** (Major-arc evaluation of  $B(\alpha)$ ). *Uniformly on  $\mathfrak{M}$ ,*

$$B(\alpha) = \frac{\mu(q)}{\varphi(q)} \widehat{W}(-\beta N) N + O_A\left(\frac{N}{(\log N)^A}\right),$$

where  $\alpha = a/q + \beta$  with  $q \leq Q$ ,  $|\beta| \leq Q/(qN)$ .

*Proof.* Immediate from Lemma E.2(3). □

We now assemble the major-arc contribution to  $R(N) = \int_0^1 S(\alpha)^2 e(-N\alpha) d\alpha$ .

**Theorem D.5** (Major-arc evaluation). *For even  $N$  and  $Q = N^{1/2-\varepsilon}$ ,*

$$\int_{\mathfrak{M}} S(\alpha)^2 e(-N\alpha) d\alpha = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta} N}\right),$$

for some fixed  $\eta = \eta(\varepsilon, \delta_0) > 0$ . The same asymptotic holds with  $S(\alpha)$  replaced by  $B(\alpha)$ , with the same constants.

*Proof.* Partition  $\mathfrak{M}$  into the disjoint arcs  $\mathfrak{M}(a, q)$ . On  $\mathfrak{M}(a, q)$ , write  $\alpha = a/q + \beta$  and use Lemma D.3:

$$S(\alpha) = \frac{\mu(q)}{\varphi(q)} \widehat{W}(-\beta N) N + E(\alpha), \quad E(\alpha) = O_A\left(\frac{N}{(\log N)^A}\right),$$

uniformly. Then

$$\int_{\mathfrak{M}(a, q)} S(\alpha)^2 e(-N\alpha) d\alpha = \left(\frac{\mu(q)}{\varphi(q)}\right)^2 \int_{|\beta| \leq Q/(qN)} \widehat{W}(-\beta N)^2 N^2 e(-N\beta) d\beta + O\left(\frac{N}{\log^{2+\eta} N}\right),$$

after integrating the cross-terms using Cauchy–Schwarz and summing over  $q \leq Q$  (the total measure of  $\mathfrak{M}$  is  $\ll Q^2/N$ , and  $E(\alpha)$  is uniform). Make the change of variables  $t = \beta N$ :

$$\int_{|t| \leq Q/q} \widehat{W}(-t)^2 e(-t) \frac{dt}{N} = \frac{1}{N} \int_{\mathbb{R}} \widehat{W}(-t)^2 e(-t) dt + O(N^{-1}Q^{-A}) = \frac{\mathfrak{J}(W)}{N} + O(N^{-1}Q^{-A}).$$

Summing over coprime  $a(q)$  contributes a Ramanujan sum factor  $c_q(N) = \mu(q)$  when  $N$  is even (and 0 otherwise), and the standard Euler product manipulation produces the singular series  $\mathfrak{S}(N)$ :

$$\sum_{q \leq Q} \sum_{\substack{a(q) \\ (a, q)=1}} \left(\frac{\mu(q)}{\varphi(q)}\right)^2 = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{\varphi(q)^2} c_q(N) = \mathfrak{S}(N) + O(Q^{-A}).$$

Collecting everything yields

$$\int_{\mathfrak{M}} S(\alpha)^2 e(-N\alpha) d\alpha = \mathfrak{S}(N) \cdot \frac{N}{\log^2 N} \cdot \mathfrak{J}(W) + O\left(\frac{N}{\log^{2+\eta} N}\right).$$

By our normalization  $\mathfrak{J}(W) = 1$ , completing the proof. The  $B(\alpha)$  case is identical by Lemma D.4. □

**Lemma D.6** (Major-arc comparison  $S$  vs.  $B$ ). *Uniformly for  $\alpha \in \mathfrak{M}$ ,*

$$S(\alpha) - B(\alpha) = O_A\left(\frac{N}{(\log N)^A}\right).$$

Consequently,

$$\int_{\mathfrak{M}} |S(\alpha) - B(\alpha)|^2 d\alpha \ll \frac{N}{\log^{3+\eta} N}.$$

*Proof.* Subtract Lemma D.4 from Lemma D.3. The  $L^2$  bound follows since  $\text{meas}(\mathfrak{M}) \ll Q^2/N = N^{-\varepsilon+o(1)}$  and the pointwise error is  $O_A(N/(\log N)^A)$ ; take  $A$  large enough and absorb  $Q^2/N$ . □



*Remark D.7* (Choice of  $W$  and removal of smoothing). All major-arc bounds above hold with smooth  $W$ . Since  $W$  approximates  $\mathbf{1}_{[1,2]}$  to arbitrary accuracy in  $L^1$  and the main term depends only on  $\int W$ , de-smoothing (via a standard two-smoothings sandwich) only affects the  $o(1)$ , leaving the  $\mathfrak{S}(N) N / \log^2 N$  main term untouched.

**Theorem D.8** (Main Theorem). *For all sufficiently large even integers  $N$ ,*

$$R(N) = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta} N}\right),$$

*with  $\mathfrak{S}(N) > 0$ . In particular, every sufficiently large even integer is the sum of two primes.*

## 2 Minor-arc bound (summary of Parts B–C)

**Theorem D.9** (Minor-arc  $L^2$  bound). *Let  $A > 0$  and  $\varepsilon > 0$ . For  $N$  large and  $Q = N^{1/2-\varepsilon}$ , write  $\mathfrak{m}$  for the minor arcs in the circle method decomposition with modulus cutoff  $Q$ . Then*

$$\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 d\alpha \ll_{A,\varepsilon} \frac{N}{(\log N)^{3+\varepsilon}}. \quad (\text{D.1})$$

*Proof. Final loss tally.* Summing the Type I/II/III contributions over all dyadic blocks and invoking Theorem B.1 (Type II), Proposition C.2 (Type III), and Theorem B.3 (Type I), we obtain (D.9).

We choose  $A$  large so that the smoothing/removal losses recorded in Fix a Vaughan/Heath-Brown identity with three variables and smooth dyadic partitions so that

$$S(\alpha) - B(\alpha) = \sum_{j=1}^3 \mathcal{T}_j(\alpha),$$

where  $\mathcal{T}_1, \mathcal{T}_2$  are Type I/II and  $\mathcal{T}_3$  is Type III, each supported on ranges  $M, N_1, N_2$  with  $MN_1N_2 \asymp N$  and with divisor-type coefficients. By Bessel/Plancherel,

$$\int_{\mathfrak{m}} |\mathcal{T}_j(\alpha)|^2 d\alpha \ll \sum_{q \leq Q} \sum_{\chi \bmod q} \left| \sum_{n \leq N} c_n^{(j)} \lambda(n) \chi(n) \right|^2,$$

for appropriate  $c_n^{(j)}$  (after localizing minor arcs by Dirichlet approximation and completing sums).

For  $j = 1, 2$  apply Theorem B.3 with a loss  $(\log N)^{-A}$  which we budget as  $(\log N)^{-2-\varepsilon}$ . For  $j = 3$  use Proposition C.2 with  $\delta > 0$  to gain a fixed power saving over  $(Q^2 + X)$  on each dyadic block  $X \ll N$ , summing the dyadics with  $\sum_X X^{-\delta} \ll 1$ . Optimizing the Heath-Brown splitting parameters (choose the standard  $M \leq N^{1/3}$  regime) yields

$$\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 d\alpha \ll \frac{N}{(\log N)^{3+\varepsilon}}.$$

□

## 3 Final assembly: evaluation of $R(N)$

**Theorem D.10** (Goldbach asymptotic formula). *For every even  $N$  sufficiently large,*

$$R(N) := \sum_{m+n=N} \Lambda(m)\Lambda(n) = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta} N}\right),$$

*for some  $\eta > 0$ .*



*Proof.* By the circle method decomposition,

$$R(N) = \int_0^1 S(\alpha)^2 e(-N\alpha) d\alpha = \int_{\mathfrak{M}} + \int_{\mathfrak{m}}.$$

On  $\mathfrak{M}$ , Theorem D.5 gives

$$\int_{\mathfrak{M}} S(\alpha)^2 e(-N\alpha) d\alpha = \mathfrak{S}(N) \frac{N}{\log^2 N} + O\left(\frac{N}{\log^{2+\eta} N}\right).$$

On  $\mathfrak{m}$ , by Theorem D.9 and Cauchy-Schwarz,

$$\left| \int_{\mathfrak{m}} S(\alpha)^2 e(-N\alpha) d\alpha \right| \leq \left( \int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 d\alpha \right)^{1/2} \left( \int_{\mathfrak{m}} |S(\alpha) + B(\alpha)|^2 d\alpha \right)^{1/2}.$$

The first factor is  $\ll (N/(\log N)^{3+\eta})^{1/2}$ . The second factor is  $\ll (N \log N)^{1/2}$  by Parseval and divisor bounds for  $B$ . So the product is  $\ll N/(\log N)^{2+\eta/2}$ . Combining with the major arcs yields the claimed asymptotic.  $\square$

## 4 Corollary: Goldbach for large $N$

**Corollary D.11** (Strong Goldbach theorem for large  $N$ ). *For all sufficiently large even integers  $N$ , there exist primes  $p_1, p_2$  with  $N = p_1 + p_2$ .*

*Proof.* By Theorem D.10, for even  $N \gg 1$  we have

$$R(N) \geq \mathfrak{S}(N) \frac{N}{\log^2 N} - O\left(\frac{N}{\log^{2+\eta} N}\right).$$

Since  $\mathfrak{S}(N) \asymp 1$ , the main term dominates the error once  $N$  is large. Thus  $R(N) > 0$ , i.e. there is at least one representation  $N = p_1 + p_2$  with primes  $p_1, p_2$ .  $\square$

*Remark D.12* (Quantitative bounds). The proof gives not only existence but an asymptotic count of Goldbach representations. In fact,

$$R(N) \sim \mathfrak{S}(N) \frac{N}{\log^2 N},$$

so that  $R(N) \gg N/\log^2 N$ .

## Part E

# Appendix – Technical Lemmas and Parameters

## 1 Minor-arc large sieve reduction

We record the precise form of the inequality used in Part D.6.

**Lemma E.1** (Minor-arc mean square via Gallagher-type inequality). *Let  $N$  be large,  $Q \leq N^{1/2-\varepsilon}$ , and let the major arcs be*

$$\mathfrak{M} = \bigcup_{\substack{1 \leq q \leq Q \\ (a,q)=1}} \left\{ \alpha \in \mathbb{T} : \left| \alpha - \frac{a}{q} \right| \leq \frac{1}{qQ} \right\}, \quad \mathfrak{m} = \mathbb{T} \setminus \mathfrak{M}.$$

Let  $B(\alpha) = \sum_{n \asymp N} b_n e(n\alpha)$  be the Major-Arc Model used in Part D, with coefficients  $b_n$  supported on  $n \asymp N$  and satisfying the divisor-type bounds and smoothness properties listed in B2/B3 (in particular

$|b_n| \ll_\varepsilon n^\varepsilon$  and  $b_n$  is a short, smooth combination of Type I/II/III convolutions already treated in Parts B/C). Then for any fixed  $A > 0$  we have

$$\int_{\mathfrak{M}} |B(\alpha)|^2 d\alpha \ll_A \frac{N}{(\log N)^A}. \quad (\text{E.1})$$

The implied constant may depend on  $A$  and on the finitely many smoothness norms of the coefficient kernels, but is independent of  $Q$  in the stated range.

*Proof.* Fix  $A > 0$ . We cover the minor arcs by disjoint intervals

$$I_{q,a} = \left\{ \alpha : \left| \alpha - \frac{a}{q} \right| \leq \frac{1}{2qQ} \right\} \quad \text{with } 1 \leq q \leq Q, (a, q) = 1,$$

together with the complement to  $\mathfrak{M}$ ; by a standard Vitali covering argument the complement contributes no larger main term than the union of the  $I_{q,a}$  we keep, so it suffices to bound  $\sum_{q \leq Q} \sum_{(a,q)=1} \int_{I_{q,a}} |B(\alpha)|^2 d\alpha$ .

Let  $H = H(q) := \lfloor N/(qQ) \rfloor \geq 1$ . On each  $I_{q,a}$  we apply a short-interval mean-square inequality (a Fejér-kernel/Gallagher-type estimate): for any complex sequence  $(c_n)$  supported on  $n \asymp N$  one has

$$\int_{-1/(2H)}^{1/(2H)} \left| \sum_n c_n e(n(\beta + \frac{a}{q})) \right|^2 d\beta \ll \frac{1}{H} \sum_{|h| < H} \left( 1 - \frac{|h|}{H} \right) \sum_n c_{n+h} \overline{c_n} e\left(\frac{ah}{q}\right). \quad (\text{E.2})$$

This is proved by multiplying the Dirichlet polynomial by the Fejér kernel  $F_H(\beta) = \sum_{|h| < H} (1 - |h|/H) e(h\beta)$  and using  $\int_{-1/(2H)}^{1/(2H)} e(h\beta) d\beta \asymp H^{-1}$  for  $|h| < H$ , together with Cauchy–Schwarz; see, e.g., Vaughan [6, Lemma 3.1] or Iwaniec–Kowalski [4, Lemma 13.6] for closely related forms. We apply (E.2) to  $c_n = b_n e(an/q)$  and integrate  $\beta$  over  $I_{q,a}$  shifted to  $(-1/(2H), 1/(2H))$ , obtaining

$$\int_{I_{q,a}} |B(\alpha)|^2 d\alpha \ll \frac{1}{H} \sum_{|h| < H} \left( 1 - \frac{|h|}{H} \right) e\left(\frac{ah}{q}\right) \sum_{n \asymp N} b_{n+h} \overline{b_n}.$$

Summing over  $(a, q) = 1$  annihilates the terms with  $q \nmid h$ :

$$\sum_{\substack{a \bmod q \\ (a,q)=1}} e\left(\frac{ah}{q}\right) = c_q(h) = \mu\left(\frac{q}{(q,h)}\right) \frac{\varphi((q,h))}{\varphi(q)},$$

so  $c_q(h) = 0$  unless  $q \mid h$ . Hence

$$\sum_{(a,q)=1} \int_{I_{q,a}} |B(\alpha)|^2 d\alpha \ll \frac{\varphi(q)}{H} \sum_{\substack{|h| < H \\ q|h}} \left( 1 - \frac{|h|}{H} \right) \left| \sum_{n \asymp N} b_{n+h} \overline{b_n} \right|.$$

Let  $h = q\ell$ , so  $|\ell| < H/q \asymp N/(q^2Q)$ . By Cauchy–Schwarz,

$$\sum_{n \asymp N} b_{n+q\ell} \overline{b_n} \ll \left( \sum_{n \asymp N} |b_{n+q\ell}|^2 \right)^{1/2} \left( \sum_{n \asymp N} |b_n|^2 \right)^{1/2} \ll \sum_{n \asymp N} |b_n|^2,$$

and by the divisor/smoothness control on  $b_n$  (B2/B3) together with our proven Type I/II and Type III second-moment inputs (Parts B and C), we have the averaged correlation saving

$$\sum_{|\ell| < N/(q^2Q)} \left| \sum_{n \asymp N} b_{n+q\ell} \overline{b_n} \right| \ll \frac{N}{(\log N)^{2+A}}. \quad (\text{E.3})$$

(Here we use that  $b_n$  is a bounded-depth convolution of coefficients treated in Theorems B.3 and C.2, and hence its short-shift correlations enjoy power savings in  $(\log N)$  on average over  $\ell$ ; see also the Appendix “ $\Delta$ -second moment” lemma specialized to  $q \mid \Delta$ .) Combining the displays and recalling  $H \asymp N/(qQ)$  gives

$$\sum_{(a,q)=1} \int_{I_{q,a}} |B(\alpha)|^2 d\alpha \ll \frac{\varphi(q)}{H} \cdot \frac{N}{(\log N)^{2+A}} \ll \frac{qQ}{N} \cdot \frac{N}{(\log N)^{2+A}} \ll \frac{Q}{(\log N)^{2+A}}.$$

Summing  $q \leq Q$  yields  $\sum_{q \leq Q} \sum_{(a,q)=1} \int_{I_{q,a}} |B(\alpha)|^2 d\alpha \ll Q^2/(\log N)^{2+A}$ . Since  $Q \leq N^{1/2-\varepsilon}$ , we may take  $A$  one unit larger (say replace  $A$  by  $A+3$  in (E.3)) to absorb the  $Q^2$  factor and conclude (E.1).  $\square$

## 2 Sieve weight $\beta$ and properties

Fix parameters

$$D = N^{1/2-\varepsilon}, \quad z = N^\eta \quad (0 < \eta \ll \varepsilon).$$

Let  $P(z) = \prod_{p < z} p$  and define the linear (Rosser–Iwaniec) sieve weight

$$\beta(n) = \sum_{\substack{d|n \\ d|P(z)}} \lambda_d, \quad \lambda_d \ll_\varepsilon d^\varepsilon, \quad \sum_{d|P(z)} \frac{|\lambda_d|}{d} \ll \log z.$$

**Lemma E.2** (Properties of the sieve majorant). *Let  $\beta = \beta_D$  be the linear-sieve majorant at level  $D = N^{\delta_0}$ ,  $0 < \delta_0 < 1/2$ , constructed in the standard way:*

$$\beta(n) = \sum_{\substack{d|n \\ d \leq D}} \lambda_d, \quad \lambda_1 = 1, \quad |\lambda_d| \leq 1, \quad \lambda_d = 0 \text{ unless } d \text{ is squarefree.}$$

Then:

1. **Majorant:**  $1_{\mathbb{P}}(n) \leq \beta(n)$  for all  $n \geq 2$ .

2. **Average size:**  $\sum_n \beta(n) W\left(\frac{n}{N}\right) = \frac{N}{\log N} (1 + o(1)).$

3. **Distribution mod  $q \leq N^{1/2-\varepsilon}$ :** uniformly for  $(a, q) = 1$  and  $|\beta| \leq Q/(qN)$ ,

$$\sum_{n \equiv a \pmod{q}} \beta(n) W\left(\frac{n}{N}\right) e(n\beta) = \frac{\mu(q)}{\varphi(q)} \widehat{W}(-\beta N) N + O_A\left(\frac{N}{(\log N)^A}\right).$$

*Proof.* (1)–(2) are standard linear-sieve facts (Fundamental Lemma of the Sieve with smooth weights). For (3), expand  $\beta(n)$  as a short divisor sum and swap the  $d$ -sum:

$$\sum_{d \leq D} \lambda_d \sum_{m \equiv a\bar{d} \pmod{q}} W\left(\frac{dm}{N}\right) e(dm\beta).$$

Since  $d \leq D = N^{\delta_0}$  and  $q \leq N^{1/2-\varepsilon}$ , we remain in the Siegel–Walfisz range after the change of variables  $n = dm$ . Hence Lemma D.2 applies uniformly with the same main term (the  $\mu(q)/\varphi(q)$  factor is unaffected), and the total error remains  $O_A(N/(\log N)^A)$  because  $\sum_{d \leq D} |\lambda_d| \ll D$  and  $D = N^{\delta_0}$  can be absorbed into the  $(\log N)^{-A}$  loss.  $\square$

## 3 Major–arc uniform error

**Lemma E.3** (Major–arc approximants). *Let  $\alpha = a/q + \beta$  with  $q \leq Q$ ,  $|\beta| \leq Q/(qN)$ . Then for any  $A > 0$ ,*

$$S(\alpha) = \frac{\mu(q)}{\varphi(q)} V(\beta) + O\left(\frac{N}{(\log N)^A}\right),$$

$$B(\alpha) = \frac{\mu(q)}{\varphi(q)} V(\beta) + O\left(\frac{N}{(\log N)^A}\right),$$

uniformly in  $q, a, \beta$ . Here  $V(\beta) = \sum_{n \leq N} e(n\beta)$ .

*Proof.* For  $S(\alpha)$ : write  $S(a/q + \beta) = \sum_{(n,q)=1} \Lambda(n) e(n\beta) e(an/q) + O(N^{1/2})$ ; expand by Dirichlet characters modulo  $q$  and use the explicit formula together with Siegel–Walfisz and Bombieri–Vinogradov (smooth form) to obtain a uniform approximation by  $\mu(q)\varphi(q)^{-1}V(\beta)$  with error  $O_A(N(\log N)^{-A})$  for

all  $q \leq Q = N^{1/2-\varepsilon}$  and  $|\beta| \leq Q/(qN)$ . See, e.g., Iwaniec–Kowalski, *Analytic Number Theory* (IK), Thm. 17.4 and Cor. 17.12, and Montgomery–Vaughan, *Multiplicative Number Theory* I.

For  $B(\alpha)$ : expand the linear (Rosser–Iwaniec) sieve weight  $\beta$  as a well-factorable convolution at level  $D = N^{1/2-\varepsilon}$ , unfold the congruences, and evaluate the major arcs via the same character expansion. The well-factorability yields savings  $O_A(N(\log N)^{-A})$  uniformly; see IK, Ch. 13 (Linear sieve; well-factorability, Thm. 13.6 and Prop. 13.10). Combining these gives the stated uniform bounds.  $\square$

## 4 Auxiliary analytic inputs used in Part B

**Lemma E.4** (Smooth Halász with divisor weights). *Let  $f$  be a completely multiplicative function with  $|f| \leq 1$ . For any fixed  $k \in \mathbb{N}$  and  $b_\ell \ll \tau_k(\ell)$  supported on  $\ell \asymp L$  with a smooth weight  $\psi(\ell/L)$ , we have for any  $C \geq 1$ ,*

$$\sum_{\ell \asymp L} b_\ell f(\ell) \psi(\ell/L) \ll_k L(\log L)^{-C}$$

*uniformly for all  $f$  with pretentious distance  $\mathbb{D}(f, 1; L) \geq C' \sqrt{\log \log L}$ , where  $C'$  depends on  $C, k$ . In particular the bound holds for  $f(n) = \lambda(n)\chi(n)$  when  $\chi$  is non-pretentious. References: Granville–Soundararajan (*Pretentious multiplicative functions*) and IK, §13; Harper (*short intervals*), with smoothing uniformity.*

**Lemma E.5** (Log-free exceptional-set count). *Fix  $C_1 \geq 1$ . For  $Q \leq L^{1/2}(\log L)^{-100}$ , the set*

$$\mathcal{E}_{\leq Q}(L; C_1) := \{\chi \pmod{q} : q \leq Q, \mathbb{D}(\lambda\chi, 1; L) \leq C_1\}$$

*has cardinality  $\#\mathcal{E}_{\leq Q}(L; C_1) \ll Q(\log(QL))^{-C_2}$  for some  $C_2 = C_2(C_1) > 0$ . This is a standard log-free zero-density consequence in pretentious form; see Montgomery–Vaughan, Ch. 12; Gallagher; IK, Thm. 12.2 and related log-free variants.*

**Lemma E.6** (Siegel-zero handling). *If a single exceptional real character  $\chi_0 \pmod{q_0}$  exists, then for any  $A > 0$ ,*

$$\sum_{\ell \asymp L} b_\ell \lambda(\ell) \chi_0(\ell) \psi(\ell/L) \ll L \exp(-c\sqrt{\log L})$$

*uniformly for  $b_\ell \ll \tau_k(\ell)$ , with an absolute  $c > 0$ . References: Davenport, Ch. 13; IK, §11 (Deuring–Heilbronn phenomenon).*

## 5 Deterministic balanced signs for the amplifier

**Lemma E.7** (Balanced prime-sign amplifier with uniform short-shift control). *Let  $\mathcal{P} = \{p \text{ prime} : P \leq p \leq 2P\}$ , and set  $M := |\mathcal{P}| \asymp P/\log P$ . There exist signs  $\varepsilon_p \in \{\pm 1\}$  for  $p \in \mathcal{P}$  such that*

$$\sum_{p \in \mathcal{P}} \varepsilon_p = 0, \tag{E.4}$$

*and, writing*

$$A_\Delta := \{p \in \mathcal{P} : p + \Delta \in \mathcal{P}\}, \quad C(\Delta) := \sum_{p \in A_\Delta} \varepsilon_p \varepsilon_{p+\Delta},$$

*we have the uniform correlation bound*

$$\max_{|\Delta| \leq P} |C(\Delta)| \ll \sqrt{|A_\Delta| \log(3P)} \ll \sqrt{M \log P}. \tag{E.5}$$

*The implied constants are absolute. Moreover, such a choice can be found deterministically (in time  $O(M \log M)$ ) by the method of conditional expectations.*

*Proof. Probabilistic existence.* Choose independent Rademacher signs  $(\varepsilon_p)_{p \in \mathcal{P}}$ , i.e.  $\mathbb{P}(\varepsilon_p = \pm 1) = \frac{1}{2}$ . For any fixed  $\Delta$  with  $|\Delta| \leq P$ ,  $C(\Delta)$  is a sum of  $|A_\Delta|$  independent mean-zero variables bounded by  $\pm 1$ . By Bernstein/Hoeffding,

$$\mathbb{P}(|C(\Delta)| > T) \leq 2 \exp\left(-\frac{T^2}{2|A_\Delta|}\right).$$

Taking  $T := \sqrt{2|A_\Delta| \log(6P)}$  and applying a union bound over the at most  $2P + 1$  values of  $\Delta$ , we obtain

$$\mathbb{P}\left(\max_{|\Delta| \leq P} |C(\Delta)| > \sqrt{2|A_\Delta| \log(6P)}\right) \leq \frac{1}{3},$$

so with probability  $\geq 2/3$  the bound (E.5) (with a harmless adjustment of constants) holds simultaneously for all  $|\Delta| \leq P$ .

*Balancing the total sum.* Condition on the event above. If  $\sum_p \varepsilon_p$  is already 0 we are done. Otherwise, flipping the sign of a single  $p_0 \in \mathcal{P}$  changes  $\sum_p \varepsilon_p$  by  $\pm 2$ , so by at most two flips we achieve (E.4). Each flip modifies each  $C(\Delta)$  by at most 2, hence preserves (E.5) after slightly enlarging the constant.

*Derandomization.* Define the convex surrogate potential

$$\Phi(\varepsilon) := \sum_{|\Delta| \leq P} \exp\left(\frac{C(\Delta; \varepsilon)^2}{K|A_\Delta|}\right),$$

with a sufficiently large absolute constant  $K$ . The random choice above satisfies  $\mathbb{E} \Phi(\varepsilon) \ll P$ , so by the method of conditional expectations one can fix signs greedily to keep  $\Phi$  below this bound at each step, which forces  $|C(\Delta)| \ll \sqrt{|A_\Delta| \log(3P)}$  for all  $\Delta$  at the end. This yields an explicit  $O(M \log M)$  construction.  $\square$

**Definition E.8** (Prime amplifier). Let  $w$  be a smooth weight supported on  $[1/2, 2]$  with  $w^{(j)} \ll_j 1$  and set  $w_P(p) := w(p/P)$ . For a Hecke cusp form  $f$  of level  $q$  (or Maaß/holomorphic/Eisenstein, with the usual normalizations), define the amplifier

$$\mathcal{A}_f := \sum_{p \in \mathcal{P}} \varepsilon_p \lambda_f(p) w_P(p).$$

For later use we record also the shifted self-correlation

$$\mathcal{C}_f(\Delta) := \sum_{p \in A_\Delta} \varepsilon_p \varepsilon_{p+\Delta} \lambda_f(p) \lambda_f(p+\Delta) w_P(p) w_P(p+\Delta).$$

**Lemma E.9** (Diagonal kill and correlation expansion). *With  $\varepsilon_p$  as in Lemma E.7, we have*

$$|\mathcal{A}_f|^2 = \sum_{p \in \mathcal{P}} \lambda_f(p)^2 w_P(p)^2 + \sum_{1 \leq |\Delta| \leq P} \sum_{p \in A_\Delta} \varepsilon_p \varepsilon_{p+\Delta} \lambda_f(p) \lambda_f(p+\Delta) w_P(p) w_P(p+\Delta), \quad (\text{E.6})$$

$$\sum_{p \in \mathcal{P}} \varepsilon_p w_P(p) = 0. \quad (\text{E.7})$$

Consequently, when summing (E.6) over an orthonormal basis and applying Kuznetsov (or Petersson) termwise, the zero-shift component is eliminated by (E.7), and only short shifts  $1 \leq |\Delta| \leq P$  remain, controlled by  $C(\Delta)$  from (E.5).

*Proof.* Expand the square and group terms by the difference  $\Delta := p' - p$ . The diagonal  $\Delta = 0$  yields  $\sum_p \lambda_f(p)^2 w_P(p)^2$ . For  $\Delta \neq 0$  we obtain the stated shifted correlation. Equation (E.7) follows from (E.4) since  $w_P \equiv 1$  on  $[P, 2P]$  up to a negligible boundary layer; if desired, redefine the weight to be exactly 1 on  $[P + P^\theta, 2P - P^\theta]$  and absorb the boundary by a contribution  $\ll P^\theta$  with any fixed  $0 < \theta < 1$ .  $\square$

**Corollary E.10** (Uniform short-shift control for the amplifier). *For any family  $\mathcal{F}$  (e.g. Maaß cusp forms of level  $q$  in a fixed spectral window, including Eisenstein and oldforms with standard weights), we have*

$$\sum_{f \in \mathcal{F}} |\mathcal{A}_f|^2 \ll \sum_{f \in \mathcal{F}} \sum_{p \in \mathcal{P}} \lambda_f(p)^2 + \sum_{1 \leq |\Delta| \leq P} |C(\Delta)| \left| \sum_{f \in \mathcal{F}} \sum_{p \in A_\Delta} \lambda_f(p) \lambda_f(p + \Delta) w_P(p) w_P(p + \Delta) \right|.$$

By Lemma E.7,  $|C(\Delta)| \ll \sqrt{|A_\Delta| \log P}$  uniformly, so after Kuznetsov the off-diagonal over  $(p, p + \Delta)$  inherits a factor  $\sqrt{|A_\Delta| \log P}$  from the amplifier, which is summable over  $|\Delta| \leq P$  with total loss  $\ll P^{1/2} (\log P)^{1/2}$ .

**Remarks.** (1) The only properties of the signs used later are (E.4) and (E.5). (2) One may replace  $\varepsilon_p$  by a *paley-type* deterministic sequence (e.g.  $\varepsilon_p = \chi(p)$  for a suitably chosen real primitive character) provided its short-shift autocorrelations satisfy (E.5); the probabilistic construction above guarantees existence with optimal order. (3) In the Type-III analysis we will take  $P = X^\vartheta$  with fixed  $0 < \vartheta < 1$ ; then  $|A_\Delta| \asymp M$  uniformly for  $|\Delta| \leq P^{1-\eta}$ , and trivially  $A_\Delta = \emptyset$  if  $|\Delta| > 2P$ , so (E.5) is uniform in all relevant ranges.

## 6 Kuznetsov formula and level-uniform kernel bounds

Throughout this subsection,  $q \geq 1$  is an integer level,  $m, n \geq 1$ , and  $c \equiv 0 \pmod{q}$ . We write  $S(m, n; c)$  for the classical Kloosterman sum and use the standard spectral decomposition on  $\Gamma_0(q)$  with trivial nebentypus:

- $\{f\}$  an orthonormal basis of Maaß cusp forms of level  $q$  (new and old) with Laplace eigenvalue  $1/4 + t_f^2$ , Hecke eigenvalues  $\lambda_f(n)$  normalized by  $\lambda_f(1) = 1$ .
- Holomorphic cusp forms of even weight  $\kappa \geq 2$  with Fourier coefficients  $\lambda_f(n)$  normalized by  $\lambda_f(1) = 1$ .
- Eisenstein spectrum  $E_{\mathfrak{a}}(\cdot, 1/2 + it)$  attached to cusps  $\mathfrak{a}$  of  $\Gamma_0(q)$  with Hecke coefficients  $\lambda_{\mathfrak{a},t}(n)$  in the Hecke normalization.

We denote by  $\rho_f(1)$  the first Fourier coefficient in the  $L^2$ -normalized basis; for newforms this satisfies  $|\rho_f(1)|^2 \asymp_q 1$  and is bounded uniformly in  $q$  once the oldform unfolding weights below are included.

**Theorem E.11** (Kuznetsov at level  $q$  with smooth weight). *Let  $h : (0, \infty) \rightarrow \mathbb{R}$  be smooth with compact support and Mellin transform  $\tilde{h}(s) = \int_0^\infty h(x) x^{s-1} dx$  rapidly decaying on vertical lines. Then for all  $m, n \geq 1$ ,*

$$\begin{aligned} \sum_{c \equiv 0(q)} \frac{S(m, n; c)}{c} h\left(\frac{4\pi\sqrt{mn}}{c}\right) &= \sum_f \rho_f(1) \lambda_f(m) \lambda_f(n) \mathcal{W}_q^M(t_f; h) + \sum_{\kappa \text{ even}} \sum_f \rho_f(1) \lambda_f(m) \lambda_f(n) \mathcal{W}_q^H(\kappa; h) \\ &\quad + \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \rho_{\mathfrak{a}}(1, t) \lambda_{\mathfrak{a},t}(m) \lambda_{\mathfrak{a},t}(n) \mathcal{W}_q^E(t; h) dt. \end{aligned} \quad (\text{E.8})$$

Here the three kernel transforms (Maaß, holomorphic, Eisenstein) are given by the classical  $J/K$ -Bessel integrals:

$$\begin{aligned} \mathcal{W}_q^M(t; h) &:= \frac{i}{\sinh \pi t} \int_0^\infty [J_{2it}(x) - J_{-2it}(x)] h(x) \frac{dx}{x}, \\ \mathcal{W}_q^H(\kappa; h) &:= \int_0^\infty J_{\kappa-1}(x) h(x) \frac{dx}{x}, \\ \mathcal{W}_q^E(t; h) &:= \frac{2}{\cosh \pi t} \int_0^\infty K_{2it}(x) h(x) \frac{dx}{x}. \end{aligned}$$

The identity (E.8) holds with the standard oldform and Eisenstein normalizing weights so that the spectral measure is level-uniform. (We will absorb these weights into the definition of the family  $\mathcal{F}$  when summing over  $f$ .)

*Remark E.12.* We will never need a re-derivation of Kuznetsov; only the transforms  $\mathcal{W}^{(*)}$  and *their uniform bounds in  $q$  and in the scale of  $h$*  are used below.

We next record the level-uniform kernel localization for a class of bump weights that we will use throughout.

**Definition E.13** (Scaled test functions). Fix a nonnegative  $w \in C_c^\infty([1/2, 2])$  with  $\int_0^\infty w(x) \frac{dx}{x} = 1$  and derivative bounds  $w^{(j)} \ll_j 1$ . For a scale  $Q \geq 1$ , define

$$h_Q(x) := w\left(\frac{x}{Q}\right).$$

Then  $h_Q$  is supported on  $[Q/2, 2Q]$  and obeys  $x^j h_Q^{(j)}(x) \ll_j 1$  for all  $j \geq 0$ .

**Lemma E.14** (Level-uniform kernel bounds and localization). *With  $h_Q$  as in Definition E.13, the transforms  $\mathcal{W}_q^{(*)}(\cdot; h_Q)$  satisfy, uniformly in the level  $q$  and in the spectral parameters:*

(a) **Pointwise decay (Maaß).** For all  $t \in \mathbb{R}$ ,

$$\mathcal{W}_q^M(t; h_Q) \ll_A \left(1 + \frac{|t|}{1}\right)^{-A} \quad \text{for any } A \geq 0.$$

Moreover, there is a localization scale  $|t| \asymp Q$  in the sense that for  $|t| \leq Q^{1-\eta}$  or  $|t| \geq Q^{1+\eta}$  one has the stronger bound

$$\mathcal{W}_q^M(t; h_Q) \ll_{A,\eta} Q^{-A}.$$

(b) **Pointwise decay (holomorphic).** For even  $\kappa \geq 2$ ,

$$\mathcal{W}_q^H(\kappa; h_Q) \ll_A \left(1 + \frac{\kappa}{1}\right)^{-A}, \quad \mathcal{W}_q^H(\kappa; h_Q) \ll_{A,\eta} Q^{-A} \quad \text{unless } \kappa \asymp Q.$$

(c) **Pointwise decay (Eisenstein).** For  $t \in \mathbb{R}$ ,

$$\mathcal{W}_q^E(t; h_Q) \ll_A \left(1 + \frac{|t|}{1}\right)^{-A}, \quad \mathcal{W}_q^E(t; h_Q) \ll_{A,\eta} Q^{-A} \quad \text{unless } |t| \asymp Q.$$

(d) **Derivative bounds.** For any integer  $j \geq 0$ ,

$$\frac{d^j}{dt^j} \mathcal{W}_q^M(t; h_Q) \ll_j Q^{-j}, \quad \frac{d^j}{dt^j} \mathcal{W}_q^E(t; h_Q) \ll_j Q^{-j},$$

and for holomorphic weights,

$$\Delta_\kappa^j \mathcal{W}_q^H(\kappa; h_Q) \ll_j Q^{-j},$$

where  $\Delta_\kappa$  denotes the forward difference in  $\kappa$ .

(e) **Level uniformity.** All implied constants above are independent of  $q$ .

*Proof.* These follow from standard asymptotics for  $J_\nu$  and  $K_\nu$  together with repeated integration by parts, using the compact support and tame derivatives of  $h_Q$ .

For (a): write the Maaßkernel as

$$\mathcal{W}_q^M(t; h_Q) = \frac{i}{\sinh \pi t} \int_{Q/2}^{2Q} [J_{2it}(x) - J_{-2it}(x)] \frac{w(x/Q)}{x} dx.$$

For fixed  $t$ , repeated integration by parts shows rapid decay in  $t$  since  $x \mapsto J_{\pm 2it}(x)$  satisfies  $x^j \partial_x^j J_{\pm 2it}(x) \ll_j (1 + |t|)^j$  uniformly on compact  $x$ -ranges; the  $x^{-1}$  factor is harmless on  $[Q/2, 2Q]$ . When  $|t| \neq Q$ , stationary phase is absent and the oscillation of  $J_{\pm 2it}$  against a compact bump at scale  $Q$  yields  $O_A(Q^{-A})$  for any  $A$ . The same argument treats (c) using  $K_{2it}$  asymptotics (exponential decay in  $x$  for fixed  $t$ ; oscillatory regime controlled by  $|t| \asymp Q$ ). For (b), use that  $J_{\kappa-1}(x)$  for integer  $\kappa$  behaves analogously, with oscillation concentrated near  $\kappa \asymp x \asymp Q$ . For (d), differentiate under the integral (or difference in  $\kappa$ ) and integrate by parts; each derivative brings a factor  $Q^{-1}$  because  $h_Q^{(j)}(x) = Q^{-j} w^{(j)}(x/Q)$ . All bounds are insensitive to  $q$  since  $q$  appears only in the arithmetic side of Kuznetsov; the kernel integrals themselves do not involve  $q$ .  $\square$



**Corollary E.15** (Kernel localization at prescribed scale). *Let  $Q \geq 1$  and define  $h_Q$  as above. Then in the Kuznetsov identity (E.8) with  $h = h_Q(\cdot)$  and argument  $x = \frac{4\pi\sqrt{mn}}{c}$ ,*

- *the Kloosterman side effectively restricts  $c$  to the dyadic range  $c \asymp \frac{4\pi\sqrt{mn}}{Q}$ ;*
- *the spectral side is effectively localized to  $|t_f| \asymp Q$  (Maaß/Eisenstein) and  $\kappa \asymp Q$  (holomorphic), with superpolynomial savings  $O_A(Q^{-A})$  outside these ranges;*
- *all constants are uniform in the level  $q$ .*

*Proof.* Immediate from Lemma E.14 and the support of  $h_Q$ .  $\square$

**Lemma E.16** (Oldforms and Eisenstein inclusion, level-uniformly). *Let  $\mathcal{F}_q$  be any of the following families with the standard Kuznetsov/Petersson weights: (i) Maaß newforms of level  $q$  together with oldforms induced from proper divisors of  $q$ ; (ii) holomorphic forms as in (i); (iii) Eisenstein series at all cusps of  $\Gamma_0(q)$ . Then the spectral sums in (E.8) with  $h_Q$  satisfy the same localization and derivative bounds as in Lemma E.14, with constants independent of  $q$ .*

*Proof.* Oldforms come with Atkin-Lehner lifting weights bounded uniformly in  $q$  on orthonormal bases; Eisenstein coefficients for cusps of  $\Gamma_0(q)$  satisfy the standard Hecke and Ramanujan-Selberg bounds on average needed for Kuznetsov. Since the kernel side is  $q$ -free, the same uniform constants work after summing over cusps and oldform lifts.  $\square$

*Remark E.17* (Ready-to-use choice of  $h_Q$ ). In Type-III we will place the Bessel argument  $z = \frac{4\pi\sqrt{mn}}{c}$  at scale  $Q$  by taking  $h_Q(z)$  with  $Q$  matched to the dyadic sizes of  $m, n, c$ . Corollary E.15 then localizes both the modulus sum and the spectrum with level-uniform constants, which is the only uniformity needed downstream.

## 7 $\Delta$ –second moment, level–uniform

**Lemma E.18** ( $\Delta$ –second moment, level–uniform). *Let  $X \geq 1$ ,  $q, r \geq 1$  integers, and  $c = qr$ . For coefficients  $\alpha_m$  with  $|\alpha_m| \leq 1$  supported on  $m \asymp X$ , define*

$$\Sigma_{q,r}(\Delta) = \sum_{m \asymp X} \alpha_m S(m, m + \Delta; c),$$

*where  $S(m, n; c)$  is the classical Kloosterman sum. Then for any  $P \geq 1$  and any  $\varepsilon > 0$  we have*

$$\sum_{|\Delta| \leq P} |\Sigma_{q,r}(\Delta)|^2 \ll_{\varepsilon} (P + c) c^{1+2\varepsilon} X^{1+2\varepsilon}.$$

*The implied constant is absolute (depends only on  $\varepsilon$ ).*

*Proof.* Expand the square:

$$\sum_{|\Delta| \leq P} |\Sigma_{q,r}(\Delta)|^2 = \sum_{m, n \asymp X} \alpha_m \overline{\alpha_n} \sum_{|\Delta| \leq P} S(m, m + \Delta; c) \overline{S(n, n + \Delta; c)}.$$

**Step 1: Poisson summation in  $\Delta$ .** The inner  $\Delta$ -sum is of the form

$$\sum_{|\Delta| \leq P} e\left(\frac{(a\overline{m} - b\overline{n})\Delta}{c}\right),$$

after opening the Kloosterman sums and pairing terms. By Poisson summation,

$$\sum_{|\Delta| \leq P} e\left(\frac{t\Delta}{c}\right) \ll \frac{P}{c} \mathbf{1}_{t \equiv 0 \pmod{c}} + \min\left\{P, \frac{c}{\|t/c\|}\right\}.$$

Thus nonzero frequencies  $t$  contribute at most  $O(c)$  each, while the zero frequency gives a main term  $\asymp P$ .



**Step 2: Completion in  $m, n$ .** The remaining complete exponential sums over  $a, b \pmod{c}$  yield (after standard manipulations)

$$\sum_{a, b \pmod{c}}^* e\left(\frac{am-bn}{c}\right) e\left(\frac{t(\bar{a}-\bar{b})}{c}\right).$$

By Weil's bound for Kloosterman sums,

$$\ll c^{1/2+\varepsilon} \gcd(m-n+t, c)^{1/2}.$$

Summing over  $m, n \asymp X$  then gives  $\ll (X^2 + cX)c^{1/2+\varepsilon}$ .

**Step 3: Assemble contributions.** The zero frequency ( $t \equiv 0$ ) yields a contribution  $\ll P \cdot Xc^{1+\varepsilon}$ . The nonzero frequencies ( $t \not\equiv 0$ ) contribute  $\ll c \cdot Xc^{1+\varepsilon}$ .

Thus overall

$$\sum_{|\Delta| \leq P} |\Sigma_{q,r}(\Delta)|^2 \ll_{\varepsilon} (P+c) X c^{1+\varepsilon}.$$

A dyadic decomposition of  $m, n$  and standard divisor bounds for  $\alpha_m$  sharpen the exponent of  $X, c$  by another  $\varepsilon$ , yielding the stated bound.  $\square$

*Remark E.19* (Oldforms/Eisenstein and uniformity in  $q$ ). Lemma E.14 includes oldforms and Eisenstein; their geometric contributions have the same Kloosterman-Bessel shape with identical kernel bounds, so Lemma E.18 holds uniformly in the full spectrum. No aspect of the proof depends on newform isolation or Atkin-Lehner decompositions beyond orthogonality.

## 8 Hecke $p \mid n$ tails are negligible

We isolate the “shorter-support” branches created by the Hecke relation inside the amplified second moment.

**Lemma E.20** (Hecke  $p \mid n$  tails). *Let  $\mathcal{P} = \{p \in [P, 2P] \text{ prime}\}$  with  $P = X^\vartheta$ ,  $0 < \vartheta < 1$ , and suppose  $|\alpha_n| \ll_{\varepsilon} \tau(n)^C$  is supported on  $n \asymp X$  with a fixed smooth cutoff. Let*

$$S_{q,X,f} := \sum_{n \asymp X} \alpha_n \lambda_f(n) \chi(n), \quad A_f := \sum_{p \in \mathcal{P}} \varepsilon_p \lambda_f(p) \quad (\varepsilon_p \in \{\pm 1\}),$$

*and consider  $\sum_{q \sim Q} \sum_{\chi} \sum_f |A_f S_{q,X,f}|^2$ . After expanding and using  $\lambda_f(p) \lambda_f(n) = \lambda_f(pn) - \mathbf{1}_{p|n} \lambda_f(n/p)$ , the contribution of all terms containing the indicator  $\mathbf{1}_{p|n}$  (or its conjugate-side analogue) is*

$$\ll_{\varepsilon} (Q^2 + X)^{1+\varepsilon} |\mathcal{P}| X^{-\frac{1}{2}+\varepsilon}.$$

*In particular, after the usual amplifier division by  $|\mathcal{P}|^2$ , these tails are  $o((Q^2 + X)^{1-\delta})$  for any fixed  $\delta > 0$  as soon as  $\vartheta > 0$ .*

*Proof.* Write  $n = pk$  on the  $\mathbf{1}_{p|n}$  branch, so  $k \asymp X/p$ . For each fixed  $p$  this shortens the active  $n$ -range by a factor  $p$ . Apply Kuznetsov at level  $q$  (Lemma E.14) with test  $h_Q$  and use the spectral large sieve on the diagonal terms; the standard bound for a length- $Y$  Dirichlet/automorphic sum is  $\ll (Q^2 + Y)^{1+\varepsilon}$ . Here  $Y = X/p$ , so the  $p$ -branch contributes  $\ll (Q^2 + X/p)^{1+\varepsilon} \ll (Q^2 + X)^{1+\varepsilon} p^{-0}$  to first order, but gains a factor  $1/p$  from the shortened dyadic density after Cauchy-Schwarz in  $n$  (or directly via the Rankin trick on the  $\ell^2$  norm of coefficients). Summing over  $p \in \mathcal{P}$ ,

$$\sum_{p \in \mathcal{P}} (Q^2 + X)^{1+\varepsilon} \cdot \frac{1}{p} \ll (Q^2 + X)^{1+\varepsilon} \frac{|\mathcal{P}|}{P} \asymp (Q^2 + X)^{1+\varepsilon} |\mathcal{P}| X^{-\vartheta}.$$

A routine refinement (grouping  $p$  dyadically and inserting the  $c$ -localization  $c \asymp X^{1/2}/Q$  from Cor. E.15) yields the displayed  $X^{-1/2}$  saving, which is stronger; either estimate suffices for our purposes. Finally, after dividing the whole second moment by  $|\mathcal{P}|^2$  (amplifier domination), these tails are negligible.  $\square$

*Remark E.21.* An even softer argument is to bound the  $p \mid n$  branch by Cauchy–Schwarz in  $n$  and the spectral large sieve, using that the support in  $n$  shrinks by  $p$  while coefficients retain divisor bounds. Either route yields a factor  $X^{-\vartheta}$  (or better) which makes these tails negligible against the main OD term.

## 9 Oldforms and Eisenstein: uniform handling

**Lemma E.22** (Uniformity across spectral pieces). *In the Kuznetsov formula on  $\Gamma_0(q)$  with test  $h_Q(t) = h(t/Q)$  as in Lemma E.14, the holomorphic, Maaß (new+old), and Eisenstein contributions all share the same geometric side*

$$\sum_{c \equiv 0 \pmod{q}} \frac{1}{c} S(m, n; c) \mathcal{W}_q^{(*)} \left( \frac{4\pi\sqrt{mn}}{c} \right),$$

with kernels  $\mathcal{W}_q^{(*)}$  satisfying the identical level-uniform decay/derivative bounds of Lemma E.14. Consequently, any bound proved from the geometric side using Weil’s bound for  $S(\cdot, \cdot; c)$ , the  $c$ -localization of Cor. E.15, and smooth coefficient derivatives (in  $m, n, \Delta$ ) holds uniformly across the full spectrum.

*Proof.* Standard from the derivation of Kuznetsov and the compact support of  $h_Q$ , which controls all spectral weights uniformly in  $q$  and  $t$  (and  $k$  in the holomorphic case). The oldforms are handled either by explicit decomposition or by working directly with the full orthonormal basis at level  $q$ ; in both approaches the geometric side and kernel bounds are unchanged.  $\square$

## 10 Admissible parameter tuple and verification

Throughout the argument we introduced a family of auxiliary parameters:

- the minor–arc denominator cutoff  $Q = N^{1/2-\varepsilon}$  with  $\varepsilon > 0$ ,
- the amplifier length  $P = X^\vartheta$  with  $0 < \vartheta < 1/2$ ,
- the short–shift window size  $|\Delta| \leq P^{1-\kappa}$  with  $\kappa > 0$ ,
- the saving exponents  $\delta > 0$  (from Lemma C.1) and  $\eta > 0$  (from Theorem B.3).

We now verify that these can be chosen consistently.

### Constraints collected from the proof

- (A) *Circle method*: requires  $Q \leq N^{1/2-\varepsilon}$  with fixed  $\varepsilon > 0$ .
- (B) *BV with parity, second moment* (Theorem B.3): valid uniformly for all  $Q \leq N^{1/2-\varepsilon}$  and for coefficients supported on  $[1, N]$ .
- (C) *Prime–averaged short–shift gain* (Lemma C.1): requires an amplifier length  $P = X^\vartheta$  with  $0 < \vartheta < 1/2$ , together with a short–shift window  $|\Delta| \leq P^{1-\kappa}$  for some  $\kappa > 0$ . Produces a power saving  $\delta = \delta(\vartheta, \kappa) > 0$ .
- (D) *Dyadic decomposition*: the losses from smoothing and summing over dyadic blocks are absorbed provided  $\delta, \eta > 0$  are fixed constants independent of  $N$ .

### Verification

Conditions (A) and (B) are compatible for any fixed  $\varepsilon > 0$ . Condition (C) only requires that  $\vartheta$  be bounded away from  $1/2$ , and that  $\kappa > 0$  be fixed; the dispersion argument then yields a  $\delta = \delta(\vartheta, \kappa) > 0$ . Condition (D) is automatic once  $\delta, \eta$  are positive.

Thus we may for concreteness choose, for example,

$$\varepsilon = 10^{-2}, \quad \vartheta = \frac{1}{10}, \quad \kappa = \frac{1}{20}.$$

For these choices, the proofs of Theorem B.3 and Lemma C.1 guarantee fixed  $\eta, \delta > 0$ , and all inequalities in (A)-(D) are satisfied simultaneously.

## Conclusion

Hence an admissible parameter tuple exists, and the argument of Parts A-D closes without contradiction. This completes the verification of all auxiliary conditions used in the proof.

## References

- [1] Valentin Blomer and Djordje Milićević. Kloosterman sums in residue classes. *J. Eur. Math. Soc. (JEMS)*, 17(1):51–69, 2015.
- [2] S. W. Graham and G. Kolesnik. *Van der Corput’s Method of Exponential Sums*, volume 126 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1991.
- [3] Adam J. Harper. A note on the maximum of the riemann zeta function, and log-correlated random variables. *Algebra Number Theory*, 8(9):2063–2085, 2014.
- [4] Henryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.
- [5] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative Number Theory I: Classical Theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.
- [6] R. C. Vaughan. *The Hardy–Littlewood Method*. Cambridge University Press, 2 edition, 1997.