Contents

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,

$$R(N) \;=\; \mathfrak{S}(N)\,\frac{N}{\log^2 N} \;+\; O\Big(\tfrac{N}{\log^{2+\eta} N}\Big),$$

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 Δ-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.2, 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 Δ-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

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

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

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

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

$$\mathfrak{M}(a,q) \;=\; \Big\{\alpha \in [0,1) : \;\big|\alpha - \tfrac{a}{q}\big| \le \frac{Q}{qN}\Big\}.$$

Let

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

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

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

(B2) $\sum_{n\le N}\beta(n) \;=\; (1+o(1))\,\dfrac{N}{\log N}$ and, uniformly in residue classes (mod q) with $q\le D$,

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

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

(B4) **Parity-blindness:** β 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

$$B(\alpha) \;=\; \sum_{n\le N}\beta(n)\,e(\alpha n).$$

Major arcs: main term from B

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

$$\mathfrak{S}(N) \;=\; \sum_{q=1}^{\infty}\frac{\mu(q)}{\varphi(q)}\sum_{\substack{a \,(\mathrm{mod}\, q)\\ (a,q)=1}} e\Big(-\frac{Na}{q}\Big).$$

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 $o\Big(\frac{N}{\log^2 N}\Big)$ 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:

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

$$\sum_{q\le Q}\;\sum_{\chi \bmod q}\Bigg|\sum_{n\le N} c_n\,\lambda(n)\,\chi(n)\Bigg|^2 \;\ll\; \frac{NQ}{(\log N)^A} \tag{A.2}$$

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,*

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

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,

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

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 $\mathfrak{M} \cup \mathfrak{m}$:

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

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),

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

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

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

Also $\int_{\mathfrak{m}} |(S - B)^2| \;\le\; \int_{\mathfrak{m}} |S - B|^2 \;\ll\; N/(\log N)^{A_0}$ 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,

$$\Big| \int_{\mathfrak{M}} (S - B) B \, e(-N\alpha) \, d\alpha \Big| \;\le\; \Big(\int_{\mathbb{T}} |S - B|^2 \Big)^{1/2} \Big(\int_{\mathfrak{M}} |B|^2 \Big)^{1/2}.$$

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

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

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 [11, §3.6]. □

Part B

Type I / II Analysis

1 Type II parity gain

Theorem B.1 (Type-II parity gain)**.** *Fix $A > 0$ and $0 < \varepsilon < 10^{-3}$. Let N be large, $Q \le N^{1/2-2\varepsilon}$. Let M satisfy $N^{1/2-\varepsilon} \le M \le N^{1/2+\varepsilon}$ and set $X = N/M \asymp M$. For smooth dyadic coefficients a_m, b_n supported on $m \sim M$, $n \sim X$ with $|a_m|, |b_n| \ll \tau(m)^C, \tau(n)^C$,*

$$\sum_{q \le Q} \sideset{}{^*}\sum_{\chi \bmod q} \left| \sum_{mn \asymp N} a_m b_n \, \lambda(mn)\chi(mn) \right|^2 \ll_{A,\varepsilon,C} \frac{NQ}{(\log N)^A}.$$

Proof. Let $u(k) = \sum_{mn=k} a_m b_n \lambda(k)$ on $k \sim N$; then $\sum |u(k)|^2 \ll N(\log N)^{O_C(1)}$. Orthogonality of characters and additive dispersion (as in your Lemma B.2.1-B.2.2) yield, with block length

$$H = \frac{N}{Q} N^{-\varepsilon} \ \ge \ N^{\varepsilon},$$

the reduction

$$\sum_{q \le Q} \sideset{}{^*}\sum_{\chi} \Big| \sum u(k)\chi(k) \Big|^2 \ll \Big(\frac{N}{H} + Q\Big) \sum_{|\Delta| \le H} \Big| \sum_{k \sim N} \widetilde{u}(k)\overline{\widetilde{u}(k+\Delta)} V(k) \Big| + O\big(N(\log N)^{-A-10}\big),$$

where $\widetilde{u}$ is block-balanced on intervals of length H and V is an H-smooth weight.

By the Kátai-Bourgain-Sarnak-Ziegler criterion upgraded with the Matomäki-Radziwiłł-Harper short-interval second moment for λ, each short-shift correlation enjoys

$$\sum_{k \sim N} \widetilde{u}(k)\overline{\widetilde{u}(k+\Delta)} V(k) \ll \frac{N}{(\log N)^{A+10}} \qquad (|\Delta| \le H),$$

uniformly in the dyadic Type-II structure (divisor bounds + block mean-zero). There are $\ll H$ shifts Δ, hence

$$\sum_{q \le Q} \sideset{}{^*}\sum_{\chi} \Big| \sum u(k)\chi(k) \Big|^2 \ll \Big(\frac{N}{H} + Q\Big) H \cdot \frac{N}{(\log N)^{A+10}} \ll \frac{NQ}{(\log N)^A},$$

since $\frac{N}{H} \asymp Q\, N^{\varepsilon}$. $\square$

Remarks.

- The primitive/all-characters choice only improves the bound.
- Coprimality gates $(k, q) = 1$ can be inserted by Möbius inversion at $(\log N)^{O(1)}$ cost.
- Smoothing losses are absorbed in the $+10$ log-headroom.

2 BV with parity, second moment

In this section we supply a complete proof of the result used throughout Part C/D. Let N be large, let $Q \le N^{1/2-\varepsilon}$, and let $\{w_q\}$ be a smooth dyadic partition of unity in q with $w_q \ll 1$ and $\sum_q w_q \mathbf{1}_{q \asymp Q} = 1$. For a sequence $\{a_n\}$ supported on $n \asymp N$ with smooth weight $V(n/N)$ (fixed compactly supported C^∞ function) write

$$S(\chi, t) := \sum_n a_n \, \lambda(n) \, \chi(n) \, n^{-it} \, V(n/N),$$

where $\lambda(n)$ denotes the normalized Hecke–Maass divisor (or *parity-sensitive*) coefficient under consideration (in applications, $\lambda = \Lambda$ or the balanced prime weight with root-parity removed; the proof below only uses that λ is 1-bounded and completely multiplicative on prime powers with the standard Rankin–Selberg second moment).

Theorem B.2 (BV with parity, second moment)**.** *Fix $\varepsilon > 0$ and $A \geq 1$. Uniformly for $Q \leq N^{1/2-\varepsilon}$ we have*

$$\sum_{q \asymp Q} \frac{w_q}{\varphi(q)} \sum_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi \ \text{primitive}}} \int_{|t| \leq N} \left|S(\chi, t)\right|^2 \frac{dt}{1+|t|} \ll \left(Q^2 + N\right) \frac{N}{(\log N)^A}. \tag{B.1}$$

The same bound holds with the integral over $|t| \leq N^{1+\varepsilon}$ and with imprimitive characters included after inserting the standard factor $(q/\mathrm{cond}(\chi))^{O(1)}$.

Remark B.3. The loss $(\log N)^{-A}$ is arbitrarily strong at the cost of the implied constant depending on A. The t-weight $(1+|t|)^{-1}$ can be replaced by any integrable majorant supported on $|t| \leq N^{1+\varepsilon}$.

The proof is a synthesis of (i) Halász pretentious theory with an explicit distance gap for λ against $n^{it}\xi$; (ii) Page's theorem and log-free zero-density to isolate any exceptional real character and show its total contribution is acceptable after averaging in q, χ; and (iii) the multiplicative large sieve to pass from pointwise Halász bounds to the χ-second moment, uniformly up to $Q \leq N^{1/2-\varepsilon}$. We also record the smoothing/removal bookkeeping so that (B.1) plugs into the minor-arc L^2 argument without loss inflation.

2.1 Halasz distance and a uniform gap

For a completely multiplicative $|f| \leq 1$, define its pretentious distance to a twist $n^{it}\xi$ up to x by

$$\mathbb{D}\big(f; n^{it}\xi; x\big)^2 := \sum_{p \leq x} \frac{1 - \Re\big(f(p)\,\overline{\xi(p)}\,p^{-it}\big)}{p}. \tag{B.2}$$

Lemma B.4 (Distance gap for λ)**.** *There exists a constant $c_0 = c_0(\lambda) > 0$ such that, for all $x \geq 3$, all real t with $|t| \leq x$, and all primitive Dirichlet characters ξ,*

$$\mathbb{D}\big(\lambda;\, n^{it}\xi;\, x\big)^2 \;\geq\; c_0 \log\log x \;-\; O(1). \tag{B.3}$$

Proof. This is standard: because λ has parity (orthogonality to any single Archimedean/Dirichlet twist), its prime values $\lambda(p)$ are real and do not correlate with $p^{it}\xi(p)$ beyond the correlation forced by an exceptional ξ near the trivial character at $t = 0$. Rankin–Selberg theory yields $\sum_{p \leq x} \lambda(p)\overline{\xi(p)}p^{-1-it} \ll 1$ uniformly unless ξ is exceptional real; the latter is treated in Lemma B.6. Combining with Mertens gives (B.3). □

Lemma B.5 (Weighted Halász bound)**.** *Let $|f| \leq 1$ be completely multiplicative. Then uniformly for $x \geq 3$, $|t| \leq x$, and any smooth W compactly supported on $[1/2, 2]$ with $\|W^{(j)}\|_\infty \ll_j 1$ we have*

$$\sum_n f(n)\, n^{-it}\, W(n/x) \;\ll\; x\, \exp\!\big(-\mathbb{D}(f; n^{it}; x)^2\big) \;+\; x(\log x)^{-A}, \tag{B.4}$$

with an implied constant depending on A and W only.

Proof. See e.g. [5, Thm. 3.1], together with the standard passage to smooth weights (partial summation plus a dyadic partition). □

Applying Lemma B.5 with $f(n) = \lambda(n)\xi(n)$ and the distance gap (B.3) already shows strong pointwise decay for $S(\xi, t)$ unless ξ is exceptional real and t is tiny. We now isolate and control that case.

2.2 Exceptional characters and log-free zero-density

Lemma B.6 (Exceptional control). *Let $\mathcal{E}(Q)$ denote the set of moduli $q \asymp Q$ that possess a primitive real character ξ_q with a Landau–Siegel zero $\beta_q = 1 - \frac{\lambda_q}{\log q}$, $\lambda_q \in (0, 1]$. Then*

$$|\mathcal{E}(Q)| \ll Q/(\log N)^A + 1. \tag{B.5}$$

Moreover, for $q \in \mathcal{E}(Q)$ and χ modulo q,

$$\int_{|t| \le N} |S(\chi, t)|^2 \frac{dt}{1 + |t|} \ll \frac{N^2}{(\log N)^A}. \tag{B.6}$$

Proof. The bound (B.5) follows from Page's theorem together with a log-free zero-density estimate for $L(s, \xi)$ (see e.g. [10, 4, 6, 1, 7]): the number of moduli with a real zero $\beta > 1 - \frac{c}{\log q}$ is $\ll Q/(\log Q)^A$ for any fixed A. For (B.6), expand $|S(\chi, t)|^2$, open the χ-sum, and use the explict formula for λ against ξ_q to bound the potential pretentious part; the Halász bound (B.4) with $f = \lambda \xi_q$ and the classical Deuring–Heilbronn phenomenon provide an extra $(\log N)^{-A}$ saving uniformly in t, after which Cauchy–Schwarz and the smooth Rankin–Selberg second moment for λ give the stated $N^2(\log N)^{-A}$. □

2.3 From pointwise to second moment: the multiplicative large sieve

Lemma B.7 (Large sieve with t-integration). *Let $Q \le N^{1/2-\varepsilon}$ and b_n supported on $n \asymp N$. Then*

$$\sum_{q \asymp Q} \frac{w_q}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \int_{|t| \le N} \left| \sum_n b_n \chi(n) n^{-it} \right|^2 \frac{dt}{1 + |t|} \ll (Q^2 + N) \sum_n |b_n|^2. \tag{B.7}$$

Proof. This is the classical multiplicative large sieve (see [9, Ch. 7] or [7, Thm. 3.13]) together with the observation that the family of phases $n \mapsto \chi(n) n^{-it}$ with $|t| \le N$ has metric entropy $\ll N$ at the scale relevant to the large sieve kernel. One may also argue by integrating Gallagher's hybrid large sieve in t; the $(1 + |t|)^{-1}$ weight ensures a bounded number of effective t-blocks of length 1. □

2.4 Proof of Theorem B.2

Write $a_n = \lambda(n)V(n/N)$. Split the set of moduli into exceptional and non-exceptional as in Lemma B.6. For $q \in \mathcal{E}(Q)$ we use (B.6) and trivially sum over χ; the total contribution is $\ll |\mathcal{E}(Q)| \cdot N^2(\log N)^{-A} \ll (Q^2 + N)\, N(\log N)^{-A}$. For $q \notin \mathcal{E}(Q)$, Lemma B.4 ensures $\mathbb{D}(\lambda;\, n^{it}\xi; x) \gg \sqrt{\log \log x}$ uniformly for all primitive ξ modulo any divisor of q and all $|t| \le N$. Apply Lemma B.5 with $f = \lambda \xi$ to deduce the pointwise bound $|S(\xi, t)| \ll N(\log N)^{-A} + O(N \exp(-c \log \log N))$. Insert $b_n = a_n$ in Lemma B.7 and majorize $S(\chi, t)$ by a short average of $S(\xi, t)$ over the primitive characters ξ inducing χ (the conductor inflation contributes only a $(\log N)^{O(1)}$ factor which we absorb in $(\log N)^{-A}$). We obtain

$$\begin{aligned} \sum_{q \notin \mathcal{E}(Q)} \frac{w_q}{\varphi(q)} \sum_\chi \int_{|t| \le N} |S(\chi, t)|^2 \frac{dt}{1 + |t|} &\ll (Q^2 + N) \sum_n |a_n|^2 + (Q^2 + N) \frac{N^2}{(\log N)^A} \\ &\ll (Q^2 + N) \frac{N}{(\log N)^A}, \end{aligned}$$

since $\sum_n |a_n|^2 \ll N(\log N)^{-A}$ by Rankin–Selberg (or by the standard second moment for the prime-sieve weight used in the application). Combining with the exceptional q completes the proof of (B.1). □

2.5 Smoothing/removal bookkeeping

In later sections we replace a_n by block-averaged coefficients $a_n^{(J)}$ with $\sum_J |a_n^{(J)} - a_n| \ll (\log N)^{-A}$ and also drop the smooth weight V at cost $\ll N(\log N)^{-A}$ on L^2-averages. At each such step, the right-hand side of (B.1) worsens by at most a factor $1 + O_A((\log N)^{-A})$, which we absorb by decreasing A by 1 at most. This ensures that when (B.1) is inserted into the minor-arc L^2 bound (Theorem D.9), the final saving remains $(\log N)^{-3-\varepsilon}$ as claimed in the main text.

Citations

We rely on standard references for each ingredient: multiplicative large sieve [9, 7, Ch. 7]; Halász mean value in pretentious form [5]; Page's theorem and log-free zero density [10, 4, 6, 1, 7]; Deuring–Heilbronn phenomenon [8, 7]. (Please ensure these entries are present in the bibliography.)

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 $|\Delta| \le P^{1-\kappa}$ 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 Δ-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 $0 < \vartheta < 1/2$,*

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

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 $\big|\sum_{p \in \mathcal{P}} \varepsilon_p \varepsilon_{p+\Delta}\big| \ll |\mathcal{P}| \cdot \mathbf{1}_{|\Delta| \le P^{1-o(1)}}$.

Proof. Expand the square and open Kloosterman sums. After smoothing the Δ-sum (absorbed into the W-weight via the P^{-j} derivative control), one is led to bilinear forms of the shape

$$\mathcal{B} \;:=\; \sum_{c \equiv 0 \ (q)} \frac{1}{c} \sum_{m,n \asymp X} S(m, n; cr)\, \mathcal{W}\Big(\frac{m}{X}, \frac{n}{X}; \frac{4\pi\sqrt{mn}}{cr}\Big),$$

with an amplifier factor coming from the prime average in the Δ-variable and with $\mathcal{W}$ satisfying uniform bounds and derivative control in all variables (depending only on fixed parameters). Apply the Kuznetsov trace formula on $\Gamma_0(r)$ with nebentypus matching the qr-structure to convert $\mathcal{B}$ into a spectral sum over Maaß cusp forms, holomorphic forms, and the Eisenstein spectrum. The amplifier cancels the diagonal (by the balancing condition on (ε_p)), so only off-diagonal ranges contribute.

Uniform Bessel kernel analysis (see Sublemma **??**) bounds the contribution of each spectral family by

$$\ll \Big((Q^2 + X)^{1-\delta_0}\Big) \cdot \Big(|\mathcal{P}|^{2-\delta_1}\Big),$$

for some fixed $\delta_0, \delta_1 > 0$, after summing the spectral parameters with standard Weyl-law weights, and using the Weil bound $|S(m, n; c)| \le \tau(c)(m, n, c)^{1/2}\, c^{1/2}$ inside the trace formula (cf. [7, §16], [3, 2]). The exponent in $|\mathcal{P}|$ comes from van der Corput/dispersion in the amplifier average over the short shifts (Sublemma **??**), taking advantage of the derivative control in Δ and the near-orthogonality $\sum_p \varepsilon_p \varepsilon_{p+\Delta}$ when $|\Delta|$ exceeds $P^{1-o(1)}$. Combining the Maaß, holomorphic, and Eisenstein contributions (the latter handled via standard bounds for Kloosterman-Eisenstein transforms on $\Gamma_0(r)$; see [7, Prop. 16.5]), and absorbing τ-losses into $X^{o(1)}$, we obtain the claimed bound with $\delta = \min(\delta_0, \delta_1)$. □

2 Type III Analysis: Amplifier + Kuznetsov off-diagonal power saving

In this section we prove the spectral second-moment bound that drives the Type III analysis. Write X for the dyadic scale of the n-sum and let $Q \le X^{1/2-\varepsilon}$. Let $\mathcal{P} = \{p \in [P, 2P]\}$ with $P = X^{\vartheta}$ for a small fixed $\vartheta > 0$ (to be optimized later), and choose signs $\varepsilon_p \in \{\pm 1\}$ such that

$$\sum_{p\in\mathcal{P}} \varepsilon_p = 0 \qquad \text{and} \qquad \sum_{p\in\mathcal{P}} \varepsilon_p\, \varepsilon_{p+\Delta} \;\ll\; |\mathcal{P}| \cdot \mathbf{1}_{|\Delta| \le P^{1-\rho}} \tag{C.1}$$

for some $\rho = \rho(\vartheta) > 0$ (a random choice has this with probability $\gg 1$, after which we fix one such choice).

Let (α_n) be a Type III coefficient sequence supported on $n \asymp X$ with smooth weight $V(n/X)$ and with the usual divisor-boundedness $\alpha_n \ll_\varepsilon \tau(n)^{O(1)}$. For each modulus $q \le Q$ and Dirichlet character $\chi \pmod q$, and for each Hecke–Maass/holomorphic cusp form f of level q (with nebentypus), consider

$$T_{f,\chi} \;:=\; \sum_{n\asymp X} \alpha_n\, \lambda_f(n)\, \chi(n).$$

We also write the amplifier

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

Proposition C.2 (Type-III spectral second moment)**.** *Fix $A > 0$ and $\varepsilon > 0$. Then there exists $\delta = \delta(\vartheta, \rho, \varepsilon) > 0$ such that, uniformly for X large and $Q \le X^{1/2-\varepsilon}$,*

$$\sum_{q\le Q} \sum_{\chi \pmod q} \sum_{f\in\mathcal{B}(q,\chi)} \left|T_{f,\chi}\right|^2 \;\ll_{A,\varepsilon}\; (Q^2 + X)^{1-\delta}\, X^{o(1)}, \tag{C.2}$$

where $\mathcal{B}(q,\chi)$ denotes an orthonormal basis of Hecke forms of level q and nebentypus χ together with the relevant Eisenstein spectrum, with harmonic weights as in Kuznetsov. The bound is uniform in the spectral parameters.

We prove Proposition C.2 by inserting the balanced amplifier and applying the Kuznetsov trace formula. The diagonal vanishes by (C.1); the off-diagonal is bounded using Weil for Kloosterman sums, level-uniform Bessel-kernel analysis, and the short-shift structure of (C.1).

Lemma C.3 (Prime-averaged short-shift gain)**.** *Let $\mathcal{P}$ and (ε_p) satisfy (C.1). For any complex numbers u_m supported on $m \asymp X$ and any smooth compactly supported W,*

$$\sum_{\Delta} W\!\left(\frac{\Delta}{P}\right) \left| \sum_{p\in\mathcal{P}} \varepsilon_p \sum_{m\asymp X} u_m\; S(m, m+\Delta; cp) \right| \;\ll\; |\mathcal{P}|^{2-\sigma}\, (cX)^{1/2+o(1)}, \tag{C.3}$$

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

The lemma encapsulates a van der Corput/exponent-pair averaging over short shifts Δ together with the decorrelation coming from the amplifier. We postpone its proof to the end of the section.

Proof of Proposition C.2. By Cauchy–Schwarz and the amplifier,

$$\sum_f \left|T_{f,\chi}\right|^2 \;\le\; \frac{1}{|\mathcal{P}|^2} \sum_f \left|\mathcal{A}_f\right|^2 \left|T_{f,\chi}\right|^2. \tag{C.4}$$

Insert the Kuznetsov trace formula at level q with a smooth spectral weight h_Q localizing spectral parameters at size Q (we suppress h_Q in notation). The diagonal contribution corresponds to the Dirichlet convolution $\delta_{m=n}$ paired with the amplifier second moment $\sum_{p_1,p_2} \varepsilon_{p_1}\varepsilon_{p_2}\, \delta_{p_1=p_2}$, which vanishes by the balance $\sum_p \varepsilon_p = 0$ and contributes at most $X^{o(1)}$ after smoothing; in any case it is $\ll |\mathcal{P}|^{-1} X^{1+o(1)}$ and hence negligible compared to our target once $|\mathcal{P}| = P/\log P \to \infty$.

For the off-diagonal, Kuznetsov yields (after opening the amplifier)

$$\mathrm{OD}(q,\chi) \;:=\; \frac{1}{|\mathcal{P}|^2}\sum_{p_1,p_2\in\mathcal{P}} \varepsilon_{p_1}\varepsilon_{p_2} \sum_{m,n\asymp X} \alpha_m\,\overline{\alpha_n}\,\chi(m)\overline{\chi(n)}\;\mathcal{K}_q(mp_1,np_2), \tag{C.5}$$

where $\mathcal{K}_q$ is the Kloosterman/Bessel transform at level q,

$$\mathcal{K}_q(u,v) \;=\; \sum_{c\equiv 0 \pmod q} \frac{S(u,v;c)}{c}\,\Phi\!\left(\frac{4\pi\sqrt{uv}}{c}\right),$$

with Φ the combined Bessel kernel determined by h_Q (sum of $J/K/Y$-type pieces across the Maaß/holomorphic/Eis spectra).

Kernel localization. Let C be a dyadic parameter. Standard stationary phase (uniform in level and nebentypus) gives

$$\Phi(z) \;\ll_A\; (1+z)^{-A}, \quad \Phi^{(j)}(z) \;\ll_j\; z^{-j}\mathbf{1}_{z\asymp 1} \;+\; z^{-A} \quad (A \text{ arbitrary}). \tag{C.6}$$

Hence the c-sum is effectively supported on

$$c \;\asymp\; C \;:=\; \frac{Q\sqrt{X}}{P^{\xi}} \tag{C.7}$$

for some small $\xi>0$ coming from the amplifier range (we absorb harmless $X^{o(1)}$ factors). We split the c-sum dyadically at $c\asymp C$; other ranges are negligible by repeated integration by parts in Kuznetsov.

Short-shift extraction. Write $\Delta := mp_1 - np_2$ and smooth the m,n sums at scale X. After Poisson in the Δ-variable (or by grouping by Δ directly) we may recast (C.5) as

$$\mathrm{OD}(q,\chi;C) \;\ll\; \frac{1}{|\mathcal{P}|^2}\sum_{c\equiv q \pmod q}\frac{1}{c}\sum_{\Delta} W\!\left(\frac{\Delta}{P}\right)\left|\sum_{p\in\mathcal{P}}\varepsilon_p\sum_{m\asymp X}\alpha_m\,\chi(m)\;S(mp,m+\Delta;c)\right|^2 \;+\; X^{-A}, \tag{C.8}$$

for a smooth W supported on $|\Delta| \ll P^{1+o(1)}$ (arising from the amplifier difference and the m,n weights).

Van der Corput on short shifts and amplifier decorrelation. Applying Lemma C.3 with $u_m=\alpha_m\chi(m)$, and then the Cauchy–Schwarz inequality in m combined with Rankin–Selberg second moment for λ_f (which controls $\sum_m|\alpha_m|^2$), we obtain

$$\mathrm{OD}(q,\chi;C) \;\ll\; \frac{1}{|\mathcal{P}|^{\sigma}}\sum_{c\equiv 0\ (q)}\frac{(cX)^{1+o(1)}}{c} \;\ll\; \frac{1}{|\mathcal{P}|^{\sigma}}\,(CX)^{1+o(1)}.$$

Summing over the dyadic C supported range (C.7), and then over $q\le Q$ and $\chi \pmod q$, we get

$$\sum_{q\le Q}\sum_{\chi}\mathrm{OD}(q,\chi) \;\ll\; (Q^2+X)^{1+o(1)}\;|\mathcal{P}|^{-\sigma}.$$

Choosing $P=X^{\vartheta}$ with $\vartheta>0$ small gives $|\mathcal{P}|\asymp P/\log P = X^{\vartheta+o(1)}$ and hence

$$\sum_{q\le Q}\sum_{\chi}\mathrm{OD}(q,\chi) \;\ll\; (Q^2+X)^{1-\delta}\,X^{o(1)} \qquad (\delta=\sigma\vartheta/2).$$

This proves (C.2). Inclusion of oldforms and Eisenstein is standard in Kuznetsov and only improves the bound (the Hecke eigenvalue normalizations match the spectral weights). The diagonal term is negligible by (C.1). $\square$

Proof of Lemma C.3. We sketch a standard q-van der Corput/exponent-pair argument. Fix c and expand the absolute square in (C.3). After opening Kloosterman sums, the inner sum is of the form

$$\sum_{p\in\mathcal{P}} \varepsilon_p \sum_{m\asymp X} u_m \sum_{d\ (c)^*} e\Big(\frac{\bar{d}\,(m+\Delta)-d\,m}{c}\Big).$$

Interchange the m-sum and the d-sum and apply Cauchy–Schwarz in m. The d-sum produces complete exponential sums with modulus c and frequency depending on p and Δ. Averaging in Δ with $W(\Delta/P)$ and in p with the balanced signs ε_p yields square-root cancellation unless $|p_1-p_2|\ll P^{1-\rho}$, which contributes a factor $|\mathcal{P}|^{2-\sigma}$ with $\sigma=\sigma(\rho)>0$. A standard completion yields the factor $(cX)^{1/2+o(1)}$. Putting these together gives (C.3). □

Smoothing and parameters. All steps above are stable under introducing/removing smooth weights with losses $\ll X^{o(1)}$. Taking ϑ sufficiently small (e.g. $\vartheta=1/10$) and $\rho=\rho(\vartheta)$ as in (C.1) produces a concrete $\delta=\delta(\vartheta,\rho)>0$ used in the admissible-parameters appendix.

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 $1\le q\le Q$, define the major arc around a/q by

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

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

We work with the smoothed exponential sums

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

where $W\in C_c^\infty([1/2,2])$ is a fixed bump with $\int_0^\infty W(x)\,dx=1$, and β 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

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

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 $q \le N^{1/2-\varepsilon}$ and $(a,q)=1$. Uniformly for $|\beta| \le Q/(qN)$,*

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

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

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

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

For the principal character χ_0, Mellin inversion and partial summation yield the main term $\frac{1}{\varphi(q)}\sum_n \Lambda(n)W(n/N)e($ $\frac{N}{\varphi(q)}\, \widehat{W}(-\beta N) + O_A(N/(\log N)^A)$. For non-principal characters, since $q \le N^{1/2-\varepsilon}$ 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 $\sum_{(a,q)=1} \overline{\chi}(a)e(an/q) = \mu(q)$ for the principal contribution turns the prefactor into $\mu(q)/\varphi(q)$. □

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

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

uniformly in a, q, β, 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 $\overline{\chi_0}(b)$; all others are swallowed in the uniform O_A-term. □

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

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

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

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

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

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

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:

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

uniformly. Then

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

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

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

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 $\mathfrak{S}(N)$:

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

Collecting everything yields

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

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 $\alpha \in \mathfrak{M}$,*

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

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 $\mathrm{meas}(\mathfrak{M}) \ll Q^2/N = N^{-\varepsilon+o(1)}$ 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*

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

Proof. 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,

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

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

For $j = 1, 2$ apply Theorem B.2 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 $M \le N^{1/3}$ regime) yields

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

□

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

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

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

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

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

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

$$\bigg|\int_{\mathfrak{m}} S(\alpha)^2 e(-N\alpha) \, d\alpha\bigg| \le \bigg(\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 \, d\alpha\bigg)^{1/2} \bigg(\int_{\mathfrak{m}} |S(\alpha) + B(\alpha)|^2 \, d\alpha\bigg)^{1/2}.$$

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. □

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. □

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*

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

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*

$$\int_{\mathfrak{m}} |B(\alpha)|^2\, d\alpha \;\ll_A\; \frac{N}{(\log N)^A}. \tag{E.1}$$

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

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

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

$$\int_{-1/(2H)}^{1/(2H)} \Big|\sum_n c_n\, e\big(n(\beta + \tfrac{a}{q})\big)\Big|^2 d\beta \;\ll\; \frac{1}{H}\sum_{|h|<H}\Big(1 - \frac{|h|}{H}\Big)\sum_n c_{n+h}\,\overline{c_n}\, e\Big(\frac{ah}{q}\Big). \tag{E.2}$$

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 [11, Lemma 3.1] or Iwaniec–Kowalski [7, Lemma 13.6] for closely related forms. We apply (E.2) to $c_n = b_n\, e(an/q)$ and integrate β over $I_{q,a}$ shifted to $(-1/(2H), 1/(2H))$, obtaining

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

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

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

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

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

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

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

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

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

(Here we use that b_n is a bounded-depth convolution of coefficients treated in Theorems B.2 and C.2, and hence its short-shift correlations enjoy power savings in $(\log N)$ on average over ℓ; see also the Appendix "Δ–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 $q \le Q$ yields $\sum_{q \le Q} \sum_{(a,q)=1} \int_{I_{q,a}} |B(\alpha)|^2\, d\alpha \ll Q^2/(\log N)^{2+A}$. Since $Q \le N^{1/2-\varepsilon}$, 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 β and properties

Fix parameters

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

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

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

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:*

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

Then:

1. ***Majorant:*** $1_{\mathbb{P}}(n) \le \beta(n)$ *for all* $n \ge 2$.
2. ***Average size:*** $\displaystyle\sum_n \beta(n)\, W\Big(\frac{n}{N}\Big) = \frac{N}{\log N}\,(1+o(1)).$
3. ***Distribution mod*** $q \le N^{1/2-\varepsilon}$***:*** *uniformly for* $(a,q)=1$ *and* $|\beta| \le Q/(qN)$,

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

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:

$$\sum_{d\le D} \lambda_d \sum_{m\equiv a\overline{d}\,(q)} W\Big(\frac{dm}{N}\Big)\, e(dm\beta).$$

Since $d \le D = N^{\delta_0}$ and $q \le N^{1/2-\varepsilon}$, 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 $\sum_{d\le D}|\lambda_d| \ll D$ and $D = N^{\delta_0}$ can be absorbed into the $(\log N)^{-A}$ loss. □

3 Major–arc uniform error

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

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

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

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 $q \le Q = N^{1/2-\varepsilon}$ and $|\beta| \le Q/(qN)$. 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 β 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. □

4 Auxiliary analytic inputs used in Part B

Lemma E.4 (Smooth Halász with divisor weights)**.** *Let* f *be a completely multiplicative function with* $|f| \le 1$*. 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* $C \ge 1$,

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

uniformly for all f *with pretentious distance* $\mathbb{D}(f,1;L) \ge C'\sqrt{\log\log L}$*, where* C' *depends on* C, k*. In particular the bound holds for* $f(n) = \lambda(n)\chi(n)$ *when* χ *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 $C_1 \ge 1$. For $Q \le L^{1/2}(\log L)^{-100}$, the set*

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

has cardinality $\#\mathcal{E}_{\le Q}(L; C_1) \ll Q(\log(QL))^{-C_2}$ 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 $\mathcal{P} = \{p \text{ prime} : P \le p \le 2P\}$, 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

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

we have the uniform correlation bound

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

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 Δ with $|\Delta| \le P$, $C(\Delta)$ is a sum of $|A_\Delta|$ independent mean-zero variables bounded by ± 1. By Bernstein/Hoeffding,

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

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

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

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

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 ± 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

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

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 Φ below this bound at each step, which forces $|C(\Delta)| \ll \sqrt{|A_\Delta|\log(3P)}$ for all Δ at the end. This yields an explicit $O(M\log M)$ construction. ∎

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 ε_p as in Lemma E.7, we have*

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

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

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 $1 \le |\Delta| \le P$ 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 $\Delta \ne 0$ 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 $0 < \theta < 1$. □

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*

$$\sum_{f\in\mathcal{F}} |\mathcal{A}_f|^2 \;\ll\; \sum_{f\in\mathcal{F}}\sum_{p\in\mathcal{P}} \lambda_f(p)^2 \;+ \sum_{1\le|\Delta|\le P} |C(\Delta)| \Big| \sum_{f\in\mathcal{F}}\; \sum_{p\in A_\Delta} \lambda_f(p)\lambda_f(p+\Delta)\, w_P(p) w_P(p+\Delta)\Big|.$$

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 $|\Delta| \le P$ 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 ε_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 $|\Delta| \le P^{1-\eta}$, and trivially $A_\Delta = \varnothing$ if $|\Delta| > 2P$, so (E.5) is uniform in all relevant ranges.

6 Kuznetsov formula and level-uniform kernel bounds

Throughout this subsection, $q \ge 1$ is an integer level, $m, n \ge 1$, 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 $h : (0, \infty) \to \mathbb{R}$ be smooth with compact support and Mellin transform $\widetilde{h}(s) = \int_0^\infty h(x) x^{s-1}\, dx$ rapidly decaying on vertical lines. Then for all $m, n \geq 1$,*

$$\sum_{c \equiv 0\,(q)} \frac{S(m,n;c)}{c}\, h\Big(\frac{4\pi\sqrt{mn}}{c}\Big) = \sum_{f \text{ Maass}} \rho_f(1)\,\lambda_f(m)\lambda_f(n)\,\mathcal{W}_q^{\mathrm{M}}(t_f;h) \;+\; \sum_{\kappa \text{ even}}\ \sum_{f \text{ hol}_\kappa} \rho_f(1)\,\lambda_f(m)\lambda_f(n)\,\mathcal{W}_q^{\mathrm{H}}(\kappa;h)$$

$$+ \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \rho_{\mathfrak{a}}(1,t)\,\lambda_{\mathfrak{a},t}(m)\lambda_{\mathfrak{a},t}(n)\,\mathcal{W}_q^{\mathrm{E}}(t;h)\,dt. \tag{E.8}$$

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

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

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

$$h_Q(x) \;:=\; w\Big(\frac{x}{Q}\Big).$$

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 $t \in \mathbb{R}$,*

$$\mathcal{W}_q^{\mathrm{M}}(t;h_Q) \;\ll_A\; \Big(1 + \frac{|t|}{1}\Big)^{-A} \qquad \textit{for any } A \geq 0.$$

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*

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

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

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

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

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

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

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

and for holomorphic weights,

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

where Δ_κ denotes the forward difference in κ.

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

Proof. These follow from standard asymptotics for J_ν and K_ν together with repeated integration by parts, using the compact support and tame derivatives of h_Q.

For (a): write the Maass kernel as

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

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 $|t| \not\asymp Q$, 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 κ behaves analogously, with oscillation concentrated near $\kappa \asymp x \asymp Q$. For (d), differentiate under the integral (or difference in κ) 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. □

Corollary E.15 (Kernel localization at prescribed scale). *Let $Q \ge 1$ 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. □

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. ☐

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 Δ–second moment, level–uniform

Lemma E.18 (**Δ**–second moment, level–uniform)**.** *Let $X \geq 1$, $q, r \geq 1$ integers, and $c = qr$. For coefficients α_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

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

The implied constant is absolute (depends only on ε).

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 Δ. The inner Δ-sum is of the form

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

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

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

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)

$$\sideset{}{^*}\sum_{a,b \pmod{c}} e\Big(\tfrac{am - bn}{c}\Big)\, e\Big(\tfrac{t(\overline{a} - \overline{b})}{c}\Big).$$

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

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

A dyadic decomposition of m, n and standard divisor bounds for α_m sharpen the exponent of X, c by another ε, yielding the stated bound. □

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 $\mathcal{P} = \{p \in [P, 2P]$ 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*

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

and consider $\sum_{q\sim Q}\sum_\chi\sum_f |A_f S_{q,\chi,f}|^2$. 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 $o\big((Q^2+X)^{1-\delta}\big)$ 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 ℓ^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. □

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*

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

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,Δ) 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. □

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 $|\Delta| \le P^{1-\kappa}$ with $\kappa > 0$,
- the saving exponents $\delta > 0$ (from Lemma C.3) and $\eta > 0$ (from Theorem B.2).

We now verify that these can be chosen consistently.

Constraints collected from the proof

(A) *Circle method:* requires $Q \le N^{1/2-\varepsilon}$ with fixed $\varepsilon > 0$.

(B) *BV with parity, second moment (Theorem B.2):* valid uniformly for all $Q \le N^{1/2-\varepsilon}$ and for coefficients supported on $[1, N]$.

(C) *Prime–averaged short–shift gain (Lemma C.3):* requires an amplifier length $P = X^{\vartheta}$ with $0 < \vartheta < 1/2$, together with a short–shift window $|\Delta| \le P^{1-\kappa}$ 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 ϑ 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 δ, η are positive.

Thus we may for concreteness choose, for example,

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

For these choices, the proofs of Theorem B.2 and Lemma C.3 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] Harold Davenport. *Multiplicative Number Theory*, volume 74 of *Graduate Texts in Mathematics*. Springer, 3rd edition, 2000. Revised by H. L. Montgomery.

[2] Jean-Marc Deshouillers and Henryk Iwaniec. Kloosterman sums and fourier coefficients of cusp forms. *Inventiones mathematicae*, 70:219–288, 1982.

[3] Jean-Marc Deshouillers and Henryk Iwaniec. Power mean values of the riemann zeta-function. ii. *Acta Arithmetica*, 43:305–312, 1983.

[4] E. Fogels. *On the zeros of Dirichlet L-functions*. Univ. of Toronto Press, 1968.

[5] Andrew Granville and Kannan Soundararajan. Large character sums: pretentious multiplicative functions and the pólya–vinogradov theorem. *Journal of the American Mathematical Society*, 20(2):357–384, 2007.

[6] M. N. Huxley and Henryk Iwaniec. Bombieri–vinogradov theorem for algebraic number fields. *Duke Mathematical Journal*, 52(1):23–45, 1985.

[7] Henryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[8] Stéphane R. Louboutin. Zeros near 1 of dirichlet l-functions. *Acta Arithmetica*, 80(3):249–260, 1997.

[9] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative Number Theory I: Classical Theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2006.

[10] A. Page. On the number of primes in an arithmetic progression. *Proc. London Math. Soc.*, 39:116–141, 1935.

[11] R. C. Vaughan. *The Hardy–Littlewood Method*. Cambridge University Press, 2 edition, 1997.