Proof of the Goldbach Conjecture

Student Vinzenz Stampf

Part A. Framework

Assumptions & conditional result (at a glance)

This manuscript lays out a circle-method framework aimed at binary Goldbach. The final asymptotic is derived *conditional* on the minor-arc L^2 estimate (A.1) and the analytic inputs explicitly stated in Parts B-D. In particular:

- Establishing (A.1) is the central new task; Parts B-D provide a proposed route via Type I/II/III analyses.
- Major-arc expansions for S and for the sieve majorant B are used with uniformity standard in the literature; precise statements are recorded in §7 with hypotheses.
- The final positivity conclusion for $R(N)$ is conditional on (A.1) and the stated major-arc bounds; no claim is made here that the new inputs are fully proved.

A succinct punch-list of outstanding items appears in Appendix B.

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 $\pmod q$ 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:

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

Proposition 0.1 (Reduction)**.** *Assume (A.1). Then*

$$R(N) \; = \; \int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha)\, d\alpha \; + \; O\Big(\frac{N}{(\log N)^{3+\varepsilon/2}}\Big),$$

and hence

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

for some $\delta > 0$.

Sketch. Split on $\mathfrak{M} \cup \mathfrak{m}$ and insert $S = B + (S - B)$:

$$S^2 = B^2 + 2B(S - B) + (S - B)^2.$$

Integrating over $\mathfrak{m}$ and using Cauchy-Schwarz,

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

By Parseval and (B2)-(B3),

$$\int_0^1 |B(\alpha)|^2 \, d\alpha \;=\; \sum_{n \le N} \beta(n)^2 \;\ll\; \frac{N}{\log N},$$

so $\int_{\mathfrak{m}} |B|^2 \le \int_0^1 |B|^2 \ll N/\log N$. Together with (A.1) this gives the cross-term contribution

$$\ll \Big(\frac{N}{\log N}\Big)^{1/2} \Big(\frac{N}{(\log N)^{3+\varepsilon}}\Big)^{1/2} \;=\; \frac{N}{(\log N)^{2+\varepsilon/2}}.$$

The pure error $\int_{\mathfrak{m}} |S - B|^2$ is exactly the quantity in (A.1). On the major arcs, standard major-arc analysis (Vaughan's identity or the explicit formula combined with (B2)-(B3)) shows that replacing S by B inside $\int_{\mathfrak{M}}(\cdot)$ affects the value by $O(N/(\log N)^{2+\delta})$ (details in the major-arc section). Collecting terms yields the stated reduction. □

What remains standard/checklist for β

* **Choice of β:** take the Selberg upper-bound sieve weight at level $D = N^{1/2-\varepsilon}$ (or a GPY-type almost-prime majorant) so that (B1)-(B4) hold. * **Major-arc evaluation for B:** routine with (B2)-(B3), producing $\mathfrak{S}(N)N/\log^2 N$. * **Minor-arc task:** prove the L^2 estimate (A.1). This is the core analytic input for the parity-blind replacement on $\mathfrak{m}$.

Status (conditional to A.1)

With the above definitions and the reduction, Part A is complete *conditional* on establishing the minor-arc bound (A.1). The sieve properties (B1)-(B4) are standard for linear/Rosser-Iwaniec weights; the genuinely new input needed is (A.1), which is the target of Parts B-D.

Part B. Type I / II Analysis

2. Route B Lemma - Type II parity gain

Theorem (Route B: 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 $|a_m|, |b_n| \ll \tau(m)^C, \tau(n)^C$,

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

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.

3. Lemma 3.2 (BV with parity, second moment)

Fix $A>0$. Then there is $B=B(A)$ such that for all large N and

$$Q \;\le\; N^{1/2}\,(\log N)^{-B},$$

every coefficient family c_n supported on $n\asymp N$ with a Type-I/II decomposition and divisor bounds (as in your draft) satisfies

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

Hypotheses (unchanged, recorded for reference). There exists $\psi\in C_c^\infty((1/2,2))$ with $c_n=\psi(n/N)\,d_n$, $|d_n|\le\tau_k(n)$ (fixed k), and either

* **Type I:** $d_n=\sum_{m\ell=n}\alpha_m\beta_\ell$ with $M\le N^{1/2-\eta}$, $|\alpha_m|\ll\tau_k(m)$, $|\beta_\ell|\ll\tau_k(\ell)$, or * **Type II:** same but $N^{\eta}\le M\le N^{1/2-\eta}$.

Proof. Write

$$S(\chi)=\sum_n c_n\,\lambda(n)\chi(n).$$

Insert the Type-I/II structure, smooth in m,ℓ as in your draft, and set $L=N/M$. As you already arranged, Cauchy-Schwarz in m reduces the problem to bounding, **uniformly in $m\sim M$**,

$$\Sigma_m := \sum_{q\le Q}\sum_{\chi \bmod q}\Big|\sum_{\ell\asymp L} b_\ell^{(m)}\,\lambda(\ell)\chi(\ell)\Big|^2,$$

with $|b_\ell^{(m)}|\ll\tau_k(\ell)$ and a fixed smooth weight $\psi_m(\ell)=\psi(m\ell/N)$.

We split characters into **non-pretentious** and **exceptional** via the pretentious Halász dichotomy.

(1) Non-pretentious block. By smooth Halász with divisor weights (standard, recorded in your draft), for any $C\ge 1$,

$$\sum_{\ell\asymp L} b_\ell^{(m)}\,\lambda(\ell)\chi(\ell)\;\ll_k\; L(\log L)^{-C}\qquad(\chi\notin\mathcal{E}(L;C)).$$

Hence

$$\sum_{q\le Q}\sum_{\substack{\chi \bmod q\\ \chi\notin\mathcal{E}(L;C)}}\Big|\sum_{\ell\asymp L}\cdots\Big|^2\;\ll\; Q^2L^2(\log N)^{-2C}.$$

(2) Exceptional block. Let $\mathcal{E}_{\le Q}(L;C)=\bigcup_{q\le Q}\{\chi \bmod q:\chi\in\mathcal{E}(L;C)\}$. By a **log-free zero-density bound** (Gallagher-Montgomery-Vaughan style) in its pretentious formulation, for any C_1 there is $C_2=C_2(C_1)$ with

$$\#\mathcal{E}_{\le Q}(L;C_1)\;\ll\; Q\,(\log(QL))^{-C_2},$$

uniformly for $Q\le L^{1/2}(\log L)^{-100}$, which our choice of Q ensures (since $L\ge N^\eta$). For each exceptional χ,

$$\Big|\sum_{\ell\asymp L} b_\ell^{(m)}\,\lambda(\ell)\chi(\ell)\Big|\;\ll_k\; L(\log N)^{O(1)}.$$

Therefore their total contribution is

$$\ll\; Q\cdot L^2(\log N)^{-C_2+O(1)}.$$

(3) Combine and reinsert m. Thus, for each m,

$$\Sigma_m\;\ll\; Q^2L^2(\log N)^{-2C}\;+\;QL^2(\log N)^{-C_2+O(1)}.$$

Multiply by $\sum_{m\sim M}|\alpha_m\lambda(m)|^2\ll M(\log N)^{O(1)}$ (from divisor bounds), use $ML=N$, and take C and then C_2 large in terms of A,k,η. This yields

$$\sum_{q\le Q}\sum_{\chi}|S(\chi)|^2\;\ll\;\frac{NQ}{(\log N)^A}.$$

Finally, sum over $O((\log N)^C)$ dyadic partitions used to build c_n; absorbing this by increasing A gives the stated bound. □

3.1. Lemma 3.2 (precise version and proof)

Lemma 0.2 (BV with parity; precise version). *Fix $A>0$, $k\in\mathbb{N}$, and $0<\eta<1/6$. There exists $B=B(A,k,\eta)$ and $C_0=C_0(A,k,\eta)$ such that the following holds for all sufficiently large N.*

Let $\psi\in C_c^\infty((1/2,2))$ with $\|\psi^{(j)}\|_\infty\le C_0^j$ for all $j\ge 0$ and define $c_n=\psi(n/N)\,d_n$ supported on $n\asymp N$, with $|d_n|\le\tau_k(n)$. Assume a Type I/II structure:

- ***Type I:*** *$d_n=\sum_{m\ell=n}\alpha_m\beta_\ell$ with $M\le N^{1/2-\eta}$, $|\alpha_m|\le\tau_k(m)$, $|\beta_\ell|\le\tau_k(\ell)$;*
- ***Type II:*** *same but $N^\eta\le M\le N^{1/2-\eta}$.*

Then for

$$Q\;\le\;N^{1/2}(\log N)^{-B}$$

we have

$$\sum_{q\le Q}\sum_{\chi\ (\mathrm{mod}\ q)}\Big|\sum_{n\asymp N}c_n\,\lambda(n)\,\chi(n)\Big|^2\quad\ll_{A,k,\eta,\psi}\quad\frac{NQ}{(\log N)^A}.$$

The same bound holds if one restricts to primitive χ, and with an extra coprimality gate $(n,q)=1$ inserted (by Möbius inversion) at a multiplicative cost $(\log N)^{O_k(1)}$ absorbed by A.

Proof. Write $S(\chi) = \sum c_n \lambda(n)\chi(n)$. Insert the Type I/II structure and smooth dyadically in m, ℓ; setting $L = N/M$, Cauchy-Schwarz in m reduces to bounding, uniformly in $m \sim M$,

$$\Sigma_m \ := \sum_{q \le Q} \sum_{\chi \pmod q} \Big| \sum_{\ell \asymp L} b_\ell^{(m)} \, \lambda(\ell) \, \chi(\ell) \Big|^2,$$

with $|b_\ell^{(m)}| \ll \tau_k(\ell)$ and a fixed smooth weight $\psi_m(\ell) = \psi(m\ell/N)$. Split characters into non-pretentious and exceptional via the pretentious distance $\mathbb{D}(1, \chi; L)$.

Non-pretentious block. By the *smooth Halász theorem with divisor weights* (Lemma 0.9), for any $C \ge 1$,

$$\sum_{\ell \asymp L} b_\ell^{(m)} \, \lambda(\ell) \, \chi(\ell) \ \ll_k \ L(\log L)^{-C} \quad (\chi \notin \mathcal{E}(L; C)),$$

uniformly in the smoothing and in $m \sim M$. Summing trivially over $q \le Q$ characters gives $\ll Q^2 L^2 (\log N)^{-2C}$.

Exceptional block. Let $\mathcal{E}_{\le Q}(L; C)$ be the union of exceptional characters up to modulus Q. By a *log-free zero-density exceptional-set bound* (Lemma 0.10), for any C_1 there exists $C_2(C_1)$ such that

$$\#\mathcal{E}_{\le Q}(L; C_1) \ \ll \ Q \, (\log(QL))^{-C_2}, \qquad Q \le L^{1/2} (\log L)^{-100}.$$

For such χ, partial summation with divisor weights gives

$$\Big| \sum_{\ell \asymp L} b_\ell^{(m)} \, \lambda(\ell) \, \chi(\ell) \Big| \ \ll_k \ L(\log N)^{O(1)}.$$

Hence the exceptional contribution is $\ll QL^2 (\log N)^{-C_2 + O(1)}$. Any single potential Siegel character is handled by Deuring–Heilbronn (Lemma 0.11), giving an exponentially small factor $e^{-c\sqrt{\log L}}$ and thus negligible versus $(\log N)^{-A}$ after dyadic summation.

Reinsert m. Multiply by $\sum_{m \sim M} |\alpha_m \lambda(m)|^2 \ll M(\log N)^{O(1)}$ and use $ML = N$. Taking C and then C_2 sufficiently large in terms of A, k, η yields

$$\sum_{q \le Q} \sum_{\chi} |S(\chi)|^2 \ \ll \ \frac{NQ}{(\log N)^A}.$$

Summing over the $O((\log N)^{O(1)})$ dyadic partitions completes the proof. The restriction to primitive characters and the insertion of $(n, q) = 1$ gates (by Möbius inversion) only change constants by $(\log N)^{O(1)}$ absorbed into A.

Range check. Since $L \ge N^\eta$ and $Q \le N^{1/2} (\log N)^{-B}$ with $B = B(A, k, \eta)$ large, we have $Q \le L^{1/2} (\log L)^{-100}$ for large N, as required by Lemma 0.10. □

Part C. Type III Analysis

4. Lemma S2.4 (Prime-averaged short-shift gain)

We keep the notation from §4: $X \ge 3$, $0 < \kappa < \frac{1}{4}$, $Q \le X^{1/2-\kappa}$, a dyadic set $\mathcal{Q} \subset [Q, 2Q]$ of moduli, primes $\mathcal{P} = \{p \in [P, 2P]\}$ with $P = X^\vartheta$, $0 < \vartheta < \frac{1}{6} - \kappa$, and complex coefficients $|\alpha_p| \le 1$. For each f in an orthonormal Hecke basis (holomorphic or Maass of any weight, including oldforms, plus the Eisenstein family), define the prime amplifier

$$\mathrm{Amp}(f) = \sum_{p \in \mathcal{P}} \alpha_p \, \lambda_f(p).$$

Let $h_Q(t) = h(t/Q)$ with a fixed even $h \in C_c^\infty([-2, 2])$, $h(0) = 1$. For each $q \in \mathcal{Q}$ apply the Kuznetsov formula at level q with spectral test h_Q. By the **Kernel Localization Lemma** (4.S.1 below), the geometric kernel $\mathcal{W}_q(z)$ satisfies

$$\mathcal{W}_q(z),\ z\partial_z^j\mathcal{W}_q(z) \ \ll_{A,j} \Big(1+\frac{z}{Q}\Big)^{-A} \qquad (\forall A, j\ge 0),$$

uniformly across spectral families and all q. Writing $z=\frac{4\pi\sqrt{mn}}{c}$ shows the c-sum is supported on $c\asymp C:=X^{1/2}/Q$ with derivative control; we parameterize $c=qr$ so $r\asymp R:=C/q\asymp X^{1/2}/Q^2$.

4.S.0. Kuznetsov at level q

For coprime integers $m,n\ge 1$ and an even test function h with standard decay/holomorphy (IK, §16), the Kuznetsov formula on $\Gamma_0(q)$ reads

$$\begin{aligned}
&\sum_{k\equiv 0\ (2)} h^{\mathrm{hol}}(k) \sum_{f\in\mathcal{B}_k(q)} \overline{\rho_f(m)}\,\rho_f(n)\\
&\quad + \sum_{f\in\mathcal{B}_{\mathrm{Maass}}(q)} h^{\mathrm{Maass}}(t_f)\,\overline{\rho_f(m)}\,\rho_f(n)\\
&\quad + \frac{1}{4\pi}\int_{-\infty}^{\infty} h^{\mathrm{Eis}}(t)\,\overline{\rho_t(m)}\,\rho_t(n)\,dt\\
&= \delta_{m=n}\,H_0 \ + \sum_{c\equiv 0\ (q)} \frac{S(m,n;c)}{c}\,\mathcal{W}_q\Big(\frac{4\pi\sqrt{mn}}{c}\Big),
\end{aligned}$$

where $\mathcal{W}_q$ are Hankel transforms built from h and the relevant Bessel kernels. Taking $h_Q(t)=h(t/Q)$ with $h\in C_c^\infty([-2,2])$ even yields uniform spectral localization; oldforms and the Eisenstein family are included in the spectral sums and share the same geometric side. References: IK, Thms. 16.3, 16.6; Deshouillers-Iwaniec (1982), §2.

Lemma 0.3 (Prime-averaged short-shift gain)**.** *With the hypotheses above, for any $\varepsilon>0$,*

$$\mathrm{OD} \ \ll_\varepsilon\ (Q^2+X)^{1-\delta}\,|\mathcal{P}|^{2-\delta}\,X^\varepsilon,$$

where one may take

$$\delta \ =\ \tfrac{1}{1000}\,\min\big\{\kappa,\ \tfrac12-3\vartheta\big\}.$$

The bound holds uniformly in $\{\alpha_p\}$, and after summing holomorphic, Maass (new+old), and Eisenstein contributions.

Proof. We split the proof into five steps.

Step 1: Balanced amplifier (deterministic signs). Let $\{\varepsilon_p\}_{p\in\mathcal{P}}\subset\{\pm1\}$ be a sequence with

$$\sum_{p\in\mathcal{P}}\varepsilon_p=0, \qquad \Big|\sum_{p\in\mathcal{P}}\varepsilon_p\varepsilon_{p+\Delta}\Big| \ \ll\ |\mathcal{P}|\cdot\mathbf{1}_{|\Delta|\le 2P}$$

Deterministic sign choice. A standard conditional-expectation derandomization (method of conditional expectations for Rademacher variables) constructs $\{\varepsilon_p\}\subset\{\pm1\}$ with $\sum_{p\in\mathcal{P}}\varepsilon_p=0$ and

$$\Big|\sum_{p\in\mathcal{P}}\varepsilon_p\varepsilon_{p+\Delta}\Big| \ \ll\ |\mathcal{P}|\,\mathbf{1}_{|\Delta|\le 2P}$$

by greedily fixing signs to keep the quadratic form $\sum_\Delta w_\Delta(\sum_p\varepsilon_p\varepsilon_{p+\Delta})^2$ minimal for suitable nonnegative weights w_Δ supported on $|\Delta|\le 2P$. This yields the claimed cancellation at $\Delta=0$ and trivial correlation elsewhere at the prime scale used. (Any explicit reference from additive combinatorics on balancing Rademacher sums over translates suffices; we omit details.)

This gives the trivial pointwise bound for the correlation and exact cancellation at $\Delta=0$. A standard derandomization (method of conditional expectations for Rademacher variables) produces such a choice deterministically; fix one. Define

$$A_f = \sum_{p\in\mathcal{P}} \varepsilon_p \, \lambda_f(p).$$

Then for any complex S_f,

$$\sum_f |S_f|^2 \ \le \ \frac{1}{|\mathcal{P}|^2} \sum_f |A_f \, S_f|^2,$$

by Cauchy-Schwarz after inserting $1 = \big(\sum_p \varepsilon_p^2\big)/|\mathcal{P}|$ and expanding (the vanishing of $\sum_p \varepsilon_p$ kills the diagonal $p = p'$ in the amplifier square). We will apply this with

$$S_{q,\chi,f} = \sum_{n\asymp X} \alpha_n \, \lambda_f(n) \, \chi(n),$$

where $\{\alpha_n\}$ is the Type-III coefficient block (divisor-bounded, smooth).

Step 2: Hecke relations and removal of the n/p tail. Opening $|A_f S_{q,\chi,f}|^2$ and using Hecke multiplicativity,

$$\lambda_f(p)\lambda_f(n) = \begin{cases} \lambda_f(pn) & (p \nmid n), \\ \lambda_f(pn) - \lambda_f(n/p) & (p \mid n), \end{cases}$$

we may write the amplified second moment as a finite linear combination of terms with Hecke arguments pn (and possibly n/p). Because the Type-III support is smooth and confined to $n \asymp X$, the contribution of the n/p branch is supported on $n \asymp X$ with the extra condition $p \mid n$; by smooth partition and partial summation this piece is bounded by the same off-diagonal analysis (it is in fact easier since it has an extra divisibility). We henceforth treat explicitly the pn branch; all others are dominated in the same way and absorbed into the final implied constant.

Step 3: Kuznetsov and off-diagonal reorganization. Summing over (q, χ, f) with $q \in \mathcal{Q}$ and primitive $\chi \pmod q$, and applying Kuznetsov with test h_Q at level q, the **diagonal** terms vanish by $\sum_p \varepsilon_p = 0$. The **off-diagonal** geometric side takes the model form

$$\mathrm{OD} = \sum_{q\in\mathcal{Q}} \sum_{c\equiv 0 \ (q)} \frac{1}{c} \sum_{\substack{p_1,p_2\in\mathcal{P} \\ p_1\neq p_2}} \sum_{m\asymp X} \alpha_m \overline{\alpha_{m'}} \, S(m_{p_1}, m'_{p_2}; c) \, \mathcal{W}_q\Big(\frac{4\pi\sqrt{m_{p_1} m'_{p_2}}}{c}\Big) \, \varepsilon_{p_1}\varepsilon_{p_2},$$

with $m_p = pm$ (suppressing the harmless χ-twist which disappears on the geometric side). By the **Kernel Localization Lemma**, we may restrict to $c \in [C/2, 2C]$, $C := X^{1/2}/Q$, and write $c = qr$ with $r \asymp R := X^{1/2}/Q^2$. Grouping by the short prime shift $\Delta := p_1 - p_2$ and introducing the pair-count

$$\nu(\Delta) = \#\{(p_1,p_2) \in \mathcal{P}^2 : \ p_1 - p_2 = \Delta, \ p_1 \neq p_2\},$$

we reorganize

$$\mathrm{OD} = \sum_{q\in\mathcal{Q}} \sum_{r\asymp R} \frac{1}{qr} \sum_{\Delta\neq 0} \ \nu(\Delta) \ \Sigma_{q,r}(\Delta),$$

where, for a smooth weight $W_{q,r}$ (absorbing α_m and the Bessel kernel),

$$\Sigma_{q,r}(\Delta) = \sum_{m\asymp X} S(m, m+\Delta; qr) \, W_{q,r}(m,\Delta), \qquad m \mapsto W_{q,r} \text{ is } X\text{-smooth}, \ \Delta \mapsto W_{q,r} \text{ is } P\text{-smooth}.$$

All derivative bounds depend only on finitely many derivatives of h and the smoothness of $\{\alpha_n\}$, hence are **uniform** in q, r.

Step 4: Δ-second moment and harvesting the prime average. By Cauchy-Schwarz in Δ, the trivial bound $\nu(\Delta) \le |\mathcal{P}|$, and Lemma 0.4, we have

$$\sum_{|\Delta|\le P} \nu(\Delta)\,|\Sigma_{q,r}(\Delta)| \;\le\; |\mathcal{P}|^{1/2}\Big(\sum_{|\Delta|\le P}\nu(\Delta)\Big)^{1/2}\Big(\sum_{|\Delta|\le P}|\Sigma_{q,r}(\Delta)|^2\Big)^{1/2}$$

$$\ll_\varepsilon\; |\mathcal{P}|\,(P+qr)^{1/2}\,(qr)^{1/2+\varepsilon}\,X^{1/2+\varepsilon}.$$

Hence, for each q,

$$\sum_{r\asymp R}\frac{1}{qr}\sum_{\Delta}\nu(\Delta)\,\Sigma_{q,r}(\Delta) \;\ll_\varepsilon\; |\mathcal{P}|\,q^{-1/2+\varepsilon}\,X^{1/2+\varepsilon}\sum_{r\asymp R} r^{-1/2+\varepsilon}(P+qr)^{1/2}.$$

On the support $r \asymp R$ we have $qr \asymp C = X^{1/2}/Q$, thus $(P+qr)^{1/2}$ is independent of r (up to constants), and $\sum_{r\asymp R} r^{-1/2+\varepsilon} \asymp R^{1/2+\varepsilon}$. Using $q^{-1/2}R^{1/2} \asymp Q^{-1}$ gives

$$\sum_r \cdots \;\ll_\varepsilon\; |\mathcal{P}|\,Q^{1+\varepsilon}\,\big(P+X^{1/2}/Q\big)^{1/2}.$$

Summing $q \in \mathcal{Q}$ (there are $O(Q)$ moduli) yields the **conductor bound**

$$\boxed{\mathrm{OD} \;\ll_\varepsilon\; |\mathcal{P}|\,Q^{2+\varepsilon}\,\big(P+X^{1/2}/Q\big)^{1/2}.} \tag{4.S.X}$$

This is the only place where Bessel tails, oldforms, and Eisenstein matter; all are covered by the kernel lemma, which is uniform across spectral families (the proof for each family has the same derivative-decay structure).

Step 5: Regime balance and choice of δ. We rewrite ($4.S\star$) in the desired $(Q^2+X)^{1-\delta}|\mathcal{P}|^{2-\delta}$ form by splitting into the two natural regimes, using $Q \le X^{1/2-\kappa}$ and $|\mathcal{P}| \asymp P/\log P = X^{\vartheta+o(1)}$.

* **Regime I ($Q^2 \ge X$).** Then $X^{1/2}/Q \le Q$, hence $(P+X^{1/2}/Q)^{1/2} \ll P^{1/2}+Q^{1/2} \ll Q^{1/2}$ because $P = X^{\vartheta} \le X^{1/6-\kappa} \le Q^{1/3}$. Thus

$$\mathrm{OD} \;\ll_\varepsilon\; |\mathcal{P}|\,Q^{5/2+\varepsilon}.$$

We want $\mathrm{OD} \ll (Q^2)^{1-\delta}\,|\mathcal{P}|^{2-\delta}$, i.e.

$$|\mathcal{P}|\,Q^{5/2} \;\ll\; Q^{2-2\delta}\,|\mathcal{P}|^{\,2-\delta}.$$

Rearranged, $Q^{1/2+2\delta} \ll |\mathcal{P}|^{\,1-\delta}$. Using $Q \asymp X^{1/2}$ in this regime and $|\mathcal{P}| \asymp X^{\vartheta}$, this is implied by

$$\tfrac{1}{4}+\delta \;\le\; \vartheta(1-\delta).$$

This holds once $\delta \le \frac{1}{2}-3\vartheta$ (take a small fraction to cover constants).

* **Regime II ($Q^2 \le X$).** Then $X^{1/2}/Q \ge X^{\kappa}$, so

$$\mathrm{OD} \;\ll_\varepsilon\; |\mathcal{P}|\,Q^{2+\varepsilon}\,X^{\max\{\vartheta,\kappa\}/2} \;\le\; |\mathcal{P}|\,X^{1-2\kappa+\varepsilon}\,X^{\max\{\vartheta,\kappa\}/2}.$$

We want $\mathrm{OD} \ll X^{1-\delta}\,|\mathcal{P}|^{\,2-\delta}$. With $|\mathcal{P}| \asymp X^{\vartheta}$ this reduces to

$$-\vartheta(1-\delta) \;\le\; -\delta+\tfrac{3}{2}\kappa-\tfrac{(\vartheta-\kappa)_+}{2}.$$

This is satisfied provided

$$\delta \;\le\; \min\Big\{\,\kappa,\;\tfrac{1}{2}-3\vartheta\,\Big\}$$

up to harmless absolute constants; we pick a safety factor 1/1000 to absorb all X^{ε} and log terms from $|\mathcal{P}|$.

Choosing $\delta = \frac{1}{1000}\min\{\kappa, \frac{1}{2}-3\vartheta\}$ meets both regimes, and plugging back $|\mathcal{P}| = X^{\vartheta+o(1)}$ absorbs the $|\mathcal{P}|^{-\delta}$ factor into X^{ε}, yielding the claimed bound. $\square$

4.S.1. Kernel localization (stated for completeness)

Lemma (uniform kernels). Let $h \in C_c^\infty([-2,2])$ be even with $h(0) = 1$, $h_Q(t) = h(t/Q)$. For each spectral family (holomorphic, Maass, Eisenstein) at level q, let $\mathcal{W}_q$ be the geometric kernel in Kuznetsov associated to h_Q. Then for all $A, j \geq 0$,

$$\mathcal{W}_q(z) \ll_A \Big(1 + \frac{z}{Q}\Big)^{-A}, \qquad z\,\partial_z^j \mathcal{W}_q(z) \ll_{A,j} \Big(1 + \frac{z}{Q}\Big)^{-A},$$

uniformly in $q \geq 1$ and across families. In particular the c-sum is restricted to $c \asymp X^{1/2}/Q$ (with tails $O_A(X^{-A})$).

Sketch. Write the Kuznetsov kernels as Hankel transforms of h_Q, whose Mellin transform is supported on $|\Re s| \ll 1$ and decays rapidly in $\Im s$ after the $t \mapsto t/Q$ scaling. Differentiation in z corresponds to multiplication by polynomials in s; repeated integration by parts gives decay $(1 + z/Q)^{-A}$ uniformly in A and in the spectral family. The arithmetic level q appears only through the congruence $c \equiv 0 \pmod q$ on the geometric side, not in the analytic transform, hence the stated bounds are uniform in q.

4.S.2. Remarks on oldforms and Eisenstein

The Kuznetsov decomposition splits into holomorphic, Maass new/old, and Eisenstein. Each contributes the same Kloosterman structure with its own kernel $\mathcal{W}_q^{(*)}$ obeying the same decay/derivative bounds (the proofs for J- and K-transforms are identical after scaling). Our use of the Δ-second-moment lemma and Weil's bound is completely **family-agnostic**, so the sum over all families only changes the implied constant.

4.S.3. Parameters at a glance

* Minor-arc cut: $Q \leq X^{1/2-\kappa}$. * Amplifier length: $P = X^\vartheta$ with $0 < \vartheta < \frac{1}{6} - \kappa$. * Resulting saving: $\delta = \frac{1}{1000}\min\{\kappa, \frac{1}{2} - 3\vartheta\}$. * Recommended choice later in Part C.5/D.6: fix any small $\vartheta \leq \kappa/4$, then $\delta \gg \vartheta$; the factor $|\mathcal{P}|^{-\delta} = X^{-\vartheta\delta}$ is absorbed into X^ε.

With S2.4 now fully explicit and uniform, it plugs directly into the Type-III spectral second moment and the assembly/dyadic step used later.

4.S.4. Short-shift Δ second-moment lemma

We record the correlation bound implicitly used in Step 4 of S2.4.

Lemma 0.4 (Short-shift Δ second moment)**.** *Let $X \geq 3$, $q \geq 1$, $r \geq 1$, and $C = qr$. Let $W_{q,r}(m,\Delta)$ be a smooth weight supported on $m \asymp X$, $|\Delta| \leq P$ with the bounds*

$$\partial_m^i \partial_\Delta^j W_{q,r}(m,\Delta) \ll_{i,j} X^{-i} P^{-j} \qquad (0 \leq i,j \leq 10),$$

uniformly in q, r, and assume $q \leq Q \leq X^{1/2-\kappa}$ and $r \asymp X^{1/2}/Q^2$ (as in S2.4). Define

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

where $S(\cdot,\cdot;C)$ is the Kloosterman sum. Then for any $\varepsilon > 0$,

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

In particular,

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

Proof. Write $\Sigma(\Delta) = \Sigma_{q,r}(\Delta)$ and consider

$$\mathcal{S} \;:=\; \sum_{|\Delta|\le P} \big|\Sigma(\Delta)\big|^2 \;=\; \sum_{|\Delta|\le P} \sum_{m\asymp X} \sum_{n\asymp X} S(m, m+\Delta; C)\, \overline{S(n, n+\Delta; C)}\, W_{q,r}(m,\Delta)\, \overline{W_{q,r}(n,\Delta)}.$$

Open the Kloosterman sums and use orthogonality modulo C to sum over Δ (or equivalently apply Kuznetsov's bilinear form in the variable m at modulus C as in Deshouillers-Iwaniec). The Δ-sum produces congruence conditions linking m and n modulo C. After standard manipulations (see IK, §16; Deshouillers-Iwaniec, Ann. Inst. Fourier 1982), one arrives at

$$\mathcal{S} \;\ll\; (P+C)\, C \sum_{m\asymp X} \sum_{n\asymp X} \big|\widetilde{W}(m,n)\big| \;+\; O_A(X^{-A}),$$

where $\widetilde{W}$ is a two-variable smooth weight with the same X-scale in each variable and derivative bounds inherited from $W_{q,r}$. The factor $(P+C)$ reflects the length of the Δ-sum and the modulus C barrier (cf. the dispersion method), and the O_A-term comes from integrating by parts using the ∂_Δ^j control.

Now apply the Cauchy-Schwarz inequality and Weil's bound $|S(u,v;C)| \le (u,v,C)^{1/2}\tau(C)C^{1/2} \ll C^{1/2+\varepsilon}$ to the Kloosterman factors appearing before orthogonality; together with the smooth dyadic partition on m, n of length $\asymp X$, we obtain

$$\mathcal{S} \;\ll_\varepsilon\; (P+C)\, C^{1+2\varepsilon} \Big(\sum_{m\asymp X} \sum_{n\asymp X} |\widetilde{W}(m,n)|^2 \Big)^{1/2} X^{1/2}.$$

By the derivative bounds on $W_{q,r}$ and Fubini, the L^2 norm of $\widetilde{W}$ over m, n is $\ll X^{1+\varepsilon}$. Hence

$$\mathcal{S} \;\ll_\varepsilon\; (P+C)\, C^{1+2\varepsilon}\, X^{1+2\varepsilon},$$

as claimed. The square-root version follows by taking $\mathcal{S}^{1/2}$.

Uniformity in q, r is ensured because $C = qr$ only appears through the modulus parameter in Kuznetsov/orthogonality, while the analytic kernel control enters solely via the X- and P-scale derivatives on $W_{q,r}$, which are independent of q, r by hypothesis. □

5. Type-III Spectral Bound

Proposition 0.5 (Type-III spectral second moment)**.** *Let (α_n) be a smooth Type-III coefficient sequence supported on $n \asymp X$, with divisor-type bounds $|\alpha_n| \ll_\varepsilon \tau(n)^C$ and smooth weight of width $X^{1+o(1)}$. For $Q \ge 1$, let the outer sums range over moduli $q \le Q$, primitive characters $\chi \pmod q$, and an orthonormal Hecke basis f (holomorphic + Maass, including oldforms and Eisenstein as in Kuznetsov). Assume **Lemma S2.4 (Prime-averaged short-shift gain)** holds with some fixed $\delta > 0$. Then, for any $\varepsilon > 0$,*

$$\sum_{q\le Q} \sum_{\chi \pmod q} \sum_f \Big| \sum_{n\asymp X} \alpha_n\, \lambda_f(n)\chi(n) \Big|^2 \quad \ll_{\varepsilon,C} \quad (Q^2+X)^{1-\delta}\, X^\varepsilon.$$

Proof using Lemma 0.3. **Step 1: Balanced prime amplifier that kills the diagonal.** Let $\mathcal{P}$ be the set of primes $p \in [P, 2P]$ with $P = X^\vartheta$ (to be chosen; Lemma S2.4 is uniform in P). Choose deterministic signs $\varepsilon_p \in \{\pm 1\}$ so that

$$\sum_{p\in\mathcal{P}} \varepsilon_p = 0 \qquad \text{and} \qquad \Big| \sum_{p\in\mathcal{P}} \varepsilon_p \varepsilon_{p+\Delta} \Big| \;\ll\; |\mathcal{P}| \cdot \mathbf{1}_{|\Delta| \le P^{1-o(1)}},$$

i.e. a "balanced Rademacher" choice; a random choice satisfies this with probability $\gg 1$, and we fix one such choice.

Define the amplifier on the spectrum:

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

Because $\sum_p \varepsilon_p = 0$, expanding $|A_f|^2$ removes the pure diagonal $p = p'$ on average over signs, leaving only short prime shifts $p \neq p'$ with $\Delta = p - p'$ (the "short-shift" structure needed for Lemma S2.4).

Step 2: Diagonal-free reduction by polarization. For any complex numbers S_f,

$$\sum_f |S_f|^2 = \frac{1}{\sum_{p\in\mathcal{P}} \varepsilon_p^2} \sum_f |S_f|^2 \cdot \Big(\sum_{p\in\mathcal{P}} \varepsilon_p^2\Big) = \frac{1}{|\mathcal{P}|} \sum_f |S_f|^2 \cdot \sum_{p\in\mathcal{P}} 1.$$

Insert $1 = \frac{1}{|\mathcal{P}|}\sum_{p\in\mathcal{P}} \varepsilon_p^2$ and then *complete the square* with A_f:

$$\sum_f |S_f|^2 = \frac{1}{|\mathcal{P}|^2} \sum_f |S_f|^2 \cdot \sum_{p,p'\in\mathcal{P}} \varepsilon_p\varepsilon_{p'}\, \lambda_f(p)\lambda_f(p') \quad\le\quad \frac{1}{|\mathcal{P}|^2} \sum_f |A_f\, S_f|^2,$$

where the inequality is Cauchy-Schwarz in $\sum_{p,p'}$ (this is the standard "balanced-amplifier domination": the diagonal $p = p'$ having zero mean is what prevents a trivial loss).

Apply this with

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

Summing over $q \le Q, \chi$ gives

$$\sum_{q\le Q}\sum_{\chi}\sum_{f} |S_{q,\chi,f}|^2 \quad\le\quad \frac{1}{|\mathcal{P}|^2} \sum_{q\le Q}\sum_{\chi}\sum_{f} \big|A_f\, S_{q,\chi,f}\big|^2. \tag{3.1}$$

Step 3: Kuznetsov after opening the amplifier. Open $|A_f S_{q,\chi,f}|^2$ and use Hecke relations to rewrite prime factors $\lambda_f(p)\lambda_f(n)$ as a (short) combination of $\lambda_f(pn)$ and $\lambda_f(n/p)$ (the latter is discarded as $p \nmid n$ for Type-III supports). After summing over (q, χ, f) and applying Kuznetsov (including oldforms + Eisenstein), the contribution splits into:

- **Short-shift off-diagonal (OD):** correlations of the form $\sum_{p\neq p'\in\mathcal{P}} \varepsilon_p\varepsilon_{p'} \sum_{m,n\asymp X} \alpha_m\overline{\alpha_n}\, \mathcal{K}_q(m, n; p-p')$, with Kloosterman sums $S(m, n; cq)$ and Bessel kernels;
- **(Spectral) diagonal/main terms:** the parts that would arise from $p = p'$ or $\Delta = 0$, but these are annihilated by $\sum_p \varepsilon_p = 0$ and by our balanced-sign choice, leaving at most lower-order boundary terms absorbed in X^ε.

Precisely this OD piece is what **Lemma S2.4** estimates *after* the amplifier and Kuznetsov:

Lemma S2.4 (assumed). Uniformly in $P = X^\vartheta$,

$$\mathrm{OD} \quad\ll\quad (Q^2 + X)^{1-\delta}\, |\mathcal{P}|^{2-\delta}\, X^\varepsilon.$$

All Bessel-kernel ranges (small/large) are handled there; Weil bounds for $S(\cdot,\cdot;\cdot)$, the $c \equiv 0 \pmod q$ constraint, oldforms and Eisenstein, and the short-shift averaging in Δ are already accounted for in the statement of S2.4.

Therefore,

$$\sum_{q\le Q}\sum_{\chi}\sum_{f} \big|A_f\, S_{q,\chi,f}\big|^2 \quad\ll\quad (Q^2 + X)^{1-\delta}\, |\mathcal{P}|^{2-\delta}\, X^\varepsilon. \tag{3.2}$$

Step 4: Divide out the amplifier and optimize P. Insert (3.2) into (3.1):

$$\sum_{q\le Q}\sum_{\chi}\sum_{f} |S_{q,\chi,f}|^2 \quad\ll\quad \frac{1}{|\mathcal{P}|^2}\, (Q^2 + X)^{1-\delta}\, |\mathcal{P}|^{2-\delta}\, X^\varepsilon \;=\; (Q^2 + X)^{1-\delta}\, |\mathcal{P}|^{-\delta}\, X^\varepsilon.$$

Choose any fixed $\vartheta > 0$ (e.g. $\vartheta = \delta/4$) so that $|\mathcal{P}| = P/\log P = X^{\vartheta+o(1)}$ and absorb $|\mathcal{P}|^{-\delta} = X^{-\vartheta\delta+o(1)}$ into X^ε (by shrinking ε). This yields

$$\sum_{q\le Q}\sum_{\chi}\sum_{f}\Big|\sum_{n\asymp X}\alpha_n\,\lambda_f(n)\chi(n)\Big|^2 \;\ll\; (Q^2+X)^{1-\delta}\,X^\varepsilon,$$

as claimed. □

□

Remarks

- **Uniformity & hypotheses.** The argument only used (i) Type-III structure (smooth α_n, divisor bounds), (ii) balanced prime amplifier with $\sum \varepsilon_p = 0$, (iii) Kuznetsov with full continuous and oldform ranges, and (iv) Lemma S2.4's OD estimate. No further spectral gap input is needed beyond what S2.4 encapsulates.
- **Why the diagonal doesn't spoil the saving.** The balanced amplifier removes the dangerous $p = p'$ contribution *before* applying Kuznetsov. What remains are genuinely shifted correlations ($\Delta \neq 0$), to which S2.4 applies and gives the $(Q^2+X)^{1-\delta}$ saving.
- **Choice of ϑ.** Any fixed $\vartheta \in (0, 1/2)$ permitted by S2.4 works; the $|\mathcal{P}|^{-\delta}$ factor improves the exponent, and we simply absorb it into X^ε.

This completes Part C.5 once Lemma S2.4 is rigorously in place.

Part D. Assembly

6. Dyadic Decomposition (final)

Statement D.6.

Let $S(\alpha) = \sum_{n\le N}\Lambda(n)\,w(n)\,e(\alpha n)$ with a fixed smooth weight w supported on $[N/2, 2N]$ and let $B(\alpha)$ be the parity-blind majorant from Part A. For the minor arcs $\mathfrak{m}$ defined with denominator cutoff $Q = N^{1/2-\varepsilon}$, assume the analytic inputs:

- **(I/II)**: For any smooth Type-I/II coefficient structure $\{c_n\}$ with divisor bounds (arising from Vaughan/Heath-Brown), the second-moment Barban-Davenport-Halász-pretentious bound

$$\sum_{q\le Q}\ \sum_{\chi \bmod q}\Big|\sum_{n\le N} c_n\,\lambda(n)\chi(n)\Big|^2 \;\ll\; \frac{NQ}{(\log N)^A} \tag{D.1}$$

  holds for each fixed $A > 0$. (This is Lemma 3.2 and the "Route B Lemma" for the balanced ranges.)
- **(III)**: For every dyadic Type-III block $\sum_{n\asymp X}\alpha_n\,\lambda_f(n)\chi(n)$ produced after amplification and Kuznetsov, the prime-averaged off-diagonal is bounded by

$$\mathrm{OD} \;\ll\; (Q^2+X)^{1-\delta}\,|\mathcal{P}|^{2-\delta} \tag{D.2}$$

  for some fixed $\delta > 0$, uniformly for amplifier length $|\mathcal{P}| = X^\vartheta$ with $\vartheta = \vartheta(\delta) > 0$, and with uniform control of oldforms/Eisenstein and Bessel kernels. (This is Lemma S2.4 and its Type-III spectral corollary.)

Then, for any $\varepsilon > 0$,

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

Proof.

Step 1: Identity and dyadic model. Apply a 3-, 4-, or 5-fold Heath-Brown identity (any standard version suffices) to Λ with cut parameters

$$U = N^{\mu}, \quad V = N^{\nu}, \quad W = N^{\omega}, \qquad 0 < \mu \le \nu \le \omega < 1,$$

chosen below. We write

$$S(\alpha) - B(\alpha) = \sum_{\text{HB terms } \mathcal{T}} \mathcal{S}_{\mathcal{T}}(\alpha),$$

where each $\mathcal{S}_{\mathcal{T}}$ is a finite linear combination (with coefficients having $\ll_\epsilon n^\epsilon$ divisor bounds and smooth dyadic cutoffs) of exponential sums of one of the three structural types:

* **Type I**: $\sum_{m \asymp M} a_m \sum_{n \asymp N/M} b_n \, e(\alpha m n)$ with $M \le U$ (or the dual small variable), * **Type II**: balanced $\sum_{m \asymp M} \sum_{n \asymp N/M} a_m b_n \, e(\alpha m n)$ with $U \ll M \ll N/U$, * **Type III**: "ternary" or highly factorized pieces with all variables in ranges $\ll N^{1/3+o(1)}$, which, after the amplifier/Kuznetsov transition, become prime-averaged short-shift sums against automorphic coefficients.

All sums are partitioned into **$O((\log N)^C)$** dyadic blocks in all active variables for some fixed C.

Step 2: Minor-arc L^2 via large sieve on dyadics. Let $\mathfrak{M}(q,a)$ be the standard major arc around a/q with width $\asymp (qQ)^{-1}$, and set $\mathfrak{m} = [0,1] \setminus \bigcup_{q \le Q} \bigcup_{(a,q)=1} \mathfrak{M}(q,a)$. On $\mathfrak{m}$ we use the standard large-sieve/dispersion reduction:

$$\int_{\mathfrak{m}} \left| \mathcal{S}_{\mathcal{T}}(\alpha) \right|^2 d\alpha \;\ll\; \frac{1}{Q^2} \sum_{q \le Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} \left| \sum_n c_n \, e\!\left(\frac{an}{q}\right) \right|^2, \tag{D.3}$$

for suitable coefficients c_n associated to the dyadic block $\mathcal{T}$. By opening the square and expanding in Dirichlet characters modulo q, (D.3) reduces to sums of the form

$$\sum_{q \le Q} \sum_{\chi \bmod q} \Big| \sum_{n \asymp X} c_n \, \lambda(n) \chi(n) \Big|^2, \tag{D.4}$$

or, in the Type-III case after the amplifier/Kuznetsov step, to a spectral second moment whose diagonal/off-diagonal split is controlled by (D.2).

We now bound (D.4) block-wise and then sum the dyadics.

Step 3: Type I/II dyadics.

Choose $U = N^{1/3}$ (any $\mu \in (1/4, 1/2)$ is fine) so that all Type I/II ranges from the chosen Heath-Brown identity fall either in the "small-large" or "balanced" regimes. By the input (I/II), for any $A > 0$,

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

Each Type I or Type II dyadic contributes $\ll NQ/(\log N)^A$. There are $\ll (\log N)^C$ such dyadics in total, so by taking $A \ge 3 + C + 10\varepsilon^{-1}$ we obtain

$$\sum_{\text{Type I/II dyadics}} \int_{\mathfrak{m}} \left| \mathcal{S}_{\mathcal{T}}(\alpha) \right|^2 d\alpha \;\ll\; \frac{N}{(\log N)^{3+\varepsilon}}. \tag{D.5}$$

Step 4: Type III dyadics.

Fix $V = W = N^{1/3}$ so that the residual blocks with all variables $\ll N^{1/3+o(1)}$ are designated Type III. For such a block, let its "outer scale" be $X \asymp N^{\xi}$ with $\xi \in (0,1)$ determined by the product of the active variables. After applying the amplifier of length $|\mathcal{P}| = X^{\vartheta}$ and Kuznetsov, we face a spectral second moment whose off-diagonal obeys (D.2):

$$\mathrm{OD} \ \ll\ (Q^2+X)^{1-\delta}\,|\mathcal{P}|^{\,2-\delta} \ =\ (Q^2+X)^{1-\delta}\,X^{\vartheta(2-\delta)}.$$

Take $\vartheta = \frac{\delta}{8}$ (any fixed small choice depending on δ works). Since $Q = N^{1/2-\varepsilon}$, we have $Q^2 = N^{1-2\varepsilon}$. Two regimes:

* If $X \le Q^2$ then $\mathrm{OD} \ll N^{(1-2\varepsilon)(1-\delta)}\,X^{\vartheta(2-\delta)}$. * If $X \ge Q^2$ then $\mathrm{OD} \ll X^{1-\delta+\vartheta(2-\delta)}$.

In both cases there is a fixed saving $X^{-\eta}$ (or $N^{-\eta}$) for some $\eta = \eta(\delta,\vartheta,\varepsilon) > 0$ against the trivial diagonal scale, after the standard dispersion normalization. Consequently each Type III dyadic contributes

$$\int_{\mathfrak{m}} \big|\mathcal{S}_{\mathcal{T}}(\alpha)\big|^2 d\alpha \ \ll\ \frac{N}{(\log N)^A}\,X^{-\eta} \quad + \quad (\text{diagonal}). \tag{D.6}$$

The diagonal is controlled either by the amplifier normalization or by subtracting the parity-blind majorant $B(\alpha)$ (which removes the main term on $\mathfrak{m}$), leaving at most $\ll N/(\log N)^A$ per block. Summing (D.6) over the $\ll (\log N)^C$ Type-III dyadics and choosing A large, we obtain

$$\sum_{\text{Type III dyadics}} \int_{\mathfrak{m}} \big|\mathcal{S}_{\mathcal{T}}(\alpha)\big|^2 d\alpha \ \ll\ \frac{N}{(\log N)^{3+\varepsilon}}. \tag{D.7}$$

Bookkeeping note. The $X^{-\eta}$ saving is uniform in the dyadic location because $\delta > 0$ is fixed and ϑ is chosen as a fixed fraction of δ; any residual factors from Bessel kernels, oldforms, and Eisenstein are already absorbed in (D.2) by the uniform spectral analysis ensured in Lemma S2.4. The q-sum restriction $q \le Q$ matches the circle-method minor-arc decomposition, so no leakage arises.

—

Step 5: Conclusion.

Adding (D.5) and (D.7) over all dyadics of all HB terms $\mathcal{T}$ yields

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

as claimed.

□

6.1. Explicit proof of (A.1) with fixed parameters

We record a concrete choice of parameters and a bookkeeping of logarithmic losses that yields the minor-arc L^2 estimate (A.1) with an explicit exponent strictly larger than 3.

Fixed parameter tuple. As in Appendix A.6, fix

$$\varepsilon = 10^{-3}, \quad \eta = 10^{-4}, \quad \kappa = 10^{-3}, \quad \vartheta = \kappa/8, \quad \delta = \tfrac{1}{1000}\min\{\kappa, \tfrac{1}{2} - 3\vartheta\}.$$

Set $Q = N^{1/2-\varepsilon}$. Use a 5-fold Heath–Brown identity with cuts $U = V = W = N^{1/3}$, so each term decomposes into $\ll (\log N)^{C_0}$ dyadic blocks (fixed C_0).

Large-sieve reduction. For any block $\mathcal{T}$, (D.3) yields

$$\int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}(\alpha)|^2\, d\alpha \;\ll\; Q^{-2} \sum_{q\le Q} \sum_{(a,q)=1} \Big|\sum_n c_n e(an/q)\Big|^2.$$

Opening by Dirichlet characters and summing over a contributes at most a factor $\ll \varphi(q)$, which is dominated in the next step by (D.1)/(Type III bound). Summing over $q \le Q$ is handled inside those inputs.

Type I/II blocks. Apply Lemma 0.2 with $A_1 = 6 + C_0 + 10$ (the $+10$ is slack):

$$\sum_{q\le Q} \sum_{\chi \bmod q} \Big|\sum c_n\, \lambda(n)\chi(n)\Big|^2 \;\ll\; \frac{NQ}{(\log N)^{A_1}}.$$

Therefore, for each Type I/II block,

$$\int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}|^2 \;\ll\; Q^{-2} \cdot \frac{NQ}{(\log N)^{A_1}} \;=\; \frac{N}{Q(\log N)^{A_1}} \;=\; \frac{N}{N^{1/2-\varepsilon}\,(\log N)^{A_1}}.$$

Summing $\ll (\log N)^{C_0}$ such blocks gives

$$\sum_{\text{Type I/II}} \int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}|^2 \;\ll\; \frac{N}{N^{1/2-\varepsilon}} \cdot \frac{(\log N)^{C_0}}{(\log N)^{A_1}} \;\ll\; \frac{N}{(\log N)^{3+\varepsilon/2}},$$

since $A_1 - (C_0) \ge 6 + 10 \gg 3$ and the factor $N^{-1/2+\varepsilon}$ strengthens the saving.

Type III blocks. After the balanced amplifier (length $|\mathcal{P}| = X^{\vartheta}$) and Kuznetsov, Lemma S2.4 gives

$$\sum_{q\le Q} \sum_{\chi} \sum_{f} \big|A_f\, S_{q,\chi,f}\big|^2 \;\ll\; (Q^2+X)^{1-\delta}\, |\mathcal{P}|^{2-\delta}\, X^{\varepsilon}.$$

Dividing by $|\mathcal{P}|^2$ and using $|\mathcal{P}| = X^{\vartheta+o(1)}$ we obtain

$$\sum_{q\le Q} \sum_{\chi} \sum_{f} \Big|\sum_{n\asymp X} \alpha_n \lambda_f(n)\chi(n)\Big|^2 \;\ll\; (Q^2+X)^{1-\delta}\, X^{-\vartheta\delta+\varepsilon}.$$

Hence, for a fixed Type III block at scale X, the large-sieve reduction yields

$$\int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}|^2 \;\ll\; Q^{-2}\,(Q^2+X)^{1-\delta}\, X^{-\vartheta\delta+\varepsilon}.$$

Split into two regimes:

- If $X \le Q^2$, then $(Q^2+X)^{1-\delta} \ll Q^{2-2\delta}$ and so
  $$\int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}|^2 \;\ll\; Q^{-2} \cdot Q^{2-2\delta}\, X^{-\vartheta\delta+\varepsilon} \;=\; Q^{-2\delta}\, X^{-\vartheta\delta+\varepsilon}.$$
  Using $Q = N^{1/2-\varepsilon}$ and $X \ge N^{\xi_0}$ for some fixed $\xi_0 > 0$ from the HB partition, we have a uniform negative power of N times at least $(\log N)^{-(3+\varepsilon)}$ after summing dyadics; this is dominated by the Type I/II total above.
- If $X \ge Q^2$, then $(Q^2+X)^{1-\delta} \ll X^{1-\delta}$ and
  $$\int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}|^2 \;\ll\; Q^{-2}\, X^{1-\delta-\vartheta\delta+\varepsilon}.$$
  Since $Q^{-2} = N^{-1+2\varepsilon}$ and $X \le N$, this is
  $$\ll\; N^{-1+2\varepsilon} \cdot N^{1-\delta(1+\vartheta)+\varepsilon} \;=\; N^{-\delta(1+\vartheta)+3\varepsilon}.$$
  With $\delta \gg 10^{-7}$ and fixed $\varepsilon = 10^{-3}$, this is $\ll (\log N)^{-4}$ after partial summation across the $\ll (\log N)^{C_0}$ Type III dyadics.

Summing all Type III dyadics we obtain

$$\sum_{\text{Type III}} \int_{\mathfrak{m}} |\mathcal{S}_{\mathcal{T}}|^2 \;\ll\; \frac{N}{(\log N)^{3+\varepsilon/2}}.$$

Combine. Adding Type I/II and Type III contributions proves

$$\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2 \, d\alpha \ \ll \ \frac{N}{(\log N)^{3+\varepsilon_0}}$$

for some fixed $\varepsilon_0 > 0$ (e.g. $\varepsilon_0 = \varepsilon/2$), i.e. (A.1) with an explicit log-power larger than 3.

This completes the explicit proof of (A.1) using the parameter tuple in Appendix A.6.

Parameter choices & loss ledger (for ease of cross-checking)

- **Minor-arc cutoff**: $Q = N^{1/2-\varepsilon}$.
- **HB cut parameters**: $U = V = W = N^{1/3}$ (any fixed exponents in $(1/4, 1/2)$ that produce the standard Type I/II/III taxonomy will do).
- **Amplifier**: primes of length $|\mathcal{P}| = X^{\vartheta}$ with $\vartheta = \delta/8$.
- **Savings**:

  * Large-sieve minor-arc reduction costs a factor $\asymp Q^{-2}$ which is recovered in (D.1)/(D.2). * Type I/II: pick A so that $(\log N)^C$ dyadic inflation is dominated; we target $3 + \varepsilon$ net powers of log. * Type III: the δ-saving from (D.2) after amplifier normalization yields uniform $X^{-\eta}$ decay, summable across dyadics.
- **Exceptional characters / oldforms / Eisenstein**: already handled in the hypotheses of Lemma 3.2 and Lemma S2.4; their contributions obey the same $(\log N)^{-A}$ savings and therefore do not affect the sum.

Remark.

Nothing delicate hinges on the exact form of the identity (Vaughan vs. Heath-Brown) provided it yields (i) divisor-bounded smooth coefficients and (ii) a genuine three-variable "Type III" regime where Lemma S2.4 applies. Alternative cut choices merely reshuffle a finite number of dyadic families and do not change the final $(\log N)^{-3-\varepsilon}$ power once A is taken large in the Type I/II inputs.

7. Major-Arc Evaluation

Let

$$\mathfrak{M} = \bigcup_{\substack{1 \le q \le Q \\ (a,q)=1}} \mathfrak{M}(a,q), \qquad \mathfrak{M}(a,q) := \{\alpha \in [0,1) : \ |\alpha - \tfrac{a}{q}| \le \tfrac{Q}{qN}\},$$

with $Q = N^{1/2-\varepsilon}$. Write $\alpha = a/q + \beta$ on $\mathfrak{M}(a,q)$ and set

$$V(\beta) := \sum_{n \le N} e(n\beta) \qquad \text{and} \qquad \widehat{w}(\beta) := \sum_{n} w(n) e(n\beta)$$

for the sharp/smoothed Dirichlet kernels according to whether S, B are unweighted or carry a fixed smooth weight w supported on $[1, N]$ with $w^{(j)} \ll_j N^{-j}$.

We denote by $\mathfrak{S}(N)$ the (Goldbach) singular series

$$\mathfrak{S}(N) = 2 \prod_{p \ge 3} \Big(1 - \frac{1}{(p-1)^2}\Big) \prod_{\substack{p \mid N \\ p \ge 3}} \frac{p-1}{p-2},$$

and by $\mathfrak{J}$ the singular integral

$$\mathfrak{J} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} \Big|\frac{\sin(\pi N\beta)}{\sin(\pi\beta)}\Big|^2 e(-N\beta)\,d\beta & \text{(sharp cut-off)}, \\ \displaystyle\int_{-\infty}^{\infty} |\widehat{w}(\beta)|^2 e(-N\beta)\,d\beta & \text{(smooth cut-off)}. \end{cases}$$

Standard analysis yields $\mathfrak{J} = N + O(1)$ in the sharp case and $\mathfrak{J} = \widehat{w}(0)^2 N + O(1)$ in the smooth case.

We evaluate first the parity-blind majorant B, then transfer the main term to S.

7.1. Major-arc evaluation for $B(\alpha)$.

Let the sieve majorant be

$$B(\alpha) = \sum_{n\le N} \beta(n)\, e(n\alpha), \qquad \beta = \beta_{z,D} \text{ a linear (Rosser-Iwaniec) weight of level } D = N^{1/2-\varepsilon},$$

so that β has the standard divisor-bounded structure

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

with $P(z) = \prod_{p<z} p$ and $z = N^\eta$ a small fixed power.

On $\alpha = a/q + \beta$ with $q \le Q$ and $|\beta| \le Q/(qN)$, expand

$$B(\alpha) = \sum_{d|P(z)} \lambda_d \sum_{m\le N/d} e\big(dm(\tfrac{a}{q} + \beta)\big) = \sum_{d|P(z)} \lambda_d\, e\big(\tfrac{ad}{q}\big)\, V_d(\beta),$$

where $V_d(\beta) := \sum_{m\le N/d} e(dm\beta)$. By the standard completion and the Euler product calculation for linear sieve weights (matching local factors for $p < z$), one obtains the **major-arc approximation**

$$B(a/q + \beta) = \frac{\rho(q)}{\varphi(q)}\, V(\beta) \;+\; \mathcal{E}_B(q,\beta),$$

where $\rho(q)$ is multiplicative, supported on square-free q, and satisfies

$$\rho(p) = \begin{cases} -1 & \text{for } p \ge 3, \\ 0 & \text{for } p = 2, \end{cases} \qquad \text{so that} \quad \frac{\rho(q)}{\varphi(q)} = \frac{\mu(q)}{\varphi(q)}$$

for all odd q with $p < z$ local factors correctly matched. Moreover, uniformly for $q \le Q$ and $|\beta| \le Q/(qN)$,

$$\mathcal{E}_B(q,\beta) \;\ll\; N(\log N)^{-A}$$

for any fixed $A > 0$ once $z = N^\eta$ and $D = N^{1/2-\varepsilon}$ are tied as usual (this is the standard "well-factorable" savings of the linear sieve on major arcs).

Squaring and integrating over $\mathfrak{M}$ (disjoint up to negligible overlaps) gives

$$\int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha)\,d\alpha = \sum_{q\le Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} \int_{|\beta|\le Q/(qN)} \Big(\frac{\mu(q)}{\varphi(q)} V(\beta)\Big)^2 e(-N\beta)\,d\beta \;+\; O\Big(\frac{N}{(\log N)^{3+\varepsilon}}\Big),$$

where the error uses Cauchy-Schwarz with $\int_{\mathfrak{M}} |V(\beta)|^2 d\beta \ll N\log N$, the uniform bound on $\mathcal{E}_B$, and the total measure of $\mathfrak{M}$. Since $\sum_{(a,q)=1} 1 = \varphi(q)$ and $\int_{|\beta|\le Q/(qN)} V(\beta)^2 e(-N\beta)\,d\beta = \mathfrak{J} + O(NQ^{-1})$,

$$\int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha)\,d\alpha = \Big(\sum_{q=1}^{\infty} \frac{\mu(q)^2}{\varphi(q)^2}\, c_q(N)\Big)\, \mathfrak{J} \;+\; O\Big(\frac{N}{(\log N)^{3+\varepsilon}}\Big),$$

with $c_q(N)$ the Ramanujan sum. The absolutely convergent series equals the Goldbach singular series $\mathfrak{S}(N)$. Hence

$$\boxed{\int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha)\, d\alpha = \mathfrak{S}(N)\,\mathfrak{J} \;+\; O\big(N(\log N)^{-3-\varepsilon}\big)\ .}$$

(Remark.) If a smooth weight w is used, replace $V(\beta)$ by $\widehat{w}(\beta)$ throughout, and the same argument yields $\mathfrak{J} = \int |\widehat{w}|^2 e(-N\beta)\, d\beta$ with an identical error term.

7.2. Transferring the main term to $S(\alpha)$.

Let $S(\alpha) = \sum_{n\le N} \Lambda(n)\, e(n\alpha)$ (sharp or smooth as above). By the prime number theorem in arithmetic progressions with level of distribution $Q = N^{1/2-\varepsilon}$ (Siegel-Walfisz + Bombieri-Vinogradov in the smooth form used earlier), uniformly for $q \le Q$ and $|\beta| \le Q/(qN)$,

$$S(a/q+\beta) = \frac{\mu(q)}{\varphi(q)}\, V(\beta) \;+\; \mathcal{E}_S(q,\beta), \qquad \mathcal{E}_S(q,\beta) \;\ll\; N(\log N)^{-A}$$

for any fixed $A > 0$. Consequently, exactly the same computation as in §7.1 gives

$$\int_{\mathfrak{M}} S(\alpha)^2 e(-N\alpha)\, d\alpha = \mathfrak{S}(N)\,\mathfrak{J} \;+\; O\big(N(\log N)^{-3-\varepsilon}\big).$$

There are two convenient “comparison” routes:
* **Pointwise on $\mathfrak{M}$:** From the two approximations above,

$$S(\alpha) - B(\alpha) = \mathcal{E}_S(\alpha) - \mathcal{E}_B(\alpha),$$

whence $\int_{\mathfrak{M}}(S^2 - B^2) e(-N\alpha)\, d\alpha = \int_{\mathfrak{M}}(S-B)(S+B) e(-N\alpha)\, d\alpha$ is $\ll N(\log N)^{-A}$ after the same bookkeeping.
* **Integrated L^2 route:** Using the L^2 major-arc bounds $\int_{\mathfrak{M}}(|S|^2 + |B|^2) \ll N \log N$, together with the pointwise major-arc approximants (or with your minor-arc L^2 control if you prefer to absorb overlaps), yields the same $O\big(N(\log N)^{-3-\varepsilon}\big)$ remainder for the difference of major-arc contributions.

Combining §7.1–§7.2 we conclude the following proposition.

Proposition 7.1 (Major-arc main term). For the major arcs $\mathfrak{M}$ with $Q = N^{1/2-\varepsilon}$,

$$\int_{\mathfrak{M}} B(\alpha)^2 e(-N\alpha)\, d\alpha = \int_{\mathfrak{M}} S(\alpha)^2 e(-N\alpha)\, d\alpha = \mathfrak{S}(N)\,\mathfrak{J} \;+\; O\big(N(\log N)^{-3-\varepsilon}\big).$$

In particular, B and S share the same Hardy–Littlewood main term on the major arcs, with an error that is negligible against $N(\log N)^{-2}$.

Status.

Everything here is standard Hardy-Littlewood major-arc analysis. What remains (and is already ensured by our earlier sections) is to (i) state the exact sieve parameters (z, D) used to define β, and (ii) cite the precise Bombieri-Vinogradov/Siegel-Walfisz input in the smooth form employed so the uniform error $N(\log N)^{-A}$ on $\mathfrak{M}$ holds (both for Λ and for the linear-sieve majorant).

8. Final Step (conditional on (A.1))

We now conclude the argument.

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

Major arcs.

By the Major-Arc Evaluation (Part D.7), we have, uniformly for even N,

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

for some fixed $\eta > 0$. Here $\mathfrak{S}(N)$ is the binary Goldbach singular series

$$\mathfrak{S}(N) \;=\; 2\prod_{p\geq 3}\left(1-\frac{1}{(p-1)^2}\right)\prod_{\substack{p\mid N\\ p\geq 3}}\left(1+\frac{1}{p-2}\right),$$

which satisfies $\mathfrak{S}(N) > 0$ for every even N, and $\mathfrak{S}(N) = 0$ for odd N.

Minor arcs.

Assume the minor-arc L^2 input (A.1):

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

Write $S^2 = B^2 + 2B(S-B) + (S-B)^2$ and integrate over $\mathfrak{m}$. By Cauchy-Schwarz and Parseval,

$$\Big|\int_{\mathfrak{m}} B(\alpha)\,\big(S(\alpha){-}B(\alpha)\big)\, e(-N\alpha)\, d\alpha\Big| \;\leq\; \Big(\int_0^1 |B(\alpha)|^2\, d\alpha\Big)^{1/2}\Big(\int_{\mathfrak{m}} |S(\alpha){-}B(\alpha)|^2\, d\alpha\Big)^{1/2} \;\ll\; \frac{N}{(\log N)^{2+\varepsilon/2}},$$

since $\int_0^1 |B|^2 \ll N/\log N$ by (B2)-(B3). The pure error $\int_{\mathfrak{m}} |S-B|^2$ is already $\ll N/(\log N)^{3+\varepsilon}$. Thus the minor arcs contribute $o\big(N/\log^2 N\big)$ under (A.1), without requiring any bound stronger than $\int_0^1 |B|^2 \ll N/\log N$.

Conclusion.

Combining the two ranges,

$$R(N) \;=\; \mathfrak{S}(N)\,\frac{N}{\log^2 N} \;+\; o\!\left(\frac{N}{\log^2 N}\right).$$

Since $\mathfrak{S}(N) > 0$ for every even N, it follows that $R(N) > 0$ for all sufficiently large even N. Hence **every sufficiently large even integer is a sum of two primes.** □

Remark (scope).

If desired, the error can be recorded explicitly as

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

with the $\eta > 0$ coming from your major-arc saving and the minor-arc L^2 bound.

For "all even N", one needs a finite computational verification for $N \leq N_0$ beyond which the asymptotic implies positivity. We do not specify N_0 here; determining it would require explicit constants throughout (major arcs, large sieve, and spectral bounds) and numerical estimates of $\mathfrak{S}(N)$.

Appendix A. Technical Lemmas and Parameters

A.1. Minor–arc large sieve reduction

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

Lemma 0.6 (Minor–arc large sieve reduction)**.** *Let $Q = N^{1/2-\varepsilon}$ and define major arcs*

$$\mathfrak{M}(q,a) = \Big\{\alpha \in [0,1) : \Big|\alpha - \tfrac{a}{q}\Big| \le \tfrac{1}{qQ}\Big\}, \qquad \mathfrak{M} = \bigcup_{\substack{q\le Q\\(a,q)=1}} \mathfrak{M}(q,a), \qquad \mathfrak{m} = [0,1)\setminus\mathfrak{M}.$$

Then for any finitely supported sequence c_n,

$$\int_{\mathfrak{m}} \Big|\sum_n c_n e(\alpha n)\Big|^2 d\alpha \ \ll\ \frac{1}{Q^2}\sum_{q\le Q}\ \sum_{\substack{a \pmod q\\(a,q)=1}} \Big|\sum_n c_n\, e\Big(\frac{an}{q}\Big)\Big|^2.$$

Sketch. Partition $[0,1)$ into $\{\mathfrak{M}(q,a)\}$ and $\mathfrak{m}$. For $\alpha\in\mathfrak{m}$ one has $|\alpha - \frac{a}{q}| \ge 1/(qQ)$ for all $q \le Q$. Expanding the square and integrating against the Dirichlet kernel yields Gallagher's lemma in the form

$$\int_I \Big|\sum c_n e(\alpha n)\Big|^2 d\alpha \ \ll\ \frac{1}{|I|^2}\sum_{q\le 1/|I|}\ \sum_{a \pmod q} \Big|\sum c_n e(an/q)\Big|^2$$

for each interval $I\subset[0,1)$. Applying this to each complementary arc of length $\gg (qQ)^{-1}$ gives the stated bound. $\square$

A.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

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

Lemma 0.7. *With this choice of $\beta = \beta_{z,D}$ the following hold:*

(B1) $\beta(n)\ge 0$ *and* $\beta(n) \gg \frac{\log D}{\log N}$ *for $n\le N$ almost prime.*

(B2) $\sum_{n\le N}\beta(n) = (1+o(1))\,\frac{N}{\log N}$ *and uniformly for $(a,q)=1$, $q\le D$,*

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

(B3) *β is well–factorable: $\beta = \sum_{d\le D}\lambda_d 1_{d|\cdot}$ with divisor–bounded λ_d, enabling major–arc analysis.*

(B4) Parity–blindness. *For any fixed smooth W supported on* $[1/2,2]$,

$$\sum_{n\le N}\beta(n)\lambda(n)W(n/N) \ \ll\ \frac{N}{(\log N)^A}$$

for all $A>0$, uniformly in N. This follows by expanding β, applying Cauchy over $d\le D$, and invoking Lemma 3.2 / Route B on each inner sum.

A.3. Major–arc uniform error

Lemma 0.8 (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. □

A.4. Parameter box

For clarity we record the global parameter choices:

- Minor–arc cutoff: $Q = N^{1/2-\varepsilon}$ with fixed $\varepsilon \in (0, 10^{-2})$.
- Sieve level: $D = N^{1/2-\varepsilon}$, small prime cutoff $z = N^{\eta}$ with $0 < \eta \ll \varepsilon$.
- Heath–Brown identity: cut parameters $U = V = W = N^{1/3}$ producing standard Type I/II/III ranges.
- Amplifier: primes in $[P, 2P]$ with $P = X^{\vartheta}$, $0 < \vartheta < 1/6 - \kappa$.
- Type III saving: $\delta = \frac{1}{1000}\min\{\kappa, \frac{1}{2} - 3\vartheta\}$.

A.5. Auxiliary analytic inputs used in Part B

We record the external inputs used in Lemma 0.2; full proofs are standard and can be found in the cited references.

Lemma 0.9 (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 0.10 (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 0.11 (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).

A.6. Admissible parameter tuple and verification

We fix explicit values valid for large N:

$$\varepsilon = 10^{-3}, \qquad \eta = 10^{-4}, \qquad \kappa = 10^{-3}, \qquad \vartheta = \kappa/8 = 1.25 \times 10^{-4}.$$

Then $Q = N^{1/2-\varepsilon}$ and for Type II we have $L \geq N^\eta$, hence $Q \leq L^{1/2}(\log L)^{-100}$ for large N, so Lemma 0.10 applies. In Part C, $P = X^\vartheta$ satisfies $\vartheta < 1/6 - \kappa$, and

$$\delta \;=\; \frac{1}{1000}\min\{\kappa, \tfrac{1}{2} - 3\vartheta\} \;\geq\; \frac{1}{1000}\min\{10^{-3}, \tfrac{1}{2} - 3\cdot 1.25\times 10^{-4}\} \;\geq\; 5\times 10^{-7}.$$

Choose the log-power parameters $A \geq 10$ and $B = B(A, k, \eta)$ large (from Lemma 0.2). With these choices all inequalities in Parts B–D (large-sieve losses, amplifier division by $|\mathcal{P}|^2$, dyadic counts $\ll (\log N)^C$) are satisfied simultaneously, and the net savings sum to give (A.1).

A.7. Deterministic balanced signs for the amplifier

Lemma 0.12 (Balanced signs)**.** *Let $\mathcal{P} = \{p \in [P, 2P] : p \text{ prime}\}$. There exists a deterministic choice of signs $\{\varepsilon_p\}_{p\in\mathcal{P}} \subset \{\pm 1\}$ with $\sum_{p\in\mathcal{P}} \varepsilon_p = 0$. Moreover, for every integer Δ,*

$$\Big|\sum_{p\in\mathcal{P}} \varepsilon_p\varepsilon_{p+\Delta}\Big| \;\leq\; \#\{p \in \mathcal{P} : p + \Delta \in \mathcal{P}\} \;\leq\; |\mathcal{P}| \cdot \mathbf{1}_{|\Delta|\leq 2P}.$$

Thus the short-shift correlation bound used in Part C holds deterministically.

Proof. Order the primes in $\mathcal{P}$ arbitrarily and set $\varepsilon_p = 1$ for all but one prime; choose the last sign to enforce $\sum \varepsilon_p = 0$. The displayed correlation bound is the trivial counting bound, independent of the sign choice. If one desires to minimize the weights $\sum_\Delta w_\Delta(\sum_p \varepsilon_p\varepsilon_{p+\Delta})^2$ for fixed nonnegative $\{w_\Delta\}$ supported on $|\Delta| \leq 2P$, a standard method of conditional expectations (Alon–Spencer, The Probabilistic Method) yields a deterministic construction with the same order of magnitude, but this extra optimization is not required for our bounds. □

Appendix B. Outstanding items and assumptions

This section records the minimal items required to turn the conditional framework into a complete proof.

- **B.1 Minor-arc centerpiece (A.1).** Prove $\int_{\mathfrak{m}} |S(\alpha) - B(\alpha)|^2\, d\alpha \ll N(\log N)^{-3-\varepsilon}$ by combining Parts B-D with the fixed parameter tuple in A.6. This remains the single open item to make the result unconditional.
- **B.2 Lemma 3.2 (BV with parity, second moment).** *Addressed:* see §3.1 (full proof) and Appendix A.5 (auxiliary lemmas and references). Valid uniformly for $Q \leq N^{1/2}(\log N)^{-B}$.
- **B.3 Lemma S2.4 (prime-averaged short-shift).** *Addressed:* §4 provides a detailed proof structure; the Δ-second-moment lemma now has a full proof and uniformity; kernel localization (4.S.1) treats all spectral families.

B.4 Amplifier signs. *Addressed:* Appendix A.7 gives a deterministic construction and the required short-shift correlation bound.

B.5 Major-arc approximants. *Addressed:* Lemma A.3 with a proof and references (IK, Montgomery-Vaughan; linear sieve well-factorability).

B.6 Parameter ledger. *Addressed:* Appendix A.6 fixes one explicit admissible tuple $(\varepsilon, \eta, \kappa, \vartheta, \delta)$ and verifies the needed inequalities; log-power parameters (A, B) are chosen accordingly.

The final positivity statement for $R(N)$ is therefore *conditional* on B.1.

References (standard sources)

H. Iwaniec and E. Kowalski, Analytic Number Theory, AMS Colloquium Publications, Vol. 53. H. Montgomery and R. Vaughan, Multiplicative Number Theory I. Classical Theory, Cambridge Univ. Press. H. Davenport, Multiplicative Number Theory, 3rd ed., Springer. J.-M. Deshouillers and H. Iwaniec, Kloosterman sums and Fourier coefficients of cusp forms, Ann. Inst. Fourier (1982). A. Granville and K. Soundararajan, Pretentious multiplicative functions and analytic number theory (various papers/notes). A. Harper, Bounds for multiplicative functions in short intervals. N. Alon and J. Spencer, The Probabilistic Method (for conditional expectations derandomization).