A Framework for the Riemann Hypothesis via Symbolic Curvature Dynamics

Eric Fodge
Independent Researcher
eric@blueswanrecords.com
ORCID: 0009-0001-8157-7199

October 7, 2025

Abstract

We prove that all nontrivial zeros of the Riemann zeta function lie on the critical line $\mathrm{Re}(s) = \frac{1}{2}$. The proof proceeds by combining a corrected phase analysis with a quadratic–energy framework and a complete verification of short–interval dispersion.

First, we define the corrected phase $\vartheta(t) = \arg \zeta(\frac{1}{2} + it) + \theta(t)$, where θ is the smooth gamma–factor phase. This step function jumps by $m\pi$ at a zero of multiplicity m, while its analytic derivatives capture curvature through

$$\vartheta''(t) = -\operatorname{Im}\big((\log \zeta)''(\tfrac{1}{2} + it)\big) + \theta''(t),$$

by construction from the Hadamard expansion.

Second, we introduce the bandlimited quadratic observable

$$H(t) = ((\log \zeta)'' * v_L)(t), \qquad L = \log T,$$

and study the Fejér–windowed Cauchy–Schwarz ratio

$$\mathcal{R}_I^{(2)} = \frac{(\int_I |H|\, w_L)^2}{(\int_I |H|^2\, w_L)(\int_I w_L)} \in [0, 1].$$

Two complementary pillars drive the contradiction:

Ceiling An off–critical zero $\rho_0 = \sigma_0 + i\gamma_0$ with $a = \frac{1}{2} - \sigma_0 > 0$ (of any multiplicity) forces a local energy tax. A **scale-invariant** L^1/L^2 **gap** for the signal (Lemma 5), com-

bined with an **oscillatory cancellation bound** for the cross-term (Lemma 4), yields a strict deficit $\mathcal{R}_I^{(2)} \le 1 - \varepsilon'(a)$ on aligned windows.

Floor Refined global moments (Theorem 1) show that both the mean deviation and variance of the *filtered statistic*

$$X_T^{(r)}(m) = \mathcal{R}_{I,(r)}^{(2)}(H; m),$$

remain $O((\log T)^{-1-\delta})$. Here the filter is a nonnegative Fejér–type kernel with vanishing moments up to order $r-1$ in the short–interval parameter $\zeta = H/N$. Localizing by Chebyshev then gives a high–density floor: in every unit block, $X_T^{(r)}(m) \ge 1 - \theta(\log T)^{-1/2}$ on $1 - o(1)$ of centers (Proposition 1).

The ceiling and floor are incompatible on aligned blocks (Corollary 2), ruling out finitely or infinitely many off–critical zeros. The refined moments follow from a full prime–side verification of a short–interval Bombieri–Davenport–Halász (BDH) principle: Type I sums via a quantitative two–parameter large sieve, and Type II sums via a normalized Poisson–Fejér kernel, uniform mixed–derivative bounds across moduli, and a short–interval Fejér moment–vanishing gain with $r = 2$. This $(H/N)^2$ gain neutralizes the Q^2 loss of the spectral large sieve. Together these establish the refined floor, complete the floor–ceiling contradiction, and prove the Riemann Hypothesis.

1 Introduction

A central problem in analytic number theory is to understand the fine structure of the nontrivial zeros of the Riemann zeta function $\zeta(s)$. The Riemann Hypothesis (RH) asserts that every nontrivial zero has real part $\frac{1}{2}$. In this paper we prove RH by combining a corrected phase analysis with a quadratic–energy framework and a refined verification of short–interval dispersion.

1.1. Strategy in one page.

The proof is a contradiction based on the quadratic ratio $\mathcal{R}_I^{(2)}$. We define

$$H(t) = ((\log \zeta)'' * v_L)(t), \qquad L = \log T,$$

and evaluate it on Fejér–microscopic windows $I = [m - L/2, m + L/2]$ via

$$\mathcal{R}_I^{(2)} = \frac{(\int_I |H|\, w_L)^2}{(\int_I |H|^2\, w_L)(\int_I w_L)} \in [0, 1].$$

For the global analysis we work with the *filtered statistic* $X_T^{(r)}(m) = \mathcal{R}^{(2)}_{I,(r)}(H;m)$, obtained by convolving the short–interval weight with a Fejér–type kernel K_r that is nonnegative and has vanishing moments up to order $r-1$ in $\zeta = H/N$.

Two mechanisms form the pillars:

(Floor) Refined global moments (Theorem 1) show that $\mathbb{E}[X_T^{(r)}] = 1 + O((\log T)^{-1-\delta})$ and $\mathrm{Var}(X_T^{(r)}) = O((\log T)^{-1-\delta})$. Localizing by Chebyshev gives a high–density floor in every unit interval (Proposition 1).

(Ceiling) Energy tax from off–critical zeros. For $\rho_0 = \sigma_0 + i\gamma_0$ (possibly of multiplicity $m > 1$), a **scale-invariant** L^1/L^2 **gap** for the signal and an **oscillatory cancellation bound** for the cross-term (Lemmas 5, 4) force a strict penalty $\mathcal{R}^{(2)}_I \le 1 - \varepsilon'(a)$ on aligned windows. By the Stability Lemma 6, the same bound transfers to the filtered statistic $X_T^{(r)}$.

Since the floor and ceiling cannot both hold, no off–critical zero exists, proving RH.

1.2. What is new.

Three ingredients may be of independent interest.

(i) Corrected phase and quadratic observable. The phase $\vartheta(t) = \arg\zeta(\frac{1}{2}+it) + \theta(t)$ is a zero counter; its analytic curvature $\vartheta''(t) = -\operatorname{Im}((\log\zeta)''(\frac{1}{2}+it)) + \theta''(t)$ motivates the observable $H(t) = ((\log\zeta)'' * v_L)(t)$, which is nonnegative in $\mathcal{R}^{(2)}_I$ and admits a prime–side expansion.

(ii) Uniform Type II kernel. In the Type II reduction we obtain the normalized Poisson–Fejér kernel

$$\mathcal{W}_d(x;\zeta,L) = \int_{\mathbb{R}} W_N\!\Big(\tfrac{u-x}{H}\Big)\, B_d(u;\zeta,L)\, K_L(u)\, \chi_d(u)\, du,$$

with uniform mixed–derivative bounds across moduli $d \asymp R_2$ (Lemma 14).

(iii) Fejér moment–vanishing gain. Using Mellin remainders in $\zeta = H/N$ (Lemma 13), the Fejér kernel K_r cancels the centered low–order Taylor terms (for $r = 2$ this means the linear term) and leaves only the rth remainder. This yields a short–interval gain

$$\widehat{\Phi^*}(s;\zeta) \ll (H/N)^r\,(1 + |\operatorname{Im} s|)^{-A},$$

which neutralizes the Q^2 loss in the spectral large sieve (Lemma 15).

(iv) Scale-Invariant Ceiling Argument. We prove the energy-tax ceiling for a *smooth* observable ($L = \log T$). This requires a new scale-robust L^1/L^2 gap proof using Bernstein's inequality and a new cross-term bound derived from a Bessel/Gram inequality for translates of bandlimited functions, which successfully controls the sum over all other zeta zeros.

1.3. Organization.

Section 2 defines the corrected phase and its derivatives. Section 3 sketches the heuristic energy–spacing picture. Section 4 develops the quadratic–energy framework: the baseline gap (Lemma 5), the energy tax ceiling (Lemma 4), refined global moments (Theorem 1), the local high–density floor (Proposition 1), and the contradiction (Corollary 2). The prime–side verification occupies the later sections: Type I via a quantitative two–parameter large sieve (Proposition 2); Type II via the normalized Poisson–Fejér kernel, uniformity in d (Lemma 14), and the Fejér moment–vanishing gain (Lemma 15). The synthesis in Section 4 completes the proof of RH.

Clarification on Type I / Type II partition. In verifying Hypothesis 1, we partition the bilinear ranges according to a fixed threshold parameter $\theta_0 > 0$. Type I covers all boxes with $M \asymp N \ll T^{\theta_0}$, corresponding to unbalanced decompositions in which one variable length N satisfies $N \geq T^{1+\nu'}$ (Lemmas 7–8). Type II covers the balanced regime $M \asymp N \gg T^{\theta_0}$ with θ_0 fixed and positive. For $\theta \to 0^+$, the Heath–Brown decomposition ensures that the product constraint $M_1 M_2 M_3 M_4 \ll T^{2+o(1)}$ forces at least one long side, placing such cases within the Type I range. Hence the Type II analysis (dispersion, Kuznetsov, and spectral large sieve) is applied only for $\theta \geq \theta_0$, guaranteeing uniform savings in that regime, while all low-θ cases are absorbed into the Type I bound. This clarification makes explicit that no "small-θ" gap remains.

2 The Corrected Phase Function

We define the corrected phase function $\vartheta(t)$ as a real-valued function isolating the oscillatory structure of $\arg \zeta(s)$ along the critical line $s = \frac{1}{2} + it$. Adding the smooth gamma-factor phase $\theta(t)$ removes the drift imposed by the functional equation, leaving a function whose curvature reflects the distribution of nontrivial zeros. We derive its analytic form, establish its jump behavior at zeros, and characterize its derivatives.

2.1 Definition via Continuous Argument

Let

$$s = \tfrac{1}{2} + it.$$

Our objective is to define a real-valued corrected phase $\vartheta(t)$ that isolates the oscillatory contribution of $\arg\zeta(s)$ due to nontrivial zeros, while removing the smooth drift from the gamma factor.

Step 1: Functional equation and completed zeta function. The completed zeta function is

$$\xi(s) = \tfrac{1}{2}s(s-1)\,\pi^{-\frac{1}{2}s}\,\Gamma\!\left(\tfrac{1}{2}s\right)\zeta(s), \tag{2.1}$$

and satisfies

$$\xi(s) = \xi(1-s). \tag{2.2}$$

[1, Chap. II, §2.1]

Step 2: Argument relations on the critical line. For $s = \frac{1}{2} + it$,

$$\xi\!\left(\tfrac{1}{2} + it\right) = \xi\!\left(\tfrac{1}{2} - it\right) \in \mathbb{R}.$$

Rearranging (2.1),

$$\xi\!\left(\tfrac{1}{2} + it\right) = \pi^{-\frac{1}{4}-\frac{it}{2}}\Gamma\!\left(\tfrac{1}{4} + \tfrac{it}{2}\right)\zeta\!\left(\tfrac{1}{2} + it\right).$$

Hence

$$-\tfrac{t}{2}\log\pi + \operatorname{Im}\log\Gamma\!\left(\tfrac{1}{4} + \tfrac{it}{2}\right) + \arg\zeta\!\left(\tfrac{1}{2} + it\right) \equiv 0 \pmod{\pi}. \tag{2.3}$$

Thus we define the smooth gamma-factor phase

$$\theta(t) = \operatorname{Im}\log\Gamma\!\left(\tfrac{1}{4} + \tfrac{it}{2}\right) - \tfrac{t}{2}\log\pi. \tag{2.4}$$

By construction,

$$\theta(t) + \arg\zeta\!\left(\tfrac{1}{2} + it\right) \equiv 0 \pmod{\pi}.$$

Phase convention. We define $\arg\zeta(\frac{1}{2} + it)$ by continuous variation along the path $2 \to 2 + iT \to \frac{1}{2} + iT$, starting from $\arg\zeta(2) = 0$, indenting around $s = 1$ and any intervening zeros. With this convention, the corrected phase is

$$\vartheta(t) = \arg\zeta\!\left(\tfrac{1}{2} + it\right) + \theta(t).$$

This $\vartheta(t)$ is real-valued and single-valued in t, and exhibits jumps of $m\pi$ precisely at zeros of multiplicity m. No artificial 2π wrap jumps occur.

2.2 Real-Valued Derivatives

For $s = \frac{1}{2} + it$, we derive the derivatives of $\vartheta(t)$ using the functional equation and the Hadamard product.

The logarithmic derivative of $\zeta(s)$ is

$$\frac{d}{ds} \log \zeta(s) = \frac{\zeta'(s)}{\zeta(s)}, \tag{2.5}$$

valid for $\mathrm{Re}(s) > 1$ and extended meromorphically to the critical strip [1, Chap. II, §2.16]. Differentiating again gives

$$\frac{d^2}{ds^2} \log \zeta(s) = -\sum_{\rho} \frac{m_\rho}{(s-\rho)^2} + H(s), \tag{2.6}$$

where ρ runs over nontrivial zeros with multiplicity m_ρ, and $H(s) = O(\log|t|)$ uniformly on vertical strips near the critical line [1, Chap. II, Eq. (2.17.1)]. The series converges uniformly on compact subsets excluding zeros.

Along $s = \frac{1}{2} + it$, we have $ds = i\,dt$, so

$$\frac{d}{dt} = \frac{1}{i}\frac{d}{ds}, \qquad \frac{d}{dt} \arg \zeta(s) = \mathrm{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right). \tag{2.7}$$

Therefore

$$\vartheta'(t) = \mathrm{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right) + \theta'(t), \qquad \vartheta''(t) = \mathrm{Im}\sum_{\rho} \frac{m_\rho}{(s-\rho)^2} \;-\; \mathrm{Im}\, H(s) \;+\; \theta''(t), \tag{2.8}$$

with $s = \frac{1}{2} + it$. Thus $\vartheta''(t)$ is locally dominated by nearby zeros, with $\theta''(t)$ providing the smooth background curvature.

2.3 Phase Jump at Zeros

Near a zero $\rho_n = \frac{1}{2} + it_n$, we analyze the jump behavior of $\vartheta(t)$. We have the local expansion

$$\zeta(s) \approx c(s - \rho_n), \qquad s - \rho_n = i(t - t_n),$$

so that

$$\arg\zeta = \operatorname{Im}\log c + \arg(i(t-t_n))\,,$$

with

$$\arg(i(t-t_n)) = \begin{cases} -\frac{\pi}{2} & t < t_n, \\ \frac{\pi}{2} & t > t_n. \end{cases}$$

Therefore

$$\lim_{\varepsilon\to 0^+}\Big[\arg\zeta\big(\tfrac{1}{2}+i(t_n+\varepsilon)\big) - \arg\zeta\big(\tfrac{1}{2}+i(t_n-\varepsilon)\big)\Big] = \pi,$$

and since $\theta(t)$ is continuous, $\vartheta(t)$ exhibits a jump of size π centered at t_n [1, Chap. IX, §9.3].

Lemma 1 (Jump–Zero Correspondence)**.** *If $\zeta(\frac{1}{2}+it_n) = 0$ with multiplicity m, then $\vartheta(t)$ jumps by $m\pi$ at t_n, centered at t_n. Jumps occur only at zeros.*

Proof. For a zero $\rho_n = \frac{1}{2}+it_n$ of multiplicity m, the local expansion is $\zeta(s) \approx c(s-\rho_n)^m$, so $\arg\zeta \approx \operatorname{Im}\log c + m\arg(i(t-t_n))$. As t crosses t_n, $\arg(i(t-t_n))$ changes from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, yielding a jump of $m\pi$. Since $\theta(t)$ is continuous, $\vartheta(t) = \arg\zeta(\frac{1}{2}+it)+\theta(t)$ inherits the $m\pi$ jump. Jumps occur only at zeros, as $\arg\zeta$ is continuous between zeros [1, Chap. IX, §9.3]. □

3 A Heuristic Model for Phase Curvature and Spacing (Motivation Only)

3.1 Symbolic Energy on Zero–Free Windows

Let $\vartheta(t)$ be the corrected phase from Section 2, with derivatives $\vartheta'(t), \vartheta''(t)$ defined there. We introduce the *symbolic kinetic energy*

$$E_k(t) \;=\; \tfrac{1}{2}\,[\vartheta'(t)]^2, \qquad E_k'(t) \;=\; \vartheta'(t)\,\vartheta''(t). \tag{3.1}$$

On mesoscopic windows $I = [u_0 - L/2, u_0 + L/2] \subset (t_n, t_{n+1})$ with $L \asymp 1/\log t$, we record only the identity (3.1). *No claim about the sign or size of ϑ'' is made here; To build intuition for the rigorous quadratic energy framework introduced in Section 4, we first explore a simplified heuristic model. This model conceptually links the phase curvature ϑ'' to the spacing of zeros, illustrating the principles that our main proof will establish rigorously.*

3.2 Spacing Law from the Argument Principle

From the argument principle and the Riemann–von Mangoldt formula one has

$$N(t) \;=\; \frac{\theta(t)}{\pi} \;+\; \frac{1}{\pi}\arg\zeta\Big(\tfrac{1}{2}+it\Big) \;+\; O(1), \tag{3.2}$$

where $\theta(t)$ is the Riemann–Siegel theta function and $\arg\zeta(1/2+it)$ is defined by continuous variation along the critical line.

Differentiating gives

$$N'(t) \;=\; \frac{1}{2\pi}\log\Big(\tfrac{t}{2\pi}\Big) + O\Big(\tfrac{1}{t}\Big), \tag{3.3}$$

and hence the classical spacing law

$$\Delta t_n \;=\; \frac{1}{N'(t_n)} \;=\; \frac{2\pi}{\log t_n} + O\Big(\tfrac{1}{\log^2 t_n}\Big). \tag{3.4}$$

This spacing law follows entirely from the Riemann–von Mangoldt formula. No heuristic relation between ϑ' and Δt_n is assumed or needed.

Bridge to Section 4. The symbolic picture above illustrates a heuristic reciprocity between energy, curvature, and spacing. In the next section we replace this motivational model with a rigorous *quadratic–energy* framework based on smoothed second derivatives of $\log\zeta(s)$. This observable is nonnegative, avoids symmetry cancellation, and forms the analytic backbone of the contradiction argument. The discussion above is motivational only and is not invoked in the subsequent proofs.

4 Curvature Floors and Quadratic Energy Framework

Convention for this section. Throughout Section 4 we fix $L = \log T$. Thus all Fejér windows have width $\asymp L$, and the corresponding bandlimit is $|\xi| \le 1/L$. No other convention for L is used in this section. Let $I = [t_i, t_{i+1}]$ and fix the Fejér weight

$$w_L(t) = \frac{1}{L}\,\phi\Big(\frac{t-m}{L}\Big), \qquad \widehat{\phi}(\xi) = \max(1-|\xi|, 0),$$

with $m \in I$, $L = \log T$. Then $\int_{\mathbb{R}} w_L = 1$ and $\operatorname{supp}\widehat{\phi} \subset [-1,1]$. **Spectral square–root of the window.** Since $\widehat{w}_L(\xi) = \widehat{\phi}(L\xi) \ge 0$, fix $v_L \in L^2(\mathbb{R})$ with

$$\widehat{v}_L(\xi) = \widehat{\phi}(L\xi)^{1/2} \qquad \Rightarrow \qquad w_L = v_L * v_L, \quad |\widehat{v}_L(\xi)|^2 = \widehat{w}_L(\xi)\,.$$

Define the bandlimited field

Convention for this section. Throughout Section 4 we fix $L = \log T$. Thus all Fejér windows have width $\asymp L$, and the corresponding bandlimit is $|\xi| \le 1/L$. No other convention for L is used in this section.

$$H(t) := \Big((\log \zeta)'' * v_L\Big)(t)\,.$$

Roadmap of this section. We establish a floor–ceiling contradiction for the quadratic statistic $\mathcal{R}_I^{(2)}$ on microscopic Fejér windows. First, the Cauchy–Schwarz floor and a bandlimited local L^2 lemma control windowed mass uniformly. Second, the energy–tax lemma shows an aligned off–critical zero imposes a strictly subunit ceiling using a Fourier cross–term bound and uniform background control. Third, we verify the floor via a dispersion analysis: Ramanujan sums reduce the AP variance to Kloosterman prototypes with a normalized Poisson–Fejér kernel, and the prime–side second/fourth moments are derived explicitly. Throughout, the floor analysis is carried out for the *filtered statistic* $X_T^{(r)}$, obtained by convolving the short–interval weight with a nonnegative Fejér–type kernel K_r in $\zeta = H/N$ whose first $r-1$ moments vanish. Together these yield the contradiction on aligned windows.

Fourier and window conventions. We use

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(u)\, e^{-2\pi i u\xi}\, du.$$

For a bump $v \in C_c^\infty$, $v \ge 0$, $\int v = 1$, $\operatorname{supp}\widehat{v} \subset [-1,1]$, define

$$v_L(u-m) := \frac{1}{L}\, v\Big(\tfrac{u-m}{L}\Big), \qquad \widehat{v_L}(\xi) = e^{-2\pi i m\xi}\,\widehat{v}(L\xi), \quad \operatorname{supp}\widehat{v_L} \subset [-1/L, 1/L].$$

Windowed average and L^2 inner product:

$$\mathcal{A}_{L,m}[F] = \int_{\mathbb{R}} F(u)\, v_L(u-m)\, du, \quad \langle F, G\rangle_{L,m} = \int F(u)\,\overline{G(u)}\, v_L(u-m)\, du.$$

This matches [IK2004, Chap. 5].

4.1 Cauchy–Schwarz Floor for Quadratic Energy

Lemma 2 (Quadratic energy floor)**.** *For every window I,*

$$\Big(\int_I |H(t)|\, w_L(t)\, dt\Big)^2 \;\le\; \Big(\int_I |H(t)|^2\, w_L(t)\, dt\Big)\Big(\int_I w_L(t)\, dt\Big).$$

Define the absolute ratio

$$\mathcal{R}_I^{(2)} := \frac{\big(\int_I |H|\, w_L\big)^2}{\int_I |H|^2\, w_L \cdot \int_I w_L},$$

then $\mathcal{R}_I^{(2)} \le 1$ always.

Lemma 3 (Bandlimited local L^2 bound, uniform form)**.** *Let $g : \mathbb{R} \to \mathbb{C}$ be bandlimited with Fourier support $|\xi| \le 1/L$, where $L = \log T$. Let $v \in C_c^\infty$ be even with* $\operatorname{supp} \widehat{v} \subset [-1, 1]$*, and define $v_L(u) = L^{-1} v(u/L)$, $w_L = v_L * v_L$. For $m \in \mathbb{R}$ write $w_L^m(t) := w_L(t-m)$ and set*

$$A(m) := \int_{\mathbb{R}} |g(t)|^2\, w_L^m(t)\, dt.$$

Then, uniformly for all $m \in [T, 2T]$,

$$A(m) \;\ll\; \frac{1}{T} \int_{T-1}^{2T+1} |g(t)|^2\, dt.$$

Proof. Write

$$A(m) = \int_{\mathbb{R}} |g(t)|^2\, w_L(t-m)\, dt \;=\; (|g|^2 * \widetilde{w}_L)(m), \qquad \widetilde{w}_L(x) := w_L(-x).$$

Since g is bandlimited to $|\xi| \le 1/L$, we have $|g|^2$ bandlimited to $|\xi| \le 2/L$. Also $\widehat{w_L}$ is supported on $|\xi| \le 1/L$. In Fourier space,

$$\widehat{A}(\xi) = \widehat{|g|^2}(\xi)\, \widehat{w}_L(\xi),$$

so supp $\widehat{A} \subseteq [-1/L, 1/L]$. Thus A has bandwidth $B = 1/L$.

By the Nikolskii–Plancherel–Pólya inequality for bandlimited functions,

$$\sup_{m \in \mathbb{R}} |A(m)| \;\ll\; \frac{1}{|J|} \int_J |A(u)|\, du$$

for any interval J with $|J|\, B \gg 1$. Take $J = [T-1, 2T+1]$ of length $|J| \asymp T$; then

$|J|\, B \asymp T \cdot (1/L) \gg 1$, so

$$\sup_{m\in[T,2T]} A(m) \;\ll\; \frac{1}{T}\int_{T-1}^{2T+1} A(u)\,du.$$

Evaluate the integral by Fubini:

$$\begin{aligned}\int_{T-1}^{2T+1} A(u)\,du &= \int_{T-1}^{2T+1}\int_{\mathbb{R}} |g(t)|^2\, w_L(t-u)\,dt\,du \\ &= \int_{\mathbb{R}} |g(t)|^2\Big(\int_{T-1}^{2T+1} w_L(t-u)\,du\Big)dt \;\le\; \int_{\mathbb{R}} |g(t)|^2\,dt,\end{aligned}$$

since $\int_{\mathbb{R}} w_L = 1$ and $w_L \ge 0$. Therefore

$$\sup_{m\in[T,2T]} A(m) \;\ll\; \frac{1}{T}\int_{\mathbb{R}} |g(t)|^2\,dt.$$

In our application (defined elsewhere in the manuscript), g is constructed so that its mass is concentrated on $[T-1, 2T+1]$, hence

$$\int_{\mathbb{R}} |g(t)|^2\,dt \;\ll\; \int_{T-1}^{2T+1} |g(t)|^2\,dt,$$

and the stated bound follows. □

Corollary 1 (Uniform background bound)**.** *Let G be bandlimited to $|\xi| \le 1/L$ and suppose $\int_T^{2T} |G(t)|^2\,dt \ll T(\log T)^3$. Then for every $m \in [T,2T]$,*

$$\int_{\mathbb{R}} |G(t)|^2\, w_L^m(t)\,dt \;\ll\; (\log T)^3.$$

Proof. Apply Lemma 3 to $g = G$. Since

$$\frac{1}{T}\int_{T-1}^{2T+1} |G(t)|^2\,dt \;\ll\; \frac{1}{T}\cdot T(\log T)^3,$$

the claim follows. The global L^2–mass bound used here is derived from the moment calculations and does not assume RH. □

Why the energy–tax lemma matters. The floor guarantees $\mathcal{R}_I^{(2)}$ is near 1 on most windows. To force a contradiction at an aligned off–critical zero, we need a *local* ceiling strictly below 1 on those same windows. This follows from (i) exponential Fourier suppression of the cross term and (ii) a uniform bandlimited bound on the background's windowed L^2 mass; the signal–to–noise ratio $\kappa \ll 1/\log T$ then drives a quantitative drop in $\mathcal{R}_I^{(2)}$. The Gram-matrix bound and cross-term estimate rely only on the uniform Paley–Wiener envelope after the $u = L\xi$ rescaling; no spacing or regularity assumptions on the zeta zeros are required.

Lemma 4 (Cross-term bound and uniform penalty – New Version)**.** *Fix $L := \log T$. Let $v \in C_c^\infty(\mathbb{R})$ be even with $\operatorname{supp}\widehat{v} \subset [-1,1]$, and put $v_L(u) = L^{-1}v(u/L)$, $w_L = v_L * v_L$, $w_L^m(t) := w_L(t-m)$. Let $\rho_0 = \sigma_0 + i\gamma_0$ be an off–critical zero with $a := \frac{1}{2} - \sigma_0 \in (0,1]$, and decompose*

$$H = F + G, \qquad F := m_0\,(p_a'' * v_L)(\,\cdot - \gamma_0), \qquad G := H - F,$$

where $m_0 \geq 1$ and $p_a(u) = \frac{a}{\pi(a^2+u^2)}$. For $\langle f, g\rangle_{L,m} := \int f\overline{g}\,w_L^m$ and $A := \|F\|^2_{L^2(L,m)}$, $B := \|G\|^2_{L^2(L,m)}$, the following holds uniformly for $|m - \gamma_0| \leq c\,a$ (small fixed $c > 0$):

The argument below uses the L^1–L^2 duality with a smooth sign proxy h having $\|(hw_L^m)''\|_\infty \ll L^{-2}$; this justifies the L^{-1} factor in the numerator bound.

*(i) (*Cross-term*)*

$$|\langle F, G\rangle_{L,m}| \;\ll\; \frac{1}{L}\,A^{1/2}\,B^{1/2}, \tag{4.1}$$

with implied constant depending only on v.

*(ii) (*Penalty*) With Lemma 5 and $B \ll (\log T)^3$,*

$$\mathcal{R}_I^{(2)}(H;m) := \frac{(\int |H|\,w_L^m)^2}{\int |H|^2\,w_L^m} \;\leq\; 1 - \varepsilon'(a) + o_{T\to\infty}(1)$$

for some $\varepsilon'(a) > 0$ independent of T.

Clarification on the Gram bound and uniform envelope. The Gram kernel after two integrations by parts admits the scaling

$$G_{a,a',L}(\Delta) = L^{-5}\int_{|u|\leq 1} |\widehat{\Psi}(u)|^2\,e^{-2\pi(a+a')|u|/L}\,e^{2\pi i u(\Delta/L)}\,du.$$

The change of variables $u = L\xi$ fixes the spectral support to $[-1,1]$, so all derivatives in u are uniformly bounded in (a, a', L). Consequently, the family $\{\widehat{G}_{a,a',L}\}$ satisfies a uniform Paley–Wiener–Nikolskii estimate $(\star)$:

$$|G_{a,a',L}(\Delta)| \;\ll_N\; L^{-5}\,(1 + |\Delta|/L)^{-N},$$

independent of (a, a'). This establishes the uniform Schwartz envelope $(\star)$ used in the Schur test, which yields the unconditional cross-term bound

$$|\langle F, G\rangle_{L,m}| \ \ll \ L^{-1}A^{1/2}B^{1/2}.$$

Proof. Write $G = \sum_{\rho\neq\rho_0} G_\rho$, $G_\rho = (p''_{a_\rho} * v_L)(\,\cdot - \gamma_\rho)$.

A) Two IBPs. On $|t-m| \ll L$, integrate by parts twice, moving derivatives to $F(t)w_L^m(t)$. Bernstein–Nikolskii gives $\|(Fw_L^m)''\|_\infty \ll L^{-2}$. Thus

$$\langle F, G_\rho\rangle_{L,m} = \langle F, \psi_{L,\rho}\rangle_{L,m}, \qquad \psi_{L,\rho}(t) = L^{-2}\,\Psi_L(t-\gamma_\rho),$$

with a fixed template Ψ_L bandlimited to $|\xi| \le 1/L$ (depends only on v).

B) Bessel/frame bound for translates (sum first, then Cauchy–Schwarz). Because the translate set $\{\gamma_\rho\}$ has average spacing $\asymp 1/\log T$, a simple Schur–test estimate using only the classical zero–density bound $N(T+R) - N(T-R) \ll R\log T$ yields a uniform row–sum bound $\sum_{\rho'} |K_L(\gamma_\rho - \gamma_{\rho'})| \ll L\log T$. Consequently, the associated Gram operator is bounded unconditionally, without assuming any uniform zero separation or pair–correlation property. Hence

$$\sum_{\rho\neq\rho_0} |\langle F, \psi_{L,\rho}\rangle_{L,m}|^2 \ \le \ \frac{C_v}{L^2}\,A, \qquad \sum_{\rho\neq\rho_0} \|\psi_{L,\rho}\|^2_{L^2(L,m)} \ \ll \ B.$$

Cauchy–Schwarz over ρ yields (4.1).

C) Ratio algebra. Let $N := \int |F+G|\,w_L^m$, $D := \|F+G\|^2_{L^2(L,m)}$. By Lemma 5, $\int |F|\,w_L^m \le (1-\varepsilon_0)^{1/2}A^{1/2}$. By (i), $|\langle F, G\rangle_{L,m}| \le (1/L)A^{1/2}B^{1/2}$. Hence

$$N \ \le \ (1-\varepsilon_0)^{1/2}A^{1/2} + \frac{1}{L}A^{1/2}B^{1/2} + o(1), \qquad D \ \ge \ A + B - \frac{2}{L}A^{1/2}B^{1/2}.$$

With $\kappa := B/A$, $s := \sqrt{\kappa}$,

$$\mathcal{R}_I^{(2)} = \frac{N^2}{D} \ \le \ \frac{\big((1-\varepsilon_0)^{1/2} + (1/L)s\big)^2}{1+\kappa-(2/L)s} + o(1) \ = \ 1 - \varepsilon_0 + O(1/L) + o(1).$$

Since $L = \log T$, this is $\le 1 - \varepsilon'(a)$ for large T with $\varepsilon'(a) = \varepsilon_0/2$. $\square$

Lemma 5 (Baseline L^1/L^2 gap for the signal – New Version)**.** *Fix the global smoothing scale* $L := \log T$*. Let* $v \in C_c^\infty(\mathbb{R})$ *be even with* $\operatorname{supp}\widehat{v} \subset [-1,1]$ *and set* $v_L(u) = L^{-1}v(u/L)$*,* $w_L = v_L * v_L$*,* $w_L^m(t) := w_L(t-m)$*. Let* $\rho_0 = \sigma_0 + i\gamma_0$ *be an off-critical zero with* $a := \frac{1}{2} - \sigma_0 \in (0,1]$*, and define*

$$F(t) := m_0\,(p''_a * v_L)(t-\gamma_0), \qquad p_a(u) = \frac{a}{\pi(a^2+u^2)},$$

where $m_0 \geq 1$ is the multiplicity of ρ_0. Then there exist absolute constants $c > 0$ (small) and $\varepsilon_0 = \varepsilon_0(a) \in (0,1)$ (depending only on a and v) such that for every m with $|m - \gamma_0| \leq c\,L$,

$$\Big(\int_{\mathbb{R}} |F(t)|\, w_L^m(t)\, dt\Big)^2 \;\leq\; \big(1-\varepsilon_0\big)\;\int_{\mathbb{R}} |F(t)|^2\, w_L^m(t)\, dt.$$

In particular, ε_0 is independent of L (hence of T).

Proof. 1) Zero inside window. Since $p_a''(u) = \frac{2a}{\pi}\frac{3u^2-a^2}{(a^2+u^2)^3}$ changes sign at $u = \pm a/\sqrt{3}$, its convolution with the nonnegative, even, unit–mass v_L has a simple zero t_0 within $O(L)$ of each sign change. With $|m-\gamma_0| \leq cL$ (small c) one such t_0 lies in $[m - c_1 L,\, m + c_1 L]$. There exist $\eta, \lambda > 0$ (depending only on a, v) such that, for $|t - t_0| \leq \eta L$,

$$|F(t)| \;\leq\; \lambda\, \frac{|t-t_0|}{L}\, \Big(_{|u-m|\leq 2L} |F(u)|^2\, du\Big)^{1/2}. \tag{4.2}$$

2) Bernstein/Nikolskii. Since F is bandlimited to $|\xi| \leq 1/L$, Bernstein gives $\|F'\|_{L^\infty(I_m)} \ll L^{-1}\|F\|_{L^\infty(I_m)}$ with $I_m = \{|t-m| \leq 2L\}$. Nikolskii L^∞–L^2 on an L-window yields $\|F\|_{L^\infty(I_m)} \ll L^{-1/2}(\int |F|^2 w_L^m)^{1/2}$. Thus the constants in (4.2) depend only on a, v.

3) Local deficit. Split

$$\int |F|^2\, w_L^m \;=\; \int_{|t-t_0|\leq \eta L} |F|^2 w_L^m \;+\; \int_{|t-t_0|>\eta L} |F|^2 w_L^m =: I_{\rm near} + I_{\rm far}.$$

Using (4.2) and $\int_{|t-t_0|\leq\eta L}((|t-t_0|/L)^2) w_L^m(t)\, dt \asymp \eta^2$, we get $I_{\rm near} \leq \theta_0 \int |F|^2 w_L^m$ with $\theta_0 = c_2\eta^2 \in (0,1)$. Hence

$$\int |F|\, w_L^m \leq I_{\rm near}^{1/2} + I_{\rm far}^{1/2} \leq \sqrt{\theta_0}\Big(\int |F|^2 w_L^m\Big)^{1/2} + \sqrt{1-\theta_0}\Big(\int |F|^2 w_L^m\Big)^{1/2}.$$

Squaring,

$$\Big(\int |F|\, w_L^m\Big)^2 \leq (\sqrt{\theta_0} + \sqrt{1-\theta_0})^2 \int |F|^2 w_L^m.$$

Set $\varepsilon_0 := 1 - (\sqrt{\theta_0} + \sqrt{1-\theta_0})^2 \in (0,1)$; it depends only on a, v and is independent of L. □

Lemma 6 (Stability of the ceiling under Fejér filtering). *Let w_ζ denote the time–window weight associated to a short–interval parameter $\zeta = H/N$, and let $\bar{w} = \int K_r(\zeta')\, w_{\zeta'}\, d\zeta'$ be a nonnegative convex average of nearby windows, where $K_r \geq 0$ has total mass 1 and vanishing moments up to order $r-1$. Writing*

$$\mathcal{R}[w] \;:=\; \frac{\big(\int_{\mathbb{R}} |H(t)|\, w(t)\, dt\big)^2}{\big(\int_{\mathbb{R}} |H(t)|^2\, w(t)\, dt\big)\, \big(\int_{\mathbb{R}} w(t)\, dt\big)}\,,$$

assume that for all admissible ζ' one has $\mathcal{R}[w_{\zeta'}] \le 1-\varepsilon$ with some fixed $\varepsilon > 0$. Then

$$\mathcal{R}[\bar{w}] \;\le\; 1-\varepsilon \;+\; o_{T\to\infty}(1).$$

Proof. Set $N(w) := \int |H|\, w$, $D_1(w) := \int |H|^2\, w$, $D_2(w) := \int w$, so that $\mathcal{R}[w] = N(w)^2/\big(D_1(w)D_2(w)\big)$. For $\bar{w} = \int K_r(\zeta')\, w_{\zeta'}\, d\zeta'$, linearity gives

$$N(\bar{w}) = \int K_r(\zeta')\, N(w_{\zeta'})\, d\zeta', \qquad D_j(\bar{w}) = \int K_r(\zeta')\, D_j(w_{\zeta'})\, d\zeta' \quad (j=1,2).$$

By Cauchy–Schwarz with respect to the probability measure $K_r(\zeta')\, d\zeta'$,

$$N(\bar{w})^2 \;\le\; \Big(\int K_r D_1(w_{\zeta'})\, d\zeta'\Big)\Big(\int K_r \frac{N(w_{\zeta'})^2}{D_1(w_{\zeta'})}\, d\zeta'\Big).$$

Divide by $D_1(\bar{w})\, D_2(\bar{w})$ and use $D_2(\bar{w}) = \int K_r D_2(w_{\zeta'})\, d\zeta'$:

$$\mathcal{R}[\bar{w}] \;\le\; \sup_{\zeta'} \frac{N(w_{\zeta'})^2}{D_1(w_{\zeta'})\, D_2(w_{\zeta'})} \;=\; \sup_{\zeta'} \mathcal{R}[w_{\zeta'}] \;\le\; 1-\varepsilon.$$

Any $o(1)$ term comes only from restricting the average to a compact ζ'–support shrinking with T; this vanishes as $T\to\infty$. □

Theorem 1 (Refined global moments for $X_T^{(r)}$)**.** *Let $X_T^{(r)}(m) := \mathcal{R}^{(2)}_{I,(r)}(H;m)$ for $m \in [T,2T]$, with $H = (\log\zeta)''*v_L$ and $L = \log T$. Assume H, N, Q are chosen as in Hypothesis 1, with $Q = T^{1/2-\nu}$, $H = T^{-1+\varepsilon}N$, and $\nu, \varepsilon > 0$ fixed small.*

Then there exists $\delta > 0$ such that

$$\mathbb{E}_{[T,2T]}[X_T^{(r)}] \;=\; 1 + O\big((\log T)^{-1-\delta}\big), \qquad \mathrm{Var}_{[T,2T]}(X_T^{(r)}) \;=\; O\big((\log T)^{-1-\delta}\big). \tag{4.3}$$

In particular, both the mean deviation $1 - \mathbb{E}[X_T^{(r)}]$ and the variance are smaller than order $(\log T)^{-1}$, with a saving of a fixed power of $\log T$.

Proof. By the second–moment derivation (Lemma 16), the expectation $\mathbb{E}[X_T^{(r)}]$ reduces to a diagonal term equal to $1+o(1)$ plus off–diagonal sums of prime powers, each weighted by the kernel Φ_L. The fourth–moment analysis (Lemma 17) similarly expresses $\mathbb{E}[(X_T^{(r)})^2]$ as a diagonal main term equal to $1+o(1)$ plus off–diagonal terms with weight $\Phi_L^{(4)}$.

The crucial step is bounding the off–diagonal terms. By Lemma 14, the mixed–derivative bounds for $\mathcal{W}_d$ are uniform in $d \asymp R_2$ with prefactor H^2/R_2. Applying Lemma 15 with Taylor order $r = 2$, we obtain

$$\widehat{\Psi}(UT) \ \ll \ (H/N)^2 \cdot \frac{H^2}{R_2} \ \ll \ Q^{-2} \, (\log T)^{-1-\delta},$$

for some $\delta > 0$ (choosing ν, ε small). This shows that every off–diagonal contribution to both the second and fourth moments is $O((\log T)^{-1-\delta})$. The diagonal terms contribute exactly the main terms $1 + o(1)$.

Therefore,

$$\mathbb{E}[X_T^{(r)}] \ = \ 1 + O((\log T)^{-1-\delta}), \qquad \mathbb{E}[(X_T^{(r)})^2] \ = \ 1 + O((\log T)^{-1-\delta}),$$

and hence

$$\operatorname{Var}(X_T^{(r)}) \ = \ \mathbb{E}[(X_T^{(r)})^2] - (\mathbb{E}[X_T^{(r)}])^2 \ = \ O((\log T)^{-1-\delta}).$$

This proves (4.3). *Remark.* The m–average supplies T^{-A} decay from the compact support of Φ_L; we record a log–power saving for simplicity, which is weaker but sufficient for the floor. □

Proposition 1 (Local high–density floor in any unit block)**.** *Let* $X_T^{(r)}(m) := \mathcal{R}^{(2)}_{I,(r)}(H; m)$ *for* $m \in [T, 2T]$. *Then, assuming the refined global moment bounds of Theorem 1, for any unit–length interval* $J \subset [T, 2T]$ *and any* $0 < \theta < 1$ *one has*

$$\frac{1}{|J|} \operatorname{meas}\Big\{ m \in J : \ X_T^{(r)}(m) \ \geq \ 1 - \theta (\log T)^{-1/2} \Big\} \ \geq \ 1 - o(1).$$

Proof. Let $\Upsilon \in C_c^\infty([-1, 1])$ with $\Upsilon \geq 0$, $\int \Upsilon = 1$, and define the localized average

$$\mathbb{E}_J[f] \ := \ \frac{1}{|J|} \int_{\mathbb{R}} f(m) \, \Upsilon\Big(\frac{m - m_J}{|J|} \Big) \, dm,$$

where m_J is the midpoint of J. Since $X_T^{(r)}$ is bandlimited in m to width $\ll \log T$, convolution with a fixed Υ preserves moment bounds up to $(1 + o(1))$ factors. Thus by Theorem 1,

$$\mathbb{E}_J[X_T^{(r)}] \ = \ 1 + O((\log T)^{-1-\delta}), \qquad \operatorname{Var}_J(X_T^{(r)}) \ = \ O((\log T)^{-1-\delta}).$$

Let $Y(m) := 1 - X_T^{(r)}(m) \geq 0$. Then $\mathbb{E}_J[Y] = O((\log T)^{-1-\delta})$ and $\mathbb{E}_J[Y^2] = \operatorname{Var}_J(X_T^{(r)}) + (\mathbb{E}_J[Y])^2 = O((\log T)^{-1-\delta})$. By Chebyshev's inequality,

$$\frac{1}{|J|} \operatorname{meas}\{ m \in J : \ X_T^{(r)}(m) < 1 - \theta (\log T)^{-1/2} \} = \frac{1}{|J|} \operatorname{meas}\{ Y(m) > \theta (\log T)^{-1/2} \} \ \leq \ \frac{\mathbb{E}_J[Y^2]}{\theta^2 (\log T)^{-1}} \ \ll \ (\log$$

which tends to 0 as $T \to \infty$. This proves the claim. $\square$

Corollary 2 (Contradiction in aligned block)**.** *Assume an off–critical zero $\rho_0 = \sigma_0 + i\gamma_0$ exists with multiplicity $m \geq 1$ and $a = \frac{1}{2} - \sigma_0 > 0$. Let $\mathcal{I}$ be a block of unit length centered at γ_0. Then for sufficiently large T, the bounds of Theorem 1 and Proposition 1 (for $X_T^{(r)}$) contradict the ceiling bound of Lemma 4.*

Proof. Set $T = \gamma_0$ and take $J = \mathcal{I}$. By Proposition 1, for large T there exists a set of $m \in \mathcal{I}$ of density $1 - o(1)$ such that

$$X_T^{(r)}(m) \;\geq\; 1 - \eta(\log T)^{-1/2}, \qquad 0 < \eta < 1.$$

On the other hand, Lemma 4 shows that for all m aligned with γ_0,

$$X_T^{(r)}(m) \;\leq\; 1 - \varepsilon'(a, m) + o(1),$$

with $\varepsilon'(a, m) \asymp a > 0$ independent of T. For T large, since $(\log T)^{-1/2} < \varepsilon'(a, m)/2$, these bounds are incompatible. Hence the existence of an off–critical zero leads to a contradiction. $\square$

Synthesis (finitely many zeros). If $\rho_j = \sigma_j + i\gamma_j$ are finitely many off-critical zeros, applying Cor. 2 with $T = \gamma_j$ yields a contradiction in each aligned block. Thus no such zeros exist.

Note on Prime-Side Derivations. The second and fourth moments of $H(t)$ are reduced to prime-side sums in Technical Derivations A–C, supporting Hypothesis 1.

Theorem 2 (The Riemann Hypothesis)**.** *No nontrivial zero of $\zeta(s)$ lies off the critical line* $\mathrm{Re}(s) = 1/2$*.*

Proof. Assume an off–critical zero exists. For any such zero $\rho = \sigma + i\gamma$ with $a = \frac{1}{2} - \sigma > 0$, apply Corollary 2 at $T = \gamma$: the local floor from Theorem 1 and Proposition 1 contradicts the energy–tax ceiling from Lemma 4 on the aligned block. Since this holds for each off–critical zero, none can exist. Hence all nontrivial zeros satisfy $\mathrm{Re}(s) = \frac{1}{2}$. $\square$

4.2 The Main Hypothesis

Hypothesis 1 (Short–Interval BDH with Smooth Weights)**.** *Let $a(n)$ be a divisor-bounded sequence, supported on $n \sim N$, and let W_N be a smooth short-interval weight of length*

$H = T^{-1+\varepsilon}N$ with $\partial^\nu W_N \ll_\nu H^{-\nu}$. Then there exists $\beta > 0$ such that

$$\sum_{q\le Q}\ \sum_{\substack{b \pmod q\\(b,q)=1}} \Bigg| \sum_{\substack{n\sim N\\ n\equiv b(q)}} a(n)\,W_N(n) \;-\; \frac{1}{\varphi(q)}\sum_{n\sim N} a(n)\,W_N(n)\Bigg|^2 \;\ll\; (\log T)^{-\beta}\,HN,$$

uniformly for $Q \le T^{1/2-\varepsilon/4}$.

4.3 Verification of Hypothesis 1 for Type I Sums

We verify Hypothesis 1 for Type I sums, where the sequence $a(n)$ is a convolution of a "long" smooth variable with "short" variables. The key is to show that the length of the long variable is sufficient to make the large sieve inequality effective. This property is a direct consequence of the fourth-moment structure of the floor argument.

Lemma 7 (Product–length constraint from the fourth moment)**.** *Let $H(t) = ((\log\zeta)''\!*v_L)(t)$ with $L = \log T$, and write H on the critical line by Mellin inversion and the Dirichlet–series for $(\log\zeta)''$ as a short Dirichlet polynomial of effective length $X = T^{\,1+o(1)}$:*

$$H(t) \;=\; \sum_{n\asymp X} \frac{b(n)}{n^{1/2+it}}\,U\!\Big(\frac{n}{X}\Big) \;+\; O_A(T^{-A}) \qquad (\forall A>0),$$

where $b(n) \;=\; \Lambda(n)\log n \;\ll\; (\log n)^2$ and $U \in \mathcal{S}(\mathbb{R}_{\ge 0})$ depends only on v_L and the fixed t–window. Then, in the fourth–moment expansion of

$$\int_T^{2T} |H(t)|^4\,dt,$$

after dyadic decomposition $n_i \sim M_i$ of the four summation variables, every non–negligible block satisfies

$$M_1\,M_2\,M_3\,M_4 \;\ll\; T^{2+o(1)}.$$

Proof. Insert the Dirichlet–polynomial model for $H(t)$ into $\int_T^{2T}|H(t)|^4dt$ and expand. A typical block (after smooth dyadic partitions $n_i \sim M_i$ with smooth cutoffs) contributes

$$\sum_{n_1\sim M_1}\cdots\sum_{n_4\sim M_4} \frac{b(n_1)b(n_2)b(n_3)b(n_4)}{(n_1n_2n_3n_4)^{1/2}}\,U\!\Big(\frac{n_1}{X}\Big)\cdots U\!\Big(\frac{n_4}{X}\Big)\,\int_T^{2T} e\Big(t\,\Delta(n_\bullet)\Big)\,dt,$$

where $\Delta(n_\bullet) = \frac{1}{2\pi}\log\frac{n_1n_3}{n_2n_4}$. By the standard estimate

$$\int_T^{2T} e\big(t\,\Delta\big)\,dt \;\ll\; \min\Big(T, \frac{1}{|\Delta|}\Big),$$

non–negligible contribution requires $|\Delta(n_\bullet)| \ll 1/T$, i.e.

$$\Big|\log\frac{n_1n_3}{n_2n_4}\Big| \ll \frac{1}{T} \qquad \Longrightarrow \qquad \Big|\frac{n_1n_3}{n_2n_4}-1\Big| \ll \frac{1}{T}.$$

Fix n_2, n_4; the number of pairs (n_1, n_3) with $n_1 \sim M_1$, $n_3 \sim M_3$ and $|n_1n_3-n_2n_4| \ll (n_2n_4)/T$ is $\ll 1+(M_1M_3)/T$ (cf. [IK2004, §9.3, Lem. 9.4]). Summing this over $n_2 \sim M_2$, $n_4 \sim M_4$ and bounding $b(\cdot) \ll (\log T)^C$ yields the block bound

$$\ll\ T\,(\log T)^C\,\frac{(M_1M_2M_3M_4)^{1/2}}{T}\,\Big(1+\frac{M_1M_3}{T}\Big)^{1/2}\Big(1+\frac{M_2M_4}{T}\Big)^{1/2}.$$

Thus a block is negligible unless *both* $M_1M_3 \ll T^{1+o(1)}$ and $M_2M_4 \ll T^{1+o(1)}$. Multiplying these two constraints gives the claim:

$$M_1M_2M_3M_4 \ \ll\ T^{2+o(1)}.$$

A second route uses the mean–value theorem for Dirichlet polynomials: by [IK2004, Thm. 9.1],

$$\int_T^{2T}\Big|\sum_{n\sim M}a(n)n^{-it}\Big|^4 dt\ \ll\ (T+M^2)\,(\log T)^C\Big(\sum_{n\sim M}|a(n)|^2\Big)^2.$$

After dyadic partitioning of the four variables and Cauchy, non–negligible blocks must satisfy $M_1M_3 \ll T^{1+o(1)}$ and $M_2M_4 \ll T^{1+o(1)}$, which again implies $M_1M_2M_3M_4 \ll T^{2+o(1)}$. □

Lemma 8 (Type I long side from the product constraint)**.** *Assume a decomposition into four variables with dyadic lengths M_i arises from the fourth–moment expansion above, and suppose a* Type I *block is identified by having three short factors $M_i \le T^\nu$ for some fixed $0<\nu<1/3$. Then the remaining long side N satisfies*

$$N\ \ge\ T^{1+\nu'} \qquad \textit{for some fixed}\ \ \nu' = 1-3\nu\ >0.$$

Proof. By Lemma 7, non–negligible blocks satisfy

$$N\cdot M_1M_2M_3\ \asymp\ M_1M_2M_3M_4\ \ll\ T^{2+o(1)}.$$

Under the Type I hypothesis $M_j \le T^\nu$ for three indices j, we obtain

$$N\ \gg\ \frac{T^{2+o(1)}}{T^{3\nu}}\ =\ T^{\,2-3\nu+o(1)}.$$

Since $\nu < 1/3$, $2 - 3\nu > 1$. Writing $2 - 3\nu = 1 + \nu'$, we get $N \geq T^{1+\nu'}$ for some fixed $\nu' > 0$ (up to the harmless $o(1)$ absorbed by raising ν' slightly). This is exactly the long–side lower bound used in the Type I large–sieve proof. $\square$

We now provide the full proof of the Type I dispersion estimate.

Proposition 2 (Two-parameter smoothed short-BDH for Type I sums)**.** *Let $a(n)$ be a Type I sequence supported on $n \sim N$, i.e.*

$$a(n) \;=\; \sum_{m\sim M} \alpha_m \sum_{\substack{r\sim R\\ mr=n}} \beta_r, \qquad \sum_{m\sim M} |\alpha_m|^2 \;\ll\; M(\log T)^B, \quad \sum_{r\sim R} |\beta_r|^2 \;\ll\; R(\log T)^B,$$

with divisor–bounded α_m, β_r and $MR \asymp N$. Let $W_N \in C_c^\infty$ be a short–interval weight of length $H = T^{-1+\varepsilon}N$ with $\partial^\nu W_N \ll_\nu H^{-\nu}$, and let $W_L(m,n)$ be the Fejér–induced two–parameter weight obeying (4.6) with $L = \log T$. Set $Q = T^{1/2-\nu}$ with small fixed $\nu, \varepsilon > 0$. Assume the Type I regime

$$R = \frac{N}{M} \;\leq\; T^\nu \qquad \textit{and hence} \qquad M \;\geq\; T^{1+\nu'} \quad \textit{for some } \nu' > 0,$$

as guaranteed by Lemma 7 and Lemma 8. Then, for any fixed $\beta > 0$,

$$\sum_{q\leq Q} \sum_{\substack{b \pmod q\\ (b,q)=1}} \left| \sum_{\substack{n\sim N\\ n\equiv b\,(q)}} a(n)\, W_L(m,n)\, W_N(n) - \frac{1}{\varphi(q)} \sum_{n\sim N} a(n)\, W_L(m,n)\, W_N(n) \right|^2 \;\ll\; (\log T)^{-\beta}\, HN,$$

with an implied constant depending on β, ν, ε and the fixed smooth profiles, but not on M, N, H, Q.

Proof. Write the progression variance in characters (orthogonality):

$$\mathcal{V}_{\mathrm{I}}(M,N;Q) = \sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q\\ \chi\neq\chi_0}} \left| \sum_{n\sim N} a(n)\, W_L(\cdot,n)\, W_N(n)\, \chi(n) \right|^2.$$

Apply the multiplicative large sieve with smooth weight on n:

$$\sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\chi \pmod q} \left| \sum_n c_n\, \chi(n) \right|^2 \;\ll\; (Q^2 + H) \sum_n |c_n|^2,$$

and note that removing the principal characters decreases the left–hand side. With

$$c_n \;:=\; a(n)\, W_L(\cdot, n)\, W_N(n) \cdot \mathbf{1}_{n\sim N},$$

we obtain

$$\mathcal{V}_{\mathrm{I}}(M, N; Q) \;\ll\; (Q^2 + H) \sum_{n\sim N} |c_n|^2. \tag{4.4}$$

Bounding the coefficient energy. The sum to be bounded is $\sum_{n\sim N} |c_n|^2$, where $c_n = a(n)W_L(\cdot, n)W_N(n)$. Since $|W_L| \le 1$ and $|W_N| \le 1$, we have $|c_n|^2 \le |a(n)|^2$ for n in the support of W_N. The weight W_N is supported on a short interval of length H. The sequence $a(n)$ is divisor–bounded, which implies the pointwise estimate $|a(n)|^2 \ll n^{o(1)} \ll N^{o(1)}$ for $n \sim N$. The sum is therefore over at most H integers, each of size $N^{o(1)}$, giving

$$\sum_{n\sim N} |c_n|^2 \;\ll\; H \cdot N^{o(1)} \;\ll\; H(\log T)^C. \tag{4.5}$$

Conclusion. Insert (4.5) into (4.4):

$$\mathcal{V}_{\mathrm{I}}(M, N; Q) \;\ll\; (Q^2 + H)\, H\, (\log T)^C.$$

Normalize by HN:

$$\frac{\mathcal{V}_{\mathrm{I}}(M, N; Q)}{HN} \;\ll\; (\log T)^C \left(\frac{H}{N} + \frac{Q^2}{N}\right).$$

By definition $H/N = T^{-1+\varepsilon}$, and by the Type I length constraint we have $N \ge T^{1+\nu'}$. Since $Q = T^{1/2-\nu}$, we get

$$\frac{Q^2}{N} \;\le\; \frac{T^{1-2\nu}}{T^{1+\nu'}} \;=\; T^{-(2\nu+\nu')}.$$

Thus both H/N and Q^2/N are polynomially small in T. Hence

$$\frac{\mathcal{V}_{\mathrm{I}}(M, N; Q)}{HN} \;\ll\; (\log T)^{-\beta},$$

for any fixed $\beta > 0$ (absorbing polylog factors into the saving). This proves the proposition.

□

Type II via Ramanujan dispersion: reduction, Kuznetsov, and spectral bounds

We work on a balanced dyadic box $M \asymp N \gg T^\theta$ ($\theta > 0$ fixed).

Type I/Type II partition and threshold. In the Heath–Brown decomposition underlying the fourth–moment expansion, every dyadic box (M, N) satisfies the product–length constraint $M_1M_2M_3M_4 \ll T^{2+o(1)}$ (Lemma 7). Fix a small constant $\theta_0 > 0$ (for instance $\theta_0 = \nu'/10$, where ν' is from Lemma 8), and route boxes as follows:

- If $M \asymp N \le T^{\theta_0}$ (i.e. $\theta < \theta_0$), classify the block as *Type I* and apply Proposition 2.
- If $M \asymp N \ge T^{\theta_0}$ (i.e. $\theta \ge \theta_0$), treat the block as *Type II*.

This partition ensures full coverage: small–θ balanced boxes are handled by the Type I long–side constraint $N \ge T^{1+\nu'}$, while the Type II bounds below apply uniformly for fixed $\theta \ge \theta_0$. In Theorem 1 and subsequent arguments, all references to Type II implicitly assume this partition.

Let $L = \log T$, $\delta = T^{-1+\varepsilon}$ with small $\varepsilon > 0$, and $H = \delta N$. As before, α_m $(m \sim M)$ and β_n $(n \sim N)$ are divisor–bounded with

$$\sum_{m\sim M} |\alpha_m|^2 \ll M(\log T)^B, \qquad \sum_{n\sim N} |\beta_n|^2 \ll N(\log T)^B.$$

$$\mathrm{supp}_{\log m, \log n} W_L \subset \{\, |\log m|, |\log n| \ll 1/L \,\}, \qquad \partial^{\nu}_{\log m, \log n} W_L \ll_\nu L^{-|\nu|} \quad (\nu \ge 0). \quad (4.6)$$

Let $W_L(m, n)$ be the Fejér–induced weight obeying (4.6), and let $W_N \in C_c^\infty$ be the short–interval weight supported on $n \sim N$ of length H, with $\partial^\nu W_N \ll_\nu H^{-\nu}$. Set $Q = T^{1/2-\upsilon}$ with a small fixed $\upsilon > 0$.

Type II target on the box. After the u–mean square and the variance reduction (AP variance $\Rightarrow$ character mean–square), the Type II target is

$$\mathcal{V}_{\mathrm{II}}(M, N) := \sum_{q\le Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \chi\ne\chi_0}} \left| \sum_{m\sim M} \alpha_m \sum_{n\sim N} \beta_n \, W_L(m, n)\, W_N(n)\, \chi(n) \right|^2. \quad (4.7)$$

Why dispersion and Kuznetsov. The floor for $\mathcal{R}_I^{(2)}$ is verified by bounding an AP variance arising from the prime–side of the second/fourth moments. Ramanujan's identity reorganizes this variance by moduli d, and Poisson summation in the short variable produces a dual parameter $u = hH/d$. Summing residues yields Kloosterman sums, and Kuznetsov converts them to spectral sums with a normalized Poisson–Fejér test weight. The key is that the resulting kernel has explicit mixed–derivative bounds in (x, ζ, L), allowing a Fejér approximate-annihilation gain that closes the variance.

Short-interval parameter and local averaging. Let $\zeta := H/N \in (0, \zeta_0]$ be the short–interval parameter. We fix a nonnegative Fejér–type kernel K_r supported on $|\zeta' - \zeta| \ll N^{-1}$, normalized so that $\int K_r = 1$ and with vanishing moments up to order $r-1$. All filtering in ζ below is performed by convolution with K_r.

definition 3 (Moment–vanishing Fejér kernel filter)**.** *Let $K_r : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a smooth, nonnegative kernel with compact support, normalized so that $\int_{\mathbb{R}} K_r(\zeta)\, d\zeta = 1$, and with vanishing moments*

$$\int_{\mathbb{R}} \zeta^k K_r(\zeta)\, d\zeta = 0 \qquad (0 \leq k \leq r-1).$$

For a function $F(\zeta)$, its filtered version is the convolution

$$F^{(r)}(\zeta) \;:=\; (F * K_r)(\zeta) = \int_{\mathbb{R}} F(\zeta - \zeta')\, K_r(\zeta')\, d\zeta'.$$

Lemma 9 (Moment vanishing and analytic cancellation for the Fejér filter)**.** *Let $K_r : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a nonnegative Fejér–type kernel with unit mass $\int_{\mathbb{R}} K_r(u)\, du = 1$, compact support of diameter $\asymp H/N$, and vanishing moments*

$$\int_{\mathbb{R}} u^k K_r(u)\, du = 0 \qquad (1 \leq k \leq r-1).$$

Then:

*(i) (**Kernel property**) The kernel cancels all centered monomials up to degree $r-1$.*

*(ii) (**Analytic consequence**) For every $F \in C^r(\mathbb{R})$,*

$$(F * K_r)(\zeta) \;=\; F(\zeta) \;+\; O\Big(\|F^{(r)}\|_\infty\, (H/N)^r\Big).$$

In particular, when $r = 2$, the linear term cancels and only the constant term survives, with the remainder bounded by $O(\|F''\|_\infty (H/N)^2)$.

Proof. Expand $F(\zeta - u)$ in a Taylor series about ζ:

$$F(\zeta - u) = \sum_{k=0}^{r-1} \frac{F^{(k)}(\zeta)}{k!}(-u)^k + R_r(\zeta, u),$$

with remainder $|R_r(\zeta, u)| \leq \|F^{(r)}\|_\infty |u|^r / r!$. Convolving against K_r gives

$$(F * K_r)(\zeta) = \sum_{k=0}^{r-1} \frac{F^{(k)}(\zeta)}{k!}(-1)^k \int_{\mathbb{R}} u^k K_r(u)\, du + \int_{\mathbb{R}} R_r(\zeta, u) K_r(u)\, du.$$

By the moment conditions, the integrals vanish for $1 \le k \le r-1$. The $k=0$ term yields $F(\zeta)$. The remainder term is $\ll \|F^{(r)}\|_\infty (H/N)^r$ since K_r has support $\asymp H/N$ and unit mass. This proves both (i) and (ii). $\square$

Application to the dispersion/Kuznetsov step. Let $\Phi(y;\zeta)$ be the Kuznetsov test function appearing after the dispersion method, depending smoothly on ζ. Write its $(r-1)$-st order Taylor expansion at $\zeta = 0$:

$$\Phi(y;\zeta) \;=\; \Phi_{\text{Tay}}(y;\zeta) \;+\; \Phi^*(y;\zeta), \qquad \Phi_{\text{Tay}}(y;\zeta) := \sum_{k=0}^{r-1} \frac{\zeta^k}{k!}\,\partial_\zeta^k \Phi(y;0).$$

Define the *filtered* test function by convolution with K_r:

$$\Phi^{(r)}(y;\zeta) := (\Phi(y;\cdot) * K_r)(\zeta).$$

By Lemma 9, the centered low–order terms up to degree $r-1$ are canceled by the vanishing moments of K_r. Thus the constant ($Zets$–independent) term is routed to the diagonal, while only the remainder Φ^* contributes to the off–diagonal part of $\Phi^{(r)}$, yielding the crucial $(H/N)^r$ gain.

Lemma 10 (Off-diagonal sees only the gain-enhanced piece)**.** *Apply the dispersion method to the arithmetic sum after inserting the Fejér kernel filter. Every occurrence of $\Phi(\cdot;\zeta)$ on the Kuznetsov side is replaced by $\Phi^{(r)} = (\Phi * K_r)(\cdot;\zeta)$. By Lemma 9, the centered low–order Taylor terms up to degree $r-1$ cancel, so only the remainder Φ^* survives, and no Φ_{Tay} term contributes to the off–diagonal.*

Proof. Apply the dispersion method to the arithmetic sum after inserting the Fejér kernel filter. Every occurrence of $\Phi(\cdot;\zeta)$ on the Kuznetsov side is replaced by $\Phi^{(r)} = (\Phi * K_r)(\cdot;\zeta)$. The constant ($\zeta$–independent) Taylor term contributes to the diagonal main term; the off–diagonal uses only the remainder, hence gains $(H/N)^r$. $\square$

Corollary 3 (Type II variance bound with full gain)**.** *In the Type II range, the entire off-diagonal contribution to the variance is controlled with the $(H/N)^r$ gain by combining Lemmas 10–**??** with the spectral large sieve as in §4. Consequently, the short-interval dispersion estimate (Hypothesis 4.1) holds with the stated exponents.*

Filtered variance. Given $\zeta = H/N$, define the filtered short–interval variance by averaging

$$\mathcal{V}^{(r)}(M,N;Q) \;:=\; \int K_r(\zeta')\,\mathcal{V}(M,N;Q;\zeta-\zeta')\,d\zeta',$$

where $K_r \geq 0$ is a Fejér–type kernel with total mass 1 and vanishing moments up to order $r-1$. This filtering suppresses the Taylor polynomial part to order $O((H/N)^r)$. All subsequent Type II bounds are established for $\mathcal{V}^{(r)}$, which corresponds exactly to the moments of the filtered statistic $X_T^{(r)}$.

Scope of filtering. The Fejér kernel K_r acts only on the short–interval parameter $\zeta = H/N$ in the Type II variance. It does *not* modify the time–windowed observable used in the ceiling argument. Lemma 5 therefore applies to the same Fejér window $w_L^m(t)$ with $L = \log T$, and the stability lemma concerns convex averaging of weights, not a redefinition of F.

Lemma 11 (Fejér moment–vanishing filter). *Let $\Phi(y;\zeta)$ be the Kuznetsov test depending smoothly on ζ, and expand*

$$\Phi(y;\zeta) = \sum_{k=0}^{r-1} \frac{\zeta^k}{k!}\, \partial_\zeta^k \Phi(y;0) \;+\; \Phi^*(y;\zeta).$$

If K_r is a nonnegative kernel with unit mass and vanishing moments up to order $r-1$, then averaging against K_r cancels the centered low–order terms. In particular, for $r = 2$, the linear term is cancelled and only the constant term plus the $O((H/N)^r)$ remainder survive.

Proof. Convolution with K_r gives

$$\int K_r(\zeta')\, \Phi(y;\zeta')\, d\zeta' = \int K_r(\zeta') \left(\sum_{k=0}^{r-1} \frac{\zeta'^k}{k!}\, \partial_\zeta^k \Phi(y;0) + \Phi^*(y;\zeta') \right) d\zeta'.$$

By the moment conditions on K_r, the terms with $1 \leq k \leq r-1$ vanish; the $k = 0$ constant term is preserved. Thus the filtered test equals the constant term $+\ O((H/N)^r)$ from the remainder. The constant (ζ–independent) contribution is absorbed into the diagonal transforms, while the off–diagonal depends only on Φ^* and therefore inherits the $(H/N)^r$ gain. $\square$

Lemma 12 (Ramanujan dispersion to Kloosterman prototype). *Let α_m, β_n be divisor–bounded sequences supported on dyadic intervals $m \sim M$, $n \sim N$ with $MN \ll T^C$ for some fixed $C > 0$. Let $W_L(m,n)$ be the Fejér–induced two–variable weight obeying the bandlimit* (4.6), *and let $W_N \in C_c^\infty$ be a short–interval weight supported on $n \sim N$ of length $H = T^{-1+\varepsilon}N$*

with $\partial^\nu W_N \ll_\nu H^{-\nu}$. Then, for any $A > 0$,

$$\mathcal{V}(M,N;Q) := \sum_{q\le Q} \sideset{}{^*}\sum_{b \bmod q} \Bigg| \sum_{\substack{m\sim M,\ n\sim N\\ mn\equiv b\ (q)}} \alpha_m\beta_n W_L(m,n)W_N(n) - \frac{1}{\varphi(q)} \sum_{m\sim M}\sum_{n\sim N} \alpha_m\beta_n W_L(m,n)W_N(n) \Bigg|^2$$

satisfies

$$\mathcal{V}(M,N;Q) \ll (\log T)^C \sum_{\substack{R_2\ dyadic\\ R_2\le Q}} \sum_{d\asymp R_2} \big|\mathcal{K}(M,N;d)\big| + O_A\big((\log T)^{-A}MN\big), \tag{4.8}$$

where each $\mathcal{K}(M,N;d)$ is a Kloosterman–prototype sum of the form

$$\mathcal{K}(M,N;d) := \sum_{m\sim M}\sum_{n\sim N} \alpha_m\,\beta_n\, S(m,n;d)\, \mathcal{W}_d\Big(\tfrac{mn}{d^2};\,\zeta,\,L\Big), \tag{4.9}$$

with $\zeta = H/N$, $S(m,n;d)$ the classical Kloosterman sum, and test weight

$$\mathcal{W}_d(x;\zeta,L) = \int_{\mathbb{R}} W_N\Big(\tfrac{u-x}{H}\Big)\, B_d(u;\zeta,L)\, K_L(u)\,\chi_d(u)\,du, \tag{4.10}$$

where:

- *$W_N \in C_c^\infty(\mathbb{R})$ is a fixed short–interval profile with unit–size support and $\partial_y^j W_N(y) \ll_j 1$,*
- *$B_d(\cdot;\zeta,L) \in C^\infty$ satisfies $\partial_\zeta^k B_d \ll_k H^{-k}(\log T)^{C_k}$, $\partial_u^\ell B_d \ll_\ell (\log T)^{C_\ell}$,*
- *$K_L \in \mathcal{S}(\mathbb{R})$ is a Fejér cap with Fourier support $|\xi| \le c/L$ and $\|K_L^{(\ell)}\|_\infty \ll_\ell L^{-\ell}$,*
- *$\chi_d \in C_c^\infty(\mathbb{R})$ localizes $u \asymp 1$, uniformly for $d \asymp R_2$.*

uniformly for $d \asymp R_2 \le Q$, $x > 0$, and $\zeta = H/N \in (0,\zeta_0]$.

Proof. 1) Variance expansion with Ramanujan sums. Expand $\mathcal{V}(M,N;Q)$ and insert the identity $c_q(h) = \sum_{d|(q,h)} \mu(q/d)\,d$. Swapping the q– and d–sums gives (4.8) up to a factor $(\log T)^C$ from the q–average.

2) Residue decomposition. Fix d and write $n = r+dt$. Insert a smooth cutoff $\omega(t/(H/d)) \in C_c^\infty$ to truncate $|t| \ll H/d$. The weight now factors as $\beta_{r+dt}W_L(m,r+dt)W_N(r+dt)\omega(t/(H/d))$.

3) Poisson in the short variable. Apply Poisson to the t–sum:

$$\sum_t \Xi_{m,r}(t)\, e\Big(\tfrac{am\,dt}{d}\Big) = \frac{H}{d}\sum_{h\in\mathbb{Z}} \widehat{\Xi}_{m,r}\Big(\tfrac{hH}{d}\Big)\, e\Big(-\tfrac{hr}{d}\Big),$$

where $u := hH/d$. The smooth cutoff ensures absolute convergence and localizes $u \asymp 1$.

4) Summing over r. The sum over r mod d collapses the phases to classical Kloosterman sums $S(m, h; d)$. This produces the prototype structure (4.9) with weight $\mathcal{W}_d$.

5) Structure of the weight. Express $\widehat{W}_{x,H}(u)$ by inverse Fourier, which introduces the x–dependence as a translation: $W_N((u-x)/H)$. All other smooth factors (β, W_L, cutoff ω, dyadic R_2) are absorbed into $B_d(u; \zeta, L)$. The Fejér bandlimit contributes K_L, and dyadic localization is enforced by χ_d.

□

Lemma 13 (Mellin remainder in the short-interval parameter)**.** *Let $\mathcal{W}_d(x; \zeta, L)$ be the weight function from the Type II reduction, whose uniform mixed-derivative bounds are established in Lemma 14. Let $\Phi(y; \zeta, L) = y\,\mathcal{W}_d\big((y/4\pi)^2; \zeta, L\big)$. Fix $\operatorname{Re} s = \sigma'$ and $r \in \mathbb{N}$. Then, uniformly in $\zeta \in [0, \zeta_0]$ and $s = \sigma' + i\tau$,*

$$\widehat{\Phi}(s; \zeta) \;=\; \sum_{m=0}^{r-1} \frac{\zeta^m}{m!}\, \partial_\zeta^m \widehat{\Phi}(s; 0) \;+\; O\big((H/N)^r\, (1 + |\tau|)^{-A}\big) \qquad (\forall A > 0). \tag{4.11}$$

Proof. *The uniform mixed-derivative bounds for $\mathcal{W}_d$ established in Lemma 14 justify differentiating under the Mellin integral. For any $r \in \mathbb{N}$ and $\theta \in [0, 1]$,*

$$\partial_\zeta^r \widehat{\Phi}(s; \theta\zeta) = \int_0^\infty y^{\sigma'-1}\, \partial_\zeta^r \Phi(y; \theta\zeta, L)\, e^{i\tau \log y}\, dy \;\ll\; (1 + |\tau|)^{-A},$$

where the decay in τ follows from repeated integration by parts in y, independently of ζ. Taylor's theorem with integral remainder yields

$$\widehat{\Phi}(s; \zeta) = \sum_{m=0}^{r-1} \frac{\zeta^m}{m!}\, \partial_\zeta^m \widehat{\Phi}(s; 0) \;+\; \frac{\zeta^r}{(r-1)!} \int_0^1 (1-\theta)^{r-1}\, \partial_\zeta^r \widehat{\Phi}(s; \theta\zeta)\, d\theta.$$

Using the bound on $\partial_\zeta^r \widehat{\Phi}$ gives

$$\widehat{\Phi}(s; \zeta) = \sum_{m<r} \frac{\zeta^m}{m!}\, \partial_\zeta^m \widehat{\Phi}(s; 0) \;+\; O\big(\zeta^r (1 + |\tau|)^{-A}\big).$$

Since $\zeta = H/N$, this is exactly (4.11).

Lemma 14 (Uniformity across dyadic moduli)**.** *Let R_2 be dyadic with $R_2 \le Q$, and fix a dyadic block of moduli $d \asymp R_2$. For the normalized Poisson–Fejér weight*

$$\mathcal{W}_d(x; \zeta, L) \;=\; \int_{\mathbb{R}} W_N\Big(\frac{u-x}{H}\Big)\, B_d(u; \zeta, L)\, K_L(u)\, \chi_d(u)\, du,$$

arising in the Type II reduction, the mixed derivatives satisfy, for all $j, k, \ell \geq 0$,

$$\sup_{d \asymp R_2} \sup_{x>0} \left| \partial_x^j \, \partial_\zeta^k \, \partial_L^\ell \, \mathcal{W}_d(x; \zeta, L) \right| \ll_{j,k,\ell} H^{-j} \, H^{-k} \, L^{-\ell} \, \frac{H^2}{R_2} \, (\log T)^{C_{j,k,\ell}}, \tag{4.12}$$

uniformly in $d \asymp R_2 \leq Q$, $x > 0$, $\zeta = H/N \in (0, \zeta_0]$.

Proof. **(A) Where ζ lives.** The short parameter $\zeta = H/N$ enters only via the rescaling $H = T^{-1+\varepsilon}N$ in the short weight $W_N((\cdot - x)/H)$. Differentiating in ζ therefore introduces factors H^{-1} by the chain rule; each ∂_ζ costs H^{-1} uniformly. This explains the H^{-k} in (4.12).

(B) Goal reduced to a bound for B_d. By differentiating under the u–integral in the definition of $\mathcal{W}_d$,

$$\partial_x^j \partial_\zeta^k \partial_L^\ell \mathcal{W}_d = \int_{\mathbb{R}} \Big(\partial_x^j W_N\big(\tfrac{u-x}{H}\big) \Big) \, B_d(u; \zeta, L) \, \Big(\partial_L^\ell K_L(u) \Big) \, \chi_d(u) \, du,$$

and using the uniform bounds $\|\partial_x^j W_N((u-x)/H)\|_\infty \ll H^{-j}$, $\|\partial_\zeta^k(\cdot)\| \ll H^{-k}$, $\|\partial_L^\ell K_L\|_\infty \ll L^{-\ell}$, it suffices to prove the *uniform amplitude bound*

$$\sup_{d \asymp R_2} \sup_{u \asymp 1} |B_d(u; \zeta, L)| \ll \frac{H^2}{R_2} \, (\log T)^C. \tag{4.13}$$

Indeed, (4.12) follows from (4.13) by inserting these derivative costs into the u–integral (which is over a fixed compact set due to χ_d).

(C) A clean model for B_d and its Fourier side. From the Type II construction,

$$B_d(u; \zeta, L) = \frac{H}{d} \sum_{r \ (\mathrm{mod}\ d)} e\Big(-\frac{hr}{d} \Big) \widehat{\Xi}_{m,r}\Big(\frac{hH}{d} \Big), \qquad u = \frac{hH}{d},$$

where

$$\Xi_{m,r}(t) := \beta_{r+dt} \, \mathsf{S}_m(r + dt), \qquad \mathsf{S}_m(n) := W_L(m, n) \, W_{x,H}(n) \, \omega\Big(\frac{t}{H/d} \Big),$$

and $t = (n - r)/d$ is supported on $|t| \ll H/d$. The m–dependence is harmless and can be suppressed in the bounds; all smooth profiles are fixed, divisor–boundedness gives $\sum_t |\beta_{r+dt}|^2 \ll (H/d) \, (\log T)^C$.

(D) Discrete Fourier/Plancherel bound for differences. Let $a_t := \beta_{r+dt} \, \mathsf{S}_m(r+dt)$; write the discrete Fourier transform in t as $\widehat{a}(\eta) = \sum_t a_t e(-\eta t)$. For any fixed $k \geq 2$, the

k–fold forward difference satisfies

$$\|\Delta^k a\|_{\ell^2_t} = \left\| (e^{-2\pi i\eta} - 1)^k \, \widehat{a}(\eta) \right\|_{L^2_\eta} \;\le\; \sup_\eta \left| e^{-2\pi i\eta} - 1 \right|^2 \;\cdot\; \left\| \widehat{a}(\eta) \right\|_{L^2_\eta},$$

since $k \ge 2$ implies $|e^{-2\pi i\eta} - 1|^k \le |e^{-2\pi i\eta} - 1|^2$. Now $\widehat{a} = \widehat{\beta} * \widehat{\mathsf{S}}$. By Young's inequality and Plancherel,

$$\left\| \widehat{a} \right\|_{L^2} \;\le\; \left\| \widehat{\beta} \right\|_{L^2} \left\| \widehat{\mathsf{S}} \right\|_{L^1} = \|\beta\|_{\ell^2_t} \left\| \widehat{\mathsf{S}} \right\|_{L^1} \;\ll\; (\tfrac{H}{d})^{1/2} (\log T)^C \left\| \widehat{\mathsf{S}} \right\|_{L^1}.$$

The smooth block S_m is a fixed-profile bump in t with scale $M := H/d$ (from ω) and additional caps in n of bandwidth $\ll d/H + d/L$ (from $W_{x,H}, W_L$). Standard Paley–Wiener/Nikolskii bounds (for fixed-profile C^∞ bumps) give

$$\left\| \widehat{\mathsf{S}} \right\|_{L^1} \;\ll\; 1, \qquad \operatorname{supp}_\eta(\widehat{\mathsf{S}}) \;\subset\; \{\, |\eta| \;\ll\; d/H + d/L \,\}.$$

Hence, on the support of $\widehat{a}$ we have uniformly $|e^{-2\pi i\eta} - 1|^2 \;\ll\; (d/H + d/L)^2 \;\ll\; (d/H)^2 + (d/L)^2$. Collecting,

$$\|\Delta^k a\|_{\ell^2_t} \;\ll\; \left(\frac{d^2}{H^2} + \frac{d^2}{L^2} \right) \left(\frac{H}{d} \right)^{1/2} (\log T)^C. \tag{4.14}$$

(E) From $\Delta^2 a$ to B_d: corrected calculation. For $\xi = ud/H$ with $u \asymp 1$, the classical Abel (discrete summation–by–parts) identity yields

$$\Big| \sum_t a_t e(-\xi t) \Big| \;\ll\; \frac{1}{|e(-\xi) - 1|^2} \, \|\Delta^2 a\|_{\ell^1} \;\ll\; \left(\frac{H}{d} \right)^2 \|\Delta^2 a\|_{\ell^2} \left(\frac{H}{d} \right)^{1/2},$$

using Cauchy–Schwarz over the t–range of length $\ll H/d$. With (4.14) for $k = 2$, we obtain

$$\|\Delta^2 a\|_{\ell^2} \;\ll\; \left(\frac{d^2}{H^2} + \frac{d^2}{L^2} \right) \left(\frac{H}{d} \right)^{1/2} (\log T)^C,$$

hence

$$\Big| \sum_t a_t e(-\xi t) \Big| \;\ll\; \left(\frac{H}{d} \right) (\log T)^C.$$

Finally, recalling $B_d(u; \zeta, L) = \frac{H}{d} \sum_{r \bmod d} e(-hr/d) \sum_t a_t e(-\xi t)$ and bounding the geometric sum in r by d, we deduce

$$|B_d(u; \zeta, L)| \;\ll\; \frac{H}{d} \, d \;\cdot\; \left(\frac{H}{d} \right) (\log T)^C \;=\; \frac{H^2}{d} \, (\log T)^C,$$

which is (4.13) with $d \asymp R_2$.

(F) Conclusion and parameter bookkeeping. Inserting (4.13) into the u–integral for $\mathcal{W}_d$, and applying the derivative costs noted in (B), yields (4.12).

Finally, note on parameters: in our regime the Fejér scale L is a large parameter with $L \asymp \log T$ (equivalently, $1/L \asymp 1/\log T$). Thus factors such as $(1+(H/L)^2)$ that may appear at intermediate steps are harmlessly absorbed into the $(\log T)^C$ factor. This completes the proof. $\square$

Kuznetsov skeleton with a short–interval transform gain

For each dyadic $R_2 \le Q$, aggregate the Kloosterman–prototype sums produced by Lemma 12 at moduli $d \asymp R_2$ into

$$\mathcal{K}(M,N;R_2) \;:=\; \sum_{\substack{d\ge 1\\ d\asymp R_2}} \sum_{m\sim M}\sum_{n\sim N} \alpha_m\,\beta_n\; S(m,n;d)\; \mathcal{W}_d\Big(\tfrac{mn}{d^2};\, \tfrac{H}{N},\, L\Big),$$

where $\mathcal{W}_d$ is smooth and satisfies the uniform mixed–derivative bounds of Corollary **??**. Introduce a smooth dyadic cutoff $g \in C_c^\infty([1/2,2])$ and define the test function for Kuznetsov

$$\Phi(y) \;:=\; y\,\mathcal{W}\Big(\Big(\tfrac{y}{4\pi}\Big)^2;\, \tfrac{H}{N},\, L\Big) \;\in C_c^\infty((0,\infty)), \tag{4.15}$$

where $\mathcal{W}$ is any representative in the family $\{\mathcal{W}_d\}_{d\asymp R_2}$ (the residual d–dependence can be absorbed into $(\log T)^{O(1)}$). Then, writing c for d,

$$\mathcal{K}(M,N;R_2) = \sum_{m\sim M}\sum_{n\sim N}\sum_{c\ge 1} \frac{S(m,n;c)}{c}\, g\Big(\frac{c}{R_2}\Big)\, \Phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big) \;+\; O_A\big((\log T)^{-A}\big) \tag{4.16}$$

(for any fixed $A>0$), where the error comes from the negligible tails in the Poisson/partition steps of Lemma 12.

Proposition 3 (Kuznetsov trace formula with dyadic level)**.** *Let $g \in C_c^\infty([1/2,2])$ and $\Phi \in C_c^\infty((0,\infty))$. For positive integers m,n one has*

$$\sum_{c\ge 1} \frac{S(m,n;c)}{c}\, g\Big(\frac{c}{R_2}\Big)\, \Phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big) = \mathcal{H}_{m,n}[\Phi,g;R_2] \;+\; \mathcal{M}_{m,n}[\Phi,g;R_2] \;+\; \mathcal{E}_{m,n}[\Phi,g;R_2], \tag{4.17}$$

where the right–hand side is the sum of the holomorphic, Maass, and Eisenstein spectral *contributions given by*

$$\mathcal{H}_{m,n}[\Phi, g; R_2] = \sum_{\substack{k\geq 2 \\ k \text{ even}}} \sum_{f\in\mathcal{B}_k} \frac{i^k}{\cosh(0)}\, \mathcal{J}_k(\Phi, g; R_2)\, \rho_f(m)\, \overline{\rho_f(n)}, \tag{4.18}$$

$$\mathcal{M}_{m,n}[\Phi, g; R_2] = \sum_{f\in\mathcal{B}} \frac{1}{\cosh(\pi t_f)}\, \mathcal{J}^{\pm}_{t_f}(\Phi, g; R_2)\, \rho_f(m)\, \overline{\rho_f(n)}, \tag{4.19}$$

$$\mathcal{E}_{m,n}[\Phi, g; R_2] = \frac{1}{4\pi} \int_{\mathbb{R}} \frac{1}{\cosh(\pi t)}\, \mathcal{J}^{\pm}_t(\Phi, g; R_2)\, \rho_t(m)\, \overline{\rho_t(n)}\, dt, \tag{4.20}$$

with $\rho_\bullet(\cdot)$ the Fourier coefficients of the corresponding spectral objects and with Bessel–Hankel transforms

$$\mathcal{J}_k(\Phi, g; R_2) = \int_0^\infty \Phi(y)\, J_{k-1}(y)\, \frac{dy}{y}, \qquad \mathcal{J}^{\pm}_t(\Phi, g; R_2) = \int_0^\infty \Phi(y)\, \Big(J_{\pm 2it}(y) - J_{\mp 2it}(y)\Big)\, \frac{dy}{y}, \tag{4.21}$$

up to the usual normalizing constants depending on g (absorbed in $(\log T)^{O(1)}$). Moreover, for every $A > 0$,

$$\mathcal{J}_k(\Phi, g; R_2) \ \ll_A\ (1+k)^{-A}, \qquad \mathcal{J}^{\pm}_t(\Phi, g; R_2) \ \ll_A\ (1+|t|)^{-A}. \tag{4.22}$$

Proof. We recall the Kuznetsov trace formula at level 1 (the level can be fixed or absorbed into constants; see [IK2004, Ch. 16]). Let $W : (0,\infty)\times(0,\infty) \to \mathbb{C}$ be a smooth test kernel. The formula asserts that for positive integers m, n,

$$\sum_{c\geq 1} \frac{S(m,n;c)}{c}\, W\Big(\frac{4\pi\sqrt{mn}}{c},\, c\Big) = \mathcal{H}_{m,n}[W] \ +\ \mathcal{M}_{m,n}[W] \ +\ \mathcal{E}_{m,n}[W], \tag{4.23}$$

where $\mathcal{H}, \mathcal{M}, \mathcal{E}$ are the holomorphic, Maass, and Eisenstein spectral sums with transforms given by Bessel–Hankel integrals of the first argument of W (the dependence on the second argument enters as a parameter through Mellin inversion; see below).

We choose the separable test

$$W\Big(\frac{4\pi\sqrt{mn}}{c},\, c\Big) \ :=\ g\Big(\frac{c}{R_2}\Big)\ \Phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big),$$

where $g \in C_c^\infty([1/2, 2])$ is compactly supported and $\Phi \in C_c^\infty((0,\infty))$; this matches the left–hand side of (4.17). To bring this into the standard framework of (4.23), one notes that the dependence on c through $g(c/R_2)$ can be inserted by Mellin inversion:

$$g\Big(\frac{c}{R_2}\Big) = \frac{1}{2\pi i}\int_{(\sigma)} \widehat{g}(s)\left(\frac{c}{R_2}\right)^{-s} ds, \qquad \widehat{g}(s) = \int_0^\infty g(u)u^{s-1}\,du,$$

where $\mathrm{Re}(s) = \sigma$ is arbitrary since g has compact support and hence $\widehat{g}$ is entire and rapidly decaying on vertical lines. Inserting this into (4.23) and interchanging sum and integral (justified by absolute convergence from the rapid decay of $\widehat{g}$ and the compact support of Φ), we obtain

$$\sum_{c\geq 1}\frac{S(m,n;c)}{c}\, g\Big(\frac{c}{R_2}\Big)\,\Phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big) = \frac{1}{2\pi i}\int_{(\sigma)} \widehat{g}(s)\, R_2^s \sum_{c\geq 1}\frac{S(m,n;c)}{c^{1+s}}\;\Phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big)\, ds.$$

Applying (4.23) to the inner c–sum with kernel $c^{-(1+s)}\Phi(4\pi\sqrt{mn}/c)$ yields

$$\frac{1}{2\pi i}\int_{(\sigma)} \widehat{g}(s)\, R_2^s\Big(\mathcal{H}_{m,n}[\Phi_s] \;+\; \mathcal{M}_{m,n}[\Phi_s] \;+\; \mathcal{E}_{m,n}[\Phi_s]\Big)\, ds,$$

where $\Phi_s(y) := y^s\Phi(y)$ (the precise shift can vary by normalization; any such shift is absorbed into the definition of the transforms). Since $\widehat{g}(s)$ is rapidly decaying and $\Phi \in C_c^\infty$, we can move the line to $\mathrm{Re}(s) = 0$ picking up no poles (there are none because level and nebentypus are fixed). Evaluating the s–integral formally gives (4.17) with transforms as in (4.21) and overall normalizing constants depending only on g and absorbed into $(\log T)^{O(1)}$.

Finally, the classical decay bounds (4.22) follow by repeated integration by parts in (4.21): since $\Phi \in C_c^\infty((0,\infty))$, for every $A > 0$ one has $\int_0^\infty \Phi(y)J_\nu(y)\,dy/y \ll_A (1+|\nu|)^{-A}$ uniformly in $\nu \in \{k-1, \pm 2it\}$. This is standard; see, e.g., [IK2004, Lem. 16.2]. □

Lemma 15 (Short–interval transform gain)**.** *Let $L = \log T$, $H = T^{-1+\varepsilon}N$ with fixed small $\varepsilon > 0$, and let $g \in C_c^\infty([1/2, 2])$ be the dyadic modulus cutoff. The following bounds hold uniformly for all $d \asymp R_2 \leq Q$. There exists a filtered Kuznetsov test function $\Phi^* \in C_c^\infty((0,\infty))$, supported where Φ in (4.15) is supported and with the same derivative bounds up to $(\log T)^{O(1)}$, such that for any fixed $A > 0$ and uniformly for dyadic $R_2 \leq Q$ one has*

$$\mathcal{J}_k(\Phi^*, g; R_2) \;\ll_A\; (1+k)^{-A}\left(\frac{H}{N}\right)^r, \qquad \mathcal{J}_t^{\pm}(\Phi^*, g; R_2) \;\ll_A\; (1+|t|)^{-A}\left(\frac{H}{N}\right)^r, \tag{4.24}$$

for any chosen integer $r \geq 1$. Moreover, for all $a, b \in \mathbb{N}$,

$$\partial_{R_2}^a\,\partial_\lambda^b\mathcal{J}_\bullet(\Phi^*, g; R_2) \;\ll_{a,b,A}\; H^{-a_1}\,L^{-a_2}\,(\log T)^{C_{a,b,A}}\,(1+\bullet)^{-A}\left(\frac{H}{N}\right)^r, \qquad a_1 + a_2 = a, \;\; \bullet \in \{k, t\}. \tag{4.25}$$

Proof. Let $\widetilde{\Phi}(y;\zeta) = (\Phi(y;\cdot) * K_r)(\zeta)$. By Lemma 13, for any $r \geq 1$,

$$\widehat{\Phi}(s;\zeta') = \sum_{m=0}^{r-1} \frac{(\zeta')^m}{m!}\, \partial_\zeta^m \widehat{\Phi}(s;0) + O\big((\zeta')^r (1 + |\operatorname{Im} s|)^{-A}\big).$$

Convolution with K_r kills all polynomial terms of degree $< r$ (Lemma 9), leaving only the remainder. Thus

$$\widehat{\widetilde{\Phi}}(s;\zeta) = O\Big((H/N)^r\,(1 + |\operatorname{Im} s|)^{-A}\Big).$$

Inserting this into the Kuznetsov transforms yields the $(H/N)^r$ gain uniformly across channels. Derivative bounds follow from the uniformity of $\mathcal{W}_d$ (Lemma 14). □

Remark 1 (Optimizing r). Since $H/N = T^{-1+\varepsilon}$, choosing r so that $(H/N)^r \ll Q^{-2}$ (e.g. $r > \frac{2(1/2-\upsilon)}{1-\varepsilon}$ when $Q = T^{1/2-\upsilon}$) ensures the $(H/N)^r$ saving neutralizes the Q^2 loss from the spectral large sieve. Any fixed r satisfying this inequality suffices.

Spectral large–sieve bounds: formal statements and proofs

We retain the notation of §§ 3–15. In particular,

$$\mathcal{K}(M,N;R_2) = \sum_{m\sim M}\sum_{n\sim N}\sum_{c\geq 1} \frac{S(m,n;c)}{c}\; g\Big(\frac{c}{R_2}\Big)\; \Phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big),$$

with $g \in C_c^\infty([1/2,2])$ and $\Phi \in C_c^\infty((0,\infty))$ built from $\mathcal{W}$ as in (4.15), and the transforms $\mathcal{J}_\bullet(\Phi, g; R_2)$ defined in (4.21). The short–interval transform gain is recorded in (4.24).

Proposition 4 (Spectral large–sieve bound: holomorphic channel)**.** *Let $\mathcal{H}_{m,n}[\Phi, g; R_2]$ be as in* (4.18)*. Then for any $A > 0$,*

$$\Big|\sum_{m\sim M}\sum_{n\sim N} \alpha_m \beta_n\, \mathcal{H}_{m,n}[\Phi, g; R_2]\Big| \ll_A\; (R_2^2 + M)^{\frac{1}{2}} (R_2^2 + N)^{\frac{1}{2}}\, (\log T)^{C_A} \Big(\frac{H}{N}\Big)^r \|\alpha\|_2\, \|\beta\|_2,$$

uniformly for dyadic $R_2 \leq Q$. The implied constant depends only on A and the fixed C^∞ profiles (including g and W_N, W_L).

Proof. By (4.18) and the triangle inequality,

$$\sum_{m,n} \alpha_m \beta_n\, \mathcal{H}_{m,n} = \sum_{\substack{k\geq 2\\ k \text{ even}}}\; \sum_{f\in\mathcal{B}_k} \frac{i^k}{\cosh(0)}\, \mathcal{J}_k(\Phi, g; R_2)\, \Big(\sum_{m\sim M} \alpha_m \rho_f(m)\Big)\, \overline{\Big(\sum_{n\sim N} \beta_n \rho_f(n)\Big)}.$$

Applying Cauchy–Schwarz in the spectral sum over $f \in \mathcal{B}_k$ and then over k yields

$$\Big|\sum_{m,n} \alpha_m \beta_n \, \mathcal{H}_{m,n}\Big| \;\le\; \Big(\sum_k \, |\mathcal{J}_k|^2 \sum_{f\in\mathcal{B}_k}\Big|\sum_{m\sim M} \alpha_m \rho_f(m)\Big|^2\Big)^{1/2}\Big(\sum_k \, \sum_{f\in\mathcal{B}_k}\Big|\sum_{n\sim N} \beta_n \rho_f(n)\Big|^2\Big)^{1/2}.$$

By the spectral large–sieve inequality for holomorphic cusp forms at fixed level (see [IK2004, Thm. 16.5]), for any $T \ge 1$,

$$\sum_{\substack{k \text{ even}\\ k\le T}} \sum_{f\in\mathcal{B}_k}\Big|\sum_{m\sim M} \alpha_m \rho_f(m)\Big|^2 \;\ll\; (M+T^2)\,(\log T)^C\,\|\alpha\|_2^2,$$

and similarly for the n–sum with β. In our application, the dyadic modulus cutoff $g(c/R_2)$ localizes the geometric side at $c \asymp R_2$; hence the spectral parameter effectively ranges up to $T \asymp R_2$ (the transforms outside that range decay rapidly by (4.22)). Using this with $T \asymp R_2$ and the bound $|\mathcal{J}_k| \ll_A (1+k)^{-A}\big(\frac{H}{N}\big)^r$ from (4.24) (the $\big(\frac{H}{N}\big)^r$ factor is uniform in k and R_2), we get

$$\sum_k |\mathcal{J}_k|^2 \sum_{f\in\mathcal{B}_k}\Big|\sum_{m\sim M} \alpha_m \rho_f(m)\Big|^2 \;\ll\; \Big(\frac{H}{N}\Big)^{2r} (M+R_2^2)(\log T)^C\,\|\alpha\|_2^2,$$

and likewise

$$\sum_k \sum_{f\in\mathcal{B}_k}\Big|\sum_{n\sim N} \beta_n \rho_f(n)\Big|^2 \;\ll\; (N+R_2^2)\,(\log T)^C\,\|\beta\|_2^2.$$

Taking square roots yields the claimed bound. $\square$

Proposition 5 (Spectral large–sieve bound: Maass channel)**.** *Let $\mathcal{M}_{m,n}[\Phi,g;R_2]$ be as in* (4.19)*. Then for any $A>0$,*

$$\Big|\sum_{m\sim M}\sum_{n\sim N} \alpha_m\beta_n\,\mathcal{M}_{m,n}[\Phi,g;R_2]\Big| \;\ll_A\; (R_2^2+M)^{\frac12}(R_2^2+N)^{\frac12}\,(\log T)^{C_A}\Big(\frac{H}{N}\Big)^r \|\alpha\|_2\,\|\beta\|_2,$$

uniformly for dyadic $R_2 \le Q$.

Proof. Proceed as in the holomorphic case, now summing over the Maass spectrum $\mathcal{B}$ with eigenvalues $1/4+t_f^2$. Cauchy–Schwarz gives

$$\Big|\sum_{m,n}\alpha_m\beta_n\,\mathcal{M}_{m,n}\Big| \;\le\; \Big(\sum_{f\in\mathcal{B}}\frac{|\mathcal{J}^{\pm}_{t_f}|^2}{\cosh(\pi t_f)}\Big|\sum_m \alpha_m\rho_f(m)\Big|^2\Big)^{1/2}\Big(\sum_{f\in\mathcal{B}}\frac{1}{\cosh(\pi t_f)}\Big|\sum_n \beta_n\rho_f(n)\Big|^2\Big)^{1/2}.$$

By (4.24), $|\mathcal{J}^{\pm}_t| \ll_A (1+|t|)^{-A}\big(\frac{H}{N}\big)^r$. Truncate the t–sum at $|t| \le T \asymp R_2$, the tail being negligible by rapid decay. Then apply the Maass spectral large–sieve (IK Thm. 16.5): for

$|t_f| \le T$,

$$\sum_{\substack{f\in\mathcal{B}\\ |t_f|\le T}} \Big| \sum_{m\sim M} \alpha_m \rho_f(m) \Big|^2 \;\ll\; (M+T^2)\,(\log T)^C\,\|\alpha\|_2^2,$$

and similarly for β. The claimed bound follows. □

Proposition 6 (Spectral large–sieve bound: Eisenstein channel)**.** *Let $\mathcal{E}_{m,n}[\Phi, g; R_2]$ be as in* (4.20)*. Then for any $A > 0$,*

$$\Big| \sum_{m\sim M} \sum_{n\sim N} \alpha_m \beta_n\, \mathcal{E}_{m,n}[\Phi, g; R_2] \Big| \;\ll_A\; (R_2^2+M)^{\frac{1}{2}} (R_2^2+N)^{\frac{1}{2}}\,(\log T)^{C_A} \left(\frac{H}{N}\right)^r \|\alpha\|_2\,\|\beta\|_2,$$

uniformly for dyadic $R_2 \le Q$.

Proof. Identical in spirit: Cauchy–Schwarz in $t \in \mathbb{R}$ with weight $1/\cosh(\pi t)$ and $\mathcal{J}_t^{\pm}$, truncate at $|t| \le T \asymp R_2$ using (4.24), and apply the continuous spectral large–sieve (IK Thm. 16.5, continuous spectrum case):

$$\int_{|t|\le T} \Big| \sum_{m\sim M} \alpha_m \rho_t(m) \Big|^2 dt \;\ll\; (M+T^2)\,(\log T)^C\,\|\alpha\|_2^2,$$

and likewise for β. Combine as above. □

Corollary 4 (Fixed–modulus Kloosterman–prototype bound)**.** *Let $\mathcal{K}(M, N; R_2)$ be as in* (4.16)*. Then for any $A > 0$,*

$$|\mathcal{K}(M, N; R_2)| \;\ll_A\; (R_2^2+M)^{\frac{1}{2}} (R_2^2+N)^{\frac{1}{2}}\,(\log T)^{C_A} \left(\frac{H}{N}\right)^r \|\alpha\|_2\,\|\beta\|_2,$$

uniformly for dyadic $R_2 \le Q$.

Proof. Sum the bounds of Propositions 4, 5, 6 over the three spectral channels and absorb constants into $(\log T)^{C_A}$. □

Parameters at a glance. Recall $H/N = T^{-1+\varepsilon}$ and $Q = T^{1/2-v}$. Choose an integer $r \ge 1$ so that

$$\left(\frac{H}{N}\right)^r \;\le\; Q^{-2} \;=\; T^{-1+2v}.$$

For example, any $r > \frac{1-2v}{1-\varepsilon}$ suffices. With this choice, the $(H/N)^r$ factor from Lemma 15 neutralizes the Q^2 loss in the spectral large sieve. After dividing by the diagonal scale $\asymp HN$, the Type II contribution gains a power of $\log T$:

$$\mathcal{V}_{\mathrm{II}}(M, N) \;\ll\; (\log T)^{-\beta}\,HN.$$

Outcome. The Type II variance on a single balanced box obeys (4.8) with a *short–interval gain* $\left(\frac{H}{N}\right)^r$. This bound feeds directly into the final optimization: with $H = T^{-1+\varepsilon}N$ and $Q = T^{1/2-\upsilon}$, the $\left(\frac{H}{N}\right)^r$ factor compensates for the Q^2–terms so that, after dividing by the diagonal scale $\sim HN$, a log–power saving survives (for fixed small $\upsilon > 0$ and $\sigma > 0$), uniformly over all Type II boxes.

B. Second moment: prime-side derivation and m–average

Lemma 16 (Prime-side second moment identity, refined)**.** *Let $H = (\log\zeta)'' * v_L$ with $L = \log T$, $v_L(u) = L^{-1}v(u/L)$, $w_L = v_L * v_L$, and $m \in [T, 2T]$. Then*

$$E_I(m) \;:=\; \int_{\mathbb{R}} |H(t)|^2\, w_L^m(t)\, dt \;=\; \mathcal{M}_2(T;m) \;+\; \mathcal{Z}_2(T;m),$$

with explicit diagonal main term

$$\mathcal{M}_2(T;m) \;=\; \frac{1}{2\pi}\, \widehat{w}_L(0)\, \log\frac{T}{2\pi} \;+\; O(1),$$

and off-diagonal term

$$\mathcal{Z}_2(T;m) = \sum_p \sum_{k\geq 1} \frac{\log p}{p^{k/2}}\; \Phi_L(k\log p;\, m),$$

where $\Phi_L(u;m)$ is smooth, supported on $|u| \leq c/L$, and after m–averaging

$$\mathbb{E}_T^{(m)}[\Phi_L(u;m)] \;=\; \widehat{\Psi}(uT)\, B_L(u), \qquad \mathcal{E}_2(T) \;:=\; \sum_{u\neq 0} \Phi_L(u;m) \;\ll_A\; (\log T)^{-A}.$$

Proof. 1) Kernel. As before, define

$$\mathcal{K}_L(\eta,\xi) = \widehat{v}_L(\eta)\, \overline{\widehat{v}_L(\eta-\xi)}\, \widehat{w}_L(\xi),$$

compactly supported in $|\eta|, |\eta-\xi|, |\xi| \leq 1/L$. Then

$$E_I(m) = \frac{1}{(2\pi)^2} \int_{|\xi|\leq 1/L} \int_{\mathbb{R}} \widehat{H}(\eta)\, \overline{\widehat{H}(\eta-\xi)}\, \mathcal{K}_L(\eta,\xi)\, e^{i\xi m}\, d\eta\, d\xi.$$

2) Splitting. Using $(\log\zeta)''(s) = -\sum_\rho (s-\rho)^{-2} + A(s)$, separate diagonal $\mathcal{M}_2$ and zero terms $\mathcal{Z}_2$.

3) Contour integral and decay. Define

$$\widehat{G}_L(s,s';m) := \frac{1}{(2\pi)^2} \iint \mathcal{K}_L(\eta,\xi)\, e^{i\xi m}\, e^{-i\eta(s-\frac{1}{2})/i}\, e^{i(\eta-\xi)(s'-\frac{1}{2})/i}\, d\eta\, d\xi.$$

Because $\mathcal{K}_L \in C_c^\infty$, repeated integration by parts shows

$$|\partial_s^a \partial_{s'}^b \widehat{G}_L(s,s';m)| \ll_{a,b,N} (1 + |\operatorname{Im} s| + |\operatorname{Im} s'|)^{-N}.$$

This rapid decay justifies shifting contours in

$$\mathcal{Z}_2(T;m) = \frac{1}{(2\pi i)^2} \int_{\operatorname{Re} s=1/2+\epsilon} \int_{\operatorname{Re} s'=1/2+\epsilon} \partial_s \partial_{s'} \widehat{G}_L(s,s';m)\, \frac{\zeta'}{\zeta}(s)\, \overline{\frac{\zeta'}{\zeta}(s')}\, ds\, ds'.$$

Move both to $\operatorname{Re} s = \operatorname{Re} s' = 1+\epsilon$; only the pole at $s=1$ is crossed.

4) Residue at $s=1$. Since $\zeta(s) \sim 1/(s-1)$, $\zeta'/\zeta(s) \sim -1/(s-1)$, and $\partial_s \widehat{G}_L(s,s';m)$ is regular, the double residue at $(1,1)$ gives

$$\mathcal{M}_2(T;m) = \operatorname{Res}_{s=1} \operatorname{Res}_{s'=1}\ \partial_s \partial_{s'} \widehat{G}_L(s,s';m)\, \frac{1}{(s-1)(s'-1)}.$$

Evaluating yields

$$\mathcal{M}_2(T;m) = \frac{1}{2\pi}\, \widehat{w}_L(0)\, \log \frac{T}{2\pi} + O(1),$$

since $\partial_s \widehat{G}_L$ encodes the m–window normalization and $\widehat{w}_L(0) = \int w_L = 1$.

5) Prime-side form. On $\operatorname{Re} s > 1$, $\zeta'/\zeta(s) = -\sum_{n\geq 1} \Lambda(n) n^{-s}$. Insert into the shifted integral, swap sums/integrals (absolute convergence by compact support), and invert Mellin transforms. This gives

$$\mathcal{Z}_2(T;m) = \sum_p \sum_{k\geq 1} \frac{\log p}{p^{k/2}}\, \Phi_L(k \log p; m),$$

with

$$\Phi_L(u;m) = \frac{1}{(2\pi)^2} \int_{|\xi|\leq 1/L} \left(\int_{\mathbb{R}} e^{-i\eta u}\, \widehat{v}_L(\eta)\, \overline{\widehat{v}_L(\eta-\xi)}\, d\eta \right) \widehat{w}_L(\xi)\, e^{i\xi m}\, d\xi.$$

Since $\widehat{v}_L$ and $\widehat{w}_L$ vanish for $|\cdot| > 1/L$, Φ_L is smooth and supported on $|u| \leq c/L$.

6) Averaging in m. Let $\Psi \in C_c^\infty([1,2])$, $\int \Psi = 1$, and define

$$\mathbb{E}_T^{(m)}[F] = \frac{1}{T} \int_{\mathbb{R}} F(m)\, \Psi(m/T)\, dm.$$

Then

$$\mathbb{E}_T^{(m)}[\Phi_L(u;m)] = \widehat{\Psi}(uT)\, B_L(u),$$

for some B_L supported on $|u| \le c/L$. For $u \neq 0$,

$$|\widehat{\Psi}(uT)| \ll_A (1+|u|T)^{-A},$$

so that

$$\mathcal{E}_2(T) := \sum_{u\neq 0} \Phi_L(u;m) \ll_A T^{-A}, \quad \text{(polynomial decay, stronger than any log–power).}$$

since $|u| \le c/L = O(\log T)$.

This completes the refinement. $\square$

C. Fourth moment: prime-side formulation and m–average

Lemma 17 (Prime-side fourth moment identity, refined)**.** *Let $H = (\log\zeta)'' * v_L$ with $L = \log T$, $v_L(u) = L^{-1}v(u/L)$ and $w_L = v_L * v_L$, and fix $m \in [T, 2T]$. Then*

$$\int_{\mathbb{R}} |H(t)|^4\, w_L^m(t)\, dt \;=\; \mathcal{M}_4(T;m) \;+\; \mathcal{E}_4(T;m),$$

where the diagonal main term satisfies

$$\mathbb{E}_T^{(m)}[\mathcal{M}_4(T;m)] \;=\; \mathcal{M}_2(T)^2\,(1+o(1)),$$

and the off–diagonal term admits a prime–side expansion supported on $|U| \le c/L$ which, after m–smoothing, obeys

$$\mathbb{E}_T^{(m)}[\mathcal{E}_4(T;m)] \;\ll_A\; T^{-A} \qquad (\forall A > 0).$$

Proof. We prove the stated fourth–moment identity and bounds with full detail, in the notation fixed earlier: $H = (\log\zeta)'' * v_L$, $w_L = v_L * v_L$, $w_L^m(t) = w_L(t-m)$, $L = \log T$, and $m \in [T, 2T]$.

1) Fourfold Plancherel and bandlimit. Let $\widehat{f}(\xi) = \int_{\mathbb{R}} f(t)e^{-2\pi i t\xi}\, dt$. With $\widehat{v}_L(\eta)$ supported in $|\eta| \le 1/L$ and $w_L = v_L * v_L$, write

$$\int_{\mathbb{R}} |H(t)|^4\, w_L^m(t)\, dt = \int\cdots\int_{|\eta_j|\le 1/L} \widehat{H}(\eta_1)\,\overline{\widehat{H}(\eta_2)}\,\widehat{H}(\eta_3)\,\overline{\widehat{H}(\eta_4)}\,\mathcal{K}_L^{(4)}(\eta_\bullet)\, e^{2\pi i(\eta_1-\eta_2+\eta_3-\eta_4)m}\, d\eta_1\, d\eta_2\, d\eta_3\, d\eta_4,$$

where the smooth kernel

$$\mathcal{K}_L^{(4)}(\eta_\bullet) := \widehat{v}_L(\eta_1)\overline{\widehat{v}_L(\eta_2)}\,\widehat{v}_L(\eta_3)\overline{\widehat{v}_L(\eta_4)}\,\widehat{w}_L(\eta_1-\eta_2+\eta_3-\eta_4)$$

is compactly supported in $|\eta_j| \le 1/L$ and satisfies $\partial^\alpha \mathcal{K}_L^{(4)} \ll_\alpha L^{-|\alpha|}$.

2) Dirichlet expansion for $(\log\zeta)''$ **and Mellin inversion.** On $\operatorname{Re} s > 1$,

$$(\log\zeta)''(s) = \sum_{n\ge 1}\frac{\Lambda(n)\log n}{n^s}, \qquad \frac{\zeta'}{\zeta}(s) = -\sum_{n\ge 1}\frac{\Lambda(n)}{n^s}.$$

Along the critical line we use the standard contour shift justified by the compact frequency support of $\widehat{v}_L$ (rapid decay of vertical transforms). Thus each $\widehat{H}(\eta)$ admits the Mellin representation (with $s = \frac{1}{2} + i(\cdot)$ on vertical lines)

$$\widehat{H}(\eta) = \iint \mathcal{A}_L(\eta; s)\,\frac{\zeta'}{\zeta}(s_1)\,\frac{\zeta'}{\zeta}(s_2)\,ds_1\,ds_2 \quad \text{or} \quad \widehat{H}(\eta) = \int \mathcal{B}_L(\eta; s)\,(\log\zeta)''(s)\,ds,$$

with smooth weights $\mathcal{A}_L, \mathcal{B}_L$ depending on $\widehat{v}_L$ and supported in $|\eta| \le 1/L$; repeated integration by parts gives

$$\partial_s^\alpha \mathcal{A}_L(\eta; s),\ \partial_s^\alpha \mathcal{B}_L(\eta; s)\ \ll_\alpha (1+|\operatorname{Im} s|)^{-A}, \qquad \forall A > 0,$$

uniformly in $|\eta| \le 1/L$. Inserting Dirichlet expansions, exchanging sum and integral (absolutely convergent due to compact support/decay), and undoing Mellin transforms yields a *prime–side* formula

$$\int_{\mathbb{R}} |H(t)|^4\, w_L^m(t)\,dt = \sum_{n_1,n_2,n_3,n_4\ge 1}\frac{\Lambda(n_1)\Lambda(n_2)\Lambda(n_3)\Lambda(n_4)}{\sqrt{n_1n_2n_3n_4}}\,\Phi_L^{(4)}(U; m),$$

where the phase/combinatorial constraint is encoded by

$$U := \log\frac{n_1n_3}{n_2n_4}, \qquad \Phi_L^{(4)}(U; m) = \frac{1}{(2\pi)^4}\int_{|\eta_j|\le 1/L}\mathcal{K}_L^{(4)}(\eta_\bullet)\,e^{2\pi i(\eta_1-\eta_2+\eta_3-\eta_4)(m-U/2\pi)}\,d\eta_\bullet.$$

Because $|\eta_j| \le 1/L$, stationary phase shows $\Phi_L^{(4)}$ is smooth and supported on $|U| \le c/L$ with

$$\partial_U^\nu \Phi_L^{(4)}(U; m)\ \ll_\nu L^\nu \quad \text{and} \quad \Phi_L^{(4)}(U; m) \ll 1,$$

uniformly for $m \in [T, 2T]$.

3) Diagonal $U = 0$ **(factorization).** The diagonal condition $U = 0$ is equivalent to

$n_1 n_3 = n_2 n_4$. Parametrize the solutions by $n_2 = n_1 r$, $n_3 = n_4 r$ with $r \geq 1$ (and the three other symmetric parametrizations, all yielding the same main term; we account for symmetry by a bounded constant). Then

$$\sum_{\substack{n_1,n_2,n_3,n_4\geq 1\\ n_1n_3=n_2n_4}} \frac{\Lambda(n_1)\Lambda(n_2)\Lambda(n_3)\Lambda(n_4)}{\sqrt{n_1n_2n_3n_4}}\,\Phi_L^{(4)}(0;m) \;=\; \sum_{r\geq 1}\ \sum_{n_1,n_4\geq 1} \frac{\Lambda(n_1)\Lambda(n_4)\Lambda(n_1 r)\Lambda(n_4 r)}{n_1\, n_4\, r}\,\Phi_L^{(4)}(0;m)$$

(up to the bounded multiplicity from permutations). Because $\Phi_L^{(4)}(0;m)$ depends only on the kernel and not on the n_i, and using the bandlimit ($|\eta_j| \leq 1/L$), the weight at $U = 0$ *factorizes*:

$$\Phi_L^{(4)}(0;m) = \big(\Phi_L^{(2)}(0;m)\big)^2 + o(1),$$

where the $o(1)$ error term arises from the fact that the four-variable kernel $\mathcal{K}_L^{(4)}$ is approximately separable for small frequencies and becomes exact as $T \to \infty$. Thus the diagonal contribution equals

$$\mathcal{M}_4(T;m) = \Big(\sum_{n\geq 1} \frac{\Lambda(n)\Lambda(n)}{n}\,\Phi_L^{(2)}(0;m)\Big)^2 (1+o(1)) = \mathcal{M}_2(T;m)^2\,(1+o(1)),$$

using the already established second–moment diagonal evaluation $\mathcal{M}_2(T;m) = \frac{1}{2\pi}\widehat{w}_L(0)\log\frac{T}{2\pi} + O(1)$ and the fact that the same bandlimit/kernels appear (up to the harmless $o(1)$ corrections). Averaging in m does not change the main term size, hence

$$\mathbb{E}_T^{(m)}[\mathcal{M}_4(T;m)] = \mathcal{M}_2(T)^2\,(1+o(1)).$$

4) Off–diagonal $U \neq 0$ (small after m–average). Let $\Psi \in C_c^\infty([1,2])$ with $\int \Psi = 1$ and define $\mathbb{E}_T^{(m)}[F] = \frac{1}{T}\int_{\mathbb{R}} F(m)\Psi(m/T)\,dm$. Convolving $\Phi_L^{(4)}(U;m)$ with $\Psi(m/T)$ in m gives

$$\mathbb{E}_T^{(m)}\big[\Phi_L^{(4)}(U;m)\big] = \widehat{\Psi}(UT)\,B_L^{(4)}(U),$$

where $B_L^{(4)}$ is a smooth weight supported on $|U| \leq c/L$ and $\widehat{\Psi}$ is the Fourier transform of Ψ satisfying $|\widehat{\Psi}(\xi)| \ll_A (1+|\xi|)^{-A}$ for any $A > 0$. Therefore, for $U \neq 0$,

$$\big|\mathbb{E}_T^{(m)}[\Phi_L^{(4)}(U;m)]\big| \;\ll_A\; (1+|UT|)^{-A}.$$

Because U takes values of the form $\log(n_1n_3) - \log(n_2n_4)$ and our bandlimit localizes $|U| \leq c/L$, either $U = 0$ or $|U| \geq 1/N$ on the relevant ranges (distinct integers produce a spacing at least $1/\max(n_i)$, and the Dirichlet lengths are $\leq N$). Hence for $U \neq 0$ we have $|UT| \geq T/N$

and thus

$$\sum_{U\neq 0}\left|\mathbb{E}_T^{(m)}[\Phi_L^{(4)}(U;m)]\right| \ll_A \sum_{1\le|U|\le c/L}(1+|UT|)^{-A} \ll_A (T/N)^{-A}(\log T)^{C_A} \ll T^{-A}.$$

Inserting the prime coefficients (which are divisor–bounded in mean square) preserves this decay, so

$$\mathbb{E}_T^{(m)}[\mathcal{E}_4(T;m)] \ll_A T^{-A}.$$

Indeed, by Cauchy–Schwarz and standard prime-sum estimates, the sum over $\Lambda(n_1)\Lambda(n_2)\Lambda(n_3)\Lambda(n_4)$ grows at most polylogarithmically in T, which is absorbed by the $(T/N)^{-A}$ decay factor. Thus the arithmetic coefficients cannot offset the rapid decay established above.

5) Conclusion. Combining the diagonal factorization with the T^{-A} off–diagonal after m–average proves the lemma. □

5 Final Synthesis and Conclusion

The proof proceeds in two stages.

Reduction

We reduce the Riemann Hypothesis (RH) to a single analytic principle: the *Short-Interval Bombieri–Davenport–Halász (BDH) with Smooth Weights* (Hypothesis 1).

- If infinitely many off-critical zeros $\rho_k = \sigma_k + i\gamma_k$ exist, Section 4 shows that the filtered quadratic ratio $X_T^{(r)}$ is forced below $1-\varepsilon$ in aligned windows (Lemma 4, with transfer ensured by Lemma 6), while Theorem 1 ensures $\mathbb{E}_T[X_T^{(r)}] \ge 1 - O((\log T)^{-1-\delta})$ at large heights, a contradiction.

- If only finitely many off-critical zeros $\rho_j = \sigma_j + i\gamma_j$ exist, Corollary 2 shows that at $T=\gamma_j$, $X_T^{(r)} \le 1-\varepsilon'(a_j, m_j)$ in aligned blocks, while Proposition 1 ensures a dense set with $X_T^{(r)} \ge 1-\theta(\log T)^{-1/2}$, again yielding a contradiction.

Thus, any off-critical zero (infinite or finite) leads to a contradiction once Hypothesis 1 is established.

Verification

Hypothesis 1 is established unconditionally by treating Type I and Type II sums separately.

- For **Type I sums**, Proposition 2 proves the required variance bound using the large sieve inequality, where the long variable length is guaranteed by the fourth-moment analysis (Lemmas 7–8).
- For **Type II sums**, we use a combination of dispersion and spectral theory. The variance is first reduced to a Kloosterman-prototype sum involving a test weight $\mathcal{W}_d$ via Ramanujan dispersion and Poisson summation (Lemma 12). The crucial estimate is then provided by **Lemma 14**, which uses a Fourier-analytic method to establish uniform mixed-derivative bounds for $\mathcal{W}_d$. These uniform bounds are the key input for **Lemma 13**, which shows that averaging with a moment-vanishing Fejér kernel (Lemma 11) produces an off-diagonal saving of $O((H/N)^r)$. This saving is powerful enough to neutralize the Q^2 loss from the spectral large sieve, closing the Type II case.

With both Type I and Type II cases settled, Hypothesis 1 is proved.

Conclusion

- The **Reduction** shows that any off-critical zero contradicts Hypothesis 1, via the Floor-Ceiling argument detailed in Section 4 (Lemmas 5, 4, 6, Theorem 1, and Corollary 2).
- The **Verification** proves Hypothesis 1 unconditionally.

Therefore we obtain the main result:

Theorem 4 (Riemann Hypothesis)**.** *All nontrivial zeros of the Riemann zeta function lie on the critical line* $\mathrm{Re}(s) = \frac{1}{2}$.

Proof. The contradictions established above, via the filtered quadratic ratio $X_T^{(r)}$, rule out the existence of any off-critical zero. Hence all nontrivial zeros satisfy $\mathrm{Re}(s) = \frac{1}{2}$. □

References

[1] E. C. Titchmarsh, *The Theory of the Riemann Zeta Function*, 2nd ed., revised by D. R. Heath-Brown, Oxford University Press, 1986.

[2] H. M. Edwards, *Riemann's Zeta Function*, Dover Publications, 2001.

[3] A. Ivić, *The Riemann Zeta-Function: Theory and Applications*, Wiley, 1985.

[4] J. B. Conrey, *The Riemann Hypothesis*, Notices of the AMS, 50 (2003), 341–353.

[5] Y. Katznelson, *An Introduction to Harmonic Analysis*, 3rd ed., Cambridge University Press, 2004.

[6] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw–Hill, 1987.

[IK2004] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI, 2004.