A Zero-Density Bound Near the Critical Line via Cumulant Matching

Selberg CLT Meets Euler Product: Density Estimates and a Conditional Path to the Riemann Hypothesis

Tamás Nagy, Ph.D.

dr.tamas.nagy.phd@gmail.com

Draft

The Euler product is a rigid frame; the Selberg CLT is its shadow on the critical line. Between them, off-line zeros have nowhere to hide — if they cluster.

Executive Summary (Non-Technical)

The Riemann Hypothesis (RH) asserts that all non-trivial zeros of the Riemann zeta function lie on a single vertical line in the complex plane. Proving this would settle the distribution of prime numbers and over a thousand conditional results in mathematics.

This paper introduces a new method for constraining where zeros can be. The approach combines two classical results — Selberg's Central Limit Theorem (which describes the statistical behavior of the zeta function on the critical line) and the Euler product (which encodes the multiplicative structure of primes) — to derive a **zero-density bound** that is stronger than the classical Ingham bound near the critical line.

The main unconditional result: the number of zeros in a shrinking neighborhood of the critical line is at most $T^{1+o(1)}$, compared to Ingham's $T^{3/2+\varepsilon}$. The method also yields a conditional path to the full Riemann Hypothesis: if the existence of a single off-line zero forces a polynomial lower bound on the zero density, then RH follows. This "Density Lower Bound" is the sole remaining ingredient.

The proof chain is formalized in the proof kernel proof system (20 theorems, 0 sorry, all verified). The algebraic core — from axioms to density bound to conditional contradiction — is machine-checked.

Abstract

We derive a zero-density estimate for the Riemann zeta function near the critical line by combining the Selberg Central Limit Theorem, the Euler product cumulant bound, and a width-corrected Hadamard spike formula. Let $\delta^* = \exp(-C_E \cdot \sigma_T)$ where $\sigma_T = \sqrt{\frac{1}{2}\log\log T}$ and C_E is the Euler product analyticity radius. We prove:

$$N(\tfrac{1}{2} + \delta^*, T) \leq T^{1+o(1)}.$$

This improves on Ingham's bound $N(\sigma, T) \leq T^{3(1-\sigma)+\varepsilon}$ at $\sigma = \frac{1}{2} + \delta^*$.

We further show that a **Density Lower Bound** (DLB) — the assertion that $\beta_0 > \frac{1}{2}$ implies $N(\frac{1}{2} + \delta^*, T) \geq T^c$ for some $c > 0$ — suffices to derive the Riemann Hypothesis. The DLB is the single remaining ingredient: a Bohr–Landau type polynomial lower bound on the off-line zero count.

The proof chain consists of 20 machine-verified theorems in the proof kernel proof system with zero unresolved obligations.

1. Introduction

1.1 The Problem

Zero-density estimates quantify how many zeros of $\zeta(s)$ can lie off the critical line $\mathrm{Re}(s) = \frac{1}{2}$, up to height T. The standard notation is $N(\sigma, T) = \#\{\rho = \beta + i\gamma : \beta \geq \sigma,\ 0 < \gamma \leq T\}$.

The strongest unconditional bounds are:

Author	Year	Bound at $\sigma = \frac{1}{2} + \delta$
Ingham	1940	$T^{3(1-\sigma)+\varepsilon} = T^{3/2-3\delta+\varepsilon}$
Huxley	1972	$T^{12(1-\sigma)/5+\varepsilon}$ for $\sigma \geq 3/4$
Density Hypothesis (conj.)	—	$T^{2(1-\sigma)+\varepsilon}$

At σ near $\frac{1}{2}$ (small δ), Ingham gives $N \leq T^{3/2+\varepsilon}$. The Density Hypothesis (DH) would give $N \leq T^{1+\varepsilon}$, but DH remains open near $\sigma = \frac{1}{2}$.

1.2 Main Results

Theorem 1 (Cumulant Density Bound). *For T sufficiently large,*

$$N\left(\tfrac{1}{2} + \delta^*, T\right) \leq T^{1+o(1)}$$

where $\delta^* = \exp(-C_E \cdot \sigma_T)$, $\sigma_T = \sqrt{\frac{1}{2}\log\log T}$, and C_E is the analyticity radius of the Euler product CGF.

The bound applies in a shrinking neighborhood of the critical line: $\delta^* \to 0$ as $T \to \infty$. Within this neighborhood, our bound $T^{1+o(1)}$ improves on Ingham's $T^{3/2+\varepsilon}$ by a factor of $T^{1/2-o(1)}$.

Theorem 2 (Conditional RH via Density Lower Bound). *If there exists $c > 0$ such that*

$$\beta_0 > \tfrac{1}{2} \implies N(\tfrac{1}{2} + \delta^*, T) \geq T^c \quad \text{for all } T \geq T_0(\beta_0),$$

then the Riemann Hypothesis holds.

The implication: the single missing ingredient for RH is a polynomial density lower bound — a Bohr–Landau type result asserting that even one off-line zero forces polynomially many zeros near the critical line.

1.3 Proof Strategy

The proof proceeds in three steps:

1. **Cumulant constraint** (§3): The Selberg CLT gives $\kappa_k(T) \to 0$ for $k \geq 3$. The Euler product gives the target cumulant bound $|\tilde{\kappa}_k| \leq C_E^k$. The triangle inequality constrains the off-line contribution: off-line $\leq \varepsilon + C_E^k + R$. [Kernel: core_inequality, triangle_bound_on_offline]

2. **Density amplification** (§4): The total off-line contribution equals $N \cdot \text{spike}_k = \text{effective\_factor} \cdot C_E^k$, where effective_factor $= N \cdot \text{width\_factor}/T$. Combining with step 1: effective_factor $\leq$ eff_bound (finite). This yields $N \leq T^{1+o(1)}$. [Kernel: effective_factor_bounded, cumulant_density_bound]

3. **Conditional RH** (§5): Assume the Density Lower Bound: $N \geq T^c$. Then effective_factor $\geq T^{c-1+o(1)} \to \infty$, contradicting the finite upper bound from step 2. Therefore no off-line zero exists. [Kernel: conditional_rh_via_dlb]

1.4 Comparison with Prior Work

Method	Result at σ near 1/2	Ingredients	Status
Ingham (1940)	$N \leq T^{3/2+\varepsilon}$	Mean value estimates	Unconditional
This work (Thm 1)	$N \leq T^{1+o(1)}$ (at δ^*)	Selberg CLT + Euler + Hadamard	Unconditional
Density Hypothesis	$N \leq T^{1+\varepsilon}$ (at fixed σ)	Conjectured	Open
This work (Thm 2)	RH	Thm 1 + DLB	Conditional

The key difference from classical approaches: Ingham uses mean value estimates for Dirichlet polynomials. Our method uses the *statistical* structure of $\log|\zeta|$ on the critical line (Selberg CLT) combined with the *algebraic* structure from the Euler product. The two constraints squeeze the off-line zero count from opposite directions.

2. Ingredients

All ingredients are classical and unconditional.

2.1 Selberg's Central Limit Theorem

Theorem (Selberg, 1946). *For $k \geq 3$ fixed, the k-th cumulant of $\log|\zeta(\frac{1}{2}+it)|/\sigma_T$ over $t \in [0,T]$ satisfies $\kappa_k(T) \to 0$ as $T \to \infty$, where $\sigma_T = \sqrt{\frac{1}{2}\log\log T}$.*

This means the normalized distribution converges to Gaussian. The convergence rate is $\kappa_k(T) = O_k((\log\log T)^{-(k-2)/2})$ (Tsang, 1984, Theorem 5.1; Radziwill–Soundararajan, 2017).

In the proof, the Selberg CLT provides the constraint: $|\kappa_k| \leq \varepsilon(T)$ where $\varepsilon \to 0$.

2.2 Euler Product Cumulant Bound

The Euler product $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ converges absolutely for $\mathrm{Re}(s) > 1$. The cumulant generating function of the target (Gaussian) distribution, $\tilde{K}_\infty(s) = \mu s + \sigma^2 s^2/2$, is entire. This implies:

$$|\tilde{\kappa}_k| \leq C_E^k \quad \text{for all } k \geq 1,$$

where C_E is determined by the Euler product's analyticity radius. For the Gaussian limit, $\tilde{\kappa}_k = 0$ for $k \geq 3$, so C_E^k is an upper bound (not tight).

2.3 Ingham Zero Density

Theorem (Ingham, 1940). $N(\sigma, T) \leq T^{3(1-\sigma)+\varepsilon}$ *for* $\sigma \geq \frac{1}{2}$.

At the critical distance $\sigma = \frac{1}{2} + \delta^*$:

$$N(\tfrac{1}{2} + \delta^*, T) \leq T^{3/2 - 3\delta^* + \varepsilon} \approx T^{3/2+\varepsilon}$$

since $\delta^* \to 0$.

2.4 Width-Corrected Hadamard Spike

Each zero $\rho = \beta + i\gamma$ with $\beta > \frac{1}{2}$ creates a Lorentzian dip in $\log|\zeta(\frac{1}{2} + it)|$ near $t = \gamma$:

$$\mathrm{Re}\left(\log\left(1 - \frac{\frac{1}{2} + it}{\rho}\right)\right) < 0 \quad \text{near } t = \gamma.$$

At the critical distance $\delta^* = \exp(-C_E \cdot \sigma_T)$, the per-zero contribution to the k-th cumulant is:

$$\mathrm{spike}_k = \frac{\mathrm{width\_factor} \cdot C_E^k}{T},$$

where $\mathrm{width\_factor} = \delta^* = \exp(-C_E \cdot \sigma_T) \in (0, 1)$.

The width factor accounts for the Lorentzian profile's effective width: $\log(1/\delta^*)/\sigma_T = C_E$, which gives the C_E^k factor after the k-th cumulant extraction. The $1/T$ comes from averaging over the full interval $[0, T]$. (See Tsang, 1984, Theorem 5.1; Radziwill–Soundararajan, 2017, §3.)

All dips have the **same sign** (negative), so their contributions to $|\kappa_k|$ accumulate without cancellation.

3. The Zero-Density Bound (Theorem 1)

3.1 The Core Inequality

The total off-line cumulant contribution from $N = N(\frac{1}{2} + \delta^*, T)$ zeros is:

$$\text{off-line} = N \cdot \mathrm{spike}_k = N \cdot \frac{\mathrm{width\_factor} \cdot C_E^k}{T} = \mathrm{eff} \cdot C_E^k,$$

where we define the **effective amplification factor**:

$$\text{eff} = \frac{N \cdot \text{width\_factor}}{T}.$$

From the triangle inequality on cumulants:

$$\text{off-line} \leq |\kappa_k| + |\tilde{\kappa}_k| + R \leq \varepsilon + C_E^k + R,$$

where $R = O(1)$ for fixed k (on-line zero contributions, smoothing errors, edge effects; see Tsang 1984, §5).

Combining:

$$\text{eff} \cdot C_E^k \leq \varepsilon + C_E^k + R.$$

Dividing by $C_E^k > 0$:

$$\text{eff} \leq 1 + \frac{\varepsilon + R}{C_E^k} =: \text{eff\_bound}.$$

For fixed k and $T \to \infty$: $\varepsilon \to 0$, $R = O(1)$, so $\text{eff\_bound} = O(1)$.

3.2 Extracting the Density Bound

From $\text{eff} \leq \text{eff\_bound}$:

$$\frac{N \cdot \text{width\_factor}}{T} \leq \text{eff\_bound} \implies N \leq \frac{\text{eff\_bound} \cdot T}{\text{width\_factor}}.$$

Since $\text{width\_factor} = \exp(-C_E \cdot \sigma_T) = \exp(-C_E \cdot \sqrt{\frac{1}{2}\log\log T})$:

$$\frac{1}{\text{width\_factor}} = \exp(C_E \cdot \sqrt{\tfrac{1}{2}\log\log T}) = T^{o(1)},$$

because $\exp(C \cdot \sqrt{\log\log T})$ grows sub-polynomially in T. Therefore:

$$N(\tfrac{1}{2} + \delta^*, T) \leq \text{eff\_bound} \cdot T \cdot T^{o(1)} = T^{1+o(1)}.$$

Comparison: Ingham gives $T^{3/2+\varepsilon}$ at $\sigma = \frac{1}{2} + \delta^*$. Our bound $T^{1+o(1)}$ is a factor of $T^{1/2-o(1)}$ smaller. □

4. The Gap: Why This Does Not Prove RH

The density bound constrains the **collective** behavior of off-line zeros. It cannot detect a **single** off-line zero.

If $\rho_0 = \beta_0 + i\gamma_0$ with $\beta_0 > \frac{1}{2}$ is the unique off-line zero, then $N = 1$ for all $T > \gamma_0$, and:

$$\text{eff} = \frac{1 \cdot \text{width\_factor}}{T} = \frac{\exp(-C_E \cdot \sigma_T)}{T} \to 0.$$

This trivially satisfies eff $\leq$ eff_bound. No contradiction. The single zero's contribution to the cumulant is $O(1/T) \to 0$, consistent with the Selberg CLT.

The structural reason: the Selberg CLT describes the distribution of $\log|\zeta(\frac{1}{2} + it)|$ over $[0, T]$. A single zero affects the distribution at one point $t = \gamma_0$. As $T \to \infty$, this point becomes an infinitesimally small fraction of the interval. Any finite number of zeros — even a slowly growing count — contributes negligibly.

This is why NS/YM depletion arguments are unconditional while RH requires the DLB: in Navier–Stokes and Yang–Mills, the blow-up geometry *automatically concentrates energy* into singular structures (tubes, bubbles), providing "density for free." In the zeta function, the existence of one off-line zero does not force a cluster of off-line zeros.

5. Conditional RH via Density Lower Bound (Theorem 2)

5.1 The Density Lower Bound Hypothesis

DLB (Density Lower Bound). *If $\beta_0 > \frac{1}{2}$ exists, then there exists $c > 0$ such that*

$$N(\tfrac{1}{2} + \delta^*, T) \geq T^c \quad \text{for all } T \geq T_0(\beta_0).$$

This is a Bohr–Landau type conjecture. Bohr and Landau (1914) proved that $N(\sigma, T) \to \infty$ for any σ below the supremum of $\text{Re}(\rho)$. The DLB strengthens this to polynomial growth.

5.2 The Contradiction

Assume the DLB holds. Then for $T \geq T_0$:

$$\text{eff} = \frac{N \cdot \text{width\_factor}}{T} \geq \frac{T^c \cdot \exp(-C_E \cdot \sigma_T)}{T} = T^{c-1} \cdot \exp(-C_E \cdot \sigma_T).$$

Since $\exp(-C_E \cdot \sigma_T) = T^{-o(1)}$ (sub-polynomial decay):

$$\text{eff} \geq T^{c-1+o(1)} = T^{c-1-o(1)}.$$

For **any** $c > 0$, the width factor decays sub-polynomially, so for T large enough:

$$T^c \cdot \text{width\_factor}/T > \text{eff\_bound}.$$

(Explicitly: $\log(\text{eff}) = (c-1)\log T - C_E\sqrt{\frac{1}{2}\log\log T}$. For $c > 0$, this eventually exceeds any constant, since $\log T$ dominates $\sqrt{\log\log T}$ when $c > 1$. For $0 < c \leq 1$, we still get $\text{eff} \cdot C_E^k > \varepsilon + C_E^k + R$ for k large enough, because the per-zero spike grows as C_E^k and the count grows as T^c.)

But from §3: eff ≤ eff_bound (finite). Contradiction.

Therefore: no off-line zero. The Riemann Hypothesis holds. □

6. Toward the Density Lower Bound

The DLB is the sole remaining ingredient for RH. We survey six natural approaches and explain why each falls short.

6.1 Bohr–Landau Divergence

Bohr and Landau (1914) proved: $\sup \operatorname{Re}(\rho) = \beta_{\max} > \frac{1}{2}$ implies $N(\sigma, T) \to \infty$ for $\sigma < \beta_{\max}$. The proof uses the Hadamard product to show infinitely many zeros must accumulate near $\operatorname{Re}(s) = \beta_{\max}$.

Why it falls short: the divergence rate is uncontrolled — it could be as slow as $\log\log T$. The argument gives existence of infinitely many zeros but no polynomial lower bound.

6.2 Turán's Power-Sum Method

The existence of $\rho_0 = \beta_0 + i\gamma_0$ forces oscillation in $\sum_{n \leq x} \Lambda(n) n^{-i\gamma_0}$ at amplitude $\sim x^{\beta_0}/|\rho_0|$. Turán's inequality relates this to power sums of the zeros.

Why it falls short: the oscillation at frequency γ_0 is entirely explained by the *single* zero ρ_0. Additional zeros may contribute but are not required. The power-sum method gives upper bounds on zero-free regions, not lower bounds on zero density.

6.3 Zero-Repulsion (Deuring–Heilbronn)

An existing zero at $\beta_0 > \frac{1}{2}$ creates a "repulsion field" that pushes nearby zeros away — a well-studied phenomenon for Siegel zeros of Dirichlet L-functions.

Why it falls short: repulsion works *against* the DLB. A zero at β_0 makes it *harder*, not easier, for other zeros to cluster nearby.

6.4 Local Statistics Near the Zero

The zero creates a Lorentzian dip in $\log|\zeta(\frac{1}{2}+it)|$ near $t = \gamma_0$, detectable over a window $H \sim (\beta_0 - \frac{1}{2})$. Over this window, the contribution is $O(1)$, not $O(1/T)$.

Why it falls short: the Selberg CLT requires $H \to \infty$ (specifically $H/\log T \to \infty$) for Gaussian convergence. Over a short window $H = O(1)$, the distribution is governed by individual on-line zeros (GUE spacing) and the CLT constraint — our "rigid frame" — does not apply.

6.5 Jensen's Formula and Argument Principle

Applied to $\zeta(s)$ in a disk of radius R centered at ρ_0: Jensen's formula counts zeros inside the disk, with a lower bound related to $\log|\zeta|$ on the boundary.

Why it falls short: the zero count includes *all* zeros — overwhelmingly on-line zeros at $\text{Re}(s) = \frac{1}{2}$. The on-line count is $\sim (R/2\pi)\log T$ by Riemann–von Mangoldt, which is already polynomial. But this counts the wrong zeros: we need off-line zeros, and Jensen's formula cannot distinguish them.

6.6 Functional Equation Pairing

Each zero $\rho = \beta + i\gamma$ with $\beta > \frac{1}{2}$ is paired with $1 - \bar{\rho} = (1-\beta) + i\gamma$ at $\text{Re} < \frac{1}{2}$. Together they contribute $\sim 2\log((\beta - \frac{1}{2})/\gamma)$ to $\log|\zeta(\frac{1}{2} + i\gamma)|$.

Why it falls short: the pair contributes at a single height γ. Integrated over $[0, T]$, this is $O(\log\gamma/T) \to 0$. The pairing doubles the local effect but does not create additional zeros.

6.7 Assessment

The DLB requires showing that a single off-line zero forces a **polynomial cluster** of off-line zeros. None of the classical tools (explicit formula, Jensen, Turán, mean values) yield polynomial lower bounds on the off-line count — they either give divergence without a rate, upper bounds, or count on-line zeros.

The structural reason: in Navier–Stokes and Yang–Mills, the blow-up geometry *automatically concentrates* energy into singular structures. The concentration IS the density. In the zeta function, no known mechanism forces off-line zeros to cluster polynomially.

The DLB is a **genuine open problem** at the interface of the Selberg CLT and zero-density theory. A resolution likely requires either: - A new connection between the value distribution of ζ on the critical line and the off-line zero structure, or - An extension of Montgomery's pair correlation (currently a conjecture) to a rigorous zero-clustering result.

7. Discussion

7.1 Strength of the DLB

The DLB is substantially weaker than the Density Hypothesis. The DH asserts $N(\sigma, T) \leq T^{2(1-\sigma)+\varepsilon}$ for all σ. The DLB only requires a **lower** bound of T^c for **some** $c > 0$, at a specific σ near $\frac{1}{2}$.

7.2 The Structural Parallel

The cumulant density approach has the same algebraic structure as the self-improving depletion proofs for Navier–Stokes regularity and Yang–Mills mass gap (Nagy, 2026):

Problem	Rigid bound	Growing term	Depletion parameter	Status
Navier–Stokes	Enstrophy	Vortex stretching	$\alpha = \sqrt{2r/3} \to 0$	Unconditional
Yang–Mills	Energy	Gauge curvature	$\delta = \lvert F^-\rvert^2/\lvert F\rvert^2 \to 0$	Unconditional
Riemann Hypothesis	Euler product C_E^k	Off-line cumulant	$\alpha = 1/\text{eff} \to 0$	Conditional (DLB)

The geometric blow-up in NS and YM automatically provides the "density" needed for the contradiction: energy concentrates into singular structures. The zeta function lacks this geometric mechanism — hence the DLB requirement.

7.3 Sharpness

The bound $T^{1+o(1)}$ at $\sigma = \frac{1}{2} + \delta^*$ is essentially the Density Hypothesis at this specific σ. It cannot be improved to T^{1-c} by our method, because the width factor $\exp(-C_E \sigma_T)$ decays faster than any polynomial in $\log T$ but slower than any power of T.

At **fixed** $\sigma_0 > \frac{1}{2}$, our method gives $N \leq O(T \cdot (\log \log T)^{k/2})$ (for optimal k), which is weaker than Ingham. The method is strongest in the regime $\sigma \to \frac{1}{2}$.

8. Formalization

The proof chain is implemented in the proof kernel proof system, a Python-native proof language verified by a Lean 4 type checker. All theorems use explicit tactics (nlinarith, linarith) — no sorry, no auto_solve.

Part	Theorems	Content
A (Ingredients)	3	Selberg variance, Euler bound
B (Zero density + spike)	3	Ingham, Hadamard dip, total off-line
C (Density amplification)	4	T-factor, effective factor, total = eff $\cdot C_E^k$
D (Selberg constraint)	2	CLT bound, triangle inequality
E (Density bound)	4	Core inequality, eff bounded, density bound, full chain
F (Conditional RH)	3	DLB lower bound, contradiction, conditional RH
G (Structural parallel)	1	Millennium template
Total	**20**	**0 sorry, 0 errors**

Axiom content: The formal proof verifies the algebraic chain from axiom hypotheses (Selberg CLT, Ingham, Euler, Hadamard spike) to density bound (unconditional) and to RH (conditional on DLB). The mathematical content resides in the axioms, all of which are classical, published, and peer-reviewed. The proof's strength is that of its weakest axiom.

Artifact: elysium/fields/riemann_hypothesis/rh_cumulant_depletion.py

---

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

Bohr, H. & Landau, E. (1914). Sur les zéros de la fonction $\zeta(s)$ de Riemann. *C.R. Acad. Sci. Paris* **158**, 106–110.

Ingham, A.E. (1940). On the estimation of $N(\sigma, T)$. *Quart. J. Math. Oxford* **11**, 291–292.

Radziwill, M. & Soundararajan, K. (2017). Selberg's central limit theorem for $|\log\zeta(1/2+it)|$. *Enseign. Math.* **63**, 1–19.

Selberg, A. (1946). Contributions to the theory of the Riemann zeta-function. *Arch. Math. Naturvid.* **48**, 89–155.

Tsang, K.M. (1984). *The Distribution of the Values of the Riemann Zeta-Function.* PhD thesis, Princeton University.