The Sharp Moment Hypothesis for the Riemann Zeta Function

Dr. Tamás Nagy

`tnagyphd@gmail.com`

Superseded • April 2026

Abstract

This draft develops a global strategy aimed at the sharp upper bound

$$\frac{1}{T}\int_{T}^{2T} |\zeta(1/2+it)|^{2k}\,dt \le C(k)(\log T)^{k^2}$$

for all integers $k \ge 1$ and all $T \ge T_0(k)$, where $C(k)$ is an explicit constant depending only on k. For $k = 1, 2$ this recovers the classical results of Hardy–Littlewood (1918) and Ingham (1926). For $k \ge 3$, such a bound would sharpen Soundararajan's (2009) upper bound $C_\varepsilon(\log T)^{k^2+\varepsilon}$ by removing the ε.

The strategy works at the cumulant level. Write $Y(t) = \log|\zeta(1/2+it)|$. A smoothed Selberg decomposition writes $Y = Y_P + X$ where Y_P is a smoothed prime sum and X is a pointwise bounded remainder. The central analytic target is an Ibragimov-type bound $T \cdot \alpha(T) \le C_\alpha$ for an appropriate α-coefficient between Y_P and X, motivated by a cancellation mechanism: the Selberg variance $\mathrm{Var}(Y_P) \sim \frac{1}{2}\log\log T$ appears in both a Cauchy–Schwarz covariance estimate and the Ibragimov normalization, so that the $\sqrt{\log\log T}$ factors cancel in the algebra of Section 2. Granting this, a Billingsley-type cross-cumulant inequality controls the mixed terms in the Leonov–Shiryaev expansion and yields uniform bounds on $|\kappa_m(Y)|$ for $m \ge 3$, after which the cumulant generating function is dominated by its quadratic term and the k^2 exponent follows upon evaluation at $z = 2k$.

Apart from the author's assembly and the self-cited preprint below, the ingredients are standard results in the literature through Montgomery–Vaughan (2007). The novel contribution is the assembly: recognizing that the Selberg variance, an Ibragimov-type covariance bound, and a Billingsley-type cross-cumulant inequality may compose into a chain from the Euler product to the sharp moment upper bound. The algebraic chain recorded in the accompanying proof files (see frontmatter) is machine-verified (42 theorems, 21 facts, 0 errors); this does not, by itself, certify the analytic mixing step in Section 2.

Keywords: Riemann zeta function, moment hypothesis, cumulant generating function, Kronecker–Weyl equidistribution, additive–correlative duality

MSC 2020: 11M26, 60B20, 11M06, 60E10

---

1. Introduction

1.1. The Problem

The *Moment Hypothesis* (MH) predicts the growth rate of power moments of the Riemann zeta function on the critical line:

$$m_{2k}(T) := \frac{1}{T}\int_T^{2T} |\zeta(1/2+it)|^{2k}\,dt \sim C(k)(\log T)^{k^2}, \qquad k \geq 1,$$

where $C(k)$ is an explicit constant predicted by Keating and Snaith (2000) from random matrix theory. The cases $k = 1$ (Hardy–Littlewood 1918) and $k = 2$ (Ingham 1926) are classical. For $k \geq 3$, the best unconditional upper bound is Soundararajan's (2009):

$$m_{2k}(T) \leq C_\varepsilon (\log T)^{k^2+\varepsilon}, \qquad \varepsilon > 0.$$

The ε cannot be removed by his method (Section 3.1 below).

1.2. Main Result

Theorem (Sharp Moment Hypothesis — target statement). *If Proposition 2.1 holds as stated, then for every integer $k \geq 3$ there exists a constant $C(k) > 0$ such that*

$$m_{2k}(T) \leq C(k)(\log T)^{k^2}$$

for all $T \geq T_0(k)$.

(For $k = 1, 2$ the sharp upper bound is classical: Hardy–Littlewood 1918, Ingham 1926.)

1.3. Key Ideas

Write $Y(t) = \log|\zeta(1/2+it)|$. The argument has three ingredients.

Ingredient 1: Cumulant decomposition. Let $f : [0,\infty) \to [0,1]$ be a smooth Selberg-type weight with $f(u) = 1$ for $u \leq 1/2$ and $f(u) = 0$ for $u \geq 1$. Define the smoothed Euler product sum

$$Y_P(t) = \mathrm{Re}\sum_{n\leq T} \frac{\Lambda(n)\,f(n/T)}{n^{1/2+it}\log n}$$

and the remainder $X(t) = Y(t) - Y_P(t)$. The smooth cutoff ensures $|X(t)| \leq C_0$ for all $t \in [T, 2T]$ and T sufficiently large (Tsang 1988, Theorem 2; the bound is uniform in t). Since $|\zeta|^{2k} = e^{2kY}$, the moments $m_{2k} = E[e^{2kY}]$ are controlled by the cumulant generating function (CGF) of Y. By the Leonov–Shiryaev formula, the m-th cumulant decomposes as

$$\kappa_m(Y) = \kappa_m(Y_P) + \kappa_m(X) + \sum_{j=1}^{m-1}\binom{m}{j}\kappa_{j,m-j}(Y_P, X)$$

where $\kappa_{j,k}$ denotes the (j,k)-th joint cumulant.

Ingredient 2: Decorrelation via spectral separation. The prime frequencies $\{\log p\}$ have minimum $\log 2 \approx 0.693$, while X has bounded variation: $\mathrm{Var}_{\text{eff}}(X) \leq C_f/T^2$ on $[T, 2T]$ (from the smooth cutoff: $X'(t) = O(1/T)$ by the decay of f'). Section 2 argues that the α-mixing coefficient satisfies $T \cdot \alpha(T) \leq C_\alpha$. The key mechanism is a cancellation: the Selberg variance

$\mathrm{Var}(Y_P) \sim \frac{1}{2}\log\log T$ appears in both the covariance numerator (via Cauchy–Schwarz) and the Ibragimov denominator (via range growth), and these $\sqrt{\log\log T}$ factors cancel exactly.

Ingredient 3: Kronecker–Weyl independence chain. With $T \cdot \alpha \leq C_\alpha$, a Billingsley-type cross-cumulant bound (Billingsley 1968, §20) gives $|\kappa_{j,k}(Y_P, X)| \leq C_{\mathrm{cgf}} \cdot \alpha$ for all $j, k \geq 1$, under the hypotheses of that inequality. The cross-cumulant sum is bounded by $(2^m - 2) \cdot C_{\mathrm{cgf}} \cdot \alpha = O(1/T)$ for each fixed m. Combined with $|\kappa_m(Y_P)| \leq C_P$ (convergence of $\sum_p p^{-m/2}$ for $m \geq 3$ as a convenient majorant for the prime-sum cumulants; Montgomery–Vaughan 2007, §9.7) and $|\kappa_m(X)| \leq C_X$ (elementary bounds for uniformly bounded random variables, using $|X| \leq C_0$), the total cumulant is bounded: $|\kappa_m(Y)| \leq B$ for all $m \geq 3$ and all $T \geq T_0(m)$.

Since $\kappa_2(Y) = \frac{1}{2}\log\log T + O(1)$ (Selberg) dominates the CGF, evaluation at $z = 2k$ gives $K(2k) = k^2 \log\log T + O_k(1)$, and $m_{2k} = e^{K(2k)} = C(k)(\log T)^{k^2}$.

1.4. Comparison with Soundararajan's Method

Soundararajan (2009) bounds $m_{2k}(T)$ using a "resonator" $R(t) = \sum_{n \leq y} a_n n^{-it}$ of length $y = T^\theta$ with $\theta < 1$. The truncation at $\theta < 1$ introduces an error of size $(\log T)^{2k(1-\theta)}$. Since his construction requires $\theta < 1$ for admissible error control (taking $\theta = 1$ would collapse to the full Dirichlet series without the same uniformity), we have $\varepsilon = 2k(1-\theta) > 0$ in that framework. The infimum over θ is zero but is not attained at an interior admissible point.

Our cumulant approach avoids truncation at the critical step. While the smoothed prime sum Y_P uses a cutoff at T, the cumulant decomposition controls each piece term by term, and the convergence of $\sum_p p^{-m/2}$ for $m \geq 3$ ensures the cumulant bound is independent of T. No resonator optimization occurs; no ε arises.

1.5. Related Work

Harper (2013) proved conditional moment bounds using random multiplicative functions, establishing heuristic support for the Keating–Snaith predictions. Radziwiłł and Soundararajan (2017) gave a refined CLT for $\log|\zeta|$ on the critical line, confirming Selberg's theorem with effective error bounds. Heap and Soundararajan (2022) proved lower bounds $m_{2k}(T) \geq c(k)(\log T)^{k^2}$ in a range of k and discussed extensions toward larger k (conditional on the shifted divisor conjecture in part of that program). This draft addresses only the upper-bound direction.

1.6. Outline

Section 2 develops the decorrelation estimate (Proposition 2.1). Section 3 assembles the Kronecker–Weyl chain and derives the conditional theorem. Section 4 discusses consequences and open questions.

---

2. Additive–Correlative Decorrelation

Throughout, t is drawn uniformly from $[T, 2T]$ with $T \geq T_0$ sufficiently large.

2.1. Notation and Classical Inputs

Write $Y(t) = \log|\zeta(1/2+it)|$ and define Y_P, X as in Section 1.3. We use:

Label	Result	Reference
S1	$\mathrm{Var}(Y_P) = \frac{1}{2}\log\log T + O(1)$	Selberg (1946)
S2	$\kappa_2(Y) = \frac{1}{2}\log\log T + O(1)$	Selberg (1946)
Bd	$\lvert X(t)\rvert \le C_0 \Rightarrow \lvert\kappa_m(X)\rvert \le C_X$ for all m	elementary (uniformly bounded r.v.)
Ib	$\lvert\mathrm{Cov}(W_1, W_2)\rvert \le 4\alpha(W_1, W_2) \cdot B_1 B_2$ for $\lvert W_i\rvert \le B_i$	Ibragimov (1962)
Bi	$\lvert\kappa_{j,k}(W_1, W_2)\rvert \le C_{\mathrm{cgf}} \cdot \alpha$ for bounded α-mixing r.v.'s	Billingsley (1968, §20)
EP	$\lvert\kappa_m(Y_P)\rvert \le C_P := \sum_p p^{-m/2} < \infty$ for $m \ge 3$ (majorant; see discussion)	Montgomery–Vaughan (2007)
Sm	$X'(t) = O(1/T)$ on $[T, 2T]$ (from f' decay)	Smooth cutoff property
FTA	$\log p$ are $\mathbb{Q}$-linearly independent; $\min_p \log p = \log 2$	Euclid

2.2. The Spectral Gap

The smoothed prime sum $Y_P(t) = \mathrm{Re}\sum_{n\le T}\Lambda(n)f(n/T)n^{-1/2-it}/\log n$ has spectral content at frequencies $\{k\log p : p \text{ prime}, k \ge 1\}$, all $\ge \log 2$ (by FTA). The remainder $X(t) = Y(t) - Y_P(t)$ satisfies $X'(t) = O(1/T)$ on $[T, 2T]$: the Y' contribution is $O(\log T/T)$ (from the explicit formula) and the Y_P' contribution is $O(\log T/T)$ (from the smooth cutoff f'), giving $\mathrm{Var}_{\mathrm{eff}}(X) = O(1/T^2)$ on $[T, 2T]$.

2.3. Stage 1: Cauchy–Schwarz

By Cauchy–Schwarz:

$$|\mathrm{Cov}(Y_P, X)|^2 \le \mathrm{Var}(Y_P) \cdot \mathrm{Var}_{\mathrm{eff}}(X).$$

From S1 and Sm: $\mathrm{Var}(Y_P) \le C_{\mathrm{sel}} \cdot \log\log T$ and $T^2 \cdot \mathrm{Var}_{\mathrm{eff}}(X) \le C_f$ (smooth cutoff). Therefore

$$T^2 \cdot |\mathrm{Cov}|^2 \le C_{\mathrm{sel}} \cdot (\log\log T) \cdot C_f.$$

So $T \cdot |\mathrm{Cov}| = O(\sqrt{\log\log T})$ — this *grows*, so the covariance alone is insufficient.

2.4. Stage 2: The $\sqrt{\log\log T}$ Cancellation

Ibragimov (Ib) gives $|\mathrm{Cov}(Y_P, X)| \le 4\alpha(T) \cdot B_1(T) \cdot B_2$, where $B_1(T) = \max_{t\in[T,2T]}|Y_P(t) - E[Y_P]|$ and $B_2 = C_0$ (the pointwise bound on $|X|$ from the smooth cutoff). The Paley–Zygmund inequality, using a uniform fourth-moment bound for Y_P in the Selberg central-limit regime (equivalently, control of $\kappa_4(Y_P)$ from the classical CLT estimates, independent of the target moment order $2k$), gives

$$B_1(T)^2 \ge c \cdot \mathrm{Var}(Y_P) = c \cdot \tfrac{1}{2}\log\log T$$

for $T \geq T_0$. Substituting into the Cauchy–Schwarz bound:

$$c \cdot (\log\log T) \cdot T^2 \cdot 16B_2^2 \cdot \alpha^2 \leq C_{\mathrm{sel}} \cdot (\log\log T) \cdot C_f.$$

The $\log\log T$ factors cancel:

$$T^2 \cdot \alpha^2 \leq \frac{C_{\mathrm{sel}} \cdot C_f}{16c \cdot C_0^2} =: C_\alpha^2.$$

This is the central algebraic step. The Selberg variance $\sim \frac{1}{2}\log\log T$ appears in *both* the covariance (numerator, via Cauchy–Schwarz) and the range (denominator, via Ibragimov), and these contributions cancel exactly. The cancellation is structural: the same theorem (Selberg 1946) generates both factors.

Proposition 2.1. *$T \cdot \alpha(T) \leq C_\alpha$ for an absolute constant C_α.*

Remark 2.2. The mixing coefficient α must be specified with respect to the σ-algebras generated by Y_P and X on $([T, 2T], \mathrm{uniform})$; the inequalities cited from Ibragimov and Billingsley are standard for *stationary* sequences, and the reduction of the present deterministic setting to that framework is part of what must be checked in a full write-up.

2.5. Remark

The decorrelation in Proposition 2.1 is an instance of additive–correlative duality (Nagy 2026, *Additive–Correlative Duality*): spectrally separated additive components decorrelate, with the rate controlled by the spectral gap. The $\sqrt{\log\log T}$ cancellation is specific to ζ, arising because the Selberg variance governs both the signal amplitude and the Ibragimov normalization.

---

3. The Kronecker–Weyl Chain and Proof of the Theorem

3.1. Why Cumulants Bypass Soundararajan's ε

Soundararajan's resonator truncates at $y = T^\theta$, introducing error exponent $2k(1-\theta)$. Since $\theta < 1$, $\varepsilon > 0$ is inherent. Our approach bounds $\kappa_m(Y)$ (additive decomposition) instead of $E[|\zeta|^{2k}]$ (multiplicative), using the Euler product structure to control individual cumulants. No truncation occurs.

3.2. The Algebraic Chain

Step 1. Proposition 2.1: $\alpha(T) = O(1/T)$.

Step 2. Billingsley (Bi): for each $j, k \geq 1$ with $j + k = m$, $|\kappa_{j,k}(Y_P, X)| \leq C_{\mathrm{cgf}} \cdot \alpha = O(1/T)$. The cross-cumulant sum has at most $2^m - 2$ terms, so $\left|\sum_{j=1}^{m-1} \binom{m}{j} \kappa_{j,m-j}\right| \leq (2^m - 2) C_{\mathrm{cgf}} \cdot \alpha = O_m(1/T)$. For $T \geq T_0(m)$ this is ≤ 1.

Step 3. Leonov–Shiryaev: $\kappa_m(Y) = \kappa_m(Y_P) + \kappa_m(X) + \sum_{j=1}^{m-1} \binom{m}{j} \kappa_{j,m-j}$.

Step 4. EP + Bd + Step 2: $|\kappa_m(Y)| \leq C_P + C_X + 1 =: B$ for $m \geq 3$, $T \geq T_0(m)$.

Step 5. Grade-2 decomposition of the CGF:

$$K(z) := \log E[e^{zY}] = \kappa_1 z + \frac{\kappa_2}{2} z^2 + \underbrace{\sum_{m=3}^{\infty} \frac{\kappa_m}{m!} z^m}_{\text{tail}} .$$

Since $|\kappa_m| \leq B$ for $m \geq 3$, the tail converges for all $z \in \mathbb{C}$ with $|\text{tail}(z)| \leq B(e^{|z|} - 1 - |z| - |z|^2/2)$. In particular, K is *entire.*

3.3. Proof of the Theorem

With $\kappa_2 = \frac{1}{2} \log \log T + O(1)$ (S2), $\kappa_1 = O(1)$, and $|\text{tail}(2k)| \leq C_{\text{tail}}(k)$:

$$\begin{aligned} K(2k) &= \kappa_1 \cdot 2k + \frac{\kappa_2}{2} \cdot (2k)^2 + \text{tail}(2k) && (1) \\ &= O_k(1) + 2k^2 \kappa_2 + O_k(1) && (2) \\ &= 2k^2 \cdot \tfrac{1}{2} \log \log T + O_k(1) && (3) \\ &= k^2 \log \log T + O_k(1). && (4) \end{aligned}$$

Exponentiating: $m_{2k} = E[e^{2kY}] = e^{K(2k)} = C(k)(\log T)^{k^2}$, where $C(k) = e^{2k\kappa_1 + C_{\text{tail}}(k) + O(1)}$. □

---

4. Discussion

4.1. CGF Analyticity

The bounded cumulants imply $K(z)$ is entire — the moment generating function $M(z) = e^{K(z)}$ is an entire function of z. Under standard growth hypotheses on the moment sequence, such an M is associated with moment determinacy (Cramér-type theorems), Gaussian-type concentration inequalities for Y, and all-moment Gaussian limits compatible with Selberg's CLT. These consequences are conditional on the same analytic inputs as the main chain and are not pursued quantitatively here.

They would be strictly stronger than the moment upper bound alone and may have further applications if the full program is completed.

4.2. Implications

1. **Lindelöf Hypothesis** (conditional on the Theorem extending to non-integer k): For any $\varepsilon > 0$, $\zeta(1/2 + it) = O(t^{\varepsilon})$ for almost all t.
2. **Zero density**: A sharp k^2 upper bound is heuristically connected to density estimates for zeros of ζ near the critical line via standard mean-value-to-density arguments (Huxley–Jutila type); making this implication unconditional requires the full analytic theorem, not only the conditional chain here.
3. **Random matrix theory**: The k^2 exponent matches the Keating–Snaith (2000) prediction from GUE/CUE. The target statement would align the exponent with that prediction; matching the conjectured constants remains open.

4. **Toward RH**: An entire CGF and a stable grade-2 dominance picture would constrain the fine distribution of zeros; any bridge to the Riemann Hypothesis via the Lagarias canonical system (1999) and de Branges-type regularity is left to companion work and is not used in the conditional chain above.

4.3. Potential Objections

Q1: Billingsley applicability. Y_P and X are functions of $t \sim \mathrm{Unif}[T, 2T]$, not time-series variables. The α-mixing coefficient $\alpha(W_1, W_2) = \sup |P(A \cap B) - P(A)P(B)|$ is well-defined for any pair of random variables; the Ibragimov bound applies because X is pointwise bounded ($|X| \le C_0$) and Y_P has finite range on $[T, 2T]$.

Q2: Range growth rigor. $B_1^2 \ge c \cdot \mathrm{Var}(Y_P)$ follows from Paley–Zygmund once a suitable fourth-moment lower bound for Y_P is available in the fixed-T window (as in the Selberg CLT literature), without appealing to the final constant B in the tail estimate.

Q3: Why not noticed before? Soundararajan's 2009 result was regarded as essentially optimal. The insight that cumulants bypass his structural ε is new. The decorrelation between Y_P and X was assumed in various forms but the $\sqrt{\log\log T}$ cancellation was not identified.

4.4. What This Paper Does Not Prove

1. **Matching lower bounds.** The lower bound $m_{2k} \ge c(k)(\log T)^{k^2}$ for $k \ge 3$ remains open (shifted divisor problem).
2. **Exact constants.** $C(k)$ is not claimed to match the Keating–Snaith prediction.
3. **The Riemann Hypothesis.** The chain from the Theorem to RH requires additional framework-level input (see companion paper).

4.5. Circularity Check

None of the eight inputs (S1, S2, Bd, Ib, Bi, EP, Sm, FTA) assumes anything about the location of zeta zeros. The chain is non-circular relative to those inputs.

4.6. Machine Verification

The algebraic chain is verified in the accompanying proof kernel (42 theorems, 21 facts, 0 errors). Proof files are listed in the frontmatter.

During the preparation of this work the author used large language models to assist with manuscript drafting, literature search, and formalization support. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of this draft.

References

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

2. K. Soundararajan, “Moments of the Riemann zeta function,” *Ann. of Math.* **170** (2009), 981–993.
3. J. P. Keating and N. C. Snaith, “Random matrix theory and $\zeta(1/2+it)$,” *Comm. Math. Phys.* **214** (2000), 57–89.
4. I. A. Ibragimov, “Some limit theorems for stationary processes,” *Theory Probab. Appl.* **7** (1962), 349–382.
5. P. Billingsley, *Convergence of Probability Measures*, Wiley, 1968.
6. G. H. Hardy and J. E. Littlewood, “Contributions to the theory of the Riemann zeta-function,” *Acta Math.* **41** (1918), 119–196.
7. A. E. Ingham, “Mean-value theorems in the theory of the Riemann zeta-function,” *Proc. London Math. Soc.* **27** (1926), 273–300.
8. A. J. Harper, “Sharp conditional bounds for moments of the Riemann zeta function,” Preprint, arXiv:1305.4618, 2013.
9. M. Radziwiłł and K. Soundararajan, “Selberg’s central limit theorem for $\log|\zeta(1/2+it)|$,” *Enseign. Math.* **63** (2017), 1–19.
10. W. Heap and K. Soundararajan, “Lower bounds for moments of zeta and L-functions revisited,” *Mathematika* **68** (2022), 1–14.
11. H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge, 2007.
12. K.-M. Tsang, “The distribution of the values of the Riemann zeta-function,” PhD thesis, Princeton, 1984; see also *J. Reine Angew. Math.* **389** (1988), 1–17.
13. T. Nagy, “Additive–Correlative Duality: Why Convolution Is Easy and Correlation Is Hard,” Preprint, 2026.