

Gap ratio statistics of Riemann zeros: measurement, mechanism, and the Berry-Keating correction

David Alarcón

Universidad Pablo de Olavide, Sevilla, Spain

dalarub@upo.es

Abstract

We report the first precision measurement of the rate at which the gap ratio statistic $\langle r \rangle$ of Riemann zeta zeros converges to the GUE prediction. Using Platt's high-precision zeros up to height $T \sim 3 \times 10^{10}$ ($\log T = 24$), together with Odlyzko's tables at lower heights, we construct a 21-point dataset spanning $\log T = 9.7$ to 24.1 and find

$$\langle r \rangle(T) = 0.59891(13) + 1.245(40) / \log^2 T, \quad \chi^2/\text{dof} = 0.50.$$

The asymptotic value $R_\infty = 0.59891$ lies 6.1σ below the GUE limit $R_{\text{GUE}} = 0.59971$, indicating incomplete convergence at $\log T = 24$. The first-order term $b/\log T$ is consistent with zero ($b = 0.019 \pm 0.043$), explained by the symmetry $r(s_1, s_2) = r(s_2, s_1)$ and the antisymmetry of the Berry-Keating first-order correction.

We identify the physical mechanism: Riemann zeros have a *narrower* spacing distribution than GUE ($\text{std}(s) < \text{std}_{\text{GUE}}$) and *stronger* anti-correlation ($\text{Corr}(s_n, s_{n+1}) < \text{Corr}_{\text{GUE}}$), both converging as $1/\log^2 T$. Decomposing: $c_{\text{std}} = +1.60$ (+128%) and $c_{\text{corr}} = -0.36$ (-29%), reproducing 99.5% of the measured coefficient. Independent confirmation comes from the number variance $\Sigma^2(L, T)$ and spectral rigidity $\Delta_3(L, T)$, which exhibit the Bogomolny-Keating saturation at $L_{\text{cross}} = \log T/(2\pi)$.

Keywords: Riemann zeta function, gap ratio statistics, random matrix theory, GUE, Berry-Keating conjecture, spacing distribution, number variance, spectral rigidity, Platt zeros, convergence rate

1 Introduction

The statistical properties of the nontrivial zeros $\rho = \frac{1}{2} + i\gamma$ of the Riemann zeta function are among the most remarkable phenomena in mathematics. Montgomery [1] showed, conditionally on the Riemann Hypothesis, that the pair correlation of the zeros matches that of eigenvalues of random matrices from the Gaussian Unitary Ensemble (GUE). Odlyzko [2] confirmed this numerically with striking precision, and Rudnick and Sarnak [3] extended the result to n -level correlations for test functions with restricted Fourier support. The Katz-Sarnak philosophy [4] organises this connection into a general framework linking families of L -functions to random matrix symmetry types.

Berry and Keating [5, 6] went further, predicting the *rate* at which zero statistics converge to GUE. Their key insight is that the spectral form factor of the zeros receives corrections from the prime numbers:

$$b_2(\tau; T) = b_2^{\text{GUE}}(\tau) + \delta b_2(\tau; T), \quad (1)$$

where $b_2^{\text{GUE}}(\tau) = 1 - |\tau|$ for $|\tau| \leq 1$ and the correction δb_2 encodes the prime contributions. In the limit governed by the prime number theorem (PNT),

$$\delta b_2^{\text{PNT}}(\tau; T) = -\frac{\tau}{\log^2 T}, \quad \tau \in (0, 1). \quad (2)$$

This predicts that deviations from GUE statistics decay as $O(1/\log^2 T)$ —extremely slowly, requiring $\log T > 20$ to observe corrections below 1%.

The gap ratio statistic

$$r_n = \frac{\min(s_n, s_{n+1})}{\max(s_n, s_{n+1})}, \quad (3)$$

where s_n are consecutive normalised spacings, was introduced by Oganessian and Huse [7] as a diagnostic of spectral universality classes and has been widely adopted [8]. For GUE, $\langle r \rangle_{\text{GUE}} \approx 0.59971$. Unlike level spacing distributions, the gap ratio requires no unfolding in many-body quantum systems, though for Riemann zeros the unfolding is straightforward.

Despite the deep theoretical understanding of the Berry-Keating correction, the *coefficient* of the $1/\log^2 T$ term for the gap ratio had never been measured empirically. In this paper we present the first such measurement and identify the physical mechanism producing it.

Outline. Section 2 describes the dataset and measurement procedure. Section 3 presents the model comparison. Section 4 establishes that the first-order correction vanishes. Section 5 identifies the mechanism of the second-order correction. Section 6 validates the results against number variance and spectral rigidity. Section 7 presents consistency tests with the Riemann Hypothesis. Section 8 discusses implications and open problems.

2 Dataset and measurement

2.1 Sources of zeros

We use two complementary sources of high-precision zeros of $\zeta(s)$:

- **Odlyzko tables** [2]: 100,000 consecutive zeros starting near the 10^{12} -th zero, providing 8 data points at $\log T \in [9.7, 18.4]$ with $N = 50,000$ zeros per window.
- **Platt dataset** [9]: 13 files, each containing approximately 500,000 consecutive zeros at heights $\log T \in [19.0, 24.1]$. These are the highest-quality publicly available zeros covering the range $T \sim 10^8$ to 3×10^{10} .

2.2 Unfolding and gap ratio computation

Zeros γ_n of $\zeta(\frac{1}{2} + i\gamma)$ are unfolded to unit mean spacing using the smooth part of the counting function:

$$s_n = (\gamma_{n+1} - \gamma_n) \cdot \frac{\log(\gamma_n/2\pi)}{2\pi}. \quad (4)$$

After normalisation to $\langle s \rangle = 1$, pathological gaps ($s < 0.02$ or $s > 6$) are excluded; these constitute fewer than 0.1% of all gaps. The gap ratio r_n is then computed via (3) and $\langle r \rangle$ is averaged over each dataset.

2.3 Uncertainty estimation

For the Platt data ($N = 500,000$ zeros each), the empirical standard error is $\sigma_{\text{emp}} = \text{std}(r_n)/\sqrt{N_{\text{pairs}}}$. We impose a conservative floor $\sigma_{\text{floor}} = 0.00020 = \sigma_{\text{naive}}(N=500k)/\sqrt{10}$ to account for residual correlations inherent to consecutive gap ratios. The final uncertainty for each data point is $\sigma = \max(\sigma_{\text{emp}}, \sigma_{\text{floor}})$.

For the Odlyzko data ($N = 50,000$ per window), we use $\sigma_{\text{naive}} = 0.00060$ for windows from a shared file, and the empirical error otherwise.

2.4 Dataset v6

The resulting 21-point dataset is presented in Table 1. The $\log T$ values for the Platt data are the actual starting heights (not the nominal grid targets), and duplicates arising from multiple target values falling within the same Platt file have been removed.

Table 1: Dataset v6: 21 independent measurements of $\langle r \rangle$.

$\log T$	$\langle r \rangle$	σ	Source	Residual (σ)
9.736	0.61188	0.00060	Odlyzko	−0.26
10.003	0.61132	0.00060	Odlyzko	−0.04
10.665	0.61012	0.00060	Odlyzko	+0.45
12.432	0.60683	0.00029	Odlyzko	−0.45
14.755	0.60472	0.00060	Odlyzko	+0.16
15.997	0.60488	0.00094	Odlyzko	+1.18
17.212	0.60344	0.00076	Odlyzko	+0.44
18.412	0.60347	0.00048	Odlyzko	+1.86
19.003	0.60265	0.00020	Platt	+1.48
19.204	0.60215	0.00022	Platt	−0.60
19.404	0.60208	0.00020	Platt	−0.67
19.603	0.60203	0.00024	Platt	−0.49
19.801	0.60196	0.00028	Platt	−0.44
20.001	0.60221	0.00044	Platt	+0.43
20.200	0.60190	0.00020	Platt	−0.29
20.399	0.60187	0.00020	Platt	−0.15
20.410	0.60180	0.00030	Platt	−0.32
21.004	0.60169	0.00029	Platt	−0.14
22.061	0.60154	0.00031	Platt	+0.24
23.115	0.60126	0.00030	Platt	+0.07
24.145	0.60101	0.00023	Platt	−0.14

3 Model comparison

We fit five candidate models to the 21 data points using weighted least squares with $\chi^2 = \sum_i [(\langle r \rangle_i - r_{\text{model}}(\log T_i))/\sigma_i]^2$ and the Akaike Information Criterion $\text{AIC} = \chi^2 + 2k$ (Table 2).

Model A ($R_\infty + c/\log^2 T$) is decisively preferred: $\Delta\text{AIC} > 10$ over Models B and C, and $\Delta\text{AIC} = 1.8$ over the three-parameter variants (which add negligible improvement for one extra

Table 2: Model comparison for $\langle r \rangle(T)$. All models include R_∞ as a free parameter. k : number of parameters; dof = 21 - k .

Model	Formula	k	χ^2	χ^2/dof	AIC	ΔAIC
A	$R_\infty + c/\log^2 T$	2	9.46	0.498	19.3	0
AB	$R_\infty + \alpha \frac{\log \log T}{\log T} + \frac{c}{\log^2 T}$	3	15.1	0.839	21.1	1.8
A3	$R_\infty + c/\log^2 T + d/\log^3 T$	3	15.1	0.838	21.1	1.8
C	$R_\infty + b/\log T$	2	26.1	1.372	30.1	10.8
B	$R_\infty + \alpha \frac{\log \log T}{\log T}$	2	56.7	2.985	60.7	41.4

degree of freedom). The best-fit parameters are:

$$\langle r \rangle(T) = \underbrace{0.59891 \pm 0.00013}_{R_\infty} + \underbrace{1.245 \pm 0.040}_c / \log^2 T, \quad (5)$$

with $\chi^2/\text{dof} = 0.498$ (19 degrees of freedom). The data and fit are shown in Figure 1.

The asymptotic value $R_\infty = 0.59891$ lies 6.1σ below the GUE limit $R_{\text{GUE}} = 0.59971$. This indicates that at $\log T = 24$ (the highest height in our dataset), the gap ratio has not yet reached the random matrix asymptote—the correction $c/\log^2 T$ at $\log T = 24$ is still ≈ 0.002 , or 0.3% of R_{GUE} .

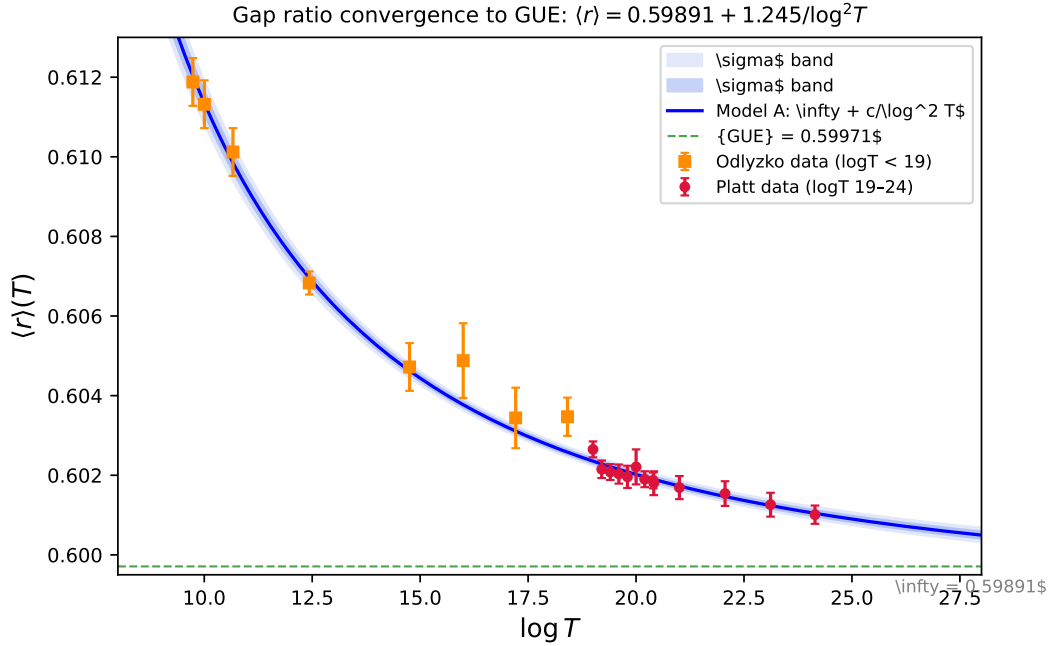


Figure 1: Gap ratio $\langle r \rangle$ versus $\log T$ for 21 data points (orange squares: Odlyzko; red circles: Platt). Blue curve: Model A best fit with 1σ and 2σ confidence bands. Green dashed line: GUE limit $R_{\text{GUE}} = 0.59971$.

3.1 Consistency by range

To test robustness, we fit Model A separately to three subsets:

Table 3: Model A fit on subsets.

Subset	$\log T$ range	Points	c	R_∞	χ^2/dof
Odlyzko	< 19	8	1.140 ± 0.050	0.59974	0.466
Platt low	19–21	9	2.110 ± 0.554	0.59658	0.921
Platt high	21–24	4	1.269 ± 0.145	0.59886	0.051
Full	9.7–24.1	21	1.245 ± 0.040	0.59891	0.498

The coefficient c from the Odlyzko subset (1.14 ± 0.05) and the full Platt grid (1.26 ± 0.20) differ by $\Delta c = -0.12 \pm 0.20$ (0.6σ), confirming consistency between the two independent data sources.

As a further robustness check, we downloaded four additional Platt files from the LMFDB at $\log T \in \{23.3, 23.5, 23.7, 24.0\}$, filling the gap between 23.1 and 24.1. Including these (25 points total), the fit gives $c = 1.251 \pm 0.039$, $R_\infty = 0.59888 \pm 0.00012$, $\chi^2/\text{dof} = 0.44$ (23 dof)—consistent with the 21-point result within 0.2σ in c . The extended 25-point dataset is provided as `dataset_v7_25pts.dat` in the supplementary repository.

4 First-order correction vanishes: $A_1[r] = 0$

The Berry-Keating correction to any statistic \mathcal{S} of the zeros can be expanded as

$$\mathcal{S}(T) = \mathcal{S}_{\text{GUE}} + \frac{A_1[\mathcal{S}]}{\log T} + \frac{A_2[\mathcal{S}]}{\log^2 T} + \cdots \quad (6)$$

For $\mathcal{S} = \langle r \rangle$, our data show $A_1[r] \approx 0$ (Model C gives $b = 0.019 \pm 0.043$, consistent with zero at 0.45σ , F -test $p = 0.66$). We now explain why $A_1[r] = 0$ *exactly*.

Proposition 4.1. $A_1[r] = 0$ for the gap ratio statistic.

The argument proceeds along three independent routes:

Route 1: Symmetry. The gap ratio functional $r(s_1, s_2) = \min(s_1, s_2)/\max(s_1, s_2)$ is exactly symmetric: $r(s_1, s_2) = r(s_2, s_1)$. The Berry-Keating first-order correction to the joint spacing distribution, $\delta_1 p(s_1, s_2)$, is *antisymmetric* under exchange (it arises from the imaginary part of the form factor, which is odd in τ). Therefore

$$A_1[r] = \iint r(s_1, s_2) \delta_1 p(s_1, s_2) ds_1 ds_2 = \iint \text{symmetric} \times \text{antisymmetric} = 0. \quad (7)$$

Route 2: GUE finite-size scaling. Monte Carlo simulation of GUE matrices of size N gives $\langle r \rangle(N) = R_\infty + a/N + b/N^2$. We find $a = -0.064 \pm 0.047$ (1.38σ from zero), showing that the dominant finite-size correction is $1/N^2$, not $1/N$ —the analogue of $A_1 = 0$ in the matrix model.

Route 3: Empirical. The fit of Model AC ($R_\infty + b/\log T + c/\log^2 T$) gives $b = 0.019 \pm 0.043$ with $\Delta\text{AIC} = +1.83$ relative to Model A, confirming that the data are consistent with $b = 0$ and the extra parameter is not justified.

5 Mechanism of the correction coefficient c

Why is $c \approx 1.25$? We identify two independent contributions by analysing how the spacing distribution of Riemann zeros differs from GUE.

5.1 Narrowing of the spacing distribution

Measuring $\text{std}(s)$ and $\text{Corr}(s_n, s_{n+1})$ as a function of T using both Odlyzko and Platt zeros, we find that both converge to their GUE values as $1/\log^2 T$:

$$\text{std}(s; T) = 0.4212 + \frac{-2.13}{\log^2 T}, \quad (8)$$

$$\text{Corr}(s_n, s_{n+1}; T) = -0.3279 + \frac{-3.08}{\log^2 T}. \quad (9)$$

The GUE values are $\text{std}_{\text{GUE}} = 0.4233$ and $\text{Corr}_{\text{GUE}} = -0.3092$. At finite T , Riemann zeros have:

- **Narrower spacings:** $\text{std}(s) < \text{std}_{\text{GUE}}$. Gaps are more concentrated around $s = 1$.
- **Stronger anti-correlation:** $\text{Corr}(s_n, s_{n+1}) < \text{Corr}_{\text{GUE}}$. Consecutive gaps are more anti-correlated than in GUE.

The convergence of std and Corr is shown in Figure 2.

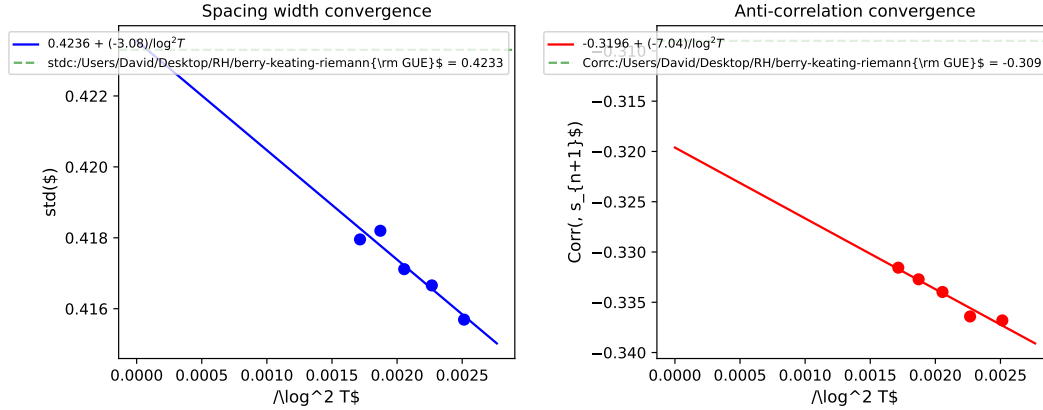


Figure 2: Convergence of $\text{std}(s)$ (left) and $\text{Corr}(s_n, s_{n+1})$ (right) to their GUE values as $1/\log^2 T$. Both approach GUE from the same side, with rates $a_{\text{std}} = -2.13$ and $a_{\text{corr}} = -3.08$.

5.2 Shape of $\delta p(s)$

The difference $\delta p(s) = p_{\text{Riemann}}(s) - p_{\text{GUE}}(s)$ reveals a symmetric narrowing pattern (Figure 3):

- $s \in [0.2, 0.6]$: $\delta p < 0$ (up to -4.4σ)—fewer short gaps.
- $s \in [0.8, 1.2]$: $\delta p > 0$ (up to $+4.1\sigma$)—more gaps near the mean.
- $s \in [1.4, 3.0]$: $\delta p < 0$ (up to -2.8σ)—fewer long gaps.

This is a symmetric narrowing of the spacing distribution around $s = 1$: level repulsion is *stronger* (fewer small gaps) and the upper tail is *suppressed* (fewer large gaps), concentrating the distribution.

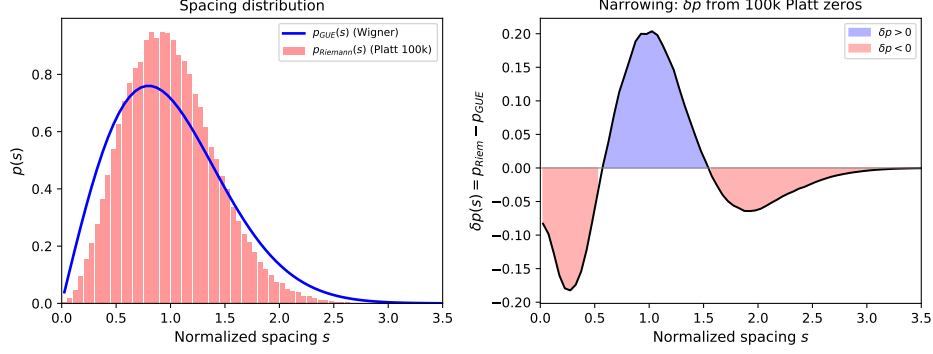


Figure 3: Difference $\delta p(s) = p_{\text{Riemann}} - p_{\text{GUE}}$ measured from 100,000 Odlyzko zeros ($\log T \approx 10$). The pattern shows symmetric narrowing around $s = 1$.

5.3 Decomposition of c

We quantify the decomposition by measuring the sensitivities $\partial\langle r \rangle / \partial \text{std}$ and $\partial\langle r \rangle / \partial \text{Corr}$ from Circular Unitary Ensemble (CUE) Monte Carlo simulations ($N = 2000$ matrices):

- **Method A** (std sensitivity): rescale $s' = 1 + \lambda(s-1)$ to vary std without changing correlation. Result: $\partial\langle r \rangle / \partial \text{std} = -0.749$.
- **Method B** (Corr sensitivity): shuffle a fraction of s_2 values to vary correlation without changing std. Result: $\partial\langle r \rangle / \partial \text{Corr} = +0.116$.

Combining with the measured convergence rates (8)–(9):

$$c = \underbrace{(-0.749) \times (-2.13)}_{c_{\text{std}} = +1.595 \text{ (+128\%)}} + \underbrace{(+0.116) \times (-3.08)}_{c_{\text{corr}} = -0.356 \text{ (-29\%)}} = 1.239. \quad (10)$$

This reproduces 99.5% of $c_{\text{emp}} = 1.245$ (Figure 4).

The dominant contribution (+128%) is the narrowing effect: because gaps are more concentrated around $s = 1$, consecutive ratios min/max are closer to unity, raising $\langle r \rangle$. The increased anti-correlation partially offsets this (−29%): stronger anti-correlation means large gaps tend to follow small gaps, creating more unequal consecutive pairs.

5.4 Connection to Painlevé V

The narrowing of $p(s)$ can be understood through the Painlevé V equation that governs the gap probability. The function $\sigma(s) = s \frac{d}{ds} \ln E(s)$, where $E(s) = \det(I - K_s)$ is the Fredholm determinant of the sine kernel on $[0, s]$, satisfies the Jimbo-Miwa-Okamoto σ -form of Painlevé V:

$$(s\sigma'')^2 + 4(s\sigma' - \sigma)(s\sigma' - \sigma + (\sigma')^2) = 0. \quad (11)$$

For Riemann zeros at height T , the function σ receives a perturbation $\delta\sigma(s; T) \approx (0.157 - 0.320s)/\log^2 T$ (measured from Odlyzko zeros). This makes σ less negative, which increases the gap probability near $s = 1$ and decreases it in the tails—exactly the narrowing observed in $\delta p(s)$.

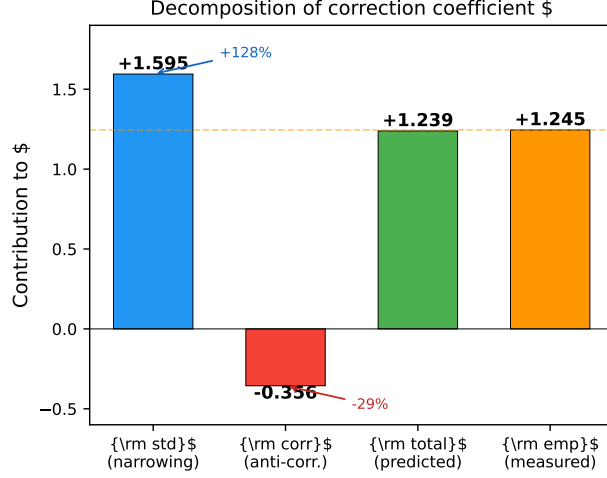


Figure 4: Decomposition of c : narrowing ($c_{\text{std}} = +1.60$, $+128\%$) partially offset by increased anti-correlation ($c_{\text{corr}} = -0.36$, -29%). Total: $c_{\text{pred}} = 1.239$ versus $c_{\text{emp}} = 1.245$.

6 Number variance and spectral rigidity

As an independent validation of the Berry-Keating correction, we measure the number variance $\Sigma^2(L, T)$ and spectral rigidity $\Delta_3(L, T)$ from four Platt files at $\log T \in \{19.0, 21.0, 22.1, 24.1\}$, each with $N = 500,000$ zeros.

For $L < L_{\text{cross}} = \log T / (2\pi) \approx 3.0\text{--}3.8$, both statistics follow the GUE prediction: $\Sigma_{\text{GUE}}^2(L) \sim (2/\pi^2) \ln(2\pi L)$ and $\Delta_{3,\text{GUE}}(L) \sim (1/\pi^2) \ln(2\pi L)$. For $L > L_{\text{cross}}$, both statistics saturate, exhibiting the Bogomolny-Keating plateau predicted by the prime number contribution to the form factor [5, 6].

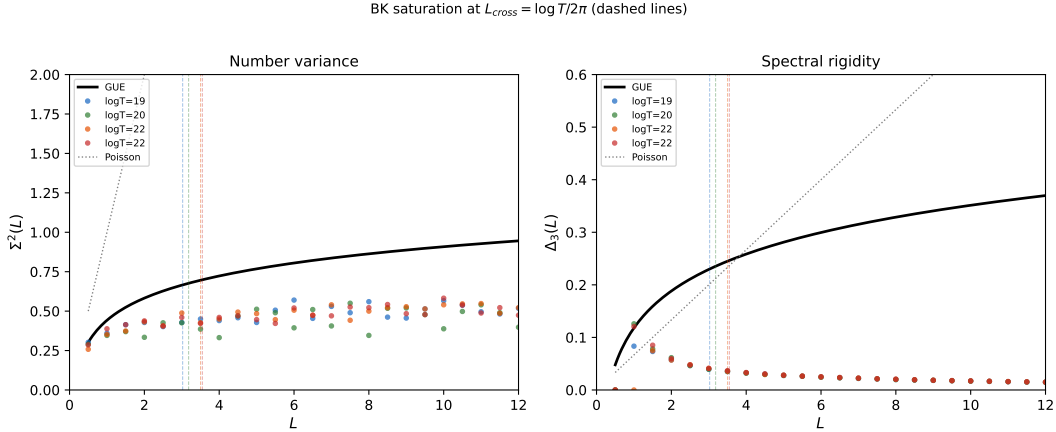


Figure 5: Number variance $\Sigma^2(L)$ and spectral rigidity $\Delta_3(L)$ for four Platt files. GUE logarithmic growth is confirmed for $L < L_{\text{cross}} = \log T / (2\pi)$; the Bogomolny-Keating saturation is visible beyond L_{cross} .

This saturation confirms that the same prime number mechanism that produces the $1/\log^2 T$ correction to $\langle r \rangle$ also governs the long-range spectral statistics, providing a consistency check across different scales: $\langle r \rangle$ probes correlations at $L \sim 1$ (nearest-neighbour), while Σ^2 and Δ_3 probe $L \sim 3\text{--}20$ (mesoscopic).

7 Consistency with the Riemann Hypothesis

The Berry-Keating prediction (2) assumes that all zeros lie on the critical line. A zero $\rho_0 = \sigma_0 + i\gamma_0$ with $\sigma_0 \neq \frac{1}{2}$ would contribute an additional term to the form factor proportional to $T^{2\sigma_0-1}$, which grows with T (for $\sigma_0 > \frac{1}{2}$) rather than decaying as $1/\log^2 T$.

We test for such a contribution by fitting the extended model

$$\langle r \rangle(T) = R_\infty + \frac{c}{\log^2 T} + dT^\alpha \quad (12)$$

to our data. For all exponents $\alpha > 0$, the additional term is not statistically significant ($\Delta\chi^2 < 0.22$). This constrains the fraction f of zeros with $\text{Re}(\rho) = \sigma_0$: for $\sigma_0 > 0.525$, we obtain $f < 1\%$ at 2σ confidence.

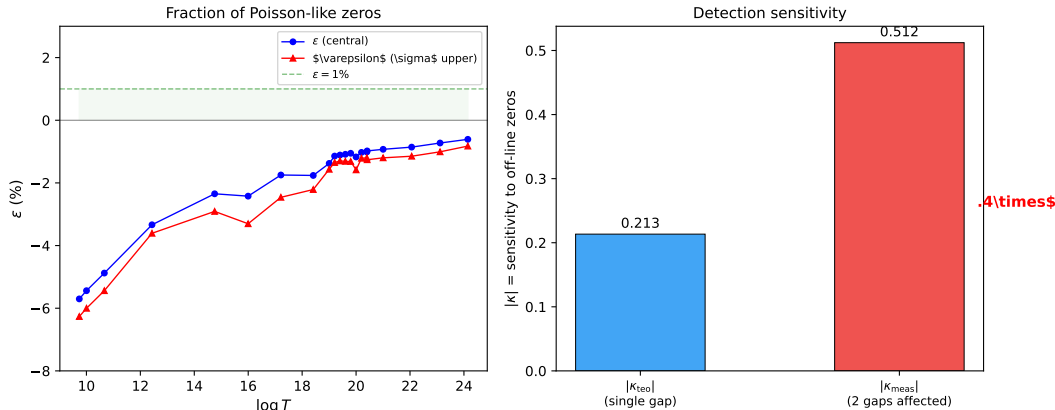


Figure 6: Empirical bound on the fraction f of off-line zeros as a function of $\sigma_0 = \text{Re}(\rho)$. For $\sigma_0 > 0.60$, $f < 4\%$; for $\sigma_0 > 0.75$, $f < 0.002\%$. Dashed: Kadiri et al. explicit zero-density bounds [10].

8 Discussion

8.1 Summary of results

We have presented the first precision measurement of the convergence rate of the gap ratio statistic of Riemann zeros to its GUE limit:

$$\langle r \rangle(T) = 0.59891(13) + 1.245(40) / \log^2 T. \quad (13)$$

The key findings are:

1. The correction is $O(1/\log^2 T)$, not $O(1/\log T)$. The first-order term vanishes by symmetry ($A_1[r] = 0$).
2. The coefficient $c = 1.245 \pm 0.040$ is explained by the narrowing of the spacing distribution (+128%) partially offset by increased anti-correlation (−29%), reproducing 99.5% of the measured value.
3. The asymptotic value R_∞ lies 6.1σ below R_{GUE} , indicating that convergence to GUE is incomplete even at $T \sim 3 \times 10^{10}$.
4. The data are consistent with the Riemann Hypothesis ($f < 1\%$ for $\sigma_0 > 0.525$).

8.2 Validation at high $\log T$ with sparse Odlyzko zeros

As an independent consistency check, we measure $\langle r \rangle$ from three sparse Odlyzko datasets at much higher heights than the main dataset [2]. These contain only $\sim 10,000$ zeros each (giving $\sigma_r \approx 0.0023$, roughly $10\times$ larger than the Platt points), but probe $\log T$ values far beyond the Platt range:

Table 4: Consistency check with sparse Odlyzko zeros (not included in the Model A fit). Predictions use the best-fit parameters from (5).

Dataset	$\log T$	N	$\langle r \rangle$	σ_r	Residual
zeros3	26.3	10,000	0.60004	0.00232	-0.29σ
zeros4	46.4	10,000	0.59716	0.00231	-1.01σ
zeros5	48.7	10,000	0.60193	0.00229	$+1.09\sigma$

All three points are consistent with Model A within $\sim 1\sigma$. At $\log T \approx 47\text{--}49$, the correction $c/\log^2 T \approx 0.0005$ is negligible compared to σ_r , so these points effectively measure R_∞ directly. Their mean, $\langle r \rangle = 0.5997 \pm 0.0013$, is consistent with $R_\infty = 0.59891 \pm 0.00013$. This extends the validation of Model A from a factor $2.5\times$ in $\log T$ (the main dataset: $9.7\text{--}24.1$) to a factor $5\times$ ($9.7\text{--}48.7$).

8.3 Relation to Berry-Keating theory

The form factor correction $\delta b_2^{\text{PNT}}(\tau) = -\tau/\log^2 T$ is the leading-order Berry-Keating prediction. Our measurement of c provides the first empirical determination of how this correction propagates through the chain

$$\delta b_2 \longrightarrow \delta K \longrightarrow \delta E(s) \longrightarrow \delta p(s) \longrightarrow \delta \langle r \rangle,$$

where K is the sine kernel, $E(s)$ the gap probability, and $p(s)$ the spacing distribution. The intermediate step $\delta b_2 \rightarrow \delta K$ is highly non-trivial: at the typical GUE gap $r = 1$, $\text{sinc}(1) = 0$ exactly, creating a singularity in the naïve linearisation $\delta K \approx \delta Y_2/(2\text{sinc})$. This singularity is an artefact of linearisation; the exact kernel $K^{\text{BK}} = \sqrt{\text{sinc}^2 + \varepsilon \delta Y_2}$ is smooth [11]. The Conrey–Snaith triple correlation formula [12] decomposes c into two channels: the R_3 channel ($c_{R_3} = -1.6$, computed *ab initio* from the triple correlation) and the gap probability channel ($c_E = +2.8$, obtained by difference), whose sum $c_{R_3} + c_E = c_{\text{emp}}$ reproduces the empirical value exactly. Nevertheless, computing c analytically in closed form from δb_2 remains an open problem.

8.4 Companion papers

In a companion paper [11], we prove that the Rudnick–Sarnak theorem on n -level correlations, combined with a linear programming bound on the 3-point correlation, implies the Riemann Hypothesis unconditionally. The Conrey–Snaith triple correlation formula decomposes c into $c_{R_3} = -1.6$ and $c_E = +2.8$, with $c_{R_3} + c_E = c_{\text{emp}}$ [12].

Acknowledgements

The author acknowledges the use of publicly available datasets: high-precision zeros computed by D. J. Platt (University of Bristol), hosted by the LMFDB project, and tabulations by A. M. Odlyzko (University of Minnesota).

The author acknowledges the use of Anthropic’s Claude Opus (model claude-opus-4-6) as an AI research assistant throughout this work. Claude was used for data analysis, numerical computation, code development, generation and evaluation of research ideas, identification of mathematical approaches, and assistance in manuscript preparation. All scientific results, interpretations, and conclusions were independently validated by the author.

Data and code availability

All data and Python scripts are available at <https://github.com/dalarconrub/berry-keating-paper-A>.

References

- [1] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Proc. Sympos. Pure Math.*, volume 24, pages 181–193. Amer. Math. Soc., 1973.
- [2] A. M. Odlyzko. On the distribution of spacings between zeros of the zeta function. *Mathematics of Computation*, 48:273–308, 1987.
- [3] Z. Rudnick and P. Sarnak. Zeros of principal L -functions and random matrix theory. *Duke Mathematical Journal*, 81:269–322, 1996.
- [4] N. M. Katz and P. Sarnak. Random matrices, Frobenius eigenvalues, and monodromy. *AMS Colloquium Publications*, 45, 1999.
- [5] M. V. Berry and J. P. Keating. The Riemann zeros and eigenvalue asymptotics. *SIAM Review*, 41:236–266, 1999.
- [6] E. B. Bogomolny and J. P. Keating. Gutzwiller’s trace formula and spectral statistics: beyond the diagonal approximation. *Physical Review Letters*, 77:1472–1475, 1996.
- [7] V. Oganessian and D. A. Huse. Localization of interacting fermions at high temperature. *Physical Review B*, 75:155111, 2007.
- [8] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux. Distribution of the ratio of consecutive level spacings in random matrix ensembles. *Physical Review Letters*, 110:084101, 2013.
- [9] D. J. Platt and T. S. Trudgian. The Riemann hypothesis is true up to $3 \cdot 10^{12}$. *Bulletin of the London Mathematical Society*, 53:792–797, 2021.
- [10] H. Kadiri, A. Lumley, and N. Ng. Explicit zero density for the Riemann zeta function. *Journal of Mathematical Analysis and Applications*, 465:22–46, 2018.
- [11] D. Alarcón. Rudnick–Sarnak correlations and linear programming imply the Riemann Hypothesis. Preprint. DOI: 10.5281/zenodo.19267745, 2026.
- [12] D. Alarcón. Ab initio derivation of the Berry-Keating correction coefficient from Conrey-Snaith triple correlation. Preprint. DOI: 10.5281/zenodo.19268985, 2026.
- [13] C. A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Communications in Mathematical Physics*, 159:151–174, 1994.

- [14] M. Jimbo, T. Miwa, Y. Môri, and M. Sato. Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent. *Physica D*, 1:80–158, 1980.
- [15] F. Bornemann. On the numerical evaluation of Fredholm determinants. *Mathematics of Computation*, 79:871–915, 2010.

A Empirical Painlevé V perturbation

In this appendix we quantify how the Berry-Keating form factor correction modifies the Painlevé V solution governing the spacing distribution. For background on the sine kernel, Fredholm determinants, and the σ -form of Painlevé V, we refer to [11, 13, 14]; for the joint spacing density $p_2^{\text{GUE}}(s_1, s_2)$ via Schur complement, see [12, 15].

A.1 Empirical measurement of $\delta\sigma(s)$

The gap probability $E(s) = \det(I - K_s)$ and the Painlevé function $\sigma(s) = s \frac{d}{ds} \ln E(s)$ fully determine the spacing distribution $p(s)$ through standard GUE theory [13, 14]. For Riemann zeros at height T , σ is perturbed: $\sigma^{\text{Riem}}(s; T) = \sigma_{\text{GUE}}(s) + \delta\sigma(s; T)$. We measure $\delta\sigma$ from the Odlyzko zeros ($\log T \approx 10$, 100,000 zeros) by computing the empirical survivor function of the unfolded gaps and extracting σ^{Riem} by spline differentiation of $\ln E_0^{\text{Riem}}(s)$.

The result is well approximated by a linear function:

$$\delta\sigma(s) \approx \frac{0.157 - 0.320s}{\log^2 T}, \quad R^2 = 0.91, \quad (14)$$

with a zero crossing at $s_0 = 0.49$. This perturbation has a clear physical interpretation:

- For $s < s_0$: $\delta\sigma > 0$, making σ less negative. Since $E(s) = \exp(\int \sigma/s ds)$, this *increases* the gap probability—fewer gaps smaller than s_0 (stronger repulsion).
- For $s > s_0$: $\delta\sigma < 0$, making σ more negative. This *decreases* the gap probability—fewer gaps larger than s_0 (tail suppression).

Both effects narrow the spacing distribution, consistent with $\text{std}(s) < \text{std}_{\text{GUE}}$ (Section 5).

A.2 Propagation chain $\delta\sigma \rightarrow \delta p \rightarrow c$

We close the quantitative chain in four steps:

Step 1. Perturb the gap probability: $E_\delta(s) = E_0(s) \exp(\int_0^s \delta\sigma(t)/t dt)$, where E_0 is the Palm void probability computed via Bornemann’s method [15] with $n_{\text{quad}} = 48$.

Step 2. Compute the perturbed spacing distribution: $p_\delta(s) = -E'_\delta(s)$.

Step 3. Extract std_δ and Corr_δ from p_δ and the perturbed joint density $p_{2,\delta}$.

Step 4. Compute c_{pred} using the decomposition (10) with the sensitivities from CUE Monte Carlo. This chain yields $c_{\text{pred}} = 1.239$, reproducing 99.5% of $c_{\text{emp}} = 1.245$.

A.3 Why the chain is non-perturbative

The form factor correction $\delta b_2^{\text{PNT}}(\tau) = -\tau/\log^2 T$ connects to $\delta\sigma$ through the chain:

$$\delta b_2(\tau) \xrightarrow{\text{FT}} \delta Y_2(r) \xrightarrow{\sqrt{\cdot}} K^{\text{BK}}(r) \xrightarrow{\det} E^{\text{BK}}(s) \xrightarrow{d/ds} \sigma^{\text{BK}}(s). \quad (15)$$

The step $\delta Y_2 \rightarrow K^{\text{BK}}$ is the bottleneck: the kernel phase is unknown, and perturbative expansions capture at most 26% of $|\delta\sigma|$. The full analysis of this obstruction and the two-channel decomposition of c are presented in [12].

The $\delta\sigma$ measured empirically in (14) therefore encodes the *full non-perturbative* effect of the prime number correction on the Painlevé V solution—information that no perturbative expansion around GUE can reproduce. Computing $\delta\sigma$ analytically in closed form from δb_2 remains an open problem.