Received: 12 April 2024 | Revised: 27 July 2024 | Accepted: 18 October 2024

DOI: 10.1112/mtk.70002

RESEARCH ARTICLE

Mathematika

Lower bounds for the large deviations of Selberg's central limit theorem

Louis-Pierre Arguin[1,2] | **Emma Bailey**[3]

[1]Mathematical Institute, University of Oxford, Oxford, UK

[2]Department of Mathematics, Baruch College and Graduate Center, City University of New York, New York, New York, USA

[3]Department of Mathematics, Graduate Center, City University of New York, New York, New York, USA

Correspondence
Emma Bailey, Department of Mathematics, Graduate Center, City University of New York, NY, USA.
Email: ebailey@gc.cuny.edu

Funding information
NSF, Grant/Award Number: DMS 2153803

Abstract

Let $\delta > 0$ and $\sigma = \frac{1}{2} + \frac{\delta}{\log T}$. We prove that, for any $\alpha > 0$ and $V \sim \alpha \log\log T$ as $T \to \infty$,

$$\frac{1}{T}\mathrm{meas}\big\{t \in [T, 2T] \,:\, \log|\zeta(\sigma + \mathrm{i}\tau)| > V\big\}$$

$$\geqslant C_\alpha(\delta)\int_V^\infty \frac{e^{-y^2/\log\log T}}{\sqrt{\pi \log\log T}}\mathrm{d}y,$$

where δ is large enough depending on α. The result is unconditional on the Riemann hypothesis. As a consequence, we recover the sharp lower bound for the moments on the critical line proved by Heap and Soundararajan and Radziwiłł and Soundararajan. The constant $C_\alpha(\delta)$ is explicit and is compared to the one conjectured by Keating and Snaith for the moments.

MSC 2020

11M06 (primary), 11M50, 60B20, 60G70 (secondary)

1 | MAIN RESULTS

The Riemann zeta function is defined on the half-plane Re $s > 1$ by

$$\zeta(s) = \sum_{n\geqslant 1} n^{-s} = \prod_{p \text{ prime}} (1 - p^{-s})^{-1}.$$

It can be continued to a meromorphic function on the whole complex plane $\mathbb{C}$ with a pole at $s = 1$. The Riemann hypothesis states that all nontrivial zeros of the function lie on the critical line Re $s = 1/2$. This paper is concerned with the distribution of large values of the function on the critical line and slightly to its right.

A good point of reference for the distribution of values of the function on the line is Selberg's Central Limit Theorem: if τ is a random point sampled uniformly in $[T, 2T]$, then we have the following convergence in distribution:

$$\frac{\log|\zeta(1/2+\mathrm{i}\tau)|}{\frac{1}{2}\sqrt{\log\log T}} \to \mathcal{N}(0,1), \quad T\to\infty, \tag{1}$$

for $\mathcal{N}(0,1)$, a standard Gaussian random variable. It is not hard to see from the proof, see, for example, [18], that the result also holds slightly off-axis, that is, Re $s = \frac{1}{2} + \frac{\delta}{\log T}$, $\delta > 0$. An important question, because of its relation to the moments of ζ and to the Lindelöf hypothesis, is to see if the distribution of the values of $\log|\zeta|$ remains Gaussian-like for values of the order of $\log\log T$, thus well beyond the standard deviation $\sqrt{\log\log T}$. It was established in [1] for $V \sim \alpha\log\log T$, $\alpha\in(0,2)$, that

$$\mathbf{P}(\log|\zeta(1/2+\mathrm{i}\tau)| > V) \leqslant K_\alpha \int_V^\infty \frac{e^{-y^2/\log\log T}}{\sqrt{\pi\log\log T}}\mathrm{d}y, \tag{2}$$

for some constant K_α that depends on α but not T, and where $\mathbf{P}$ stands for the uniform probability on τ. This, in turn, implied another proof for a sharp upper bound for the fractional moments of ζ between 0 and 4 as first proved in [9]. The large deviations of the imaginary part of $\log\zeta$ for a wide range values are investigated in [7].

Here, we prove an unconditional lower bound off-axis matching (2) up to constant.

Theorem 1. *Let $\alpha > 0$ and $V \sim \alpha\log\log T$. Consider τ sampled uniformly on $[T, 2T]$. There exists $\delta^\star(\alpha)$ such that if $\delta > \delta^\star(\alpha)$, we have*

$$\mathbf{P}(\log|\zeta(1/2+\delta/\log T+\mathrm{i}\tau)| > V) \geqslant C_\alpha(\delta) \int_V^\infty \frac{e^{-y^2/\log\log T}}{\sqrt{\pi\log\log T}}\mathrm{d}y,$$

for some $C_\alpha(\delta) > 0$.

The constant $C_\alpha(\delta)$ is explicit, see Equation (19). So is $\delta^\star$, as given in Equation (33). For α tending to ∞, $\delta^\star(\alpha)$ is of order $\log\alpha$, whereas $\delta^\star(\alpha)$ remains bounded away from 0 as $\alpha\to 0$. (For technical reasons, δ is redefined in (13), but this does not change the asymptotic behavior.) Our method of proof does not seem to apply at $\delta = 0$, although it does imply sharp lower bound on the moments on the critical line. More precisely, it was conjectured by Keating and Snaith [13] that for all $k \geqslant 0$

$$\frac{1}{T}\int_T^{2T} |\zeta(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t = \mathbf{E}\big[|\zeta(1/2+\mathrm{i}\tau)|^{2k}\big] \sim C_k\,(\log T)^{k^2}, \tag{3}$$

where $C_k = a_k f_k$. The *arithmetic part* a_k is of the form

$$a_k = \prod_p (1-p^{-1})^{k^2} \sum_{m\geqslant 0}\left(\frac{\Gamma(k+m)}{m!\Gamma(k)}\right)^2 p^{-m} \tag{4}$$

(see also Equation (B.14)), and the *random matrix part* is given by

$$f_k = \frac{G^2(1+k)}{G(1+2k)}, \tag{5}$$

where G is the Barnes G-function. Theorem 1 yields the following.

Corollary 1. *For $k \geqslant 0$ fixed, we have*

$$\mathbf{E}\left[|\zeta(1/2+\mathrm{i}\tau)|^{2k}\right] \geqslant c_k(\log T)^{k^2},$$

where the constant c_k is proportional to

$$a_k e^{\gamma k^2}(k+1)^{-38k^2}. \tag{6}$$

Sharp lower bounds of this kind were proved in [10] and [17]. The constant obtained by Heap and Soundararajan is better for large k as it behaves like k^{-2k^2}. For k approaching 0, their constant tends to 0, whereas the one in Corollary 1 is bounded away from it, as it should. Although, as mentioned in their paper, a modification of their argument would lead to the correct behavior for small k. Note that the constant c_k includes explicitly the arithmetic part a_k, and this is, in fact, necessary in our argument. However, this does not change the asymptotic behavior for large k, since as shown in Conrey and Gonek [6], $\log a_k \sim -k^2 \log\log k$ as $k \to \infty$. The correct asymptotic behavior of f_k is $\log f_k \sim -k^2 \log k$. It is unclear to us at this time how to improve the constant c_k, though a better approximation of indicator functions in terms of polynomials seems necessary, cf. Lemma 2 and the discussion following Equation (19). It is important to note that the same constant C_k as for the moments should appear in the large deviation of $\log|\zeta|$. Indeed, it was conjectured by Radziwiłł that for any $\alpha > 0$ and $V \sim \alpha \log\log T$ as $T \to \infty$,

$$\mathbf{P}(\log|\zeta(1/2+\mathrm{i}\tau)| > V) \sim C_\alpha \int_V^\infty \frac{e^{-y^2/\log\log T}}{\sqrt{\pi \log\log T}}\mathrm{d}y. \tag{7}$$

The paper is organized as follows. The proof of Theorem 1 is explained in Section 2. It is divided in four propositions that are proved in Section 3. Corollary 1 is proved in Section 4. Finally, the Appendix contains results needed in the proof that have appeared elsewhere in the literature. One notable exception is the mollification result stated in Lemma A.1 that may be of independent interest. It was developed with one of the authors and Bourgade and Radziwiłł in an earlier version of [4].

Notation. We write $f(T) = \mathrm{O}(g(T))$ whenever $\limsup_{T\to\infty} |f(T)/g(T)| < \infty$ with other parameters fixed, such as α and k. We also use Vinogradov's notation $f(T) \ll g(T)$ whenever $f(T) = \mathrm{O}(g(T))$, and $f(T) \asymp g(T)$ if $f(T) \ll g(T)$ and $f(T) \gg g(T)$.

2 STRUCTURE OF THE PROOF OF THEOREM 1

2.1 Some definitions

As in [1], the proof relies on approximating $\log|\zeta|$ by Dirichlet polynomials of the form:

$$S_\ell = \sum_{T_0 < p \leqslant T_\ell} \frac{\mathrm{Re}\, p^{-\mathrm{i}\tau}}{p^\sigma} + \frac{\mathrm{Re}\, p^{-2\mathrm{i}\tau}}{2p^{2\sigma}}, \quad \ell > 0. \tag{8}$$

Throughout we use the probability convention of dropping τ in the notation. Also, σ will be fixed and its dependence often omitted. The small primes need to be treated with care to get a good estimate on the constant. For the primes below T_0, we consider $|\mathcal{M}_0^{-1}|$ where

$$\mathcal{M}_0 = \mathcal{M}_0(\sigma + \mathrm{i}\tau) = \prod_{p \leqslant T_0} \left(1 - p^{-\sigma - \mathrm{i}\tau}\right) = \sum_{p | m \Rightarrow p \leqslant T_0} \frac{\mu(m)}{m^{\sigma + \mathrm{i}\tau}}. \tag{9}$$

The parameter T_0 cannot be too large to ensure that $\mathcal{M}_0$ stays relatively short. An admissible choice turns out to be:

$$T_0 = \log\log T. \tag{10}$$

The parameters T_ℓ are chosen to approach T fairly quickly:

$$T_\ell = T^{\frac{1}{(\log_{\ell+1} T)^{\mathfrak{s}}}}, \; 1 \leqslant \ell \leqslant \mathcal{L}. \tag{11}$$

The constant $\mathcal{L}$ is the largest ℓ for which $\log_{\ell+1} T > e$. By definition, we thus have $e < \log_{\mathcal{L}+1} T \leqslant e^2$ and $1 < \log_{\mathcal{L}+2} T \leqslant e$. The parameter $\mathfrak{s}$ plays an important role in the analysis to estimate the constant $C_\alpha(\delta)$. The technical analysis requires $\mathfrak{s} = \mathfrak{s}(\alpha)$ to be not too small, to be precise:

$$\frac{(\log_{\mathcal{L}+1} T)^{\mathfrak{s}}}{(\mathfrak{s} \log_{\mathcal{L}+2} T)^{35}} > K(\alpha + 1)^{35}, \tag{12}$$

where K is a large absolute constant independent of α. The power of 35 is imposed during the proof of Proposition 2 (in particular, in Lemma 2), see (57) and the surrounding discussion for further details.

It is convenient for the analysis to parametrize $\sigma > 1/2$ in the following form:

$$\sigma = \frac{1}{2} + \frac{\delta}{\log T_{\mathcal{L}}}. \tag{13}$$

2.2 | Heuristic

By informally expanding the Euler product in the definition of ζ, we expect that $\log|\zeta|$ at $\sigma + \mathrm{i}\tau$ should behave like

$$\log|\zeta(\sigma + \mathrm{i}\tau)| \approx S_{\mathcal{L}} + \log|\mathcal{M}_0|^{-1}.$$

Moreover, since S_ℓ is a sum of weakly correlated terms, it should behave like a Gaussian random variable for large T. With this in mind, we define the random variables

$$\mathcal{Z}_\ell = \sum_{T_0 < p \leqslant T_\ell} \frac{Z_p}{p^\sigma} + \frac{Z_p^2}{2p^{2\sigma}}, \quad 1 \leqslant \ell \leqslant \mathcal{L}. \tag{14}$$

We effectively replaced $p^{-\mathrm{i}\tau}$ by Z_p where $(Z_p, \mathrm{p\ prime})$ are centered Gaussian random variables with variance $1/2$. The contribution of very small primes is not Gaussian even in the limit $T \to \infty$. Instead, for the small primes, we substitute the phase $p^{-\mathrm{i}\tau}$ in by $e^{\mathrm{i}\theta_p}$, where $(\theta_p, \mathrm{p\ prime})$ are IID uniform on $[0, 2\pi]$:

$$\mathcal{Z}_0 = \log \prod_{p \leqslant T_0} \left|1 - \frac{e^{\mathrm{i}\theta_p}}{p^\sigma}\right|^{-1} = -\sum_{p \leqslant T_0} \sum_{k \geqslant 1} \frac{\cos k\theta_p}{k p^{k\sigma}}. \tag{15}$$

We thus expect for $V \sim \alpha \log\log T$, $\alpha > 0$, that

$$\mathbf{P}(\log|\zeta(\sigma + \mathrm{i}\tau)| > V) \approx \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V). \tag{16}$$

The variance of $\mathcal{Z}_{\mathcal{L}}$ is not exactly $1/2(\log\log T_\ell - \log\log T_0)$ as one would expect from Mertens's theorem, since $\sigma > 1/2$. In fact, we have for any $1 \leqslant \ell \leqslant \mathcal{L}$ by the prime number theorem

$$\begin{aligned} v_\ell^2 = \mathbf{E}[\mathcal{Z}_\ell^2] = \sum_{T_0 < p \leqslant T_\ell} \left(\frac{1}{2p^{2\sigma}} + \frac{1}{8p^{4\sigma}}\right) &= \frac{1}{2}\int_{\log T_0}^{\log T_\ell} \frac{e^{-(2\sigma-1)u}}{u}\mathrm{d}u + \mathrm{O}(e^{-c\sqrt{\log T_0}}) \\ &= \frac{1}{2}\left(\log\log T_\ell - \log\log T_0\right) - \delta\frac{\log T_\ell}{\log T_{\mathcal{L}}} + \mathrm{O}(e^{-c\sqrt{\log T_0}}). \end{aligned} \tag{17}$$

The conclusion of Theorem 1 is then simply the result of a Gaussian estimate, conditioning on $\mathcal{Z}_0$, and of expanding $v_{\mathcal{L}}$:

$$\begin{aligned} \mathbf{P}(\mathcal{Z}_{\mathcal{L}} + \mathcal{Z}_0 > V) &\asymp \mathbf{E}\left[\frac{e^{-(V-\mathcal{Z}_0)^2/(2v_{\mathcal{L}}^2)}}{\alpha\sqrt{\log\log T}}\right] \\ &\asymp \left(\frac{\mathbf{E}[e^{2\alpha\mathcal{Z}_0}]}{(\log T_0)^{\alpha^2}} \cdot e^{-2\alpha^2\delta} \cdot (\log_{\mathcal{L}+1} T)^{-\alpha^2\mathfrak{s}}\right)\frac{e^{-V^2/\log\log T}}{\alpha\sqrt{\log\log T}}. \end{aligned} \tag{18}$$

The term $(\log T_0)^{-\alpha^2}\mathbf{E}[e^{2\alpha\mathcal{Z}_0}]$ turns out to be exactly the arithmetic factor $e^{\gamma\alpha^2}a_\alpha$ by a short computation on the random model, cf. Lemma B.7. We therefore obtain the lower bound claimed in Theorem 1 with

$$C_\alpha(\delta) = e^{\gamma\alpha^2}a_\alpha \cdot e^{-2\alpha^2\delta} \cdot (\alpha+1)^{-36\alpha^2}, \tag{19}$$

by taking $(\log_{\mathcal{L}+1} T)^{\mathfrak{s}} = K(\alpha+1)^{36}$, which satisfies the bound (12). The significance of each factor constituting $C_\alpha(\delta)$ is clear from the heuristic. The arithmetic factor a_α is produced by the small primes. The term $e^{-2\alpha^2\delta}$ is due to the shift off-axis and should be 0 on the critical line (perhaps by assuming the Riemann hypothesis). Finally, the factor $(\alpha+1)^{-36\alpha^2}$ is due to the cutoff of the large prime at $T^{1/(\log_{\mathcal{L}+1} T)^{\mathfrak{s}}}$. The optimal cutoff should be expected to be $T^{c/\alpha}$ for some small constant $c > 0$. Such a cutoff would yield a factor of $(\alpha+1)^{-\alpha^2}$ corresponding to the asymptotics of f_α in (5) predicted by random matrix theory.

2.3 | Proof of Theorem 1

We now outline the roadmap to establish a rigorous lower bound for (16). An approximation of $\log|\zeta|$ in terms of the partial sums S_ℓ and of the Gaussian sums $\mathcal{Z}_\ell$ can be made precise on a *good event* where the sums are constrained. Specifically, we define

$$G_0 = \{|\mathcal{M}_0|^2 \in [L_0, U_0]\}, \tag{20}$$

and

$$G_\ell = G_{\ell-1} \cap \{L_\ell(S_0) \leqslant S_\ell \leqslant U_\ell(S_0)\}, 1 \leqslant \ell \leqslant \mathcal{L}, \tag{21}$$

where the upper and lower *barriers* $L_\ell(z), U_\ell(z)$ are defined and discussed below in (25), and $S_0 = -\log|\mathcal{M}_0|$, following the definition (9). The initial good event G_0 is defined in terms of $\mathcal{M}_0$, since it is a short Dirichlet polynomial, cf. (43). We take

$$L_0 = \exp(-2\alpha \log\log T_0) \qquad U_0 = \exp(-2\alpha \log\log T_0 + 2\sqrt{\log\log T_0}). \tag{22}$$

The factor of 2 in (22) is a direct consequence of taking the squared polynomial in (20). The final event of interest is

$$G(V) = \{\log|\mathcal{M}_0|^{-1} + S_\mathcal{L} > V\} \cap G_\mathcal{L}. \tag{23}$$

We define the corresponding events for the sums $\mathcal{Z}_\ell$,

$$\mathcal{G}_0 = \{\mathcal{Z}_0 - \alpha \log\log T_0 \in [0, \sqrt{\log\log T_0}]\}$$
$$\mathcal{G}_\ell = \mathcal{G}_{\ell-1} \cup \{L_\ell(\mathcal{Z}_0) \leqslant \mathcal{Z}_\ell \leqslant U_\ell(\mathcal{Z}_0)\},\ 1 \leqslant \ell \leqslant \mathcal{L},$$

and

$$\mathcal{G}(V) = \{\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > V\} \cap \mathcal{G}_\mathcal{L}. \tag{24}$$

In G_0 and $\mathcal{G}_0$, the upper bound is natural, coming from the standard deviation in both cases. It is convenient to assume that the recentered sum in case is positive, and we use this fact in the proof of Proposition 1, cf. (35). The (random) upper and lower barriers $U_\ell(z)$ and $L_\ell(z)$ are defined by

$$\begin{aligned} L_\ell(z) &= m(z)(\log\log T_\ell - \log\log T_0) - \mathcal{C}\log_{\ell+2} T, \\ U_\ell(z) &= m(z)(\log\log T_\ell - \log\log T_0) + \mathcal{B}\log_{\ell+2} T, \end{aligned} \tag{25}$$

where the *slope* $m(z)$ is defined as

$$m(z) = \frac{\alpha \log\log T - z}{\log\log T_\mathcal{L} - \log\log T_0}, \tag{26}$$

where $\log_\ell$ stands for the logarithm iterated ℓ times. In each case, the slope $m(z)$ will be evaluated at the (random) initial position, that is, S_0 or $\mathcal{Z}_0$. Taking the barriers randomly is necessitated by

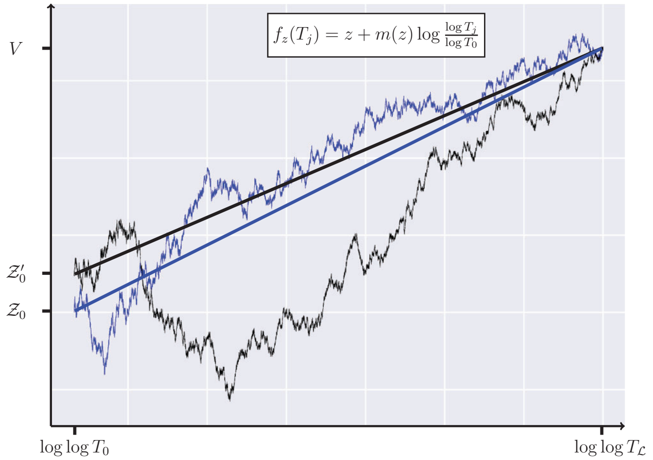

FIGURE 1 Illustration of the random sums starting at two distinct initial values, both ending at the same point V. The lines $f_z(T_j)$ have (distinct) slopes $m(z)$ for $z = \mathcal{Z}_0, \mathcal{Z}_0'$.

separating out the contribution from the small primes. On first reading, one can think of the slope $m(z) \approx \alpha$ (since typically z is like $\alpha \log\log T_0$ in each case). However, we need to take in to account the variance coming from the initial random sums that informs the slope depending on the random position of $\mathcal{Z}_0$ or S_0, see, in particular, the remark following (36). See also Figure 1, where two different initial values $\mathcal{Z}_0$ and $\mathcal{Z}_0'$ both reach level V but elicit two different slopes.

The constants $\mathcal{B}$ will need to be larger than $\alpha\mathfrak{s}$ and $\mathcal{C}$ a sufficiently large constant

$$\mathcal{B} = \alpha\mathfrak{s} + \frac{1}{\alpha} \qquad \mathcal{C} = 100. \tag{27}$$

We include the term $1/\alpha$ in $\mathcal{B}$ to allow for small α, cf. Section 3.1.

The rationale behind the choice of restrictions is that either sum S_ℓ or $\mathcal{Z}_\ell$, whose total value is $V \sim \alpha \log\log T$, should stay close to the linear interpolation with slope $m(z)$ with some fluctuations that are getting smaller with ℓ. It is not hard to first establish this fact for the random model.

Proposition 1. *For fixed $\alpha > 0$ and $V \sim \alpha \log\log T$, the probability of the event $\mathcal{G}(V)$ defined in (23) is*

$$\mathbf{P}(\mathcal{G}(V)) \asymp \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V).$$

To prove this for the partial sums S_ℓ is much trickier. To do so, we follow the idea of [3] and [4] and approximate the events involving S_ℓ and $\mathcal{M}_0$ in terms of short Dirichlet polynomials. This allows us to explicitly compare the probability of the event $G(V)$ with the event $\mathcal{G}(V)$ of the random model. We ultimately show the following.

Proposition 2. *For fixed $\alpha > 0$ and $V \sim \alpha \log\log T$, the probability of the event $\mathcal{G}(V)$ defined in (24) is*

$$\mathbf{P}(G(V)) \asymp \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V). \tag{28}$$

It remains to connect the event $G(V)$ to $\log|\zeta|$. This is done by defining the following short mollifiers:

$$\mathcal{M}_\ell = \sum_{\substack{p|m \Rightarrow p\in(T_{\ell-1},T_\ell] \\ \Omega_\ell(m)\leqslant \mu_\ell}} \frac{\mu(m)}{m^{\sigma+\mathrm{i}\tau}}, \quad 1 \leqslant \ell \leqslant \mathcal{L}, \tag{29}$$

where $\Omega_\ell(m)$ is the number of prime factors of m in the interval $(T_{\ell-1}, T_\ell]$, and we choose

$$\mu_\ell = 100(\alpha \vee 1)(\log\log T_\ell - \log\log T_{\ell-1}), \quad 1 \leqslant \ell \leqslant \mathcal{L}. \tag{30}$$

This is to ensure that the length of the mollifiers is not too long. Indeed, taking the product over ℓ, we set (recalling the definition of $\mathcal{M}_0$ from (9))

$$\mathcal{M} = \prod_{0\leqslant \ell \leqslant \mathcal{L}} \mathcal{M}_\ell. \tag{31}$$

Then, the definition of μ_ℓ and the choice of $\mathfrak{s}$ in (12) imply that $\mathcal{M}$ is supported on integers smaller than

$$\prod_{0\leqslant \ell\leqslant \mathcal{L}} (T_\ell)^{\mu_\ell} \ll T^{200(\alpha+1)\mathfrak{s}(\log_{\mathcal{L}+1} T)^{1-\mathfrak{s}}} \leqslant T^{\frac{1}{K(\alpha+1)^{34}}} < T^{1/100}, \tag{32}$$

for a sufficiently large K. The next proposition shows that $\zeta\mathcal{M}$ is close to 1 on the event $G(V)$.

Proposition 3. *Let $\varepsilon \in (0,1)$. For fixed $\alpha > 0$ and $V \sim \alpha\log\log T$, we have*

$$\mathbf{P}(\log|\zeta| > V) \gg \mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V - \log(1-\varepsilon)\} \cap G(V)\big) \\ - \frac{1}{\varepsilon^2}\left(\frac{e^{-2\delta}}{\delta} + e^{-\mu_{\mathcal{L}}/10}\right)\mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V).$$

The proposition establishes an interplay between ε and δ: if ε is close to 0 or 1, then the expression could be negative. In the proof of Theorem 1, we find an optimal ε and derive the threshold $\delta^\star$ from it. The last proposition relates the inverse of the mollifier to the partial sums, which can then be compared to $\mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V)$ using Propositions 2 and 1.

Proposition 4. *For fixed $\alpha > 0$ and $V \sim \alpha\log\log T$, we have for the event $G(V)$ in (23)*

$$\mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V\} \cap G(V)\big) \gg \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V).$$

The above propositions combine to give a proof of the theorem.

Proof of Theorem 1. Consider V with $V \sim \alpha\log\log T$. For some $\varepsilon \in (0,1)$ to be picked below, define $V' = V - \log(1-\varepsilon)$. Then, Propositions 3 and 4 directly imply

$$\mathbf{P}(\log|\zeta| > V) \gg \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V - \log(1-\varepsilon)) - \frac{1}{\varepsilon^2}\left(\frac{e^{-2\delta}}{\delta} + e^{-\mu_{\mathcal{L}}/10}\right)\mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V).$$

We have from (19) that this is

$$\gg C_\alpha(\delta)\frac{e^{-V^2/\log\log T}}{\alpha\sqrt{\log\log T}}\left\{(1-\varepsilon)^{2\alpha}-\frac{1}{\varepsilon^2}\left(\frac{e^{-2\delta}}{\delta}+e^{-\mu_{\mathcal{L}}/10}\right)\right\}.$$

For a fixed α, we are looking for the smallest δ that makes the quantity in the parenthesis positive. The function $\varepsilon^2(1-\varepsilon)^{2\alpha}$ is maximal at $\varepsilon^\star = 1/(\alpha+1)$. We thus pick $\delta^\star$ to be equal to a fraction (say e^{-1}) of $\varepsilon^2(1-\varepsilon)^{2\alpha}$:

$$\frac{e^{-2\delta^\star}}{\delta^\star}=\frac{1}{e}{\varepsilon^\star}^2(1-\varepsilon^\star)^{2\alpha}=\frac{1}{e}\frac{\alpha^{2\alpha}}{(\alpha+1)^{2(\alpha+1)}},$$

which gives

$$\delta^\star+\frac{1}{2}\log\delta^\star=\log(\alpha+1)+\alpha\log(\alpha^{-1}+1)+\frac{1}{2}. \tag{33}$$

Since $\frac{e^{-2\delta}}{\delta}$ is a decreasing function of δ and that $(1-\varepsilon^\star)^{2\alpha}\asymp 1$, we get that for $\delta>\delta^\star$, the above is

$$\gg C_\alpha(\delta)\frac{e^{-V^2/\log\log T}}{\alpha\sqrt{\log\log T}}\left(\frac{\alpha}{\alpha+1}\right)^{2\alpha}\gg C_\alpha(\delta)\frac{e^{-V^2/\log\log T}}{\alpha\sqrt{\log\log T}}.$$

The term $e^{-\mu_{\mathcal{L}}}$ does not contribute by the choice of μ_ℓ in (30) by taking $\mathfrak{s}$ large enough independently of α, as mentioned below (12). □

3 | PROOF OF THE PROPOSITIONS

Throughout the proofs, we use the following notations for conciseness:

$$t=\log\log T,$$

and

$$t_0=\log\log T_0=\log_4 T \qquad t_\ell=\log\log T_\ell=t-\mathfrak{s}\log_\ell t,\ 1\leqslant\ell\leqslant\mathcal{L}.$$

3.1 | Proof of Proposition 1

The upper bound is immediate by dropping the event $\mathcal{G}_{\mathcal{L}}$, so we focus on the lower bound. Decomposing the event $\mathcal{G}_{\mathcal{L}}$, we inductively have

$$\mathbf{P}(\mathcal{G}(V))=\mathbf{P}(\{\mathcal{Z}_0+\mathcal{Z}_{\mathcal{L}}>V\}\cap\mathcal{G}_0)-\sum_{\ell=0}^{\mathcal{L}-1}\mathbf{P}(\{\mathcal{Z}_0+\mathcal{Z}_{\mathcal{L}}>V\}\cap\mathcal{G}_\ell\cap\mathcal{G}_{\ell+1}^c). \tag{34}$$

We subsequently bound the summands on the right-hand side of (34) for $V \sim \alpha \log\log T = \alpha t$. Separately considering the two possibilities for $\mathcal{G}^c_{\ell+1}$, we have for the upper barrier

$$\mathbf{P}(\{\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V\} \cap \mathcal{G}_\ell \cap \{\mathcal{Z}_{\ell+1} > U_{\ell+1}(\mathcal{Z}_0)\}) \leqslant \mathbf{P}\{\mathcal{Z}_{\ell+1} + \mathcal{Z}_0 > U_{\ell+1}(\alpha t_0) + \alpha t_0\}) \qquad (35)$$
$$\asymp \frac{1}{\alpha\sqrt{t}}\mathbf{E}\left[\exp\left(-\frac{(U_{\ell+1}(\alpha t_0) + \alpha t_0 - \mathcal{Z}_0)^2}{2v^2_{\ell+1}}\right)\right]$$

akin to (18). The first bound uses positivity of the recentered sum. Unfolding and using the definition of $U_{\ell+1}(z)$ in (25), we have that the above is bounded by

$$\ll \frac{1}{\alpha\sqrt{t}}\mathbf{E}\left[\exp\left(-\frac{1}{2v^2_{\ell+1}}\left(\frac{\alpha}{t_{\mathcal{L}} - t_0}(t_{\ell+1}(t-t_0) - t_0(t - t_{\mathcal{L}})) + \mathcal{B}\log_{\ell+1} t - \mathcal{Z}_0\right)^2\right)\right]$$
$$\ll \frac{1}{\alpha\sqrt{t}} e^{-\alpha^2 t} e^{-\alpha^2 t_0}\mathbb{E}\left[e^{2\alpha\mathcal{Z}_0}\right] e^{-2\alpha^2\delta e^{t_{\ell+1}-t_{\mathcal{L}}} - 2\alpha^2 \mathfrak{s}\log_{\mathcal{L}} t} \cdot e^{(\alpha^2\mathfrak{s} - 2\alpha\mathcal{B})\log_{\ell+1} t}$$
$$\ll \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > \alpha t) \cdot e^{(\alpha^2\mathfrak{s} - 2\alpha\mathcal{B})\log_{\ell+1} t},$$

the final bound achieved by using Lemma B.7 and comparing to (18). The coefficient in the exponent involving $\log_{\ell+1} t$ is strictly negative by choosing $\mathcal{B} = \alpha\mathfrak{s} + 1/\alpha$ (cf. Equation (27)), and the implicit constant can be taken independent of α.

To complete the bound, we estimate the exceedance of the lower barrier

$$\mathbf{P}(\{\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > \alpha t\} \cap \mathcal{G}_0 \cap \{\mathcal{Z}_{\ell+1} < L_{\ell+1}(\mathcal{Z}_0)\}). \qquad (36)$$

If one uses a deterministic barrier $L_{\ell+1} \approx \alpha(t_{\ell+1} - t_0) - C\log_{\ell+1} t$ here (corresponding to the typical position of $\mathcal{Z}_0$), then the position of the process at $\ell+1$ is at most $L_{\ell+1} + \mathcal{Z}_0$. This position could be as large as $\approx \alpha t_{\ell+1} + \sqrt{t_0} + C\log_{\ell+1} t$. This additional term coming from the variance of the initial contribution does not permit a strong enough bound in the calculation below.

First conditioning on the value of $\mathcal{Z}_0$ and decomposing on the value at $t_{\ell+1}$, we find

$$\mathbf{P}(\{\mathcal{Z}_{\mathcal{L}} - \mathcal{Z}_{\ell+1} > \alpha t - \mathcal{Z}_{\ell+1} - w\} \cap \{\mathcal{Z}_{\ell+1} < L_{\ell+1}(w)\})$$
$$\leqslant \sum_{u \leqslant L_{\ell+1}(w)} \mathbf{P}(\mathcal{Z}_{\mathcal{L}} - \mathcal{Z}_{\ell+1} > \alpha t - (u+1) - w)\mathbf{P}(\mathcal{Z}_{\ell+1} \in (u, u+1]),$$

using independence. By the Gaussian decay, this is therefore bounded by

$$\ll \sum_{u \leqslant L_{\ell+1}(w)} e^{-\frac{(\alpha t - (u+1) - w)^2}{2(v^2_{\mathcal{L}} - v^2_{\ell+1})}} \frac{v_{\ell+1}}{u} e^{-\frac{u^2}{2v^2_{\ell+1}}}$$
$$\ll \frac{1}{\alpha\sqrt{t}} e^{-\alpha^2\frac{(t_{\mathcal{L}} - t_{\ell+1})^2}{2(v^2_{\mathcal{L}} - v^2_{\ell+1})}} e^{-\alpha^2\frac{t^2}{2v^2_{\ell+1}}} e^{2\alpha t\frac{\alpha(t_{\mathcal{L}} - t_{\ell+1}) + w}{2v^2_{\ell+1}}} \sum_{u > C\log_{\ell+1} t} e^{-2\alpha\frac{(u-1)(t_{\mathcal{L}} - t_{\ell+1})}{2(v^2_{\mathcal{L}} - v^2_{\ell+1})}} e^{-\frac{(u-1)^2}{2(v^2_{\mathcal{L}} - v^2_{\ell+1})}} e^{2\alpha t\frac{u}{2v^2_{\ell+1}}},$$

achieved by shifting the sum by $(t_{\ell+1} - t_0)m(w)$ and expanding. The prefactor simplifies to

$$\frac{1}{\alpha\sqrt{t}} \exp\left(-\alpha^2 t\right) \cdot \exp\left((-\alpha^2 t_0 + 2\alpha w) - 2\alpha^2\delta - \alpha^2 \mathfrak{s} \log_{\mathcal{L}} t\right).$$

After unconditioning on the value of $\mathcal{Z}_0$, this term is like $\mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > \alpha t)$. Similarly, after simplifying and estimating at the dominant u, the sum is bounded by

$$\ll \exp\left(2\alpha - 4\alpha C\delta \frac{\log_{\ell+1} t}{t_{\mathcal{L}} - t_{\ell+1}}(1 - e^{t_{\ell+1} - t_{\mathcal{L}}}) - \frac{(C\log_{\ell+1} t - 1)^2}{2(v_{\mathcal{L}}^2 - v_{\ell+1}^2)}\right) \ll c_\alpha \exp\left(-\frac{C^2}{\mathfrak{s}} \log_{\ell+1} t\right).$$

Again, the coefficient in the final exponent involving $\log_{\ell+1} t$ is strictly negative (and we implicitly use that C is strictly positive when estimating the sum above), resulting in a subdominant contribution for most ℓ. For the largest[†] ℓ, namely, $\mathcal{L} - 2$, notice that $e \leqslant \log_{\ell+1} t < e^2$. The term c_α is an explicit coefficient dependent on α,

$$c_\alpha = \exp\left(2\alpha - \frac{C}{\mathfrak{s}}\left(4\alpha\delta + C\log_{\mathcal{L}} t + 2C\delta - 2\right)\right).$$

The coefficient will appear in the sum over ℓ below. Recall $\mathfrak{s} = 35\log(\alpha + 1) + \log(2 \cdot 10^6)$ and $\delta > \delta^\star(\alpha)$ (cf. (57) and (33), respectively). Therefore, for small α, both $C/\mathfrak{s}, 2\delta > 1$ so $c_\alpha < 1/100$ say. For large α, we have $\delta^\star/\mathfrak{s} \approx 1$ so again $c_\alpha < \exp(-10\alpha)$ say.

Therefore, returning to the original probability, unconditioning, and comparing to (18), we have for $0 \leqslant \ell \leqslant \mathcal{L} - 2$

$$\mathbf{P}(\{\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > \alpha t\} \cap \mathcal{G}_0 \cap \{\mathcal{Z}_{\ell+1} < L_{\ell+1}(\mathcal{Z}_0)\}) \ll \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > \alpha t) \cdot c_\alpha e^{-\frac{C^2}{\mathfrak{s}}\log_{\ell+1} t}.$$

Overall, from (34) (using the estimate (B.16) to remove $\mathcal{G}_0$ in the first instance), we find

$$\begin{aligned}\mathbf{P}(\mathcal{G}(V)) \gg \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V)&\left(1 - \sum_{j=0}^{\mathcal{L}-1} e^{-(2\alpha\mathcal{B} - \alpha^2\mathfrak{s})\log_{\ell+1} t} - c_\alpha \sum_{j=0}^{\mathcal{L}-2} e^{-\frac{C^2}{\mathfrak{s}}\log_{\ell+1} t}\right)\\ &\asymp \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V),\end{aligned}$$

where the final bound follows from the rapid convergence of the sum and taking $\mathcal{B}, C$ as in (27) (note that c_α is also small according to the discussion above) so that the term in brackets is strictly positive and can be taken independent of α.

3.2 | Proof of Proposition 2

We first need two lemmas whose goal is to approximate quantitatively an indicator function by a polynomial. The content of these lemmas can be found in [4, Lemma 3]. We reproved them here for completeness, and also to get a clear dependence on the parameters, which is key in estimating the constant $C_\alpha(\delta)$.

[†] For $\ell = \mathcal{L} - 1$, the probability involved is zero since the events $\{\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > \alpha t\}$ and $\mathcal{Z}_{\mathcal{L}} < L_{\mathcal{L}}(\mathcal{Z}_0)$ are incompatible.

Lemma 1. *Let $c > b > a > 1$ and Δ large enough. There exist entire functions h^+ and h^- depending on these parameters such that for all $x \in \mathbb{R}$,*

(i) $0 \leqslant h^- \leqslant h^+ \leqslant 1$,
(ii) $\mathbf{1}(x \in [0, \Delta^{-1}])\left(1 - e^{-\Delta^{c-b-1}}\right) \leqslant h^+(x) \leqslant \mathbf{1}(x \in [-\Delta^{-a}, \Delta^{-1} + \Delta^{-a}]) + e^{-\Delta^{c-a-1}}$,
(iii) $\mathbf{1}(x \in [\Delta^{-a}, \Delta^{-1} - \Delta^{-a}])\left(1 - e^{-\Delta^{c-a-1}}\right) \leqslant h^-(x) \leqslant \mathbf{1}(x \in [0, \Delta^{-1}]) + e^{-\Delta^{c-b-1}}$,
(iv) *for all $\ell \in \mathbb{N}$, the Fourier transform $\widehat{h^\pm}$ is such that*

$$\left| \int_{\mathbb{R}} \xi^\ell \widehat{h^\pm}(\xi) \mathrm{d}\xi \right| \leqslant \Delta^{c(\ell+2)}.$$

Proof. We define $h^\pm$ as convolutions with a suitable kernel. Following [3, Lemma 5] based on a classical construction of [11], we know that there exists a real function F such that $0 \leqslant F \leqslant 1$ with decay

$$F(x) \ll \exp(-|x|/\log^2(|x| + 10)),$$

for which the Fourier transform is supported on $[-1, 1]$. With this in mind, consider the corresponding PDF $\varphi(x) = F(x)/\|F\|_1$, and the scaled version:

$$\varphi_c(y) = \Delta^c \varphi(\Delta^c y).$$

Since $c > 1$, the scale of φ_c is much smaller than the width of the interval $[0, \Delta^{-1}]$ to be approximated. The function $h^\pm$ is defined by the convolution of an indicator function with φ_c:

$$h^+(x) = \int \mathbf{1}(y \in [-\Delta^{-b}, \Delta^{-1} + \Delta^{-b}])\varphi_c(x - y)\mathrm{d}y, \tag{37}$$

and

$$h^-(x) = \int \mathbf{1}(y \in [\Delta^{-b}, \Delta^{-1} - \Delta^{-b}])\varphi_c(x - y)\mathrm{d}y. \tag{38}$$

The functions $h^\pm$ should be thought of as smoothings of the indicator function $1(x \in [0, \Delta^{-1}])$. The slight overspill Δ^{-b} must be added to properly bound the indicator function close to 0 and Δ^{-1}. It is necessary to take $c > b > 1$, as this ensures that the effective width of the kernel is much smaller than the overspill. Moreover, to bound h^+ above (and h^- below) by an indicator function, the interval has to be enlarged (or reduced) by Δ^{-a}. For this, we need $b > a$ to ensure that the indicator function in the definition of h^+ is bounded by $\mathbf{1}(x \in [-\Delta^{-a}, \Delta^{-1} + \Delta^{-a}])$, and similarly for h^-.

Property (i) is obvious from the definitions (37) and (38). The first inequality in Property (ii) is also clear when $x \notin [0, \Delta^{-1}]$. The case $x \in [0, \Delta^{-1}]$ becomes

$$1 - e^{-\Delta^{c-b-1}} \leqslant h^+(x).$$

To prove this, note that for any $0 \leqslant x \leqslant \Delta^{-1}$,

$$h^+(x) = \int_{-\Delta^{-1}-\Delta^{-b}}^{\Delta^{-b}} \varphi_c(u + x)\mathrm{d}u \geqslant \int_{-\Delta^{-b}}^{\Delta^{-b}} \varphi_c(u)\mathrm{d}u \geqslant 1 - e^{-\Delta^{c-b-1}},$$

where the last inequality follows by the decay of φ. The right-hand side inequality of Property (ii) is also clear when $x \in [-\Delta^{-a}, \Delta^{-1} + \Delta^{-a}]$). If $x < -\Delta^{-a}$, then we have

$$h^{+}(x) = \int_{-\Delta^{-1}-\Delta^{-b}}^{\Delta^{-b}} \varphi_c(u+x)\mathrm{d}u \leqslant \int_{-\infty}^{\Delta^{-b}-\Delta^{-a}} \varphi_c(u)\mathrm{d}u \leqslant e^{-\Delta^{c-a-1}},$$

for Δ large enough and by the decay of φ (since $b > a$, we can ask for $\Delta^{-b} - \Delta^{-a} \leqslant -\frac{1}{2}\Delta^{-a}$ say). Similarly, if $x > \Delta^{-1} + \Delta^{-a}$, we get

$$h^{+}(x) = \int_{-\Delta^{-1}-\Delta^{-b}}^{\Delta^{-b}} \varphi_c(u+x)\mathrm{d}u \leqslant \int_{\Delta^{-a}-\Delta^{-b}}^{\infty} \varphi_c(u)\mathrm{d}u \leqslant e^{-\Delta^{c-a-1}}.$$

Property (iii) is proved exactly as Property (ii) by using the definition of h^{-} in (38) instead.

For Property (iv), note that

$$\widehat{h^{\pm}}(\xi) = \widehat{\varphi}_c(\xi) \cdot \widehat{\mathbf{1}_{[0,\Delta^{-1}]}}(\xi).$$

On the one hand, by integration, we have the simple bound

$$\widehat{\mathbf{1}_{[0,\Delta^{-1}]}}(\xi) \ll \Delta^{-1}.$$

This smaller than 1 for Δ large enough. On the other hand, since φ is supported on $[-1, 1]$, we have

$$\left| \int \xi^{\ell} \widehat{\varphi}_c(\xi)\mathrm{d}\xi \right| \leqslant \Delta^{c\ell} \int |\widehat{\varphi}_c(\xi)|\mathrm{d}\xi.$$

By the Cauchy–Schwarz inequality and Parseval's identity, we can bound the integral by

$$\int |\widehat{\varphi}_c(\xi)|\mathrm{d}\xi \leqslant \sqrt{2\Delta^c}\left(\int \varphi_c(x)^2 \mathrm{d}x \right)^{1/2} \ll \Delta^c,$$

since $\varphi_c(x) \ll \Delta^c$. This can be made $\leqslant \Delta^{2c}$ by taking Δ large enough to absorb the implicit constant. This gives the desired bound. □

In a second step, we approximate $h^{\pm}$ by polynomials.

Lemma 2. *Let $a > 2$, $X > 1$ and Δ large enough. There exist polynomials $\mathcal{D}^{+}$ and $\mathcal{D}^{-}$ of degree smaller than $100X\Delta^{3a}$ with ℓth coefficient bounded by $\frac{(2\pi)^{\ell}}{\ell!}\Delta^{2a(\ell+2)}$ such that for $|x| \leqslant X$, we have*

$$\mathbf{1}(x \in [0, \Delta^{-1}])\left(1 - e^{-\Delta^{a-2}}\right) \leqslant |\mathcal{D}^{+}(x)|^2 \leqslant \mathbf{1}(x \in [-\Delta^{-a}, \Delta^{-1} + \Delta^{-a}]) + e^{-\Delta^{a-2}}, \tag{39}$$

and

$$\mathbf{1}(x \in [\Delta^{-a}, \Delta^{-1} - \Delta^{-a}])\left(1 - e^{-\Delta^{a-2}}\right) \leqslant |\mathcal{D}^{-}(x)|^2 \leqslant \mathbf{1}(x \in [0, \Delta^{-1}]) + e^{-\Delta^{a-2}}. \tag{40}$$

Proof. We consider the entire functions $h^{\pm}$ from Lemma 1. We pick $b = 2a$ and $c = 3a$. Expressing $h^{\pm}$ in terms of their Fourier transform, we get

$$h^{\pm}(x) = \int e^{2\pi i\xi x}\widehat{h^{\pm}}(\xi)d\xi = \sum_{\ell \geqslant 0} \frac{(2\pi i x)^{\ell}}{\ell!}\int \xi^{\ell}\widehat{h^{\pm}}d\xi = \sum_{\ell<\nu} \frac{(2\pi i x)^{\ell}}{\ell!}\int \xi^{\ell}\widehat{h^{\pm}}(\xi)d\xi + \mathcal{R}(x).$$

The cutoff ν will need to be suitably picked. By Taylor's theorem, the remainder term satisfies the bound

$$|\mathcal{R}| \leqslant \frac{(2\pi X)^{\nu}}{\nu!}\left|\int \xi^{\nu}\widehat{h^{\pm}}(\xi)d\xi\right| \leqslant \frac{(2\pi X)^{\nu}}{\nu!}\Delta^{c(\nu+2)},$$

where we used Property (iv) of Lemma 1. Using Stirling's formula, we see that by picking

$$\nu = 100X\Delta^{c}, \tag{41}$$

we have

$$|\mathcal{R}| \leqslant e^{-100X\Delta^{c}}. \tag{42}$$

Thus far, we have established the following decomposition

$$h^{\pm}(x) = \mathcal{D}^{\pm}(x) + \mathcal{R}(x),$$

where $\mathcal{D}^{\pm}$ is a polynomial of degree at most $100X\Delta^{c}$. The ℓth coefficient of $\mathcal{D}^{\pm}$, $\ell < \nu$, satisfies the bound

$$\left|\frac{(2\pi i)^{\ell}}{\ell!}\int \xi^{\ell}\widehat{h^{\pm}}(\xi)d\xi\right| \leqslant \frac{(2\pi)^{\ell}}{\ell!}\Delta^{c(\ell+2)},$$

by Property (iv) of Lemma 1.

It remains to prove the inequalities (39) and (40). We prove the left-hand side inequality of (39), the other ones are proved the same way. The case $x \notin [0, \Delta^{-1}]$ is clear. When $x \in [0, \Delta^{-1}]$, we have from Property (ii) of Lemma 1 and (42) that

$$|\mathcal{D}^{+}(x)| = |h^{+}(x) - \mathcal{R}(x)| \geqslant 1 - e^{-\Delta^{2a-1}} - e^{-100X\Delta^{c}}.$$

This is $1 - e^{-\Delta^{a-2}}$ for Δ large enough. □

We now apply the last two lemmas to the proof of Proposition 2.

Proof of Proposition 2. For simplicity, we prove the case $\ell = \mathcal{L}$. The proof when $\ell < \mathcal{L}$ is the same. Also, without loss of generality, we can assume $\alpha \geqslant 1$ by replacing α by $\alpha \vee 1$ in what follows.

Define the increments

$$Y_{\ell} = S_{\ell} - S_{\ell-1},\ 2 \leqslant \ell \leqslant \mathcal{L},$$

and $Y_1 = S_1$ for $\ell = 1$. For $\ell = 0$, we set

$$Y_0 = |\mathcal{M}_0|^2. \tag{43}$$

The technical reason behind this is that $\mathcal{M}_0$ is a short Dirichlet polynomial, but $|\mathcal{M}_0^{-1}|$ and $\log|\mathcal{M}_0^{-1}|$ are not. Therefore, it is technically easier to approximate events involving $|\mathcal{M}_0|^2$.

Consider the events G_ℓ as in (21) and $G(V)$ as in (20). We want to express these in terms of the increments

$$\{Y_\ell \in [u_\ell, u_\ell + \Delta_\ell^{-1}]\},\ 0 \leqslant \ell \leqslant \mathcal{L}.$$

The set $\mathcal{I} = \mathcal{I}(V)$ of tuples $\mathbf{u} = (u_\ell, 0 \leqslant \ell \leqslant \mathcal{L})$ with $u_\ell \in \Delta_\ell^{-1} \cap \mathbb{Z}$ must be picked to mimic the restrictions in the G_ℓ's and $G(V)$. For $\ell = 0$, from the definition of G_0 in (20), we take $u_0 \in [L_0, U_0]$. Clearly, $|u_0| \leqslant 1$. We discretize this interval with a step size of

$$\Delta_0^{-1} = e^{-100\alpha t_0}. \tag{44}$$

Note that for $Y_0 \in [u_0, u_0 + \Delta_0^{-1}]$, we then have by expanding the logarithm

$$-\frac{1}{2}\log u_0 - \frac{1}{2}\Delta_0^{-1/2} \leqslant -\frac{1}{2}\log Y_0 \leqslant -\frac{1}{2}\log u_0.$$

For the u_ℓ, $1 \leqslant \ell \leqslant \mathcal{L}$, and $s_0 = -\frac{1}{2}\log u_0$, we ask that the u's satisfy the "barrier" condition,

$$L_\ell(s_0) \leqslant s_0 + \sum_{1\leqslant k\leqslant \ell} u_k \leqslant U_\ell(s_0). \tag{45}$$

In the preceding and following, we write $L_\ell = L_\ell(s_0)$. The condition (45) implies that for $1 \leqslant \ell \leqslant \mathcal{L}$:

$$L_\ell - U_{\ell-1} \leqslant u_\ell \leqslant U_\ell - L_{\ell-1}.$$

For $\ell \geqslant 2$, this yields the bounds

$$\begin{aligned} u_\ell &\leqslant U_\ell - L_{\ell-1} = (\mathfrak{s}\, m(s_0) + \mathcal{C})\log_{\ell-1} t - (\mathfrak{s}\, m(s_0) - \mathcal{B})\log_\ell t, \\ u_\ell &\geqslant L_\ell - U_{\ell-1} = (\mathfrak{s}\, m(s_0) - \mathcal{B})\log_{\ell-1} t - (\mathfrak{s}\, m(s_0) + \mathcal{C})\log_\ell t. \end{aligned}$$

From the choice of slope $m(z)$ in (26) and $\mathcal{B}$ and $\mathcal{C}$ in (27), we can assume that

$$|u_\ell| \leqslant 2\alpha\mathfrak{s}\log_{\ell-1} t, \quad 2 \leqslant \ell \leqslant \mathcal{L} \tag{46}$$

and $|u_1| \leqslant 2\alpha t_1$. This is the important input for the choice of parameters. From this, we take

$$\Delta_\ell = 10\alpha\mathfrak{s}\log_{\ell-1} t, \quad 2 \leqslant \ell \leqslant \mathcal{L}; \qquad \Delta_\ell = 10\alpha t_1.$$

This ensures that

$$\sum_{1\leqslant k\leqslant \ell} \Delta_k^{-1} \leqslant \alpha^{-1}, \quad 1 \leqslant \ell \leqslant \mathcal{L}. \tag{47}$$

This choice is also motivated by the condition $|u_\ell \Delta_\ell^{-1}| \leqslant 1$ arising later in (61). Finally, the restriction $\{\mathcal{S}_\mathcal{L} + \log|\mathcal{M}_0^{-1}| > V\}$ imposes the following condition on $\mathcal{I}$:

$$\sum_{1\leqslant \ell \leqslant \mathcal{L}} u_\ell - \frac{1}{2}\log u_0 > V\,. \tag{48}$$

With these choices, it is not hard to check that we have the following inclusions:

$$\begin{aligned}&\bigcup_{\mathbf{u}\in\mathcal{I}}\{Y_\ell \in [u_\ell, u_\ell+\Delta_\ell^{-1}], \ell\leqslant\mathcal{L}\}\subset\\ &\left\{S_\mathcal{L} - \frac{1}{2}\log Y_0 > V-\alpha^{-1}, S_\ell\in[L_\ell, U_\ell+\alpha^{-1}], 1\leqslant\ell\leqslant\mathcal{L}, Y_0\in[L_0, U_0+\Delta_0^{-1}]\right\},\end{aligned} \tag{49}$$

and

$$\begin{aligned}&\bigcup_{\mathbf{u}\in\mathcal{I}}\{Y_\ell \in [u_\ell, u_\ell+\Delta_\ell^{-1}], \ell\leqslant\mathcal{L}\}\supset\\ &\left\{S_\mathcal{L} - \frac{1}{2}\log Y_0 > V+\alpha^{-1}, S_\ell\in[L_\ell+\alpha^{-1}, U_\ell], 1\leqslant\ell\leqslant\mathcal{L}, Y_0\in[L_0+\Delta_0^{-1}, U_0]\right\}.\end{aligned} \tag{50}$$

Note that the events on the left are disjoint for two different tuples $\mathbf{u}$.

The lower and upper bounds in the statement of Proposition 2 are proved separately using the two above inclusions. Both proofs consist in reexpressing the probabilities of the increments Y_k in terms of the ones of a random model where we effectively replace $p^{-i\tau}$ by a uniform random variable on the unit circle. This will be achieved by converting the indicator functions in terms of polynomials with Lemma 2 and using mean-value theorems for Dirichlet polynomials. The random model is defined as follows: for $(\theta_p, p$ primes) be IID random variables uniformly distributed on $[0, 2\pi]$, we consider the random model for the Dirichlet sums S_ℓ

$$\mathcal{S}_\ell = \sum_{T_0<p\leqslant\exp(e^{t_\ell})} \frac{\mathrm{Re}\, e^{\mathrm{i}\theta_p}}{p^\sigma} + \frac{\mathrm{Re}\, e^{2\mathrm{i}\theta_p}}{2p^{2\sigma}}, \quad 1\leqslant\ell\leqslant\mathcal{L}, \tag{51}$$

and their corresponding increments

$$\mathcal{Y}_\ell = \mathcal{S}_\ell - \mathcal{S}_{\ell-1},\ 2\leqslant\ell\leqslant\mathcal{L}, \tag{52}$$

with $\mathcal{Y}_1 = \mathcal{S}_1$. For $\ell = 0$, keeping in line with the choice in (43), we define

$$\mathcal{Y}_0 = \prod_{p\leqslant T_0}\left|1-\frac{e^{\mathrm{i}\theta_p}}{p^\sigma}\right|^2.$$

Proof of $\mathbf{P}(G(V)) \gg \mathbf{P}(\mathcal{Z}_0+\mathcal{Z}_\mathcal{L} > V)$.

It suffices to prove that

$$\mathbf{P}(G(V)) \gg \mathbf{P}(\mathcal{G}(V+\alpha^{-1})) \tag{53}$$

for the event $\mathcal{G}(V)$ given in (23). Indeed, note that the added terms $+\alpha^{-1}$ to the barrier U_ℓ and L_ℓ are irrelevant by Proposition 1. Moreover, the same proposition implies that $\mathbf{P}(\mathcal{G}(V+\alpha^{-1})) \gg \mathbf{P}(\mathcal{Z}_0+\mathcal{Z}_\mathcal{L} > V+\alpha^{-1})$. This is $\gg \mathbf{P}(\mathcal{Z}_0+\mathcal{Z}_\mathcal{L} > V)$ by (18).

To prove (53), we use the inclusion (49) to get

$$\mathbf{P}(G(V)) \gg \sum_{\mathbf{u}\in\mathcal{I}} \mathbf{P}(Y_\ell \in [u_\ell, u_\ell+\Delta_\ell^{-1}], 0 \leqslant \ell \leqslant \mathcal{L}).$$

To apply Lemma 2 to the indicator function $[0,\Delta_\ell^{-1}]$, we introduce the restriction $|Y_\ell - u_\ell| \leqslant X_\ell$ for a parameter $X_\ell \leqslant 100\Delta_\ell^{4a}$. The justification of the constraint will be specified in (75). We will also need to pick $a = 5$, to guarantee that $\Delta_\ell^{a-2} > \Delta_\ell^2$. Together with (57), this informs the choice (12). With this, we have for $1 \leqslant \ell \leqslant \mathcal{L}$

$$\begin{aligned}\mathbf{1}(Y_\ell \in [u_\ell, u_\ell+\Delta_\ell^{-1}]) &\geqslant \mathbf{1}_{[0,\Delta_\ell^{-1}]}(Y_\ell-u_\ell)\mathbf{1}(|Y_\ell-u_\ell|\leqslant X_\ell)\\ &\geqslant \left(|\mathcal{D}_\ell^-(Y_\ell-u_\ell)|^2 - e^{-\Delta_\ell^{a-2}}\right)\mathbf{1}(|Y_\ell-u_\ell|\leqslant X_\ell),\end{aligned}\tag{54}$$

where $\mathcal{D}_\ell^-(x)$ is a polynomial of degree $\leqslant 100X_\ell\Delta_\ell^{3a}$ as guaranteed by Lemma 2. For $\ell=0$, there is no need to introduce the indicator function. Indeed, we have the pointwise bound

$$|Y_0| = \prod_{p\leqslant T_0}\left|1-\frac{p^{-\mathrm{i}\tau}}{p^\sigma}\right|^2 \leqslant \exp(10T_0^{1/2}).$$

Therefore, by picking $X_0 = \exp(10T_0^{1/2}) = \exp(10\sqrt{\log\log T})$, we get

$$\mathbf{1}(Y_0 \in [u_0, u_0+\Delta_0^{-1}]) \geqslant |\mathcal{D}_0^-(Y_0-u_0)|^2 - e^{-\Delta_0^{a-2}}.\tag{55}$$

This means that $\prod_{0\leqslant\ell\leqslant\mathcal{L}}\mathcal{D}_\ell^-(Y_\ell-u_\ell)$ is a Dirichlet polynomial of length less than

$$\exp(2\cdot 100e^{t_\mathcal{L}}X_\mathcal{L}\Delta_\mathcal{L}^{3a}) = \exp\left(2\cdot 10^4\frac{(\alpha\vee 1)^{7a}(\mathfrak{s}\log_\mathcal{L} t)^{7a}}{e^{\mathfrak{s}\log_\mathcal{L} t}}\right).\tag{56}$$

To ensure that the length is less than $T^{1/100}$, we are led to impose

$$\boxed{\mathfrak{s}\log_\mathcal{L} t - 7a\log(\mathfrak{s}\log_\mathcal{L} t) > 7a\log(\alpha\vee 1)+\log(2\cdot 10^6).}\tag{57}$$

For α large, this essentially imposes $\mathfrak{s}\log_\mathcal{L} t$ larger than $7a\log(\alpha\vee 1)$. Our goal is now to get rid of the indicator function for $1\leqslant\ell\leqslant\mathcal{L}$. In this case, writing $\mathbf{1}(|Y_\ell-u_\ell|\leqslant X_\ell) = 1-\mathbf{1}(|Y_\ell-u_\ell|>X_\ell)$, we get by expanding the product

$$\begin{aligned}&\mathbf{E}\left[\prod_{0\leqslant\ell\leqslant\mathcal{L}}\left(|\mathcal{D}_\ell^-(Y_\ell-u_\ell)|^2-e^{-\Delta_\ell^{a-2}}\right)\prod_{1\leqslant\ell\leqslant\mathcal{L}}\mathbf{1}(|Y_\ell-u_\ell|\leqslant X_\ell)\right]\\ &=\sum_{J\subseteq\{1,\dots,\mathcal{L}\}}(-1)^{|J|}\mathbf{E}\left[\prod_{0\leqslant\ell\leqslant\mathcal{L}}\left(|\mathcal{D}_\ell^-(Y_\ell-u_\ell)|^2-e^{-\Delta_\ell^{a-2}}\right)\prod_{j\in J}\mathbf{1}(|Y_j-u_j|>X_j)\right].\end{aligned}\tag{58}$$

The dominant term will be the summand corresponding to $J = \emptyset$. It is good to start with estimating this term. By the mean-value theorem of Dirichlet polynomials, see Lemma B.2, the first term can be expressed in terms of the random increments $\mathcal{Y}_\ell$, which are now independent:

$$\mathbf{E}\left[\prod_\ell \left(|\mathcal{D}_\ell^-(Y_\ell - u_\ell)|^2 - e^{-\Delta_\ell^{a-2}}\right)\right] = (1+\mathrm{O}(T^{-1/100}))\prod_\ell \mathbf{E}\left[\left(|\mathcal{D}_\ell^-(\mathcal{Y}_\ell - u_\ell)|^2 - e^{-\Delta_\ell^{a-2}}\right)\right]. \tag{59}$$

We can apply Lemma 2 again to convert the polynomial back to an indicator function

$$\begin{aligned}\mathbf{E}\left[|\mathcal{D}_\ell^-(\mathcal{Y}_\ell - u_\ell)|^2\right] &\geqslant \mathbf{E}\left[|\mathcal{D}_\ell^-(\mathcal{Y}_\ell - u_\ell)|^2 \mathbf{1}(|\mathcal{Y}_\ell - u_\ell| \leqslant X_\ell)\right] \\ &\geqslant (1 - e^{-\Delta_\ell^{a-2}})\mathbf{P}(\mathcal{Y}_\ell - u_\ell \in [\Delta_\ell^{-a}, \Delta_\ell^{-1} - \Delta_\ell^{-a}]),\end{aligned}$$

where we noticed that $\mathcal{Y}_\ell - u_\ell \in [\Delta_\ell^{-a}, \Delta_\ell^{-1} - \Delta_\ell^{-a}]$ implies $|\mathcal{Y}_\ell - u_\ell| \leqslant X_\ell$. The goal is to rewrite the probability in terms of Gaussian increments and without the overspill Δ_ℓ^{-a}. This is evaluated differently for $\ell = 0$, $\ell = 1$ and $1 < \ell \leqslant \mathcal{L}$.

For the range $1 < \ell \leqslant \mathcal{L}$, by Lemma B.3, the probability is

$$\mathbf{P}(\mathcal{N}_\ell - u_\ell \in [\Delta_\ell^{-a}, \Delta_\ell^{-1} - \Delta_\ell^{-a}]) + e^{-ce^{t_\ell}/2}, \tag{60}$$

for the Gaussian increment $\mathcal{N}_\ell = \mathcal{Z}_\ell - \mathcal{Z}_{\ell-1}$. To get rid of the overspills Δ_ℓ^{-a}, note that by the choice of u_ℓ and Δ_ℓ, we have

$$|u_\ell|\Delta_\ell^{-a} \leqslant 1, \tag{61}$$

so that

$$\mathbf{P}(\mathcal{N}_\ell - u_\ell \in [\Delta_\ell^{-a}, \Delta_\ell^{-1} - \Delta_\ell^{-a}]) = (1+\mathrm{O}(\Delta_\ell^{-a}))\mathbf{P}(\mathcal{N}_\ell - u_\ell \in [0, \Delta_\ell^{-1}]). \tag{62}$$

Moreover, we have trivially

$$\mathbf{P}(\mathcal{N}_\ell - u_\ell \in [0, \Delta_\ell^{-1}]) \gg \Delta_\ell^{-1} e^{-u_\ell^2/2v_\ell} \gg e^{-\Delta_\ell^2}, \tag{63}$$

by the choice of Δ_ℓ. This is much larger than the additive error in (60). We conclude that for each $1 < \ell \leqslant \mathcal{L}$

$$\mathbf{E}\left[|\mathcal{D}_\ell^-(\mathcal{Y}_\ell - u_\ell)|^2\right] \geqslant (1+\mathrm{O}(\Delta_\ell^{-a}))\mathbf{P}(\mathcal{N}_\ell \in [u_\ell, u_\ell + \Delta_\ell^{-1}]) \gg e^{-\Delta_\ell^2}. \tag{64}$$

Putting this together with the choice $a = 5$ allows us to rewrite (59) as

$$\begin{aligned}\mathbf{E}\left[\left(|\mathcal{D}_\ell^-(\mathcal{Y}_\ell - u_\ell)|^2 - e^{-\Delta_\ell^{a-2}}\right)\right] &\geqslant \left(1 - e^{-\Delta_\ell^{a-2}+\Delta_\ell^2}\right)\mathbf{E}\left[|\mathcal{D}_\ell^-(\mathcal{Y}_\ell - u_\ell)|^2\right] \\ &\geqslant (1+\mathrm{O}(\Delta_\ell^{-a}))\mathbf{P}(\mathcal{N}_\ell \in [u_\ell, u_\ell + \Delta_\ell^{-1}]).\end{aligned} \tag{65}$$

For $\ell = 1$, the additive error in Lemma B.3 is too big compared to the probability. Taking a larger t_0 to reduce the error would cause the polynomial $\mathcal{M}_0$ to be too long. To circumvent this issue, we estimate the probability using Lemma B.4. This is done at the cost of a multiplicative constant, which is harmless when it appears for a single term in the product over ℓ. We ultimately get

$$\mathbf{E}\Big[|\mathcal{D}_1^-(\mathcal{Y}_1 - u_1)|^2 - e^{-\Delta_1^{a-2}}\Big] \gg \mathbf{P}(\mathcal{N}_1 \in [u_1, u_1 + \Delta_1^{-1}]), \tag{66}$$

where $\mathcal{N}_1$ is a centered Gaussian with variance v_1^2. We now establish a similar bound for $\mathcal{Y}_0$. Lemma 2 implies that

$$\mathbf{E}\big[|\mathcal{D}_0^-(\mathcal{Y}_0 - u_0)|^2\big] \gg \mathbf{P}(\mathcal{Y}_0 \in [u_0 + \Delta_0^{-a}, u_0 + \Delta_0^{-1} - \Delta_0^{-a}]) \geqslant \frac{1}{2}\mathbf{P}(\mathcal{Y}_0 \in [u_0 + \Delta_0^{-1}, u_0 + 2\Delta_0^{-1}]).$$

To bound the probability from below, note that

$$\mathbf{E}[\mathcal{Y}_0] \geqslant 1, \qquad \mathbf{E}[\mathcal{Y}_0^2] = \mathbf{E}[|\mathcal{M}_0|^4] = \prod_{p \leqslant T_0}\left(1 + \frac{2}{p^{2\sigma}} + \mathrm{O}(p^{-3\sigma})\right) \ll \exp(2t_0).$$

This implies

$$\mathbf{P}(\mathcal{Y}_0 \in [u_0 + \Delta_0^{-1}, u_0 + 2\Delta_0^{-1}]) \gg \Delta_0^{-1}\mathbf{P}(\mathcal{Y}_0 > u_0 + 2\Delta_0^{-1}) \gg \Delta_0^{-1}\frac{\mathbf{E}[\mathcal{Y}_0]^2}{\mathbf{E}[\mathcal{Y}_0^2]} \gg e^{-\Delta_0^{a-2}}. \tag{67}$$

This ensures that in (59)

$$\mathbf{E}\Big[|\mathcal{D}_0^-(\mathcal{Y}_0 - u_0)|^2 - e^{-\Delta_0^{a-2}}\Big] \gg \mathbf{E}\big[|\mathcal{D}_0^-(\mathcal{Y}_0 - u_0)|^2\big] \gg \mathbf{P}(\mathcal{Y}_0 \in [u_0 + \Delta_0^{-1}, u_0 + 2\Delta_0^{-1}]). \tag{68}$$

Altogether, Equations (65), (66), and (68) show that the term for $J = \emptyset$ in (58) is

$$\gg \mathbf{P}(\mathcal{Y}_0 \in [u_0 + \Delta_0^{-1}, u_0 + 2\Delta_0^{-1}]) \prod_{1 \leqslant \ell \leqslant \mathcal{L}} \mathbf{P}(\mathcal{N}_\ell \in [u_\ell, u_\ell + \Delta_\ell^{-1}]). \tag{69}$$

It remains to estimate the summands in (58) with $J \neq \emptyset$. For each J, we get that each summand is bounded by (using that $\mathbf{1}\{a < b\} \leqslant (b/a)^{2q}$)

$$\leqslant \mathbf{E}\left[\prod_{0 \leqslant \ell \leqslant \mathcal{L}} (d_\ell(Y_\ell) - \varepsilon_\ell) \prod_{j \in J} X_j^{-2q_j} |Y_j - u_j|^{2q_j}\right], \tag{70}$$

for an appropriate choice of q_j, $j \in J$, and writing $d_\ell(y) = |D_\ell^-(y - u_\ell)|^2$, $\varepsilon_\ell = e^{-\Delta_\ell^{a-2}}$. We will show that this is subdominant to the contribution from $J = \emptyset$. Expanding the first factor, we get

$$\sum_{K \subseteq \{0,\ldots,\mathcal{L}\}} \prod_{k \in K^c} \varepsilon_k \mathbb{E}\left[\prod_{\ell \in K} d_\ell(Y_\ell) \prod_{j \in J} X_j^{-2q_j} |Y_j - u_j|^{2q_j}\right]. \tag{71}$$

The length of the polynomial in the expectation is then at most

$$\exp\left(\sum_{0\leqslant \ell\leqslant \mathcal{L}} 2\cdot 10^2 e^{t_\ell} X_\ell \Delta_\ell^{3a} + \sum_{j\leqslant \mathcal{L}} 2e^{t_j}(2q_j)\right). \tag{72}$$

The mean-value theorem of Dirichlet polynomials can be applied if q_j is not too large, for example, $q_j \leqslant X_j\Delta_j$ suffices. Passing to the random model, Equation (70) is

$$\leqslant \sum_{K\subseteq\{0,\dots,\mathcal{L}\}} \prod_{k\in K^c} \varepsilon_k \prod_{j\in J} X_j^{-2q_j} \prod_{\ell\in K\cap J^c} \mathbb{E}\big[d_\ell(\mathcal{Y}_\ell)\big] \prod_{j\in K\cap J} \mathbb{E}\big[d_j(\mathcal{Y}_j)|\mathcal{Y}_j-u_j|^{2q_\ell}\big] \prod_{j\in J\cap K^c} \mathbb{E}\big[|\mathcal{Y}_j-u_j|^{2q_j}\big]. \tag{73}$$

For the term with $j\in K\cap J$, we may split the expectation using

$$\ll \mathbb{E}\Big[|\mathcal{D}_j^-(\mathcal{Y}_j-u_j)|^4\Big]^{1/2}\mathbb{E}\big[\mathcal{Y}_j-u_j|^{4q_j}\big]^{1/2}. \tag{74}$$

We now estimate the moments of $\mathcal{D}_j^-(\mathcal{Y}_j-u_j)$ and $|\mathcal{Y}_j-u_j|$ separately. The fourth moment for the former, for $j\geqslant 1$, is estimated using the bounds on the coefficients in Lemma 2 and the elementary estimates on the random model, see Equation (B.12) in Lemma B.5,

$$\begin{aligned}
\mathbf{E}[|\mathcal{D}_j^-(\mathcal{Y}_j-u_j)|^4] &\leqslant \Delta_j^{4a}\mathbf{E}\left[\left(\sum_{\ell\geqslant 0}\frac{(2\pi\Delta_j^{2a})^\ell}{\ell!}|\mathcal{Y}_j-u_j|^\ell\right)^4\right]\\
&\leqslant \Delta_j^{4a}\mathbf{E}[\exp(8\pi\Delta_j^{2a}|\mathcal{Y}_j-u_j|)]\\
&\leqslant \Delta_j^{4a}e^{8\pi\Delta_j^{2a}|u_j|+8^2\pi^2\Delta_j^{4a}v_j^2/2}\\
&\leqslant \exp(100\Delta_j^{4a}),
\end{aligned}$$

since $|u_j|\leqslant\Delta_j$. For the second factor in (74), Minkowski's inequality implies

$$\mathbf{E}[|\mathcal{Y}_j-u_j|^{4q_j}]^{1/(4q_j)} \leqslant \mathbf{E}[|\mathcal{Y}_j|^{4q_j}]^{1/(4q_j)} + |u_j|.$$

Moreover, the bound (B.10) yields $\mathbf{E}[|\mathcal{Y}_j|^{4q_j}] \leqslant \frac{(4q_j)!}{2^{2q_j}(2q_j)!}v_j^{4q_j}$, which by Stirling gives the estimate

$$\mathbf{E}[|\mathcal{Y}_j|^{4q_j}]^{1/(4q_j)} \leqslant q_j^{1/2}.$$

We are now in position to pick $q_j = X_j$. Note that this satisfies the restriction $q_j\leqslant X_j\Delta_j^{3a}$ imposed by (72). With this choice, we have

$$X^{-4q_j}\mathbf{E}[|\mathcal{Y}_j-u_j|^{4q_j}] \leqslant \left(\frac{1}{X_j^{1/2}}+\frac{|u_j|}{X_j}\right)^{4q_j} \leqslant e^{-4X_j}.$$

Putting all these observations back in Equation (73) gives the upper bound

$$
\begin{aligned}
&\leqslant \sum_{K\subseteq\{0,\dots,\mathcal{L}\}} \prod_{k\in K^c} \varepsilon_k \prod_{j\in J} e^{50\Delta_j^{4a}-2X_j} \prod_{j\in K\cap J^c} \mathbb{E}\left[d_j(\mathcal{Y}_j)\right] \\
&= \left(\prod_{0\leqslant \ell\leqslant \mathcal{L}} \mathbb{E}\left[d_\ell(\mathcal{Y}_\ell)\right]\right) \sum_{K\subseteq\{0,\dots,\mathcal{L}\}} \prod_{j\in K\cap J} \frac{e^{50\Delta_j^{4a}-2X_j}}{\mathbb{E}[d_j(\mathcal{Y}_j)]} \prod_{j\in K^c\cap J} \varepsilon_j \cdot \frac{e^{50\Delta_j^{4a}-2X_j}}{\mathbb{E}[d_j(\mathcal{Y}_j)]} \prod_{j\in K^c\cap J^c} \frac{\varepsilon_j}{\mathbb{E}\left[d_j(\mathcal{Y}_j)\right]} \\
&= \left(\prod_{0\leqslant \ell\leqslant \mathcal{L}} \mathbb{E}\left[d_\ell(\mathcal{Y}_\ell)\right]\right) \prod_{j\in J} \frac{e^{50\Delta_j^{4a}-2X_j}}{\mathbb{E}[d_j(\mathcal{Y}_j)]} \left(\sum_{K\subseteq\{0,\dots,\mathcal{L}\}} \prod_{j\in K^c\cap J} \varepsilon_j \prod_{j\in K^c\cap J^c} \frac{\varepsilon_j}{\mathbb{E}\left[d_j(\mathcal{Y}_j)\right]} \right) \\
&\ll \left(\prod_{0\leqslant \ell\leqslant \mathcal{L}} \mathbb{E}\left[d_\ell(\mathcal{Y}_\ell)\right]\right) \prod_{j\in J} \frac{e^{50\Delta_j^{4a}-2X_j}}{\mathbb{E}[d_j(\mathcal{Y}_j)]},
\end{aligned}
$$

where we used the bound (64) and the value of ε_j in the last inequality. This leads to the choice

$$X_j = 100\Delta_j^{4a}. \tag{75}$$

This shows that the terms with $J \neq \emptyset$ are indeed subdominant to (69). The claim (53) follows after summing over $\mathbf{u} \in \mathcal{I}$ and appealing to the inclusion (49) with $\mathcal{Y}_0$ and $\mathcal{Z}_\ell$.

Proof of $\mathbf{P}(G(V)) \ll \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > V)$. The proof of the upper bound follows the same method using the polynomial $\mathcal{D}^+$ instead of $\mathcal{D}^-$. The proof is, in fact, slightly easier as we do not need to bound a long polynomial as in (70). The corresponding quantity needs to be estimated in the random model where the factors are decoupled. □

We will need the following spinoff of Proposition 1 in the proof of Proposition 4.

Corollary 2. *Consider the complex partial sums* $\widetilde{S}_k = \sum_{T_0<p\leqslant T_k} \frac{p^{-i\tau}}{p^\sigma} + \frac{p^{-2i\tau}}{2p^{2\sigma}}$, $1 \leqslant \ell \leqslant \mathcal{L}$. *For* $1 < \ell \leqslant \mathcal{L}$ *and* $q \leqslant 100\alpha^2(t_\ell - t_{\ell-1})$, *we have*

$$\mathbf{E}[|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}|^{2q}\mathbf{1}(G_{\ell-1})] \ll \mathbf{E}[|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}|^{2q}]\mathbf{P}(\{\mathcal{Z}_{\ell-1} \geqslant L_{\ell-1} - (\alpha \vee 1)^{-1}\} \cap \mathcal{G}_0),$$

Proof. This is done the same way as the proof of Proposition 2 by approximating the indicator function $\mathbf{1}(G_{\ell-1})$ by polynomials. The mean-value theorem for Dirichlet polynomial can be applied with the additional $(\widetilde{S}_\ell - \widetilde{S}_{\ell-1})^q$ since the total length is still admissible, as in Equations (56) and (72), by the bound on q. Since we only need an upper bound, one can ultimately drop all restrictions on $\mathcal{Z}_\ell$, $1 \leqslant \ell \leqslant \mathcal{L}$, except for the lower bound on $\mathcal{Z}_{\ell-1}$. The restriction on $\mathcal{Z}_0$ also remains. □

3.3 | Proof of Proposition 3

The main tool to prove the proposition is Lemma A.1 from the Appendix. The following lemma adapts it to our setting.

Lemma 3. *Consider the mollifier $\mathcal{M}$ defined in* (31)*. For $\alpha > 0$, let $G(V)$ be the event in* (23) *for $V \sim \alpha \log\log T$ and $\mathcal{Z}_\ell$, $0 \leqslant \ell \leqslant \mathcal{L}$, the random variables defined in* (15) *and* (14)*. We have for $\delta > \delta^\star(\alpha)$ given in* (33)

$$\mathbf{E}\left[\left|\zeta\mathcal{M}-1\right|^2 \mathbf{1}(G(V))\right] \ll \left(\frac{e^{-2\delta}}{\delta} + e^{-\mu_\mathcal{L}/10}\right)\cdot \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > V).$$

Proof. The result follows from Lemma A.1 after properly approximating the indicator function. To do this, we proceed as in the proof of Proposition 2 by writing $\mathbf{1}(G_\mathcal{L})$ as a Dirichlet polynomial. Writing the event in terms of the increments Y_ℓ and using the inclusion (49) yields

$$\begin{aligned}\sum_{\mathbf{u}\in\mathcal{I}} \mathbf{E}\left[\left|\zeta\mathcal{M}-1\right|^2 \mathbf{1}(Y_\ell \in [u_\ell, u_\ell + \Delta_\ell^{-1}], \ell \leqslant \mathcal{L})\right] \\ \leqslant 2\sum_{\mathbf{u}\in\mathcal{I}} \mathbf{E}\left[\left|\zeta\mathcal{M}-1\right|^2 \prod_{\ell\leqslant\mathcal{L}} |\mathcal{D}_+(Y_\ell - u_\ell)|^2\right],\end{aligned}$$

where the inequality follows by an application of Lemma 2. In view of the choice of $\mathfrak{s}$ in (57), the product is a well-factorable polynomial and of length admissible to apply Lemma A.1, cf. Equation (56). The above is then

$$\ll \left(\frac{e^{-2\delta}}{\delta} + e^{-\mu_\mathcal{L}/10}\right)\sum_{\mathbf{u}\in\mathcal{I}} \mathbf{E}\left[\prod_{\ell\leqslant\mathcal{L}} |\mathcal{D}_+(Y_\ell - u_\ell)|^2\right].$$

We can also replace Y_ℓ by $\mathcal{Y}_\ell$ at a cost of a small multiplication constant $1 + \mathrm{O}(T^{-99})$. It remains to reexpress the polynomial in terms of the $\mathcal{Y}_\ell$'s as in the proof of Proposition 2. This proves the claim. □

Proof of Proposition 3. By introducing the event that $\zeta\mathcal{M}$ is close to 1, we have

$$\begin{aligned}\mathbf{P}(\log|\zeta| > V) &\geqslant \mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V - \log|\zeta\mathcal{M}|\} \cap \{|\zeta\mathcal{M}-1| < \varepsilon\big) \\ &\geqslant \mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V - \log(1-\varepsilon)\} \cap \{|\zeta\mathcal{M}-1| < \varepsilon\}\big).\end{aligned}$$

If we also intersect with $G(V)$, we get the lower bound

$$\geqslant \mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V - \log(1-\varepsilon)\} \cap G(V)\big) - \mathbf{P}(\{|\zeta\mathcal{M}-1| > \varepsilon\} \cap G(V)).$$

The second term is bounded by Markov's inequality

$$\mathbf{P}(\{|\zeta\mathcal{M}-1| > \varepsilon\} \cap G(V)) \leqslant \frac{1}{\varepsilon^2}\mathbf{E}\big[|\zeta\mathcal{M}-1|^2 \mathbf{1}(G(V))\big].$$

The claimed bound then follows from Lemma 3. □

3.4 | Proof of Proposition 4

The first step consists in applying Lemma B.1. For this, it is necessary to consider the complex partial sums

$$\widetilde{S}_\ell = \sum_{T_0<p\leqslant T_\ell} \frac{p^{-\mathrm{i}\tau}}{p^\sigma} + \frac{p^{-2\mathrm{i}\tau}}{2p^{2\sigma}}, \quad \ell \geqslant 0. \tag{76}$$

We consider the decreasing events

$$A_\ell = A_{\ell-1} \cap \left\{|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}| \leqslant 10\alpha(t_\ell - t_{\ell-1})\right\},\ 1 \leqslant \ell \leqslant \mathcal{L},$$

with A_0 being the whole sample space. On $A_{\mathcal{L}}$, Lemma B.1 applied with $A = 10(\alpha \vee 1)$ implies

$$\prod_{1\leqslant\ell\leqslant\mathcal{L}} |\mathcal{M}_\ell^{-1}| \geqslant \exp\left(S_{\mathcal{L}} + \mathrm{O}^\star(5T_0^{-1/2}) - \sum_{\ell=1}^{\mathcal{L}} e^{-10(\alpha\vee1)(t_\ell - t_{\ell-1})}\right) \geqslant \exp(S_{\mathcal{L}} - (\alpha\vee1)^{-1}),$$

where we use the notation of Lemma B.1, writing $\mathrm{O}^\star$ for an error term with implicit constant smaller than 1. These observations together mean that

$$\begin{aligned}\mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V\} \cap G(V)\big) &\geqslant \mathbf{P}\big(S_{\mathcal{L}} + \log|\mathcal{M}_0|^{-1} > V + (\alpha\vee1)^{-1}\} \cap G(V)\big) - \mathbf{P}\big(G_{\mathcal{L}} \cap A_{\mathcal{L}}^c\big)\\ &\geqslant \mathbf{P}\big(G(V + (\alpha\vee1)^{-1})\big) - \mathbf{P}\big(G_{\mathcal{L}} \cap A_{\mathcal{L}}^c\big)\\ &\gg \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_{\mathcal{L}} > V) - \mathbf{P}\big(G_{\mathcal{L}} \cap A_{\mathcal{L}}^c\big).\end{aligned}$$

The second inequality is by the definition of $G(V)$ and the third is by Propositions 1 and 2. It remains to bound the second probability. It is smaller than

$$\sum_{\ell=1}^{\mathcal{L}} \mathbf{P}\big(G_{\ell-1} \cap \{|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}| > 10\alpha(t_\ell - t_{\ell-1})\}\big),$$

with the convention that G_0 is the whole sample space $[T, 2T]$. Markov's inequality for fixed $q \in \mathbb{N}$ can be applied to each summand to yield

$$\leqslant \frac{\mathbf{E}[|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}|^{2q}\mathbf{1}(G_{\ell-1})]}{|10\alpha(t_\ell - t_{\ell-1})|^{2q}} \ll \frac{\mathbf{E}[|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}|^{2q}]}{|10\alpha(t_\ell - t_{\ell-1})|^{2q}} \cdot \mathbf{P}\big(\mathcal{Z}_{\ell-1} \geqslant L_{\ell-1} - (\alpha\vee1)^{-1}, \mathcal{Z}_0 > \alpha t_0\big),$$

for $\ell > 1$ by Corollary 2. The factor $(\alpha\vee1)^{-1}$ is irrelevant by (18). For $\ell = 1$, we simply use the bound $\frac{\mathbf{E}[|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}|^{2q}]}{|10\alpha(t_\ell - t_{\ell-1})|^{2q}}$. The $2q$-moment can be estimated using Lemma B.6. Picking $q = \lfloor \alpha^2(t_\ell - t_{\ell-1})\rfloor$, this yields the upper bound

$$\ll \alpha\sqrt{t_\ell - t_{\ell-1}} \exp\left(\frac{-100\alpha^2(t_\ell - t_{\ell-1})^2}{t_\ell - t_{\ell-1}}\right)\mathbf{P}(\mathcal{Z}_{\ell-1} \geqslant L_{\ell-1}, \mathcal{Z}_0 > \alpha t_0)$$

$$\ll e^{-99\alpha^2(t_\ell - t_{\ell-1})} \frac{e^{-2\alpha^2\delta}}{\alpha\sqrt{t}} \exp(-\alpha^2(t_{\ell-1} - t_0) + 2\alpha C \log_{\ell-1} t)\mathbf{P}(\mathcal{Z}_0 > \alpha t_0)$$

$$\ll \frac{e^{-2\alpha^2\delta} e^{-\alpha^2 t}}{\alpha\sqrt{t}} e^{-10\alpha^2 \mathfrak{s} \log_{\ell-1} t} \cdot e^{\alpha^2 t_0}\mathbf{P}(\mathcal{Z}_0 > \alpha t_0).$$

The second inequality follows by a Gaussian estimate and the variance estimate (17). The factor $e^{\alpha^2 t_0}\mathbf{P}(\mathcal{Z}_0 > \alpha t_0)$ is $\ll a_\alpha e^{\gamma\alpha^2}$ by (B.15). Summing the above bound on $1 \leqslant \ell \leqslant \mathcal{L}$ yields

$$\mathbf{P}\big(G_\mathcal{L} \cap A_\mathcal{L}^c\big) \ll \frac{C_\alpha(\delta)}{\alpha\sqrt{t}} e^{-\alpha^2 t} e^{-10\alpha^2 \mathfrak{s} \log_\mathcal{L} t}.$$

By the choice of $\mathfrak{s}$, we thus have

$$\mathbf{P}\big(\{\log|\mathcal{M}^{-1}| > V\} \cap G(V)\big) \gg \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > V) - \mathbf{P}\big(G_\mathcal{L} \cap A_\mathcal{L}^c\big) \gg C_\alpha(\delta)\frac{e^{-\alpha^2 t}}{\alpha\sqrt{t}},$$

as claimed

4 | PROOF OF COROLLARY 1

We first establish that, for $k \geqslant 0$, $\sigma = 1/2 + \delta e^{-t_\mathcal{L}}$, and δ satisfying the conditions of Theorem 1,

$$\mathbf{E}[|\zeta(\sigma + \mathrm{i}\tau)|^{2k}] \geqslant c_k (\log T)^{k^2}, \tag{77}$$

with the coefficient proportional to

$$a_k e^{\gamma k^2}(k+1)^{-38k^2}. \tag{78}$$

First, we have

$$\mathbf{E}[|\zeta(\sigma + \mathrm{i}\tau)|^{2k}] = x^{2k}\mathbf{P}(|\zeta(\sigma + \mathrm{i}\tau)| > x)\Big|_0^\infty + 2k\int_\mathbb{R} e^{2ky}\mathbf{P}(|\zeta(\sigma + \mathrm{i}\tau)| > e^y)dy$$

$$\geq 2k\int_{\alpha^-}^{\alpha^+} e^{2k\alpha t}\mathbf{P}(\log|\zeta(\sigma + \mathrm{i}\tau)| \geq \alpha t)t \,\mathrm{d}\alpha$$

$$\gg 2k\int_{\alpha^-}^{\alpha^+} \frac{C_\alpha(\delta)}{\alpha} e^{2k\alpha t} e^{-\alpha^2 t}\sqrt{t}\,\mathrm{d}\alpha,$$

by localizing to a window around αt and employing Theorem 1. The values α^-, α^+ can be taken to be

$$\alpha^- = k - \frac{1}{\sqrt{t}}, \qquad \alpha^+ = k.$$

For $\alpha > 0$, $C_\alpha(\delta)$ is decreasing and therefore the above is greater than

$$\gg C_k(\delta)e^{k^2 t}\int_{\alpha^-}^{\alpha^+} e^{-(\alpha-k)^2 t}\sqrt{t}\,\mathrm{d}\alpha \gg C_k(\delta)e^{k^2 t}.$$

As noted in (19) the coefficient grows as (78) using the optimal $\delta \sim \log(\alpha+1)$, completing the proof of (77).

The corollary follows by demonstrating that the moments on-axis are an upper bound for the moments off-axis. To do so, we employ a result from [5], a generalization of a convexity result of Gabriel [8].

Proposition 5 [5, Corollary 3.4]. *Write $z = v + it$. Suppose that F is a complex valued function, analytic in the half-plane $v \geqslant 1/2$, such that $|F(z)| \to 0$ uniformly in $v > 1/2$ as $t \to \infty$ and*

$$\int_{\mathbb{R}} |F(v+\mathrm{i}t)|^k \mathrm{d}t \to 0$$

as $v \to \infty$, for $k > 0$. Then, for any $v \geqslant 1/2$ and $k > 0$

$$\int_{\mathbb{R}} |F(v+\mathrm{i}t)|^k \mathrm{d}t \leqslant \int_{\mathbb{R}} |F(1/2+\mathrm{i}t)|^k \mathrm{d}t.$$

The purpose of the following lemma, an adaptation of [5, Lemma 3.5], is to construct an analytic approximation for the indicator function of the rectangle

$$\{\nu + \mathrm{i}t\ :\ \nu \in [\tfrac{1}{2}, \sigma], t \in [T, 2T]\}.$$

Lemma 4 (Adapted from [5]). *Let $b_1 \in (0,1)$ and $\Delta, T, A, b_2 > 0$. There exists an entire function $\Phi_{\Delta,T}(z) \equiv \Phi(z)$ such that, for $z = \nu + \mathrm{i}t$ with $\nu \geqslant \frac{1}{2}$ and $t \in \mathbb{R}$,*

(i) *for* $|t - \frac{3}{2}T| \geqslant \frac{3}{4}(1+b_2)T$, *uniformly in* $\nu \geqslant \frac{1}{2}$, $\Phi(z) \ll_A b_2^{-A}\Delta^{1-A}$,
(ii) *for* $|t - \frac{3}{2}T| \leqslant \frac{3}{4}(1-b_1)T$, $|\Phi(z)| = 1 + \mathrm{O}_{b_1,A}(\Delta^{-A}) + \mathrm{O}((\nu - \frac{1}{2})\frac{\Delta^2}{T})$,
(iii) *for* $|t - \frac{3}{2}T| \leqslant \frac{3}{4}(1+b_2)T$, $|\Phi(z)| \ll 1 + (\nu - \frac{1}{2})\frac{\Delta^2}{T}$,
(iv) $\Phi(z) \to 0$ *uniformly in t as $\nu \to \infty$.*

In order to apply Proposition 5, we use a Dirichlet polynomial approximation of $\zeta(\sigma + \mathrm{i}t)$. Define for $\sigma \geqslant \frac{1}{2}$, $t \asymp T$

$$D(\sigma+\mathrm{i}t) = \begin{cases} \sum_{n \leqslant T^2} n^{-\sigma-\mathrm{i}t} w\left(\log \frac{e^{100}n}{T^2}\right), & \frac{1}{2} \leqslant \sigma \leqslant 2 \\ \sum_{n \leqslant T} n^{-\sigma-\mathrm{i}t}, & \sigma > \frac{1}{2}, \end{cases}$$

where the smoothing $w(x)$ is

$$w(x) = \frac{1}{2\pi\mathrm{i}}\int_{2-\mathrm{i}\infty}^{2+\mathrm{i}\infty} e^{-xz}\left(\frac{e^z - 1}{z}\right)^{100}\frac{\mathrm{d}z}{z}.$$

By [5, Lemma 2.8], we then have

$$\zeta(\sigma+\mathrm{i}t)=D(\sigma+\mathrm{i}t)+\mathrm{O}\left(\tfrac{1}{T}\right). \tag{79}$$

Set $F(z)=D(z)\Phi(z)$. By properties (i) and (iv) of Lemma 4, the requirements of Proposition 5 are satisfied. Thus, for any $k>0$, we have the "smoothed" version of the required bound:

$$\int_{\mathbb{R}}|D(\sigma+\mathrm{i}t)\Phi_{\Delta,T}(\sigma+\mathrm{i}t)|^{2k}\mathrm{d}t \leqslant \int_{\mathbb{R}}|D(1/2+\mathrm{i}t)\Phi_{\Delta,T}(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t. \tag{80}$$

The aim is now to unsmooth both sides using Lemma 4.

By property (ii), with $\Delta\asymp T\log T$ and $b_1=\frac{1}{3}$, and the approximation (79), we get the lower bound

$$\int_T^{2T}|\zeta(\sigma+\mathrm{i}t)|^{2k}\mathrm{d}t \ll \int_T^{2T}|D(\sigma+\mathrm{i}t)|^{2k}\mathrm{d}t \ll \int_{\mathbb{R}}|D(\sigma+\mathrm{i}t)\Phi_{\Delta,T}(\sigma+\mathrm{i}t)|^{2k}\mathrm{d}t. \tag{81}$$

The upper bound follows similarly, using properties (i), (iii), and (iv) of Lemma 4,

$$\begin{aligned}\int_{\mathbb{R}}|D(1/2+\mathrm{i}t)\Phi_{\Delta,T}(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t &\ll_A \Delta^{2k(1-A)}b_2^{-2kA}\int_{|t-\frac{3}{2}T|>\frac{3}{4}T(1+b_2)}|D(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t\\ &\quad+\int_{|t-\frac{3}{2}T|\leqslant\frac{3}{4}T(1+b_2)}|D(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t\\ &\ll \int_{|t-\frac{3}{2}T|\leqslant T}|D(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t\\ &\ll \int_T^{2T}|\zeta(1/2+\mathrm{i}t)|^{2k}\mathrm{d}t\end{aligned}$$

choosing A large enough and $b_2=1/3$. Combining with (77), (80), and (81), we arrive at the claimed statement.

APPENDIX A: A MOLLIFICATION LEMMA

This section contains a mollification lemma that was first developed by one of the authors together with Bourgade and Radziwiłł in conjunction with the work [4]. The lemma was in the end no longer needed in the proof and was therefore unpublished. We include it here as we believe that it is of independent interest. The content of the lemma is to show that ζ can be well approximated by $\mathcal{M}^{-1}$ for some appropriate mollifier $\mathcal{M}$. The error of the approximation can be made smaller if this approximation is made on an event represented by a Dirichlet polynomial $\mathcal{Q}$. In that regard, the lemma can be seen as a substantial extension of [2, Lemma 4.2], with additional technical difficulties coming from the presence of the arbitrary Dirichlet polynomial $\mathcal{Q}$. To state the lemma, we first recall the following definition from [3].

We say a Dirichlet polynomial $\mathcal{Q}$ is *degree-$\mathcal{L}$ well-factorable* if it can be written as

$$\prod_{0\leqslant\ell\leqslant\mathcal{L}}\mathcal{Q}_\ell(s),$$

where

$$\mathcal{Q}_\ell(s) = \sum_{\substack{p|m \Rightarrow p \in (T_{\ell-1}, T_\ell] \\ \Omega_\ell(k) \leqslant \nu_\ell}} \frac{\gamma_\ell(k)}{k^s}, \quad \nu_\ell = \frac{1}{10}\mu_\ell,$$

and the coefficients $\gamma_\ell(k)$ are arbitrary. Here, $\Omega_\ell(k)$ is the number of prime factors of k in the interval $(T_{\ell-1}, T_\ell]$, and μ_ℓ defines the length of the mollifier in (29). Note that at $\ell = 0$, there are no restrictions on the number of prime factors in $\mathcal{M}_0$ given in (9), so $\mu_0 = \infty$ effectively.

A degree-ℓ well-factorable Dirichlet polynomial $\mathcal{Q}(s) = \sum_k \frac{\gamma(k)}{k^s}$ is short for our purpose since the length is

$$\leqslant \prod_{\ell \leqslant \mathcal{L}} T_\ell^{\nu_\ell} \leqslant \exp\left(e^t \sum_{\ell \leqslant \mathcal{L}} \frac{10(\alpha \vee 1)\mathfrak{s} \log_{\ell-1} t)}{(\log_{\ell-1} t)^{\mathfrak{s}}}\right) \leqslant \exp\left(e^t 20(\alpha \vee 1)\mathfrak{s}(\log_{\mathcal{L}-1} t)^{1-\mathfrak{s}}\right) \leqslant T^{1/100}, \tag{A.1}$$

for any $\ell \leqslant \mathcal{L}$ by the choice of $\mathfrak{s}$ in (12).

Lemma A.1. *Consider the mollifiers $\mathcal{M}_\ell$, $\ell \leqslant \mathcal{L}$ (see (29) and (31)) and write*

$$\sigma_{\mathcal{L}} = \frac{1}{2} + \frac{\delta}{e^{t_{\mathcal{L}}}}$$

for some fixed $\delta > 0$. For any degree-$\mathcal{L}$ well-factorable $\mathcal{Q}$, we have for T large enough

$$\mathbf{E}\left[\left|\zeta\mathcal{M} - 1\right|^2 |\mathcal{Q}|^2(\sigma_{\mathcal{L}} + \mathrm{i}\tau)\right] \leqslant \left(\frac{e^{-2\delta}}{\delta} + e^{-\mu_{\mathcal{L}}/10}\right) \cdot \mathbf{E}\left[|\mathcal{Q}|^2(\sigma_{\mathcal{L}} + \mathrm{i}\tau)\right].$$

Proof. To lighten the notation, we write throughout the proof $\sigma = \sigma_{\mathcal{L}}$ and recall that $\mathcal{M} = \prod_{0 \leqslant \ell \leqslant \mathcal{L}} \mathcal{M}_\ell$. We also sometimes drop the dependence on $\sigma_{\mathcal{L}} + \mathrm{i}\tau$ in the notation. The proof is by evaluating the expectations for each term in the expansion

$$|\zeta\mathcal{M} - 1|^2 = |\zeta\mathcal{M}|^2 - \zeta\mathcal{M} - \overline{\zeta\mathcal{M}} + 1. \tag{A.2}$$

Estimating $\mathbf{E}[\zeta\mathcal{M}|\mathcal{Q}|^2]$.
We prove

$$\mathbf{E}[\zeta\mathcal{M}|\mathcal{Q}|^2] = \mathbf{E}[|\mathcal{Q}|^2] \cdot (1 + \mathrm{O}(T^{-1/2})). \tag{A.3}$$

The same obviously holds for $\mathbf{E}[\overline{\zeta\mathcal{M}}|\mathcal{Q}|^2]$. We use the approximation (see, e.g., [12, Theorem 1.8])

$$\zeta(\sigma + \mathrm{i}\tau) = \sum_{n \leqslant T} \frac{1}{n^{\sigma + \mathrm{i}\tau}} + \mathrm{O}(T^{-1/2}). \tag{A.4}$$

We define $a_\ell(n) = 1$ if $\Omega_\ell(n) \leqslant \mu_\ell$ and all prime factors of n are in $(T_{\ell-1}, T_\ell]$, and 0 otherwise. For $n \in \mathbb{N}$, we also write $n_\ell = \prod_{p|n, p \in (T_{\ell-1}, T_\ell]} p^{v_p(n)}$, where $v_p(n)$ stands for the p-valuation of n. We then get a multiplicative function $a(n) = \prod_{\ell \leqslant \mathcal{L}} a_\ell(n_\ell)$. As μ is also multiplicative, we can now write the mollifier in a compact form

$$\mathcal{M}(s) = \sum_m \frac{\mu(m)a(m)}{m^s}.$$

By redefining the coefficient γ in the same way to include the restriction on prime factors, we can also write $\mathcal{Q}(s) = \sum_k \frac{\gamma(k)}{k^s}$, so that

$$\mathcal{M}|\mathcal{Q}|^2(s) = \sum_{m,k_1k_2} \frac{\mu(m)a(m)\gamma(k_1)\overline{\gamma}(k_2)}{(mk_1k_2)^s}. \tag{A.5}$$

The contribution from the error term in (A.4) is

$$\ll T^{-1/2} \sum_{m,k_1,k_2} \frac{|\gamma(k_1)\gamma(k_2)|}{(mk_1k_2)^\sigma} \ll T^{-\frac{1}{2}+\frac{1}{100}} \left(\sum_k \frac{|\gamma(k)|}{k^\sigma} \right)^2 \ll T^{-\frac{1}{2}+\frac{2}{100}} \sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}}, \tag{A.6}$$

where we have used the bound (32) on m in the second inequality, and the Cauchy–Schwarz inequality in the third. We now observe that $\mathbf{E}[|\mathcal{Q}|^2]$ is close to $\sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}}$ since

$$\begin{aligned} \mathbf{E}[|\mathcal{Q}|^2] - \sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}} = \sum_{k_1 \neq k_2} \frac{\gamma(k_1)\overline{\gamma}(k_2)}{(k_1k_2)^\sigma} \mathbf{E}\left[\left(\frac{k_2}{k_1} \right)^{\mathrm{i}\tau} \right] &\ll \sum_{k_1,k_2} \frac{|\gamma(k_1)\gamma(k_2)|}{(k_1k_2)^\sigma} \frac{k_2}{T} \\ &\ll T^{-1+\frac{1}{100}} \sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}}. \end{aligned}$$

We used $\mathbf{E}[(\frac{k_2}{k_1})^{\mathrm{i}\tau}] \ll 1/(T|\log(k_1/k_2)|) \ll \frac{k_2}{T}$, (A.1) to bound k_2, and the Cauchy–Schwarz inequality. This shows that

$$\sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}} = (1 + \mathrm{O}(T^{-1/2}))\mathbf{E}[|\mathcal{Q}|^2]. \tag{A.7}$$

The above equation together with (A.6) then yields

$$\mathbf{E}\left[\zeta\mathcal{M}|\mathcal{Q}|^2(\sigma + \mathrm{i}\tau)\right] = \sum_{n \leqslant T,m,k_1k_2} \frac{\mu(m)a(m)\gamma(k_1)\overline{\gamma}(k_2)}{(nmk_1k_2)^\sigma} \mathbf{E}\left[\left(\frac{k_2}{nmk_1} \right)^{\mathrm{i}\tau} \right] + \mathrm{O}(T^{-1/2})\mathbf{E}[|\mathcal{Q}|^2]. \tag{A.8}$$

We first bound the terms $nm \neq k_2/k_1$, based on $\mathbf{E}[(\frac{k_2}{nmk_1})^{\mathrm{i}\tau}] \ll 1/(T|\log \frac{nmk_1}{k_2}|) \ll \frac{k_2}{T}$, and (A.1) and (32) as before:

$$\begin{aligned} \sum_{nm \neq k_2/k_1} \frac{\mu(m)a(m)\gamma(k_1)\overline{\gamma}(k_2)}{(mn)^\sigma} \mathbf{E}\left[\left(\frac{k_2}{k_1nm} \right)^{\mathrm{i}\tau} \right] &\ll T^{\frac{1}{100}} \sum_{n \leqslant T,k_1,k_2} \frac{\gamma(k_1)\overline{\gamma}(k_2)}{(mn)^\sigma} \frac{k_2}{T} \\ &\ll T^{-1+\frac{2}{100}} \sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}} \\ &\ll T^{-1+\frac{2}{100}} \cdot \mathbf{E}[|\mathcal{Q}|^2], \end{aligned} \tag{A.9}$$

where (A.7) was used for the last bound.

It remains to estimate the contribution of the terms $nm = k_2/k_1$. Note that the restriction $n \leqslant T$ is no longer needed here since it is implied by $n = k_2/(k_1 m)$. Moreover, $m = k_2/(k_1 n)$ and $\Omega_\ell(k_2) \leqslant \mu_\ell$ imposes $a(m) = 1$. With these simple observations, we are left to estimate, for fixed k_1 dividing k_2, the sum $\sum_{nm=k_2/k_1} \mu(m) = \mathbf{1}_{k_2=k_1}$. With Equations (A.7)–(A.9), this gives (A.3).

Estimating $\mathbf{E}[|\zeta \mathcal{M} \mathcal{Q}|^2]$.

We will now prove the more delicate asymptotics

$$\mathbf{E}[|\zeta \mathcal{M}\mathcal{Q}|^2] = \mathbf{E}[|\mathcal{Q}|^2] \cdot \left(1 + \mathrm{O}\left(\frac{e^{-2\delta}}{\delta} + e^{-\mu_{\mathcal{L}}}\right)\right), \tag{A.10}$$

relying on the expansion

$$\mathbf{E}[|\zeta \mathcal{M}\mathcal{Q}|^2] = \sum_{k_1,k_2,m_1,m_2} \frac{\mu(m_1)\mu(m_2)a(m_1)a(m_2)\gamma(k_1)\overline{\gamma(k_2)}}{(m_1 m_2 k_1 k_2)^\sigma} \mathbf{E}\left[|\zeta|^2 \left(\frac{k_1 m_1}{k_2 m_2}\right)^{\mathrm{i}\tau}\right], \tag{A.11}$$

and on the estimate (see [18, Lemma 4])

$$\begin{aligned}
&\int_T^{2T} \left(\frac{n_1}{n_2}\right)^{\mathrm{i}t} |\zeta(\sigma + \mathrm{i}t)|^2 \mathrm{d}t \\
&= \int_T^{2T} \left(\zeta(2\sigma)\left(\frac{(n_1,n_2)^2}{n_1 n_2}\right)^\sigma + \left(\frac{t}{2\pi}\right)^{1-2\sigma} \zeta(2-2\sigma)\left(\frac{(n_1,n_2)^2}{n_1 n_2}\right)^{1-\sigma}\right)\mathrm{d}t \\
&\quad + \mathrm{O}(T^{1-\sigma+\varepsilon}\min(n_1,n_2)).
\end{aligned} \tag{A.12}$$

As $k_1, m_1, k_2, m_2 \leqslant T^{1/100}$, and $|a(m_1)|, |a(m_2)| \leqslant 1$, the contribution in (A.11) from the error term in (A.12) is

$$\ll T^{-\frac{1}{4}} \sum_{k_1,k_2} \frac{|\gamma(k_1)\gamma(k_2)|}{(k_1 k_2)^\sigma} \ll T^{-\frac{1}{10}} \sum_k \frac{|\gamma(k)|^2}{k^{2\sigma}} \leqslant T^{-\frac{1}{100}} \cdot \mathbf{E}[|\mathcal{Q}|^2],$$

where we have used (A.7) in the last inequality.

Contribution from the right of Re $s = 1$

We now consider the contribution to (A.11) from the term $\zeta(2\sigma)\left(\frac{(h,k)^2}{hk}\right)^\sigma$ in (A.12):

$$\zeta(2\sigma) \sum_{k_1,k_2} \frac{\gamma(k_1)\overline{\gamma(k_2)}}{(k_1 k_2)^{2\sigma}} \sum_{m_1,m_2} \frac{\mu(m_1)\mu(m_2)a(m_1)a(m_2)}{(m_1 m_2)^{2\sigma}} (k_1 m_1, k_2 m_2)^{2\sigma}. \tag{A.13}$$

We first notice that we can estimate the sums for each interval of primes $(T_{\ell-1}, T_\ell]$ and take the product on ℓ at the very end, by the definition of the a's and the fact that $\mathcal{Q}$ is well-factorable.

With this in mind, let us first estimate the sum without the restriction on Ω_ℓ. Fix k_1 and k_2. Assume first that $k_1 \neq k_2$. Without loss of generality, we can assume that $(k_1, k_2) = 1$; otherwise, we can factor it out. We write $\mathcal{P}_1$ for the set of prime factors of k_1 and similarly for k_2. The sum over m_1 and m_2 without the restriction can then be expressed as an Euler product as follows:

$$\sum_{\substack{m_1,m_2\\ p|m_i \Rightarrow p\in(T_{\ell-1},T_\ell], i=1,2}} \frac{\mu(m_1)\mu(m_2)}{(m_1m_2)^{2\sigma}}(k_1m_1,k_2m_2)^{2\sigma}$$
$$= \prod_{p\in(\mathcal{P}_1\cup\mathcal{P}_2)^c}(1-p^{-2\sigma}-p^{-2\sigma}+p^{-2\sigma})\prod_{p\in\mathcal{P}_1}(1-1+p^{-2\sigma}-p^{-2\sigma})\prod_{p\in\mathcal{P}_2}(1-1+p^{-2\sigma}-p^{-2\sigma}). \tag{A.14}$$

This is obviously 0. With this crucial observation, we have that

$$\sum_{p|k_i \Rightarrow p\in(T_{\ell-1},T_\ell], i=1,2} \frac{\gamma(k_1)\overline{\gamma(k_2)}}{(k_1k_2)^{2\sigma}} \sum_{p|m_i \Rightarrow p\in(T_{\ell-1},T_\ell], i=1,2} \frac{\mu(m_1)\mu(m_2)}{(m_1m_2)^{2\sigma}}(k_1m_1,k_2m_2)^{2\sigma}$$
$$= \sum_{p|k \Rightarrow p\in(T_{\ell-1},T_\ell]} \frac{|\gamma(k)|^2}{k^{2\sigma}} \sum_{p|m_i \Rightarrow p\in(T_{\ell-1},T_\ell], i=1,2} \frac{\mu(m_1)\mu(m_2)}{(m_1m_2)^{2\sigma}}(m_1,m_2)^{2\sigma}$$
$$= (1+\mathrm{O}(T^{-1/2}))\mathbf{E}[|\mathcal{Q}_\ell|^2]\prod_{p\in(T_{\ell-1},T_\ell]}\left(1-\frac{1}{p^{2\sigma}}\right), \tag{A.15}$$

where we used (A.7).

We now bound the contribution from integers m_1 with $\Omega_\ell(m_1) > \mu_\ell$ for a fixed $\ell \geqslant 1$. We use Rankin's trick, that is, $\mathbf{1}(\Omega_\ell(m) > \mu_\ell) \leqslant e^{-\mu_\ell}e^{\Omega_\ell(m)}$. We get that the sum is bounded by

$$e^{-\mu_\ell}\sum_{\substack{m_1,m_2,k_1,k_2\\ p|k_i,m_i \Rightarrow p\in(T_{\ell-1},T_\ell], i=1,2}} \frac{\gamma(k_1)\overline{\gamma(k_2)}}{(k_1k_2)^{2\sigma}} \cdot \frac{|\mu(m_1)\mu(m_2)|e^{\Omega_\ell(m_1)}}{(m_1m_2)^{2\sigma}}(k_1m_1,k_2m_2)^{2\sigma}. \tag{A.16}$$

The sum over m_1, m_2 is computed similarly to (A.14). It is

$$= (k_1,k_2)^{2\sigma}\prod_{\mathcal{P}_1}\left(2+\frac{2e}{p^{2\sigma}}\right)\prod_{\mathcal{P}_2}\left(1+e+\frac{1+e}{p^{2\sigma}}\right)\prod_{(\mathcal{P}_1\cup\mathcal{P}_2)^c}\left(1+\frac{2e+1}{p^{2\sigma}}\right)$$
$$\leqslant (k_1,k_2)^{2\sigma}(1+e)^{\omega_\ell(k_1)+\omega_\ell(k_2)}\prod_{p\in(T_{\ell-1},T_\ell]}\left(1+\frac{2e+1}{p^{2\sigma}}\right),$$

where $\omega_\ell(m)$ stands for the number of distinct prime factors of m in $(T_{\ell-1}, T_\ell]$. Using the fact that $|\gamma(k_1)\gamma(k_2)| \leqslant (|\gamma(k_1)|^2 + |\gamma(k_2)|^2)/2$, the sum in (A.16) is bounded by

$$\prod_{p\in(T_{\ell-1},T_\ell]}\left(1+\frac{2e+1}{p^{2\sigma}}\right)\sum_{\substack{k_1\\ p|k_1 \Rightarrow p\in(T_{\ell-1},T_\ell]}} \frac{|\gamma(k_1)|^2}{k_1^{2\sigma}}(1+e)^{\omega_\ell(k_1)}$$
$$\times \sum_{\substack{k_2\\ p|k_2 \Rightarrow p\in(T_{\ell-1},T_\ell]}} \frac{(k_1,k_2)^{2\sigma}}{k_2^{2\sigma}}(1+e)^{\omega_\ell(k_2)}. \tag{A.17}$$

For k_1 fixed, we can express the sum over k_2 as an Euler product

$$
\begin{aligned}
&= \prod_{p|k_1}\left(1+(1+e)v_p(k_1)+(1+e)\sum_{j>v_p(k_1)} p^{-2\sigma(j-v_p(k_1))}\right)\prod_{p\nmid k_1}\left(1+(1+e)\sum_{j\geqslant 1} p^{-2\sigma j}\right)\\
&\leqslant (1+e)^{\omega_\ell(k_1)}\prod_{p|k_1}(1+v_p(k_1))\prod_{p\nmid k_1}\left(1+\frac{2e+1}{p^{2\sigma}}\right)\\
&\leqslant (1+e)^{\Omega_\ell(k_1)}e^{\Omega_\ell(k_1)}\prod_{p\in(T_{\ell-1},T_\ell]}\left(1+\frac{2e+1}{p^{2\sigma}}\right),
\end{aligned}
$$

where we use the fact that $1+x\leqslant e^x$ to estimate the product over $p|k_1$. Using the assumption that $\Omega_\ell(k_1)\leqslant \nu_\ell$, we can finally bound (A.16) by

$$
\leqslant (1+\mathrm{O}(T^{-1/2}))(1+e)^{3\nu_\ell-\mu_\ell}\prod_{p\in(T_{\ell-1},T_\ell]}\left(1+\frac{2e+1}{p^{2\sigma}}\right)^2\mathbf{E}[|\mathcal{Q}_\ell|^2], \tag{A.18}
$$

where we used (A.7). By the choice of ν_ℓ, we have that $3\nu_\ell-\mu_\ell<\mu_\ell/2$.

To estimate (A.13), it remains to take the product over ℓ and multiply by $\zeta(2\sigma)=\prod_p(1-p^{-2\sigma})^{-1}$. We now apply the identity

$$
\prod_{\ell\leqslant\mathcal{L}}(1-\mathbf{1}(\Omega_\ell(m)>\mu_\ell))=\sum_{I\subseteq\{1,\dots,\mathcal{L}\}}\prod_{\ell\in I}(-1)^{|I|}\mathbf{1}(\Omega_\ell(m)>\mu_\ell)
$$

to the sums in m_1 and m_2 in (A.13) to get that it is equal to

$$
\begin{aligned}
&\mathbf{E}[|\mathcal{Q}|^2]\Bigg\{(1+\mathrm{O}(T^{-1/2}))\prod_{p>T_\mathcal{L}}\left(1-\frac{1}{p^{2\sigma}}\right)^{-1}-\\
&\qquad \mathrm{O}^\star\left(\sum_{I_1\cup I_2\neq\emptyset}\prod_{\ell\in I_1\cup I_2}e^{-\mu_\ell/2}\prod_{p\in(T_{\ell-1},T_\ell]}\left(1+\frac{2e+1}{p^{2\sigma}}\right)^4\left(1-\frac{1}{p^{2\sigma}}\right)^{-1}\right)\Bigg\}.
\end{aligned}\tag{A.19}
$$

Here, the sum is over subsets I_1 and I_2 whose union is not empty. They represent the ℓ's for which $\Omega_\ell(m)>\mu_\ell$ for m_1 and m_2, respectively. As in Lemma B.1, $\mathrm{O}^\star$ represents an error term with implicit constant less than 1. We then use the estimate (A.15) when $\ell\in(I_1\cup I_2)^c$ and the estimate (A.18) when $\ell\in I_1\cup I_2$. Note that in the case where $\Omega_\ell(m)>\mu_\ell$ for both m_1 and m_2, we need to square the term $\left(1+\frac{2e+1}{p^{2\sigma}}\right)^2$. The first product in the braces is estimated by the prime number theorem:

$$
\begin{aligned}
\sum_{p>T_\mathcal{L}}\frac{1}{p^{2\sigma}}\leqslant\int_{T_\mathcal{L}}^\infty\frac{1}{u^{2\sigma}\log u}\mathrm{d}u+\mathrm{O}(e^{-c\sqrt{T_\mathcal{L}}})&=\int_{\log T_\mathcal{L}}^\infty\frac{e^{-v(2\sigma-1)}}{v}\mathrm{d}v+\mathrm{O}(e^{-c\sqrt{T_\mathcal{L}}})\\
&\leqslant\frac{e^{-2\delta}}{2\delta}+\mathrm{O}(e^{-c\sqrt{T_\mathcal{L}}}),
\end{aligned}
$$

by the definition of σ and δ. Therefore, $\prod_{p>T_{\mathcal{L}}}\left(1-p^{-2\sigma}\right)^{-1}$ is $1+e^{-2\delta}/\delta$ for T large enough.

Write $\rho_\ell = \prod_{p\in(T_{\ell-1},T_\ell]}\left(1+(2e+1)p^{-2\sigma}\right)^4\left(1-p^{-2\sigma}\right)^{-1}$. We have that locally the factors are bounded by $(1+100p^{-2\sigma})$. Moreover, Mertens's theorem implies $e^{-\mu_\ell/2}\rho_\ell \leqslant e^{-\mu_\ell/4}$ (for $\mathfrak{s} > 100$). The $\mathrm{O}^\star$-term is then

$$\sum_{I_1\cup I_2\neq\emptyset}\prod_{\ell\in I_1\cup I_2} e^{-\mu_\ell/2}\rho_\ell \leqslant \prod_{\ell\leqslant\mathcal{L}}(1+e^{-\mu_\ell/4})^2 - 1 \leqslant e^{-\mu_\ell/10}.$$

This concludes the proof of the contribution in (A.11) from the term $\zeta(2\sigma)\left(\frac{(n_1,n_2)^2}{n_1n_2}\right)^\sigma$ in (A.12).

Contribution from the left of Re $s = 1$

We now consider the contribution from $\left(\frac{t}{2\pi}\right)^{1-2\sigma}\zeta(2-2\sigma)\left(\frac{(n_1,n_2)^2}{n_1n_2}\right)^{1-\sigma}$ in (A.12). The integral factor there is

$$\frac{1}{T}\int_T^{2T}\left(\frac{t}{2\pi}\right)^{1-2\sigma}\mathrm{d}t \ll T^{1-2\sigma} = \exp(-2\delta e^{t-t_{\mathcal{L}}}). \tag{A.20}$$

For the other terms, we can work on each interval $(T_{\ell-1},T_\ell]$ for ℓ fixed as before. Without the restriction on the number of prime factors, the sum over k's and m's is

$$\sum_{\substack{m_1,m_2,k_1,k_2\\ p|k_i,m_i\Rightarrow p\in(T_{\ell-1},T_\ell],i=1,2}}\frac{\gamma(k_1)\overline{\gamma(k_2)}}{k_1k_2}\frac{\mu(m_1)\mu(m_2)}{m_1m_2}(k_1m_1,k_2m_2)^{2-2\sigma}. \tag{A.21}$$

We write $k_1 = (k_1,k_2)l_1$, $k_2 = (k_1,k_2)l_2$ where l_1 and l_2 are coprime with corresponding prime divisors sets $\mathcal{P}_1$ and $\mathcal{P}_2$. The above sum over m_1, m_2 is equal to

$$\begin{aligned}&\sum_{\substack{m_1,m_2\\ p|m_1,m_2\Rightarrow p\in(T_{\ell-1},T_\ell]}}\frac{\mu(m_1)\mu(m_2)}{m_1m_2}(k_1m_1,k_2m_2)^{2-2\sigma}\\ &\quad= (k_1,k_2)^{2-2\sigma}\prod_{p\in(\mathcal{P}_1\cup\mathcal{P}_2)^c}(1-2p^{-1}+p^{-2\sigma})\prod_{p\in\mathcal{P}_1}(1-p^{-1}-p^{1-2\sigma}+p^{-2\sigma})\\ &\quad\times\prod_{p\in\mathcal{P}_2}(1-p^{-1}-p^{1-2\sigma}+p^{-2\sigma}).\end{aligned} \tag{A.22}$$

As opposed to (A.14), the above product does not vanish when $k_1\neq k_2$. Therefore, estimating the sum is slightly more involved. We define

$$c_\ell = (2\sigma-1)\log T_\ell.$$

On $(\mathcal{P}_1\cup\mathcal{P}_2)^c$, we bound

$$1-p^{-1}-p^{-1}+p^{-2\sigma}\leqslant 1-p^{-1},$$

and on $\mathcal{P}_1\cup\mathcal{P}_2$, we have when $p\leqslant T_\ell$,

$$1-p^{-1}-p^{1-2\sigma}+p^{-2\sigma} = (1-p^{-1})\cdot(1-p^{1-2\sigma})\leqslant(1-p^{-1})c_\ell.$$

This shows that the sum in (A.22) is

$$\leqslant (k_1,k_2)^{2-2\sigma} \prod_{p\in(T_{\ell-1},T_\ell]} \left(1-\frac{1}{p}\right)\cdot c_\ell^{\omega_\ell(k_1)+\omega_\ell(k_2)}. \tag{A.23}$$

Moreover, bounding the γ's as before, we get that (A.22) is

$$\leqslant \sum_{\substack{k_1,k_2\\ p|k_1,k_2\Rightarrow p\in(T_{\ell-1},T_\ell]}} \frac{|\gamma(k_1)|^2}{k_1k_2}(k_1,k_2)^{2-2\sigma} c_\ell^{\omega_\ell(k_1)+\omega_\ell(k_2)}.$$

The sum over k_2 can be evaluated for fixed k_1 by reexpressing it as an Euler product. We get that

$$\begin{aligned}
&\sum_{\substack{k_2\\ p|k_2\Rightarrow p\in(T_{\ell-1},T_\ell]}} c_\ell^{\omega_\ell(k_2)}\frac{(k_1,k_2)^{2-2\sigma}}{k_2}\\
&= \prod_{p\nmid k_1}\left(1+\sum_{j=1}^{v_p(k_1)} c_\ell p^{j(1-2\sigma)} + \sum_{j>v_p(k_1)} c_\ell p^{v_p(k_1)(2-2\sigma)}p^{-j}\right)\prod_{p|k_1}\left(1+\sum_{j\geqslant 1}\frac{c_\ell}{p^j}\right)\\
&\leqslant \prod_{p|k_1}\left(1+c_\ell v_p(k_1)p^{1-2\sigma}+c_\ell p^{v_p(k_1)(1-2\sigma)}\right)\prod_{p\nmid k_1}\left(1+\frac{c_\ell}{p-1}\right).
\end{aligned}$$

We factor $p^{(1-2\sigma)v_p(k_1)}$ from the first product to get $k_1^{1-2\sigma}$. Combined with k_1, this will get the factor $k_1^{-2\sigma}$ needed to recover Q_ℓ^2. The second product is bounded by $\exp(2c_\ell(t_\ell - t_{\ell-1}))$ by Mertens's theorem. Putting this together yields the bound

$$\begin{aligned}
&k_1^{1-2\sigma}e^{2c_\ell(t_\ell-t_{\ell-1})}\prod_{p|k_1}\left(p^{v_p(k_1)(2\sigma-1)}+c_\ell v_p(k_1)p^{(v_p(k_1)-1)(2\sigma-1)}+c_\ell\right)\\
&\leqslant k_1^{1-2\sigma}e^{2c_\ell(t_\ell-t_{\ell-1})}(c_\ell+c_\ell\nu_\ell e^{c_\ell\nu_\ell}+e^{c_\ell\nu_\ell})^{\nu_\ell}\leqslant k_1^{1-2\sigma}e^{2c_\ell(t_\ell-t_{\ell-1})+c_\ell\nu_\ell^2},
\end{aligned}$$

where we used the definition of c_ℓ and the bound $\Omega_\ell(k_1)\leqslant \mu_\ell$ to get the inequality. Taking the product over ℓ and including the integral factor (A.20), we can now conclude that the contribution to (A.21) of sums on m's and k's without the restriction on the number of factors of the m's is

$$\zeta(2-2\sigma)\prod_{p\leqslant T_{\mathcal{L}}}(1-p^{-1})e^{\sum_{\ell\leqslant\mathcal{L}}2c_\ell\nu_\ell^2-2\delta e^{t-t_{\mathcal{L}}}}\sum_{k_1}\frac{|\gamma(k_1)|^2}{k_1^{2\sigma}}.$$

The first two factors are of order 1. The sum over k_1 is $\ll \mathbf{E}[|\mathcal{Q}|^2]$ by (A.7). Since $c_{\mathcal{L}}=2\delta$, the exponential factor is

$$\leqslant \exp\left(4c_{\mathcal{L}}\nu_{\mathcal{L}}^2-2\delta e^{t-t_{\mathcal{L}}}\right)\leqslant \exp\left(8\delta(10\mathfrak{s}\log_{\mathcal{L}} t)^2-2\delta e^{\mathfrak{s}\log_{\mathcal{L}} t}\right)\leqslant e^{-4\delta},$$

under the condition that $\mathfrak{s}\log_{\mathcal{L}} t > 100$. This is the dominant contribution coming from $\mathrm{Re}\, s = 2 - 2\sigma$. Finally, we need to bound the contribution to the sum over m's of the restrictions $\Omega_\ell(m) > \mu_\ell$ for each ℓ. This is done exactly as before starting from (A.16). $\square$

APPENDIX B: AUXILIARY RESULTS

The following lemma is an adaption of [4, Lemma 23] that gives upper and lower bound in approximating $\mathcal{M}_\ell$ by $e^{-(S_\ell - S_{\ell-1})}$ with precise error.

Lemma B.1. *Consider the partial sums (76). Let $\ell \geqslant 1$ and $A > 0$. If $|\widetilde{S}_\ell - \widetilde{S}_{\ell-1}| \leqslant A(t_\ell - t_{\ell-1})$, then the mollifier $\mathcal{M}_\ell$ defined in (29) is satisfies*

$$|\mathcal{M}_\ell| = \left(1 + \mathrm{O}^\star\left(\frac{5}{\sqrt{T_{\ell-1}}}\right)\right) e^{-(S_\ell - S_{\ell-1})} + e^{-\mu_\ell + 5(A+1)(t_\ell - t_{\ell-1}) + 1},$$

where $\mathrm{O}^\star$ stands for O with implicit constant smaller than 1.

Proof. We set $s = \sigma + \mathrm{i}\tau$. We have by expanding the logarithm

$$\prod_{p \in (T_{\ell-1}, T_\ell]} \left(1 - \frac{1}{p^s}\right) = e^{-(S_\ell - S_{\ell-1}) - R_\ell},$$

where

$$R_\ell = \sum_{a \geqslant 3} \sum_{p \in (T_{\ell-1}, T_\ell]} \frac{1}{a} \mathrm{Re}\, p^{-as}.$$

We estimate R_ℓ as follows:

$$|R_\ell| \leqslant \sum_{a \geqslant 3} \sum_{p \in (T_{\ell-1}, T_\ell]} \frac{1}{a} p^{-a/2} \leqslant \frac{1}{3} \frac{\sqrt{2}}{\sqrt{2} - 1} \int_{T_{\ell-1}}^{\infty} y^{-3/2} \mathrm{d}y \leqslant 2.5 (T_{\ell-1})^{-1/2}.$$

Using the fact that $e^x \leqslant 1 + 2x$ for $0 < x < 1/2$ and $e^{-x} \geqslant 1 - x$ for $x < 0$, we get that

$$1 - (T_{\ell-1})^{-1/2} \leqslant e^{-R_\ell} \leqslant 1 + 5(T_{\ell-1})^{-1/2}. \tag{B.1}$$

It remains to relate the product over primes to the mollifier. By definition, we have

$$\prod_{p \in (T_{\ell-1}, T_\ell]} \left(1 - \frac{1}{p^s}\right) = \mathcal{M}_\ell + \sum_{\substack{p | n \Rightarrow p \in (T_{\ell-1}, T_\ell] \\ \Omega_{\ell-1}(n) > \mu_\ell}} \frac{\mu(n)}{n^s}.$$

The last term is equal to

$$\sum_{j > \mu_\ell} (-1)^j \left(\sum_{T_{\ell-1} < p_1 < \cdots < p_j \leqslant T_\ell} \frac{1}{(p_1 \cdots p_j)^s} \right). \tag{B.2}$$

The Girard–Newton identities (see, e.g., Equation 2.14' in [15]) show that the inner sum is

$$\sum_{T_{\ell-1}<p_1<\cdots<p_j\leqslant T_\ell} \frac{1}{(p_1 \dots p_j)^s} = (-1)^j \sum_{\substack{m_1,\dots,m_k,\dots\geqslant 0\\ m_1+2m_2+\cdots+km_k+(k+1)m_{k+1}+\cdots=j}} \prod_{k\geqslant 1} \frac{(-\mathcal{P}(ks))^{m_k}}{m_k!\, k^{m_k}},$$

where the sum is over the partitions of j and

$$\mathcal{P}(s) = \sum_{T_{\ell-1}<p\leqslant T_\ell} \frac{1}{p^s}.$$

Using this, we can bound the absolute value of (B.2) by

$$\begin{aligned} &\sum_{m_1,\dots,m_k,\dots\geqslant 0} \exp\left(-\mu_\ell + m_1 + 2m_2 + \cdots + km_k + \cdots\right) \prod_{k\geqslant 1} \frac{|\mathcal{P}(ks)|^{m_k}}{m_k!\, k^{m_k}} \\ &= e^{-\mu_\ell} \prod_{k\geqslant 1}\left(\sum_{m_k\geqslant 0} \frac{(e^k|\mathcal{P}(ks)|/k)^{m_k}}{m_k!}\right) \qquad \text{(B.3)} \\ &= \exp\left(-\mu_\ell + \sum_{k\geqslant 1} e^k|\mathcal{P}(ks)|/k\right). \end{aligned}$$

For $\ell = 1$, the assumption implies $|\mathcal{P}(s)| \leqslant A(t_\ell - t_{\ell-1})$, giving a contribution to the sum of

$$\leqslant eA(t_\ell - t_{\ell-1}). \tag{B.4}$$

For $k = 2$, Mertens's theorem (see [20] for a precise bound) also implies

$$|\mathcal{P}(2s)| \leqslant t_\ell - t_{\ell-1} + \mathrm{O}^\star\left(\tfrac{4}{\log^3 T_{\ell-1}}\right) \leqslant t_\ell - t_{\ell-1} + 0.001,$$

by the definition of T_0. This gives an error of

$$\tfrac{e^2}{2}|\mathcal{P}(2s)| \leqslant 5(t_\ell - t_{\ell-1}) + 0.01. \tag{B.5}$$

For $k \geqslant 3$, we have

$$|\mathcal{P}(ks)| \leqslant \sum_{p\in(T_{\ell-1},T_\ell]} p^{-k/2} \leqslant 2(T_{\ell-1})^{-k/2+1}.$$

This implies that

$$\sum_{k\geqslant 3} e^k|\mathcal{P}(ks)|/k \leqslant 2T_0 \sum_{k\geqslant 3}\left(\frac{e}{T_0^{1/2}}\right)^k \leqslant 0.001. \tag{B.6}$$

The claim follows from (B.1), (B.4), (B.5), and (B.6). □

The mean-value theorem for Dirichlet polynomials, see [16, Corollary 3], takes the following form when expressed in terms of the random model 51. Let $(\theta_p, p \text{ prime})$ a sequence of IID random variables, uniformly distributed on $[0, 2\pi]$. For an integer n with prime factorization $n = p_1^{a_1} \dots p_k^{a_k}$ with $p_1, \dots, p_k$ all distinct, consider

$$\theta_n = \sum_{j=1}^{k} \theta_{p_j}^{a_j}. \tag{B.7}$$

The orthogonality relation $\mathbf{E}[e^{\mathrm{i}\theta_n}\overline{e^{\mathrm{i}\theta_m}}] = \mathbf{1}_{n=m}$ is straightforward from the definition. This implies that for an arbitrary sequence $a(n)$ of complex numbers,

$$\mathbf{E}\left[\Big|\sum_{n\leqslant N} a(n)e^{\mathrm{i}\theta_n}\Big|^2\right] = \sum_{n\leqslant N} |a(n)|^2.$$

The mean-value theorem for Dirichlet polynomials is as follows:

Lemma B.2 [16, Corollary 3]. *Let $(a(n), n \geqslant 1)$ be any sequence of complex numbers. Then, for $N \geqslant 1$ and $T \geqslant 1$, we have with the notation above:*

$$\mathbf{E}\left[\Big|\sum_{n\leqslant N} a(n)n^{\mathrm{i}\tau}\Big|^2\right] = \left(1 + \mathrm{O}\left(\frac{N}{T}\right)\right)\mathbf{E}\left[\Big|\sum_{n\leqslant N} a(n)e^{\mathrm{i}\theta_n}\Big|^2\right].$$

We need several lemmas controlling the distribution of the random model (51). The first one compares it explicitly to the Gaussian random variables $\mathcal{Z}_\ell$.

Lemma B.3. *Consider the increments $\mathcal{Y}_\ell = \mathcal{S}_\ell - \mathcal{S}_{\ell-1}$, $1 \leqslant \ell \leqslant \mathcal{L}$, defined in (52), and the corresponding Gaussian variables $\mathcal{N}_\ell = \mathcal{Z}_\ell - \mathcal{Z}_{\ell-1}$ as in (14). There exists a constant $c > 0$ such that, for any intervals A,*

$$\mathbf{P}\big(\mathcal{Y}_\ell \in A\big) = \mathbf{P}\big(\mathcal{N}_\ell \in A\big) + \mathrm{O}(e^{-ce^{t_{\ell-1}/2}}).$$

Proof. This follows similarly to [3, Lemma 20], based on the Berry–Esseen estimate as stated in [3, Lemma 19]. The proof is actually more immediate because the covariances of $(\mathcal{Y}, \mathcal{Y}')$ and $(\mathcal{N}, \mathcal{N}')$ exactly coincide. $\square$

The error $\mathrm{O}(e^{-ce^{t_{\ell-1}/2}})$ is too large for $\ell = 1$ for our purpose. In this case, we resort to the weaker estimate given in [3, Lemma 18].

Lemma B.4. *Consider the random variables $\mathcal{S}_\ell$ defined in (51). Let $\alpha \geqslant 1$. Then, for all $\Delta \geqslant \alpha$ and $|u| \leqslant 100\alpha\ell$, we have*

$$\mathbb{P}(\mathcal{S}_\ell \in [u, u + \Delta^{-1}]) \asymp \frac{1}{\Delta} \cdot \frac{e^{-u^2/2v_\ell^2}}{\sqrt{v_\ell}}.$$

When an exact Gaussian comparison is not needed, it is sometimes useful to bound the moments of the random model by Gaussian ones. The following statement is a direct consequence of the fact that

$$\mathbf{E}\left[\exp\left(\lambda\sum_{x<p\leqslant y}\frac{\operatorname{Re} a(p)X(p)}{p^{\sigma}}\right)\right]=\prod_{x<p\leqslant y}I_0\left(\frac{|a(p)|\lambda}{p^{\sigma}}\right), \tag{B.8}$$

where

$$I_0(w)=\sum_{k=0}^{\infty}\frac{w^{2k}}{2^{2k}(k!)^2}. \tag{B.9}$$

Lemma B.5. *Let $\sigma\geqslant 1/2$ and $1\leqslant x<y$. For any complex numbers $a(p)$, we have for $k\in\mathbb{N}$,*

$$\mathbf{E}\left[\left(\sum_{x<p\leqslant y}\frac{\operatorname{Re} a(p)e^{\mathrm{i}\theta_p}}{p^{\sigma}}\right)^{2k}\right]\leqslant\frac{(2k!)}{2^k k!}\mathfrak{s}(\sigma)^{2k}, \tag{B.10}$$

where

$$\mathfrak{s}^2(\sigma)=\frac{1}{2}\sum_{x<p\leqslant y}\frac{|a(p)|^2}{p^{2\sigma}}. \tag{B.11}$$

In particular, we have for any $\lambda\in\mathbb{R}$,

$$\mathbf{E}\left[\exp\left(\lambda\sum_{x<p\leqslant y}\frac{\operatorname{Re} a(p)e^{i\theta_p}}{p^{\sigma}}\right)\right]\leqslant\exp(\lambda^2\mathfrak{s}^2(\sigma)/2). \tag{B.12}$$

To bound the moments of actual Dirichlet polynomials, we have the useful [19, Lemma 3]:

Lemma B.6. *Let $2\leqslant x\leqslant T$. For any complex numbers $a(p)$ and $k\in\mathbb{N}$ with $x^k\leqslant T/\log T$, we have for T large enough*

$$\mathbf{E}\left[\left|\sum_{p\leqslant x}\frac{a(p)}{p^{\sigma+\mathrm{i}\tau}}\right|^{2k}\right]\ll k!\left(\sum_{p\leqslant x}\frac{|a(p)|^2}{p^{2\sigma}}\right)^{k}.$$

Using Stirling, we get by picking $k=\lfloor V^2/v_\ell^2\rfloor$ (assuming that V is not too large so the moment condition is satisfied) that

$$\mathbf{P}\big(S_\ell-S_{\ell-1}>V\big)\ll\sqrt{\frac{V^2}{v_\ell^2-v_{\ell-1}^2}}\exp\left(-\frac{V^2}{2(v_\ell^2-v_{\ell-1}^2)}\right).$$

The asymptotics of the moments of the random model with all the prime powers present, as in the definition of $\mathcal{Z}_0$ in (15), can be calculated explicitly.

Lemma B.7. *Let $\delta > 0$ and $\sigma = \frac{1}{2} + \frac{\delta}{\log T}$. Let T_0 such that $\log T_0 = \mathrm{o}(\log T)$. Then, as $T \to \infty$, we have for any $k > 0$*

$$(\log T_0)^{-k^2}\mathbf{E}\left[\left|\prod_{p\leqslant T_0}\left(1-\frac{e^{\mathrm{i}\theta_p}}{p^\sigma}\right)^{-1}\right|^{2k}\right] \sim e^{\gamma k^2}\prod_p (1-p^{-1})^{k^2}\sum_{m\geqslant 0}\frac{|d_k(p^m)|^2}{p^{2m}}, \tag{B.13}$$

where $d_k(n)$ is the k-divisor function defined through the Taylor series

$$(1-x)^{-k} = \sum_{n\geqslant 0} d_k(n)x^n = \sum_{n\geqslant 0}\frac{\Gamma(n+k)}{n!\Gamma(k)}x^n. \tag{B.14}$$

(For an integer k, $d_k(n)$ is the number of ways of writing n as the product of k factors.) The Euler product on the right-hand side of (B.13) is a_k, cf. (4).

Proof. Let $2 \leqslant X < T_0$. We will take the limit $T \to \infty$ then $X \to \infty$. We have by expanding the logarithm and the exponential

$$\begin{aligned}
\mathbf{E}\left[\left|\prod_{X<p\leqslant T_0}\left(1-\frac{e^{\mathrm{i}\theta_p}}{p^\sigma}\right)^{-1}\right|^{2k}\right] &= \prod_{X<p\leqslant T_0}\mathbf{E}\left[\exp\left(-2k\mathrm{Re}\log(1-e^{\mathrm{i}\theta_p}p^{-\sigma})\right)\right]\\
&= \prod_{X<p\leqslant T_0}\mathbf{E}\left[\exp\left(2k\left(\frac{\cos\theta_p}{p^\sigma}+\frac{\cos 2\theta_p}{2p^{2\sigma}}+\mathrm{O}(p^{-3\sigma})\right)\right)\right]\\
&= \prod_{X<p\leqslant T_0}\left(1+\frac{k^2}{p^{2\sigma}}+\mathrm{O}(p^{-3\sigma})\right)\\
&= \exp\left(\sum_{X<p\leqslant T_0}\frac{k^2}{p^{2\sigma}}+\mathrm{O}(X^{-1/2})\right).
\end{aligned}$$

We can evaluate the sum over p as in (17) using the prime number theorem. This gives as $T \to \infty$ and $X \to \infty$

$$\mathbf{E}\left[\left|\prod_{X<p\leqslant T_0}\left(1-\frac{e^{\mathrm{i}\theta_p}}{p^\sigma}\right)^{-1}\right|^{2k}\right] \sim (\log T_0)^{k^2}(\log X)^{-k^2},$$

since $\delta\frac{\log T_0}{\log T} \to 0$ by assumption. It remains to evaluate $(\log X)^{-k^2}\mathbf{E}[|\prod_{p\leqslant X}(1-e^{\mathrm{i}\theta_p}p^{-\sigma})^{-1}|^{2k}]$ in the limit $T \to \infty$ and $X \to \infty$. Taking $T \to \infty$ simplifies $\sigma \to 1/2$. Hence, writing the random Euler product as a random Dirichlet series yields

$$\mathbf{E}\left[\left|\prod_{p\leqslant X}\left(1-\frac{e^{\mathrm{i}\theta_p}}{p^\sigma}\right)^{-k}\right|^{2}\right] \sim \mathbf{E}\left[\left|\sum_{\substack{n\geqslant 1\\ p|n\Rightarrow p\leqslant X}}\frac{d_k(n)}{n^{1/2}}e^{\mathrm{i}\theta_n}\right|^2\right],$$

where θ_n is defined by $\theta_n = \theta_{p_1}^{a_1} \dots \theta_{p_r}^{a_r}$ if the prime factorization of n is $p_1^{a_1} \dots p_r^{a_r}$. The right-hand side is easily evaluated by the orthogonality of the $e^{\mathrm{i}\theta_n}$ and equals

$$\sum_{\substack{n \geqslant 1 \\ p|n \Rightarrow p \leqslant X}} \frac{|d_k(n)|^2}{n} = \prod_{p \leqslant X} \sum_{m \geqslant 0} \frac{|d_k(p^m)|^2}{p^m}.$$

Mertens's third theorem, see, for example, [14], states that $\log X \prod_{p \leqslant X}(1-p^{-1}) \sim e^{-\gamma}$ This implies that as $X \to \infty$

$$(\log X)^{-k^2} \mathbf{E}\left[\left|\prod_{p \leqslant X}\left(1 - \frac{e^{\mathrm{i}\theta_p}}{p^{1/2}}\right)^{-1}\right|^{2k}\right] \sim e^{\gamma k^2} \prod_p (1-p^{-1})^{k^2} \sum_{m \geqslant 0} \frac{|d_k(p^m)|^2}{p^m},$$

as claimed. □

The lemma implies some useful estimates. First, Markov's inequality yields

$$\mathbf{P}(\mathcal{Z}_0 > \alpha t_0) \leqslant \mathbf{E}[e^{2\alpha \mathcal{Z}_0}](\log T_0)^{-2\alpha^2} \ll e^{\gamma\alpha^2} a_\alpha (\log T_0)^{-\alpha^2}. \tag{B.15}$$

Second, the heuristic given in (18) remains almost unchanged on $\mathcal{G}_0 = \{\mathcal{Z}_0 - \alpha t_0 \in [0, \sqrt{t_0}\}$. Indeed, let us consider the change of probability defined by $\frac{\mathrm{d}\widetilde{\mathbf{P}}}{\mathrm{d}\mathbf{P}} = \frac{e^{2\alpha\mathcal{Z}_0}}{\mathbf{E}[e^{2\alpha\mathcal{Z}_0}]}$. We have similar to (18):

$$\begin{aligned}
\mathbf{P}(\{\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > \alpha t\} \cap \mathcal{G}_0) &\asymp \frac{e^{-\alpha^2 t}}{\alpha\sqrt{t}} e^{-\alpha^2 \mathfrak{s} \log_\mathcal{L} t - 2\delta\alpha^2} \cdot e^{-\alpha^2 t_0} \mathbf{E}[e^{2\alpha\mathcal{Z}_0}] \cdot \widetilde{\mathbf{P}}(\mathcal{Z}_0 - \alpha t_0 \in [0, \sqrt{t_0}]) \\
&\asymp \frac{e^{-\alpha^2 t}}{\alpha\sqrt{t}} a_\alpha e^{-\alpha^2 \mathfrak{s} \log_\mathcal{L} t - 2\delta\alpha^2 + \gamma\alpha^2} \cdot \widetilde{\mathbf{P}}(\mathcal{Z}_0 - \alpha t_0 \in [0, \sqrt{t_0}]),
\end{aligned}$$

by Lemma B.7. Under $\widetilde{\mathbf{P}}$, the random variables $e^{\mathrm{i}\theta_p}$ remain independent and $\widetilde{\mathbf{E}}[\mathcal{Z}_0] = \alpha t_0 + \mathrm{o}(1)$. Therefore, by the Berry–Esseen theorem, we have that $\widetilde{\mathbf{P}}(\mathcal{Z}_0 - \alpha t_0 \in [0, \sqrt{t_0}]) \asymp 1$. We conclude that

$$\mathbf{P}(\{\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > \alpha t\} \cap \mathcal{G}_0) \asymp \mathbf{P}(\mathcal{Z}_0 + \mathcal{Z}_\mathcal{L} > \alpha t). \tag{B.16}$$

ACKNOWLEDGEMENTS

The authors are grateful to Winston Heap for his comments on the first version of this paper. L.-P. A. thanks Paul Bourgade and Maksym Radziwiłł for numerous insightful discussions on the subject, especially on the proof of Lemma A.1. L.-P. A. is supported in part by the NSF award DMS 2153803.

JOURNAL INFORMATION

Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society's research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

ORCID

Emma Bailey https://orcid.org/0000-0003-3558-1602

REFERENCES

1. L.-P. Arguin and E. Bailey, *Large deviation estimates of Selberg's central limit theorem and applications*, Int. Math. Res. Not. **2023** (2023), no. 23, 20574–20612.
2. L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan, *Maximum of the Riemann zeta function on a short interval of the critical line*, Comm. Pure Appl. Math. **72** (2019), no. 3, 500–535.
3. L.-P. Arguin, P. Bourgade, and M. Radziwiłł, *The Fyodorov-Hiary-Keating conjecture. I*, Preprint arXiv:2007.00988, 2020.
4. L.-P. Arguin, P. Bourgade, and M. Radziwiłł, *The Fyodorov-Hiary-Keating conjecture. II*, Preprint arXiv:2307.00982, 2023.
5. L.-P. Arguin, F. Ouimet, and M. Radziwiłł, *Moments of the Riemann zeta function on short intervals of the critical line*, Ann. Probab. **49** (2021), no. 6, 3106–3141.
6. J. B. Conrey and S. M. Gonek, *High moments of the Riemann zeta-function*, Duke Math. J. **107** (2001), no. 3, 577–604.
7. A. Dobner, *Large deviations of the argument of the Riemann zeta function*, Preprint arXiv:2101.01747, to appear in Mathematika, 2024.
8. R. M. Gabriel, *Some results concerning the integrals of moduli of regular functions along certain curves*, J. London Math. Soc. **2** (1927), no. 2, 112–117.
9. W. Heap, M. Radziwiłł, and K. Soundararajan, *Sharp upper bounds for fractional moments of the Riemann zeta function*, Q. J. Math. **70** (2019), no. 4, 1387–1396.
10. W. Heap and K. Soundararajan, *Lower bounds for moments of zeta and L-functions revisited*, Mathematika **68** (2022), no. 1, 1–14.
11. A. E. Ingham, *A note on Fourier transforms*, J. London Math. Soc. **s1–9** (1934), no. 1, 29–32.
12. A. Ivic, *Riemann zeta-function*, A Wiley-Interscience Publication. John Wiley & Sons, New York, 1985. (reissue, Dover, Mineola, New York, 2003).
13. J. P. Keating and N. C. Snaith, *Random matrix theory and* $\zeta(1/2+it)$, Comm. Math. Phys. **214** (2000), no. 1, 57–89.
14. D. Koukoulopoulos, *The distribution of prime numbers*, Graduate Studies in Mathematics, vol. 203, American Mathematical Society, Providence, RI, [2019] ©2019.
15. I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.
16. H. L. Montgomery and R. C. Vaughan, *Hilbert's inequality*, J. London Math. Soc. (2) **8** (1974), 73–82.
17. M. Radziwiłł and K. Soundararajan, *Continuous lower bounds for moments of zeta and L-functions*, Mathematika **59** (2013), no. 1, 119–128.
18. M. Radziwiłł and K. Soundararajan, *Selberg's central limit theorem for* $\log|\zeta(1/2+it)|$, Enseign. Math. **63** (2017), no. 1-2, 1–19.
19. K. Soundararajan, *Moments of the Riemann zeta function*, Ann. of Math. (2) **170** (2009), no. 2, 981–993.
20. R. Vanlalngaia, *Explicit Mertens sums*, Integers **17** (2017), A11.