The Fyodorov-Hiary-Keating Conjecture. I.

Louis-Pierre Arguin, Paul Bourgade, and Maksym Radziwiłł

Abstract. By analogy with conjectures for random matrices, Fyodorov-Hiary-Keating and Fyodorov-Keating proposed precise asymptotics for the maximum of the Riemann zeta function in a typical short interval on the critical line. In this paper, we settle the upper bound part of their conjecture in a strong form. More precisely, we show that the measure of those $T \leq t \leq 2T$ for which

$$\max_{|h|\leq 1} |\zeta(\tfrac{1}{2} + \mathrm{i}t + \mathrm{i}h)| > e^y \frac{\log T}{(\log\log T)^{3/4}}$$

is bounded by $Cye^{-2y}T$ uniformly in $y \geq 1$ with $C > 0$ an absolute constant. This is expected to be optimal for $y = \mathrm{O}(\sqrt{\log\log T})$. This upper bound is sharper than what is known in the context of random matrices, since it gives (uniform) decay rates in y. In a subsequent paper we will obtain matching lower bounds.

Contents

1. Introduction

Motivated by the problem of understanding the global maximum of the Riemann zeta function on the critical line, Fyodorov-Keating [15] and Fyodorov-Hiary-Keating [14] raised the question of understanding the distribution of the local maxima of the Riemann zeta function on the critical line. They made the following conjecture.

Conjecture 1 (Fyodorov-Hiary-Keating)**.** *There exists a cumulative distribution function F such that, for any y, as $T \to \infty$,*

$$\frac{1}{T}\operatorname{meas}\Big\{T \le t \le 2T : \max_{0\le h\le 1} |\zeta(\tfrac12 + \mathrm{i}t + \mathrm{i}h)| \le e^{y}\frac{\log T}{(\log\log T)^{3/4}}\Big\} \sim F(y).$$

Moreover, as $y \to \infty$, the right-tail decay is $1 - F(y) \sim Cye^{-2y}$ for some $C > 0$.

The striking aspect of this conjecture is the exponent $\frac34$ on the $\log\log T$ and the decay rate $1 - F(y) \ll ye^{-2y}$. This suggests that around the local maximum there is a significant degree of interaction between nearby shifts of the Riemann zeta function (on the scale $1/\log T$). If there were no interactions, one would expect an exponent of $\frac14$ on the $\log\log T$ and a decay rate e^{-2y} (see [19]).

This paper settles the upper bound part of the Fyodorov-Hiary-Keating conjecture in a strong from, with uniform and sharp decay in y.

Theorem 1. *There exists $C > 0$ such that for any $T \ge 3$ and $y \ge 1$, we have*

$$\frac{1}{T}\operatorname{meas}\Big\{T \le t \le 2T : \max_{|h|\le 1} |\zeta(\tfrac12 + \mathrm{i}t + \mathrm{i}h)| > e^{y}\frac{\log T}{(\log\log T)^{3/4}}\Big\} \le Cye^{-2y}.$$

Theorem 1 is expected to be sharp in the range $y = \mathrm{O}(\sqrt{\log\log T})$. For larger y in the range $y \in [1, \log\log T]$, it is expected that the sharp decay rate is

$$\ll ye^{-2y}\exp\Big(-\frac{y^2}{\log\log T}\Big).$$

Conjecture 1 emerges in [14, 15] from the analogous prediction for random matrices, according to which

$$\sup_{|z|=1}\log|X_n(z)| = \log n - \frac34\log\log n + M_n, \tag{1}$$

with $X_n(z)$ the characteristic polynomial of a Haar-distributed $n\times n$ unitary matrix, and with M_n converging to a random variable M in distribution. Progress on (1) was accomplished by Arguin-Belius-Bourgade [2] and Paquette-Zeitouni [28], culminating in the work of Chhaibi-Madaule-Najnudel [12]. In [12] it was established for the circular beta ensemble that the sequence of random variables M_n is tight. The convergence of M_n in distribution to a limiting random variable M and the decay rate of $\mathbb{P}(M > y)$ as y increases remain open. In this regard, Theorem 1 is a rare instance of a result obtained for the Riemann zeta function prior to the analogue for random matrices. This type of decay is expected by analogy with branching random walks, but has only been proved for a few processes, notably for the two-dimensional Gaussian free field [13, 10].

Previous results in the direction of Conjecture 1 were more limited than for unitary matrices. The first order, that is,

$$\max_{|h|\leq 1}\log|\zeta(\tfrac12+\mathrm{i}t+\mathrm{i}h)|\sim\log\log T\ ,\ T\to\infty,$$

for all $t\in[T,2T]$ outside of an exceptional set of measure $\mathrm{o}(T)$, was established conditionally on the Riemann Hypothesis by Najnudel [27], and unconditionally by the authors with Belius and Soundararajan [3]. Harper [18] subsequently obtained the upper bound up to second order. More precisely, Harper showed that for $t\in[T,2T]$ outside of an exceptional subset of measure $\mathrm{o}(T)$, and for any $g(T)\to\infty$,

$$\max_{|h|\leq 1}\log|\zeta(\tfrac12+\mathrm{i}t+\mathrm{i}h)|\leq\log\log T-\frac34\log\log\log T+\frac32\log\log\log\log T+g(T). \quad (2)$$

Progress towards Conjecture 1 has been made by observing that the large values of $\log|\zeta(\frac12+\mathrm{i}t+\mathrm{i}h)|$ on a short interval indexed by $h\in[-1,1]$ are akin to the ones of an approximate branching random walk, see for example [1]. This is because, the average of $\log|\zeta(\frac12+\mathrm{i}t+\mathrm{i}h)|$ over a neighborhood of h of width e^{-k} for $k\leq\log\log T$ can be thought of as a Dirichlet sum S_k of $p^{-1/2+\mathrm{i}t+\mathrm{i}h}$ up to $p\leq\exp e^k$, see Equation (4) below. The partial sums S_k, $k\leq\log\log T$, for different h's have a correlation structure that is approximately the one of a branching random walk.

For branching random walks, the identification of the maximum up to an error of order one relies on a precise upper barrier for the values of the random walks S_k at every $k\leq\log\log T$, as introduced in the seminal work of Bramson [9]. This approach cannot work directly for $\log|\zeta|$ as one needs to control large deviations for Dirichlet polynomials involving prime numbers close to T. This amounts to computing large moments of long Dirichlet sums, and current number theory techniques cannot access these with a small error.

To circumvent this problem, the proof of Theorem 1 is based on an iteration scheme that recursively constructs *upper* and *lower barrier* constraints for the values of the partial sums as the scales k approaches $\log\log T$. Each step of the iteration relies on elaborate second and twisted fourth moments of the Riemann zeta function, which may be of independent interest. The lower barrier reduces in effect the number of h's to be considered for the maximum of $\log|\zeta|$. One upshot is that smaller values for the Dirichlet sums are needed, and thus only moments with good errors are necessary. Furthermore, the reduction of the number of h's improves the approximation of $\log|\zeta|$ in terms of Dirichlet sums for the subsequent scales in the iteration. Lower constraints have appeared before in [4] to study correlations between extrema of the branching Brownian motion. There, they were proved *a posteriori* based on the work of Bramson on the maximum.

The paper is organized as follows. The iterative scheme is described in details in Section 3. Its initial condition, induction and final step are proved in Sections 4, 5 and 6. The number-theoretic input of the recursion using second and twisted fourth moments of the Riemann zeta function is the subject of Sections 7 and 8.

In a subsequent paper we will complement the upper bound in Theorem 1 with matching lower bounds, for fixed $y>1$. This will also rely on the multiscale analysis and on twisted moments.

Notations. We use Vinogradov's notation and write $f(T) \ll g(T)$ to mean $f(T) = \mathrm{O}(g(T))$ as $T \to \infty$. If the O-term depends on some parameter A, we write $\ll_A$ or O_A to emphasize the dependence. We write $f(T) \asymp g(T)$ when $f(T) \ll g(T)$ and $g(T) \ll f(T)$.

Acknowledgments. The authors are grateful to Frederic Ouimet for several discussions, and to Erez Lapid and Ofer Zeitouni for their careful reading, pointing at a mistake in the initial proof of Lemma 23. The research of LPA was supported in part by NSF CAREER DMS-1653602. PB acknowledges the support of NSF grant DMS-1812114 and a Poincaré chair. MR acknowledges the support of NSF grant DMS-1902063 and a Sloan Fellowship.

2. Initial Reductions

Throughout the paper we will adopt probabilistic notations and conventions. In particular τ will denote a random variable uniformly distributed in $[T, 2T]$ and $\mathbb{P}, \mathbb{E}$ the associated probability and expectation. Furthermore we set throughout

$$n = \log\log T.$$

This notation will become natural later when S_k, $k \leq n$, given in Equation (4) will be thought of as a random walk. We will find it convenient to consider $\zeta(\frac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)$ as a random variable and write for short $\zeta_\tau(h) = \zeta(\frac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)$. In this notation, Theorem 1 can be restated as follows.

Theorem. *Let τ be a uniformly distributed random variable in $[T, 2T]$. Then uniformly in $T \geq 3$, $y \geq 1$, one has*

$$\mathbb{P}\Big(\max_{|h|\leq 1} |\zeta_\tau(h)| > e^y \frac{e^n}{n^{3/4}}\Big) \ll y e^{-2y} \ .$$

Along the proof, we will refer to well-known results, or variations of well-known results. To emphasize the core ideas of the proof, we chose to gather these in the appendix. Appendix A deals with estimates on sums of primes and on moments of Dirichlet polynomials. Appendix B presents a version of the ballot theorem for random walks. Finally, tools for discretizing the maximum of Dirichlet polynomial on a short interval are presented in Appendix C. With this in mind, we first observe that it is easy to establish Theorem 1 for $y > n$.

Lemma 1. *Uniformly in $y > n$ we have*

$$\mathbb{P}\Big(\max_{|h|\leq 1} |\zeta_\tau(h)| > e^y \frac{e^n}{n^{3/4}}\Big) \ll y e^{-2y}. \tag{3}$$

Proof. By Chebyshev's inequality, the probability in (3) is

$$\leq e^{-4y} e^{-4n} n^3 \mathbb{E}\Big[\max_{|h|\leq 1} |\zeta_\tau(h)|^4\Big].$$

By Lemma 28 in Appendix C, the above is

$$\ll e^{-4y} e^{-4n} n^3 e^{5n} = n^3 e^n e^{-4y}.$$

Since $y > n$, this is $\ll y e^{-2y}$ and the claim follows. $\square$

To handle the remaining values $1 \leq y \leq n$ it will be convenient to discretize the maximum over $|h| \leq 1$ into a maximum over a set

$$\mathcal{T}_n = e^{-n-100}\mathbb{Z} \cap [-2, 2].$$

To accomplish this, we use the following simple lemma.

Lemma 2. *There exists an absolute constant $C > 1$ such that for any $V > 1$ and $A > 100$,*

$$\mathbb{P}\Big(\max_{|h|\leq 1} |\zeta_\tau(h)| > V\Big) \leq \mathbb{P}\Big(\max_{h\in\mathcal{T}_n} |\zeta_\tau(h)| > V/C\Big) + \mathrm{O}_A(e^{-An}).$$

Proof. This is Lemma 26 in Appendix C. □

Combining the above lemma with Lemma 1, it suffices to prove the following result to establish Theorem 1. Without loss of generality, we state the result for $T \geq \exp(e^{1000})$ and $y > 4000$, which is more convenient for further estimates.

Theorem 2. *Let τ be a random variable, uniformly distributed in $[T, 2T]$. Then, uniformly in $T \geq \exp(e^{1000})$, $4000 \leq y \leq n$, we have*

$$\mathbb{P}\Big(\max_{h\in\mathcal{T}_n} |\zeta_\tau(h)| > e^y \frac{\log T}{(\log\log T)^{3/4}}\Big) \ll y e^{-2y}.$$

3. Iteration Scheme

3.1. **Notations.** In this section, we explain the structure of the proof of Theorem 2. We start by defining the main objects of study. Consider first the time scales

$$T_{-1} = \exp(e^{1000}), \qquad T_0 = \exp(\sqrt{\log T}), \qquad T_\ell = \exp\Big(\frac{\log T}{(\log_{\ell+1} T)^{10^6}}\Big),$$

where $\ell \geq 1$ and $\log_\ell$ stands for the logarithm iterated ℓ times. We adopt the convention that $\log_0 n = n$ and $\log_{-1} n = e^n$. It is convenient to write the above in the log log-scale, denoting (remember $n = \log\log T$)

$$n_{-1} = 1000, \qquad n_0 = \tfrac{n}{2}, \qquad n_\ell = \log\log T_\ell = n - 10^6 \log_\ell n.$$

Consider the Dirichlet polynomial

$$S_k(\tfrac12 + \mathrm{i}\tau + \mathrm{i}h) := S_k(h) = \sum_{e^{1000}\leq \log p \leq e^k} \mathrm{Re}\Big(p^{-(\frac12+\mathrm{i}\tau+\mathrm{i}h)} + \tfrac12 p^{-2(\frac12+\mathrm{i}\tau+\mathrm{i}h)}\Big), \quad k \leq n, \tag{4}$$

with $S_{n_{-1}}(h) = 0$. The above summand consists in the first two terms in the expansion of $-\log|1-p^{-s}|$. The second order may be essentially ignored on a first reading; however this additional term is necessary to handle the maximum of $|\zeta|$ up to tightness, due to the contribution of the small primes to $|\zeta(s)|$. Moreover, starting the sum in (4) at e^{1000} will be convenient for some estimates in Section 8. We also define,

$$\widetilde{S}_k(\tfrac12 + \mathrm{i}\tau + \mathrm{i}h) := \widetilde{S}_k(h) = \sum_{e^{1000}\leq \log p \leq e^k} \Big(p^{-(\frac12+\mathrm{i}\tau+\mathrm{i}h)} + \tfrac12 p^{-2(\frac12+\mathrm{i}\tau+\mathrm{i}h)}\Big), \quad k \leq n, \tag{5}$$

so that $S_k(h) = \mathrm{Re}\, \widetilde{S}_k(h)$ and $|S_k(h)| \leq |\widetilde{S}_k(h)|$.

We use the probabilistic notation of omitting the dependence on the random τ, and think of $(S_k(h))_{h\in[-2,2]}$ as a stochastic process. The dependence in h will sometimes be omitted when there is no ambiguity.

It will be necessary to control the difference $\log|\zeta| - S_k$ which represents the contribution of primes larger than e^{e^k}. To do so, given $\ell \geq 0$, we define the following random mollifiers,

$$\mathcal{M}_\ell(h) = \sum_{\substack{p|m \Rightarrow p\in(T_{\ell-1},T_\ell] \\ \Omega_\ell(m)\leq(n_\ell-n_{\ell-1})^{10^5}}} \frac{\mu(m)}{m^{\frac{1}{2}+\mathrm{i}\tau+\mathrm{i}h}},$$

where $\Omega_\ell(m)$ stands for the number of prime factors of m in the interval $(T_{\ell-1}, T_\ell]$, counted with multiplicity, and μ denotes the Möbius function[1]. Furthermore we set $\mathcal{M}_{-1}(h) = 1$ for all $h \in \mathbb{R}$. Given $\ell \geq 0$ and $k \in [n_{\ell-1}, n_\ell]$, we define the mollifier up to k as

$$\mathcal{M}_{\ell-1}^{(k)}(h) = \sum_{\substack{p|m \Rightarrow p\in(T_{\ell-1},\exp(e^k)] \\ \Omega_\ell(m)\leq(n_\ell-n_{\ell-1})^{10^5}}} \frac{\mu(m)}{m^{\frac{1}{2}+\mathrm{i}\tau+\mathrm{i}h}}.$$

This way we have $\mathcal{M}_{\ell-1}^{(n_{\ell-1})} = 1$ and $\mathcal{M}_{\ell-1}^{(n_\ell)} = \mathcal{M}_\ell$. The product $\mathcal{M}_{-1}\dots\mathcal{M}_{\ell-1}\mathcal{M}_{\ell-1}^{(k)}$ will be a good proxy for $\exp(-S_k)$ for most τ, cf. Lemma 23 in Appendix A.

Finally, the deterministic centering of the maximum is denoted

$$m(k) = k\left(1 - \frac{3}{4}\frac{\log n}{n}\right).$$

For a fixed $y \geq 1$, we set the following upper and lower barriers for the values of S_k:

$$U_y(k) = y + \begin{cases} \infty & \text{for } 1 \leq k < \lceil y/4 \rceil, \\ 10^3 \log k & \text{for } \lceil y/4 \rceil \leq k \leq n/2, \\ 10^3 \log(n-k) & \text{for } n/2 < k < n, \end{cases} \tag{6}$$

$$L_y(k) = y - \begin{cases} \infty & \text{for } 1 \leq k < \lceil y/4 \rceil, \\ 20k & \text{for } \lceil y/4 \rceil \leq k \leq n/2, \\ 20(n-k) & \text{for } n/2 < k < n. \end{cases} \tag{7}$$

Note that $U_y(k) - L_y(k)$ is independent of y and that $L_y(y/4) = -4y$ is negative.

[1]We could have also counted the prime factors of m without multiplicity because m has to be square-free, but $\Omega_\ell(m)$ will be more consistent with other constraints appearing along the proof.

3.2. **Iterated good sets.** The proof of Theorem 1 progressively reduces the set of h's for which ζ is large. We define iteratively the following decreasing subsets for $\ell \geq 0$:

$$\begin{aligned} A_\ell &= A_{\ell-1} \cap \{h \in \mathcal{T}_n : |\widetilde{S}_k(h) - \widetilde{S}_{n_{\ell-1}}(h)| \leq 10^3(n_\ell - n_{\ell-1}) \text{ for all } k \in (n_{\ell-1}, n_\ell]\} \\ B_\ell &= B_{\ell-1} \cap \{h \in \mathcal{T}_n : S_k(h) \leq m(k) + U_y(k) \text{ for all } k \in (n_{\ell-1}, n_\ell]\} \\ C_\ell &= C_{\ell-1} \cap \{h \in \mathcal{T}_n : S_k(h) > m(k) + L_y(k) \text{ for all } k \in (n_{\ell-1}, n_\ell]\} \\ D_\ell &= D_{\ell-1} \cap \{h \in \mathcal{T}_n : |(\zeta_\tau e^{-S_k})(h)| \leq c_\ell |(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_{\ell-1} \mathcal{M}_{\ell-1}^{(k)})(h)| + e^{-10^4(n-n_{\ell-1})} \\ &\qquad \text{for all } k \in (n_{\ell-1}, n_\ell]\}, \end{aligned}$$

where $c_\ell := \prod_{i=0}^{\ell}(1 + e^{-n_{i-1}})$, and where we set $A_{-1} = B_{-1} = C_{-1} = D_{-1} = [-2, 2]$. Define the "good" sets

$$G_\ell = A_\ell \cap B_\ell \cap C_\ell \cap D_\ell, \quad \ell \geq -1,$$

and the set of interest in Theorem 2

$$H(y) = \Big\{h \in \mathcal{T}_n : |\zeta_\tau(h)| > e^y \frac{e^n}{n^{3/4}}\Big\},$$

where $\zeta_\tau(h)$ stands for $\zeta(\frac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)$ as before. We will call the points $h \in \mathcal{T}_n$ belonging to $H(y)$ the "high points". The subsets A_ℓ and D_ℓ will be needed as auxiliary steps towards the proof that high points are in C_ℓ, and C_ℓ will be needed for the proof of B_ℓ.

3.3. **Induction steps.** Theorem 2 follows from three propositions. The first one proves that most high points are in the good set G_0. This control for small primes up to n_0 is simple, because the barrier U_y is quite high and the $p^{\mathrm{i}\tau}$'s show strong decoupling (i.e "quasi-random" behavior) for primes small enough with respect to T.

Proposition 1. *There exists $K > 0$ such that for any $4000 \leq y \leq n$, one has*

$$\mathbb{P}(\exists h \in H(y) \cap G_0^{\mathrm{c}}) \leq K e^{-2y}.$$

Second, the proposition below gives a precise control of the large values of $(S_k(h))_{h\in[-2,2]}$ for all k up to n_ℓ. This proposition is the most involved part of the proof.

Proposition 2. *There exists $K > 0$ such that for any $4000 \leq y \leq n$, and $\ell \geq 0$ such that $\exp(10^6(n - n_\ell)^{10^5} e^{n_{\ell+1}}) \leq \exp(\frac{1}{100}e^n)$, one has*

$$\mathbb{P}\Big(\exists h \in H(y) \cap G_\ell\Big) \leq \frac{Kye^{-2y}}{\log_{\ell+1} n} + \mathbb{P}\Big(\exists h \in H(y) \cap G_{\ell+1}\Big).$$

Finally, one has the following estimate for the remaining points of the set.

Proposition 3. *There exists $K > 0$ such that for any $4000 \leq y \leq n$, and $\ell \geq 0$ such that $\exp(10^6(n - n_\ell)^{10^5} e^{n_{\ell+1}}) \leq \exp(\frac{1}{100}e^n)$, one has*

$$\mathbb{P}\Big(\exists h \in H(y) \cap G_\ell\Big) \leq Kye^{-2y} e^{10^3(n-n_\ell)}.$$

Theorem 2 can be proved assuming Propositions 1, 2 and 3 .

Proof of Theorem 2. Let L be the largest index ℓ such that

$$\exp(10^6(n-n_\ell)^{10^5}e^{n_{\ell+1}}) \leq \exp\left(\frac{1}{100}e^n\right),$$

so that in particular $n-n_L = \mathrm{O}(1)$. We clearly have

$$\mathbb{P}(\exists h \in H(y)) \leq \mathbb{P}(\exists h \in H(y)\cap G_0^{\mathrm{c}}) + \mathbb{P}(\exists h \in H(y)\cap G_0).$$

By Proposition 1 and iterating Proposition 2 up to L, the above is

$$\leq Ke^{-2y} + \sum_{1\leq \ell\leq L}\frac{Kye^{-2y}}{\log_\ell n} + \mathbb{P}(\exists h\in H(y)\cap G_L) \ll ye^{-2y} + \mathbb{P}(\exists h \in H(y)\cap G_L),$$

since the sum over ℓ is rapidly convergent. Finally, Proposition 3 implies

$$\mathbb{P}(\exists h \in H(y)\cap G_L) \leq Ke^{(n-n_L)^{10^3}}ye^{-2y} \ll ye^{-2y},$$

since $n-n_L = \mathrm{O}(1)$. All the above steps together yield $\mathbb{P}(\exists h \in H(y)) \ll ye^{-2y}$, as expected. □

We note that to obtain $\mathbb{P}(\max_{|h|\leq 1}|\zeta(\frac{1}{2}+\mathrm{i}\tau+\mathrm{i}h)| > e^y(\log T)/(\log\log T)^{3/4}) = \mathrm{o}(1)$ for large y, the number of steps in the induction can be lower than L. For example if y is of order $\log_2 n$ as in (2), iterating up to $\ell = 3$ suffices. Further iterations improves the error by extra logarithms.

4. Initial Step

This section proves Proposition 1. Notice that by a union bound

$$\mathbb{P}(\exists h\in H(y)\cap G_0^{\mathrm{c}}) \leq \mathbb{P}(\exists h \in A_0^{\mathrm{c}}) + \mathbb{P}(\exists h \in D_0^{\mathrm{c}}\cap A_0) + \mathbb{P}(\exists h\in C_0^{\mathrm{c}}) + \mathbb{P}(\exists h \in B_0^{\mathrm{c}}).$$

The first two probabilities on the right-hand side will be bounded by $\ll e^{-7n}$, and the last two by $\ll e^{-2y}$. This will imply the claim.

For the first probability, a union bound on h and $k\leq n_0$ together with the Gaussian tail (83) yield

$$\mathbb{P}(\exists h\in A_0^{\mathrm{c}}) \ll e^n\, n_0 \exp(-10^2 n) \ll e^{-7n}.$$

We now show that $\mathbb{P}(\exists h \in B_0^{\mathrm{c}}) \ll e^{-2y}$. A union bound on $y/4 < k \leq n_0$ implies that for any sequence of integers $q_k \geq 1$,

$$\begin{aligned}\mathbb{P}(\exists h\in B_0^{\mathrm{c}}) &\leq \sum_{y/4<k\leq n_0}\mathbb{P}\Big(\max_{|h|\leq 2} S_k(h) > U_y(k) + k - \frac{3}{4}\log k\Big)\\ &\leq \sum_{y/4<k\leq n_0}\mathbb{E}\Big[\max_{|h|\leq 2}\frac{|S_k(h)|^{2q_k}}{(y+k+10\log k)^{2q_k}}\Big],\end{aligned}\tag{8}$$

where we use the fact that $m(k)\geq k-\frac{3}{4}\log k$ for $k>e$. We discretize the maximum over q_ke^k points using Lemma 27 in Appendix C with $N=\exp(2q_ke^k)$ and $A=1000$. We can also apply (80) on each of these terms, taking $q_k = \lceil (y+k+10\log k)^2/(k+C)\rceil$ with $C>0$ an absolute constant. It is easily checked that the condition $2q_k \leq e^{n-k}$ is fulfilled here to get a Gaussian tail, as $y\leq n$ and $k\leq n_0$.

Note that the second sum on the right-hand side of (104) is negligible. To see this, all terms up to $\frac{2\pi j}{8e^k} = T/2$ yield the same moment, as the average over τ could be replaced by an average over $[T/2, 2T]$ which yields the same bounds. The prefactor $1/(1+j^{1000})$ then makes the contribution negligible. For larger j's, that is $j > \frac{2}{\pi}\frac{T}{\sqrt{\log T}}$, we use the deterministic bound $|S_k|^{2q_k} \leq \exp(q_k \cdot (\log T)^{1/2})$, so that the corresponding sum is at most $\sum_{|j|>\frac{T}{\sqrt{\log T}}} |j|^{-1000} \exp(q_k \cdot (\log T)^{1/2}) \ll T^{-10} e^{(y+n)^2\sqrt{\log T}} \ll e^{-2y}$ for $y < n$.

Putting this together yields

$$\mathbb{P}(\exists h \in B_0^{\rm c}) \ll \sum_{y/4<k\leq n_0} e^k \frac{(k+y)^3}{k^{3/2}} \exp\big(-(k+10\log k+y)^2/(k+C)\big)$$
$$\ll e^{-2y} \sum_{y/4<k\leq n_0} (k^{3/2}+y^3k^{-3/2})\, k^{-20} \ll e^{-2y}.$$

To bound the probability $\mathbb{P}(\exists h \in C_0^{\rm c})$ we note that if there exists h in $C_0^{\rm c}$ then $S_k(h) \leq y - 20k$ for some $h \in \mathcal{T}_n$ and some $y/4 < k \leq n_0$. Therefore we obtain the bound,

$$\mathbb{P}(\exists h \in C_0^{\rm c}) \leq \sum_{y/4<k\leq n_0} \mathbb{P}\Big(\max_{|h|\leq 2} |S_k(h)| > 20k - y\Big)$$
$$\leq \sum_{y/4\leq k\leq n_0} \mathbb{E}\Big[\max_{|h|\leq 2} \frac{|S_k(h)|^{2q_k}}{(20k-y)^{2q_k}}\Big] \tag{9}$$

for any choice of $q_k \geq 1$. We choose $q_k = \lceil (20k-y)^2/k \rceil$. The length of $S_k(h)^{q_k}$ is $\exp(2q_k e^k)$. We discretize the maximum over $q_k e^k$ points using Lemma 27 in Appendix C with $N = \exp(2q_k e^k)$ and $A = 1000$. This shows that (9) is

$$\ll \sum_{y/4\leq k\leq n_0} q_k\, e^k\, \mathbb{E}\Big[\frac{|S_k(0)|^{2q_k}}{(20k-y)^{2q_k}}\Big].$$

By Equation (80) from Lemma 16 in Appendix A, and the bound $q_k \ll k$ valid in the range $y/4 \leq k \leq n_0$, this is

$$\ll \sum_{y/4\leq k\leq n_0} k^{3/2} e^k \exp\Big(-\frac{(20k-y)^2}{k+C}\Big) \leq \sum_{y/4\leq k\leq n_0} k^{3/2} e^k\, e^{-400k+20y} \leq e^{-2y},$$

with C an absolute constant. This is the expected result.

Finally, we show that $\mathbb{P}(\exists h \in D_0^{\rm c} \cap A_0) \ll e^{-100n}$. Suppose that we are placed on a τ for which for all $h \in \mathcal{T}_n$ we have

$$|\zeta_\tau(h)| \leq e^{100n}. \tag{10}$$

Then for all $h \in A_0$ we have by Lemma 23 in Appendix A

$$|e^{-S_k(h)}| \leq (1+e^{-n_{-1}})|\mathcal{M}_{-1}^{(k)}(h)| + e^{-10^5(n_0-n_{-1})}.$$

It follows that for such τ's we have for all $h \in A_0$,

$$|(\zeta_\tau e^{-S_k})(h)| \leq (1+e^{-n_{-1}})\, |(\zeta_\tau \mathcal{M}_{-1}^{(k)})(h)| + e^{100n-10^5(n_0-n_{-1})}$$
$$\leq (1+e^{-n_{-1}})\, |(\zeta_\tau \mathcal{M}_{-1}^{(k)})(h)| + e^{-10^4(n-n_{-1})},$$

as claimed. Therefore, we are left with the elementary bound

$$\mathbb{P}(\exists h \in D_0^c \cap A_0) \le \mathbb{P}(\exists h : |\zeta_\tau(h)| \ge e^{100n}) \le \sum_{h\in\mathcal{T}_n} \mathbb{E}\Big[\frac{|\zeta_\tau(h)|^2}{e^{200n}}\Big] \ll e^{-100n},$$

by the second moment bound for the zeta function (Lemma 21, Appendix A).

5. Induction

We now prove Proposition 2. The subsets A, B, C and D's need to be refined to account for the intermediate increments in the interval $(n_\ell, n_{\ell+1}]$: For $k \in [n_\ell, n_{\ell+1}]$, define

$$\begin{aligned}
A_\ell^{(k)} &= A_\ell \cap \{h \in \mathcal{T}_n : |\widetilde{S}_j(h) - \widetilde{S}_{n_\ell}(h)| \le 10^3(n-n_\ell) \text{ for all } n_\ell < j \le k\},\\
B_\ell^{(k)} &= B_\ell \cap \{h \in \mathcal{T}_n : S_j(h) \le m(j) + U_y(j) \text{ for all } n_\ell < j \le k\},\\
C_\ell^{(k)} &= C_\ell \cap \{h \in \mathcal{T}_n : S_j(h) > m(j) + L_y(j) \text{ for all } n_\ell < j \le k\},\\
D_\ell^{(k)} &= D_\ell \cap \{h \in \mathcal{T}_n : |(\zeta_\tau e^{-S_k})(h)| \le c_{\ell+1}|(\zeta_\tau \mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(h)| + e^{-10^4(n-n_\ell)}\\
&\qquad\qquad \text{for all } n_\ell < j \le k\},
\end{aligned}$$

where $c_{\ell+1} := \prod_{i=0}^{\ell+1}(1+e^{-n_{i-1}})$. Note that with this notation $A_\ell^{(n_{\ell+1})} = A_{\ell+1}$. We also take as a convention that $A_\ell^{(n_\ell)} = A_\ell$. The same holds for $B_\ell^{(k)}$, $C_\ell^{(k)}$ and $D_\ell^{(k)}$.

The proof of Proposition 2 is based on the following two lemmas. We defer the proofs of these lemma to later sections.

Lemma 3. *Let $\ell \ge 0$ be such that $\exp(10^6(n-n_\ell)^{10^5}e^{n_{\ell+1}}) \le \exp(\frac{1}{100}e^n)$. Let $k \in (n_\ell, n_{\ell+1}]$. Let $\mathcal{Q}$ be a Dirichlet polynomial of length $N \le \exp(\frac{1}{100}e^n)$. Suppose that $\mathcal{Q}$ is supported on integers all of whose prime factors are $> \exp(e^k)$. Then, for $4000 \le y \le n$ and $L_y(k) < w - m(k) < U_y(k)$, one has*

$$\begin{aligned}
&\mathbb{E}\Big[\max_{|h|\le 2}|\mathcal{Q}(\tfrac12 + \mathrm{i}\tau + \mathrm{i}h)|^2 \cdot \mathbf{1}\Big(h \in B_\ell^{(k)} \cap C_\ell^{(k)} \text{ and } S_k(h) \in (w, w+1]\Big)\Big]\\
&\qquad \ll \mathbb{E}\Big[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2\Big]\Big(e^{-k}\log N + (n-k)^{800}\Big)\, y\,(U_y(k) - w + m(k) + 2)\, e^{-2(w-m(k))},
\end{aligned}$$

where the implicit constant in $\ll$ is absolute and in particular independent of ℓ and k.

Lemma 4. *Let $\ell \ge 0$ with $\exp(10^6(n-n_\ell)^{10^5}e^{n_{\ell+1}})) \le \exp(\frac{1}{100}e^n)$. Let $k \in [n_\ell, n_{\ell+1}]$. Let $\gamma(m)$ be a sequence of complex coefficients with $|\gamma(m)| \ll \exp(\frac{1}{1000}e^n)$ for all $m \ge 1$. Let*

$$\mathcal{Q}_\ell^{(k)}(h) := \sum_{\substack{p|m \Rightarrow p\in(T_\ell, \exp(e^k)]\\ \Omega_{\ell+1}(m) \le (n_{\ell+1}-n_\ell)^{10^4}}} \frac{\gamma(m)}{m^{\frac12 + \mathrm{i}\tau + \mathrm{i}h}}. \tag{11}$$

Then, for any $h \in [-2,2]$, $4000 \le y \le n$ and $L_y(n_\ell) < u - m(n_\ell) \le U_y(n_\ell)$,

$$\begin{aligned}
&\mathbb{E}\Big[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(h)|^4 \cdot |\mathcal{Q}_\ell^{(k)}(h)|^2 \cdot \mathbf{1}\Big(h \in B_\ell \cap C_\ell \text{ and } S_{n_\ell}(h) \in (u, u+1]\Big)\Big]\\
&\qquad \ll e^{4(n-k)}\,\mathbb{E}\Big[|\mathcal{Q}_\ell^{(k)}(h)|^2\Big]\, e^{-n_\ell}\, y\,(U_y(n_\ell) - u + m(n_\ell) + 2)\, e^{-2(u-m(n_\ell))},
\end{aligned}$$

where the implicit constant in $\ll$ is absolute and in particular independent of ℓ and k.

Note that we allow $k = n_\ell$ in which case $\mathcal{Q}^{(k)} = 1$. Some explanations on the heuristics of Lemmas 3 and 4 might be in order. First, one expects the partial sums $S_k(h)$ to be approximately Gaussian. In fact, one can see $S_k(h)$ for a fixed h as a Gaussian random walk of mean 0 and variance 1/2 for each of its increment. For such a random walk, the endpoint S_k is independent of the "bridge" $S_j - \frac{j}{k} S_k$ for all $j \leq k$. Since $S_k \approx m(k)$, the latter is approximately $S_j - m(j)$. With this in mind, the indicator function can be thought of as the restriction of the endpoint S_k being in w and that the walk $S_j - m(j)$ starting at 0 and ending at $w - m(k)$ stays below the barrier $y + U_y(k)$. Using the ballot theorem, Proposition 4 from Appendix B, the probability of this happening for a fixed h is

$$\frac{y(U_y(k) - w + m(k))}{k^{3/2}} e^{-\frac{w^2}{k}} \ll y(U_y(k) - w + m(k)) e^{-k} e^{-2(w - m(k))}.$$

Since $S_k(h)$ has length $\exp(e^k)$ as a Dirichlet polynomial, one expects that there are approximately e^k independent random walks as h varies in $[-2, 2]$. Moreover, the Dirichlet polynomial $\mathcal{Q}$ is supported on primes larger than $\exp(e^k)$, so its value should be independent of the e^k walks, as they are "supported" on different primes. Also, due to the greatest frequency $\log N$ in the summands of $\mathcal{Q}$, there should be $\log N$ independent values when discretizing the maximum. The factor $(n-k)^{800}$ comes from the process of approximating the indicator function by a Dirichlet polynomial. These factors together reproduce the result of Lemma 3. The heuristics for Lemma 4 is the same with the extra fourth moment. Again, one expects $\log \zeta_\tau(h) - S_k(h)$ to be independent of $\mathcal{Q}_\ell^{(k)}$ and S_{n_ℓ}. Therefore, the expectation of the fourth moment could formally be factored out. The variable $\log \zeta_\tau(h) - S_k(h)$ should be approximately Gaussian with variance $n - k$. Therefore, $\mathbb{E}[e^{4(\log \zeta_\tau(h) - S_k(h))}] \approx e^{4(n-k)}$. The mollifiers $\mathcal{M}$ are designed to approximate $e^{-S_{n_\ell}}$.

We are now ready to begin the proof of Proposition 2. Notice that by a union bound,

$$\mathbb{P}(\exists h \in H(y) \cap G_\ell) \leq \mathbb{P}(\exists h \in H(y) \cap G_\ell \cap G_{\ell+1}^{\mathrm{c}}) + \mathbb{P}(\exists h \in H(y) \cap G_{\ell+1}).$$

The first term can be further split by another union bound,

$$\begin{aligned}
\mathbb{P}(\exists h \in H(y) \cap G_\ell \cap G_{\ell+1}^{\mathrm{c}}) \leq\ & \mathbb{P}(\exists h \in A_{\ell+1}^{\mathrm{c}} \cap H(y) \cap G_\ell) \\
& + \mathbb{P}(\exists h \in D_{\ell+1}^{\mathrm{c}} \cap A_{\ell+1} \cap H(y) \cap G_\ell) \\
& + \mathbb{P}(\exists h \in C_{\ell+1}^{\mathrm{c}} \cap D_{\ell+1} \cap A_{\ell+1} \cap H(y) \cap G_\ell) \\
& + \mathbb{P}(\exists h \in B_{\ell+1}^{\mathrm{c}} \cap C_{\ell+1} \cap A_{\ell+1} \cap H(y) \cap G_\ell).
\end{aligned}$$

It will be shown that each of the above probabilities is bounded by

$$\ll \frac{y e^{-2y}}{(\log_{\ell+1} n)^{100}}.$$

This will conclude the proof. The proof of each bound is broken down into a separate subsection. The estimate with $B_{\ell+1}^c$ is the tightest. We will sometimes drop some events that are not needed to achieve the bound.

5.1. **Bound on increments.** We first consider $A_{\ell+1}^{\mathrm{c}}$. This is the simplest bound. We show by a Markov-type inequality that

$$\mathbb{P}(\exists h \in A_{\ell+1}^{\mathrm{c}} \cap G_\ell) \ll y e^{-2y} (\log_{\ell-1} n)^{-1}.$$

(Recall our convention that $\log_{-1} n = e^n$ and $\log_0 n = n$.) If there is a $k \in (n_\ell, n_{\ell+1}]$ and an h such that $|\widetilde{S}_k(h) - \widetilde{S}_{n_\ell}(h)| > 10^3(n - n_\ell)$, then one has that

$$\sum_{k \in (n_\ell, n_{\ell+1}]} \max_{|h| \le 2} \frac{|\widetilde{S}_k(h) - \widetilde{S}_{n_\ell}(h)|^{2q}}{(10^3(n-n_\ell))^{2q}} \ge 1, \text{ for all } q \ge 1.$$

Therefore, for any choice of $q \ge 1$, the following bound holds

$$\mathbb{P}(\exists h \in A_{\ell+1}^{\mathrm{c}} \cap G_\ell) \le \sum_{k \in (n_\ell, n_{\ell+1}]} \mathbb{E}\Big[\max_{|h| \le 2} \frac{|(\widetilde{S}_k - \widetilde{S}_{n_\ell})(h)|^{2q}}{(10^3(n-n_\ell))^{2q}} \cdot \mathbf{1}\Big(h \in G_\ell\Big)\Big]. \tag{12}$$

We pick $q = \lfloor 10^6 (n - n_\ell)^2 / (k - n_\ell) \rfloor$. The Dirichlet polynomial $(\widetilde{S}_k - \widetilde{S}_{n_\ell})^q$ is then of length at most $\exp(2qe^k) \ll \exp(2 \cdot 10^6 (n-n_\ell)^2 e^{n_{\ell+1}})$, which is much smaller than $\exp(e^n/100)$ by the definition of n_ℓ. Lemma 3 thus bounds the right-hand side of (12) with

$$y e^{-2y} \sum_{k \in (n_\ell, n_{\ell+1}]} (q + (n-n_\ell)^{800})\, e^{100(n-n_\ell)}\, \mathbb{E}\Big[\frac{|\widetilde{S}_k - \widetilde{S}_{n_\ell}|^{2q}}{(10^3(n-n_\ell))^{2q}}\Big].$$

For our choice of q, we have $2q \ll (n - n_\ell)^2 \le e^{n-k}$, so that the estimate in Lemma 17 applies. Together with Stirling's approximation as in (83) we conclude that the above is

$$\ll y e^{-2y} e^{-(n-n_\ell)} \ll y e^{-2y} (\log_{\ell-1} n)^{-1}.$$

5.2. **Bound with mollifiers.** We now estimate $D_{\ell+1}^{\mathrm{c}}$. In this section, we obtain

$$\mathbb{P}(\exists h \in D_{\ell+1}^{\mathrm{c}} \cap A_{\ell+1} \cap G_\ell) \ll y e^{-2y} (\log_{\ell-1} n)^{-1}. \tag{13}$$

For h in $A_{\ell+1} \cap D_\ell$, we have

$$|(\widetilde{S}_k - \widetilde{S}_{n_\ell})(h)| < 10^3(n_{\ell+1} - n_\ell), \tag{14}$$

$$|(\zeta_\tau e^{-S_{n_\ell}})(h)| < c_\ell |(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell)(h)| + e^{-10^4(n-n_{\ell-1})}, \tag{15}$$

where $c_\ell = \prod_{i=0}^{\ell} (1 + e^{-n_{i-1}})$. If we additionally assume that, for all $h \in A_{\ell+1} \cap D_\ell$, both

$$|(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell)(h)| < e^{10^3(n-n_\ell)} \tag{16}$$

$$|(e^{-(S_k - S_{n_\ell})})(h)| \le (1 + e^{-n_\ell})\, |\mathcal{M}_\ell^{(k)}(h)| + e^{-10^5(n_{\ell+1}-n_\ell)}, \tag{17}$$

hold for all $k \in (n_\ell, n_{\ell+1}]$, then we obtain (where each of the expression below is evaluated at h),

$$\begin{aligned}
|\zeta_\tau e^{-S_k}| &= |\zeta_\tau e^{-S_{n_\ell}}|\, e^{-(S_k - S_{n_\ell})}\\
&< \Big(c_\ell |\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell| + e^{-10^4 (n - n_{\ell-1})} \Big) e^{-(S_k - S_{n_\ell})}\\
&\leq c_\ell |\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell| e^{-(S_k - S_{n_\ell})} + e^{-10^3 (n - n_{\ell-1})}\\
&\leq c_{\ell+1} |\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell \mathcal{M}_\ell^{(k)}| + c_\ell |\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell|\, e^{-10^5 (n_{\ell+1} - n_\ell)} + e^{-10^3 (n - n_{\ell-1})}\\
&\leq c_{\ell+1} |\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell \mathcal{M}_\ell^{(k)}| + e^{-10^4 (n - n_\ell)}.
\end{aligned}$$

Here, we have successively used the estimates (15), (14), (17), (16), and the fact that the sequence c_ℓ, $\ell > -1$, is rapidly convergent. It remains to verify that the bounds (16) and (17) hold with high probability for $h \in A_{\ell+1} \cap D_\ell$. The bound (17) holds pointwise for all $h \in A_{\ell+1}$ by Lemma 23 in Appendix A. As for Equation (16), the probability of the complement of the event is

$$\sum_{h \in \mathcal{T}_n} \mathbb{P}\Big(|(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell)(h)| \geq e^{10^3 (n - n_\ell)}, h \in G_\ell \Big) \tag{18}$$

$$\ll e^{-4 \cdot 10^3 (n - n_\ell)}\, e^n\, \mathbb{E}\Big[|(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell)(0)|^4 \cdot \mathbf{1}(0 \in G_\ell) \Big]. \tag{19}$$

Lemma 4 applied for $\mathcal{Q} \equiv 1$ and $k = n_\ell$ then implies the expected bound,

$$\ll y e^{-2y - 4 \cdot 10^3 (n - n_\ell)}\, e^{100 (n - n_\ell)} \ll y e^{-2y} (\log_{\ell-1} n)^{-1}.$$

Note that (18) can be made small, because the union bound on the random variables $\log |(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell)(h)|$ (which are approximately Gaussian of variance $n - n_\ell$) is effectively on the h's in $G_\ell(0)$. The number of such h's is small enough, of order $e^{n - n_\ell}$.

5.3. **Extension of the lower barrier.** We now want to prove the following bound on $C^{\mathrm{c}}_{\ell+1}$:

$$\mathbb{P}(\exists h \in H(y) \cap C^{\mathrm{c}}_{\ell+1} \cap D_{\ell+1} \cap A_{\ell+1} \cap G_\ell) \ll y e^{-2y} (\log_\ell n)^{-1}. \tag{20}$$

Here, we explicitly make use of the fact that ζ_τ is large. Let $h \in C^{\mathrm{c}}_{\ell+1} \cap D_{\ell+1} \cap G_\ell \cap H(y)$. By definition of $C^{\mathrm{c}}_{\ell+1}$, there must be a k such that $S_k(h) \leq m(k) - 20(n-k) + y$. We split $S_k(h)$ according to the value of $S_{n_\ell}(h) \in [u, u+1]$ and $(S_k - S_{n_\ell})(h) \in [v, v+1]$, where $u, v \in \mathbb{Z}$, $|v| \leq 10^3 (n - n_\ell)$ and $u + v \leq m(k) - 20(n-k) + y$. We notice that since $h \in H(y)$,

$$|(\zeta_\tau\, e^{-S_k})(h)| > V e^{-u-v}, \tag{21}$$

where $V = e^y e^n n^{-3/4}$. Since $h \in D_{\ell+1}$ also, we either have

$$|(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell \mathcal{M}_\ell^{(k)})(h)| \gg V e^{-u-v}$$

or $\frac{1}{2} V e^{-u-v} \leq e^{-10^4 (n - n_\ell)}$. However, the second possibility cannot occur since it implies that $e^{u+v} > e^y e^n e^{10^4 (n - n_\ell)} n^{-3/4}$ and hence $e^u > e^y e^n e^{10^3 (n - n_\ell)} n^{-3/4}$. This means that $S_{n_\ell}(h)$ is above the barrier, and this is impossible because $h \in G_\ell$.

Therefore, with a union bound and (21), the left-hand side of (20) is bounded for any $q \geq 1$ by

$$\sum_{\substack{k\in(n_\ell,n_{\ell+1}]\\ h\in\mathcal{T}_n}} \sum_{\substack{u+v\leq m(k)-20(n-k)+y\\ |v|\leq 10^3(n_{\ell+1}-n_\ell)\\ L_y(n_\ell)\leq u-m(n_\ell)\leq U_y(n_\ell)}} \frac{e^{4u+4v}}{V^4}\cdot\mathbb{E}\Big[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(h)|^4\cdot\frac{|(S_k-S_{n_\ell})(h)|^{2q}}{(1+v^{2q})} \tag{22}$$
$$\times\mathbf{1}\Big(S_{n_\ell}(h)\in[u,u+1]\text{ and }h\in A_\ell\cap B_\ell\cap C_\ell\Big)\Big].$$

Pick $q=\lfloor v^2/(k-n_\ell)\rfloor$. Since $q\leq 10^7(n-n_\ell)^2$, the Dirichlet polynomial $(S_k-S_{n_\ell})^q$ can be written in the form (11). In particular, Lemma 4 with $\mathcal{Q}=(S_k-S_{n_\ell})^q$ is applicable. Lemma 16 and Stirling's approximation also imply

$$\mathbb{E}\Big[\frac{|(S_k-S_{n_\ell})(h)|^{2q}}{(1+v^{2q})}\Big]\ll e^{-q}\ll\exp\Big(-\frac{v^2}{k-n_\ell}\Big).$$

Therefore, Lemma 4 and the above computation show that (22) is

$$\ll e^n\sum_{\substack{k\in(n_\ell,n_{\ell+1}]\\ u+v\leq m(k)-20(n-k)+y\\ |v|\leq 10^3(n-n_\ell)\\ L_y(n_\ell)\leq u-m(n_\ell)\leq U_y(n_\ell)}}\frac{e^{4u+4v}}{V^4}e^{4(n-k)}e^{-\frac{v^2}{k-n_\ell}}\ \frac{y(U_y(n_\ell)-u+m(n_\ell)+2))}{e^{n_\ell}}\ e^{-2(u-m(n_\ell))}.$$

We use the restriction $u-m(n_\ell)\in[L_y(n_\ell),U_y(n_\ell)]$ to bound

$$0\leq U_y(n_\ell)-u+m(n_\ell)\leq U_y(n_\ell)-L_y(n_\ell)\ll(n-n_\ell)\ll\log_\ell n\text{ for all }y.$$

Subsequently we remove this restriction on u. After replacing V by $e^y e^n n^{-3/4}$, the above sum is thus bounded by

$$ye^{-4y}\sum_{k\in(n_\ell,n_{\ell+1}]}e^{n-4k-n_\ell}\,n^3\sum_{\substack{u+v\leq m(k)-20(n-k)+y\\ |v|\leq 10^3(n-n_\ell)\\ u,v\in\mathbb{Z}}}e^{2u+2m(n_\ell)+4v}\,(\log_\ell n)\,\exp\Big(-\frac{v^2}{k-n_\ell}\Big).$$

Performing the summation over u, we get

$$\ll ye^{-4y}\sum_{k\in(n_\ell,n_{\ell+1}]}e^{n-4k-n_\ell}\,n^3\sum_{\substack{|v|\leq 10^3(n-n_\ell)\\ v\in\mathbb{Z}}}e^{2m(k)+2m(n_\ell)-40(n-k)+2v+2y}\,(\log_\ell n)\,\exp\Big(-\frac{v^2}{k-n_\ell}\Big).$$

The sum over v can then be performed and yields the bound

$$\begin{aligned}&ye^{-2y}\sum_{k\in(n_\ell,n_{\ell+1}]}e^{n-4k-n_\ell}\,n^3\,e^{2m(k)+2m(n_\ell)+(k-n_\ell)}\,(\log_\ell n)\,e^{-40(n-k)}(k-n_\ell)^{1/2}\\ &\ll ye^{-2y}\sum_{k\in(n_\ell,n_{\ell+1}]}(\log_\ell n)^{3/2}e^{-9(n-k)}\ll ye^{-2y}(\log_\ell n)^{-1},\end{aligned}\tag{23}$$

since $n-k\geq n-n_{\ell+1}=10^6\log_{\ell+1}n$. Notice that in the case $\ell=0$ we use the fact that we save a large power of n in $e^{-(n-k)}$ to offset the term n^3, whereas in the case $\ell\geq 1$ we use the fact that $e^{4m(k)}n^3\asymp e^{4k}$ for $k\in(n_\ell,n_{\ell+1}]$.

5.4. **Extension of the upper barrier.** We need the following bound on $B_{\ell+1}^{\rm c}$:

$$\mathbb{P}(\exists h\in H(y)\cap B_{\ell+1}^{\rm c}\cap A_{\ell+1}\cap C_{\ell+1}\cap G_\ell)\ll \frac{ye^{-2y}}{(\log_{\ell+1} n)^{100}}.$$

In fact, we show the stronger estimate

$$\mathbb{P}(\exists h\in (B_\ell\setminus B_{\ell+1})\cap C_{\ell+1})\ll \frac{ye^{-2y}}{(\log_{\ell+1} n)^{100}}. \tag{24}$$

We write $\overline{S}_j=S_j-m(j)$ for simplicity.

By considering a union bound on $k\in[n_\ell,n_{\ell+1})$ and by partitioning the values of $S_k(h)$ according to $S_k(h)\in[w,w+1]$ with $w\in\mathbb{Z}$, the above reduces to

$$\ll \sum_{k\in[n_\ell,n_{\ell+1})}\mathbb{P}(\exists h\in (B_\ell^{(k)}\setminus B_\ell^{(k+1)})\cap C_\ell^{(k)})$$
$$\ll \sum_{\substack{k\in[n_\ell,n_{\ell+1})\\ w\in[L_y(k),U_y(k))}}\mathbb{P}(\exists h:\overline{S}_j(h)<U_y(j)\ \forall j\leq k, \overline{S}_{k+1}(h)>U_y(k+1),\overline{S}_k(h)\in(w,w+1]).$$

Note that the condition $\overline{S}_{k+1}>U_y(k+1)$ under the restriction $\overline{S}_k(h)\in(w,w+1]$ can be rewritten as

$$S_{k+1}-S_k>U_y(k+1)+m(k+1)-m(k)-\overline{S}_k>U_y(k+1)-w+{\rm o}(\log n/n).$$

Write $V_{w,k}=U_y(k+1)-w$. By Markov's inequality, the above sum is bounded by

$$\ll \sum_{\substack{k\in[n_\ell,n_{\ell+1})\\ w\in[L_y(k),U_y(k))}}\mathbb{E}\Big[\max_{|h|\leq 2}\frac{|(S_{k+1}-S_k+1)(h)|^{2q}}{(V_{w,k}+1)^{2q}}\mathbf{1}\Big(\overline{S}_j(h)<U_y(j)\ \forall j\leq k, \overline{S}_k(h)\in(w,w{+}1]\Big)\Big].$$

We pick $q=(V_{w,k}+1)^2/10=(U_y(k+1)-w+1)^2/10\leq 400(n-k)^2$, by the bounds on w. For this choice, note that the Dirichlet $(S_{k+1}-S_k+1)^q$ has length at most $\exp(2qe^{k+1})\leq\exp(1000(n-k)^2e^{k+1})$. In particular, Lemma 3 can be applied (note that the Dirichlet polynomial $S_{k+1}-S_k+1$ is supported on integers all of whose prime factors are $>\exp(e^k)$ since 1 is not a prime!). This yields the bound

$$\ll \sum_{k\in[n_\ell,n_{\ell+1})}(n-k)^{800}\sum_{w\in[L_y(k),U_y(k))}\frac{\mathbb{E}[|(S_{k+1}-S_k+1)(0)|^{2q}]}{(V_{w,k}+1)^{2q}}\,y\,e^{-2w}\,(U_y(k)-w+1).$$

The expectation is $\ll (2q)!/q!+4^q\ll 100^q(q/e)^q$ by Equation (79) of Lemma 16 in Appendix A. We then find using Stirling's formula (similarly as in (82), but the optimal exponent is not needed here), that

$$\frac{\mathbb{E}[|(S_{k+1}-S_k+1)(0)|^{2q}]}{(V_{w,k}+1)^{2q}}\ll e^{-(V_{w,k}+1)^2/10}.$$

Putting this back in the estimate gives the bound

$$\ll y\sum_{k\in[n_\ell,n_{\ell+1})}(n-k)^{800}e^{-2U_y(k)}\sum_{w\in[L_y(k),U_y(k))}(U_y(k)-w+1)e^{-\frac{1}{10}(U_y(k+1)-w+1)^2+2(U_y(k+1)-w+1)},$$

where we added $U_y(k+1)$ and subtracted $U_y(k)$ which is allowed since $U_y(k+1)-U_y(k) = \mathrm{O}(1)$. Finally $-(U_y(k+1)-w+1)^2/10+2(U_y(k+1)-w+1) = -(1/10)(U_y(k+1)-w-9)^2+10$ so the last sum over w is finite. It remains to recall that $U_y(k) = y+10^3\log(n-k)$ to conclude that

$$\mathbb{P}(\exists h \in (B_\ell \setminus B_{\ell+1}) \cap C_{\ell+1}) \ll ye^{-2y} \sum_{k\in[n_\ell,n_{\ell+1})} (n-k)^{800} e^{-10^3\log(n-k)} \ll \frac{ye^{-2y}}{(\log_{\ell+1} n)^{100}}.$$

6. Final Step

This short section proves Proposition 3. We notice that if $h \in H(y) \cap G_\ell$, then $S_{n_\ell}(h) \in [v, v+1]$ with $|v-y-m(n_\ell)| \le 20(n-n_\ell)$, and $|(\zeta_\tau e^{-S_{n_\ell}})(h)| \ge Ve^{-v}$ where $V = e^y e^n/n^{3/4}$. We wish to apply Markov's inequality and Lemma 4. We first need to compare the expression to the one with mollifiers. To this end, note that since $h \in D_\ell$ and $Ve^{-v} > 2e^{-10^4(n-n_{\ell-1})}$, we have $Ve^{-v} \ll |(\zeta_\tau \mathcal{M}_{-1}\dots\mathcal{M}_\ell)(h)|$. Therefore, Markov's inequality implies that

$$\mathbb{P}(\exists h \in H(y) \cap G_\ell)$$
$$\ll \sum_{\substack{h\in\mathcal{T}_n \\ |v-y-m(n_\ell)|\le 20(n-n_\ell)}} \frac{e^{4v}}{V^4}\,\mathbb{E}\Big[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell)(h)|^4 \cdot \mathbf{1}\Big(S_{n_\ell}(h) \in [v,v+1] \text{ and } h \in B_\ell \cap C_\ell\Big)\Big].$$

By Lemma 4, this is

$$\ll e^{-4y}e^{-4n}n^3\, e^n \sum_{|v-y-m(n_\ell)|\le 20(n-n_\ell)} e^{4(n-n_\ell)} e^{2v} e^{2m(n_\ell)}\, y\,(n-n_\ell)e^{-n_\ell} \ll ye^{-2y}e^{100(n-n_\ell)}.$$

The last inequality is obtained similarly as in Equation (23): when $\ell = 0$ the term n^3 is included in $e^{100(n-n_0)}$, while for $\ell \ge 1$ we have $e^{4m(n_\ell)}n^3 \ll e^{4n_\ell}$. This concludes the proof of (24).

7. Decoupling and Second Moment

7.1. Lemmas from harmonic analysis.

We will need the following lemmas from harmonic analysis.

Lemma 5. *There exists a smooth function F_0 such that*

(1) For all $x \in \mathbb{R}$, we have $0 \le F_0(x) \le 1$ and $\widehat{F}_0(x) \ge 0$.
(2) $\widehat{F}_0$ is compactly supported on $[-1, 1]$.
(3) Uniformly in $x \in \mathbb{R}$, we have

$$F_0(x) \ll e^{-|x|/\log^2(|x|+10)}.$$

Proof. This follows from the sufficient part of the main theorem of [22]. Note that this theorem does not state the positivity conditions on F_0 and $\hat{F}_0$ but these can be obtained from the explicit construction in [22]. □

The above lemma allows us to construct a convenient approximation to the indicator function of a small interval $[0, \Delta^{-1}]$.

Lemma 6. *There exists an absolute constant $C > 0$ such that for any $\Delta, A \geq 3$ there exists an entire function $G_{\Delta,A}(x) \in L^2(\mathbb{R})$ such that*

(1) The Fourier transform $\widehat{G}_{\Delta,A}(x)$ is supported on $[-\Delta^{2A}, \Delta^{2A}]$.
(2) We have, $0 \leq G_{\Delta,A}(x) \leq 1$ for all $x \in \mathbb{R}$.
(3) We have $\mathbf{1}(x \in [0, \Delta^{-1}]) \leq G_{\Delta,A}(x) \cdot (1 + Ce^{-\Delta^{A-1}})$.
(4) We have, $G_{\Delta,A}(x) \leq \mathbf{1}(x \in [-\Delta^{-A/2}, \Delta^{-1} + \Delta^{-A/2}]) + Ce^{-\Delta^{A-1}}$.
(5) We have, $\int_{\mathbb{R}} |\widehat{G}_{\Delta,A}(x)| \mathrm{d}x \leq 2\Delta^{2A}$.

Proof. Let $F = F_0/\|F_0\|_1$ so that $\int_{\mathbb{R}} F(x)\mathrm{d}x = 1$, where F_0 is the function of Lemma 5. Consider

$$G_{\Delta,A}(x) = \int_{-\Delta^{-A}}^{\Delta^{-1}+\Delta^{-A}} \Delta^{2A} F(\Delta^{2A}(x-t))\mathrm{d}t. \tag{25}$$

Notice that the Fourier transform of $F(\Delta^{2A}(x-t))$ is compactly supported on $[-\Delta^{2A}, \Delta^{2A}]$, and therefore so is the Fourier transform of $G_{\Delta,A}$. Clearly, $G_{\Delta,A}$ is non-negative. By completing the integral to infinity and a change of variables, $G_{\Delta,A}$ is bounded by 1. This proves the first two assertions.

For a given $x \in [0, \Delta^{-1}]$, the right-hand side of (25) is at least

$$C_{\Delta,A} = \int_{-\Delta^{-A}}^{\Delta^{-A}} \Delta^{2A} F(\Delta^{2A}x)\mathrm{d}x = \int_{-\Delta^{A}}^{\Delta^{A}} F(x)\mathrm{d}x = 1 + \mathrm{O}(e^{-\Delta^{A-1}}).$$

Hence, for $x \in [0, \Delta^{-1}]$, we have $1 \leq G_{\Delta,A}(x)/C_{\Delta,A} = G_{\Delta,A}(x)\,(1 + \mathrm{O}(e^{-\Delta^{A-1}}))$, thus proving the third assertion.

For $x \in [-\Delta^{-A/2}, \Delta^{-1} + \Delta^{-A/2}]$, the upper bound $G_{\Delta,A}(x) \leq 1$ is immediate from completing the integral in (25) to all $t \in \mathbb{R}$. Thus we can assume that $x \notin [-\Delta^{-A/2}, \Delta^{-1} + \Delta^{-A/2}]$. We want to show that for such x we have $G_{\Delta,A}(x) \ll e^{-\Delta^{A-1}}$. Assuming first that $x < -\Delta^{-A/2}$ we get

$$G_{\Delta,A}(x) = \int_{-\Delta^{-A}}^{\Delta^{-1}+\Delta^{-A}} \Delta^{2A} F(\Delta^{2A}(x-t))\mathrm{d}t \ll e^{-\Delta^{A-1}},$$

using the decay bound $F(x) \ll e^{-|x|/\log^2(10+|x|)}$. The bound for $x > \Delta^{-1} + \Delta^{-A/2}$ is obtained in the same way.

Finally, to prove the last claim, we first notice that, since $\widehat{G}_{\Delta,A}(x)$ is supported on $[-\Delta^{2A}, \Delta^{2A}]$, the Cauchy-Schwarz inequality and the Plancherel theorem imply that

$$\int_{\mathbb{R}} |\widehat{G}_{\Delta,A}(x)|\mathrm{d}x \leq \sqrt{2}\Delta^{A}\Big(\int_{\mathbb{R}} |G_{\Delta,A}(x)|^2\mathrm{d}x\Big)^{1/2}. \tag{26}$$

Second, the Cauchy-Schwarz inequality also implies, taking $u = \Delta^{2A}t$ in (25),

$$|G_{\Delta,A}(x)|^2 \leq \Delta^{2A} \int_{-\Delta^{A}}^{\Delta^{2A-1}+\Delta^{A}} F^2(\Delta^{2A}x - u)\mathrm{d}u \leq \Delta^{2A} \int_{-\Delta^{A}}^{\Delta^{2A-1}+\Delta^{A}} F(\Delta^{2A}x - u)\mathrm{d}u,$$

since $0 \leq F \leq 1$. Thus, we have by integrating

$$\int_{\mathbb{R}} |G_{\Delta,A}(x)|^2\mathrm{d}x \leq 2\Delta^{2A},$$

giving the desired bound in Equation (26). □

7.2. **Approximation of indicators by Dirichlet polynomials.** We will work throughout with the increments

$$Y_j := S_j - S_{j-1}\ ,\ j \geq 1,$$

with S_j as in Equation (4). For $\ell \geq -1$ and $k \in (n_\ell, n_{\ell+1}]$, consider the discretization parameter

$$\Delta_j = (\min(j, n-j))^4, \qquad j \in (n_\ell, k].$$

We approximate indicator functions of Y_j on intervals of width Δ_j^{-1}. This choice for Δ_j is guided by two constraints. First, some summability is used, in particular in (48). From the proof it will be clear that we could choose any exponent strictly greater than 1 instead of 4. Second, the Gaussian approximation of the Dirichlet sums gets worse for very small primes, imposing a decrease down to $\Delta_j \asymp 1$ for $j \asymp 1$, see Equation (43) below.

Set $r = r(y) = \lceil y/4 \rceil$. Since $y > 4000$ we have $r > n_{-1} = 1000$. For $L_y(r) \leq v - m(r) \leq U_y(r)$ and $L_y(k) \leq w - m(k) \leq U_y(k)$, define the set $\mathcal{I}_{r,k}(v,w) \subset \mathbb{R}^{k-r}$ of $(k-r)$-tuples $(u_{r+1}, \ldots, u_k)$ with $u_j \in \Delta_j^{-1}\mathbb{Z}$, $r < j \leq k$ such that

$$\begin{aligned} &\text{for all } j \in (r,k]\text{: } L_y(j) - 1 \leq v + \sum_{i=r+1}^{j} u_i - m(j) \leq U_y(j) + 1, \\ &\Big| \sum_{i=r+1}^{k} u_i + v - w \Big| \leq 1 \end{aligned} \tag{27}$$

Note that since $U_y(j) - L_y(j) \leq 40 \min(j, n-j)$ the first restriction on the u_j's imply that $|u_j| \leq 100 \Delta_j^{1/4}$ for every $j \in (r,k]$.

Given $\Delta, A > 1$, we define the following truncated polynomial,

$$\mathcal{D}_{\Delta,A}(x) = \sum_{\ell \leq \Delta^{10A}} \frac{(2\pi \mathrm{i} x)^\ell}{\ell!} \int_{\mathbb{R}} \xi^\ell \widehat{G}_{\Delta,A}(\xi) \mathrm{d}\xi. \tag{28}$$

We will be approximating the indicator function $\mathbf{1}(Y_j(h) \in [u_j, u_j + \Delta_j^{-1}])$ by the Dirichlet polynomial $\mathcal{D}_{\Delta_j,A}(Y_j - u_j)$. The following properties of $\mathcal{D}_{\Delta_j,A}(Y_j - u_j)$ are straightforward from the definition of $\mathcal{D}_{\Delta,A}$ and Y_j:

(1) It is is supported on integers n whose prime factors lie in $(\exp e^{j-1}, \exp(e^j)]$ and such that $\Omega(n) \leq \Delta_j^{10A}$.
(2) The length of the Dirichlet polynomial $\mathcal{D}_{\Delta_j,A}(Y_j - u_j)$ is at most $\exp(2\Delta_j^{10A} e^j)$ (the factor 2 in the exponential is due to the second order term $p^{-1-2\mathrm{i}h}$ in the summands of S_k).
(3) We have

$$\int_{\mathbb{R}} |\xi|^\ell |\widehat{G}_{\Delta,A}(\xi)| \mathrm{d}\xi \leq \Delta^{2A\ell} \int_{\mathbb{R}} |\widehat{G}_{\Delta,A}(\xi)| \mathrm{d}\xi \leq 2\Delta^{2A\ell}\Delta^{2A}, \tag{29}$$

by properties *(1)* and *(5)* of Lemma 6. In particular, the coefficients of $\mathcal{D}_{\Delta_j,A}(Y_j - u_j)$ are bounded by $\ll \Delta_j^{2A(\ell+1)}$.

The first lemma successively approximates the indicator functions $\mathbf{1}(Y_j(h) \in [u_j, u_j + \Delta_j^{-1}])$ by the polynomials $\mathcal{D}_{\Delta_j,A}(Y_j(h) - u_j)$.

Lemma 7. *Let $A > 10$. Let $y > 4000$, $\ell \geq -1$ and $k > r$. Let w be such that $L_y(k) \leq w - m(k) \leq U_y(k)$. Then, for any fixed τ, one has*

$$\mathbf{1}\Big(h \in B_\ell^{(k)} \cap C_\ell^{(k)} : S_k(h) \in [w, w+1]\Big)$$
$$\leq C \sum_{\substack{v \in \Delta_r^{-1}\mathbb{Z} \\ L_y(r) \leq v - m(r) \leq U_y(r) \\ \mathbf{u} \in \mathcal{I}_{r,k}(v,w)}} |\mathcal{D}_{\Delta_r,A}(S_r(h) - v)|^2 \prod_{j=r+1}^{k} |\mathcal{D}_{\Delta_j,A}(Y_j(h) - u_j)|^2.$$

with $C > 0$ an absolute constant.

The proof of the above lemma is split in two parts. We will first rely on the following claim: For every $j \in (n_\ell, k]$ and any $|u_j| \leq 100 \min(j, n-j)$, we have

$$\mathbf{1}(Y_j(h) \in [u_j, u_j + \Delta_j^{-1}]) \leq |\mathcal{D}_{\Delta_j,A}(Y_j(h) - u_j)|^2 \,(1 + Ce^{-\Delta_j^{A-1}}), \tag{30}$$

and for $|v| \leq 100 \min(j, n-j)$

$$\mathbf{1}(S_r(h) \in [v, v + \Delta_r^{-1}]) \leq C|\mathcal{D}_{\Delta_r,A}(S_r(h) - v)|^2, \tag{31}$$

with $C > 0$ an absolute constant.

7.2.1. *Proof of Equations* (30) *and* (31). We prove Equation (30). Equation (31) is done the same way. Lemma 6 implies

$$\begin{aligned}\mathbf{1}(Y_j(h) \in [u_j, u_j + \Delta_j^{-1}]) &\leq |G_{\Delta_j,A}(Y_j - u_j)|^2 \,(1 + Ce^{-\Delta_j^{A-1}}) \\ &= \Big| \int_{\mathbb{R}} e^{2\pi \mathrm{i} \xi (Y_j(h) - u_j)} \widehat{G}_{\Delta_j,A}(\xi) \mathrm{d}\xi \Big|^2 \,(1 + Ce^{-\Delta_j^{A-1}}) \ ,\end{aligned}$$

with $C > 0$ an absolute constant. Expanding the exponential up to $\nu = \Delta_j^{10A}$, the integral in the absolute value is equal to

$$\sum_{\ell \leq \nu} \frac{(2\pi \mathrm{i})^\ell}{\ell!} (Y_j(h) - u_j)^\ell \int_{\mathbb{R}} \xi^\ell \widehat{G}_{\Delta,A}(\xi) \mathrm{d}\xi + \mathrm{O}^\star\Big(\frac{(2\pi)^\nu}{\nu!} |Y_j(h) - u_j|^\nu \int_{\mathbb{R}} |\xi|^\nu |\widehat{G}_{\Delta_j,A}(\xi)| \mathrm{d}\xi\Big) \tag{32}$$

where $\mathrm{O}^\star$ means that the implicit constant in the O is ≤ 1.

To bound the error term, observe that, since $h \in B_\ell^{(k)} \cap C_\ell^{(k)}$, the restriction on u_j and on $Y_j(h)$ imposed by the upper and lower barriers imply $|Y_j(h) - u_j| \leq 10^4 \Delta_j^{1/4}$. Together with (29), this implies the bound

$$\frac{(2\pi)^\nu}{\nu!} |Y_j(h) - u_j|^\nu \int_{\mathbb{R}} |\xi^\nu| |\widehat{G}_{\Delta_j,A}(\xi)| \mathrm{d}\xi \leq \frac{(10^6)^\nu}{\nu!} \Delta_j^{\nu/4} \, \Delta_j^{2A(\nu+1)} \leq \frac{(10^6)^\nu}{\nu!} \Delta_j^{3A\nu}, \tag{33}$$

provided that $A > 5$.

The choice $\nu = \Delta_j^{10A}$ ensures that altogether the error is of order $\leq e^{-\Delta_j^{4A}}$. Thus we have shown that,

$$\mathbf{1}(Y_j(h) \in [u_j, u_j + \Delta_j^{-1}]) \leq |\mathcal{D}_{\Delta_j,A}(Y_j(h) - u_j) + \mathrm{O}^\star(e^{-\Delta_j^{4A}})|^2 \,(1 + Ce^{-\Delta_j^{A-1}}).$$

Notice that if the left-hand side is equal to one, then $\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)$ is at least $1/2$ in absolute value, therefore we can re-write the above as (30) for some absolute constant $C>0$, establishing the claim.

7.2.2. *Conclusion of the proof of Lemma 7.* We partition the event $L_y(r)\le S_r(h)-m(r)\le U_y(r)$ into the union of events $S_r(h)\in[v,v+\Delta_r^{-1}]$ with

$$v-m(r)\in[L_y(r),U_y(r)]\cap\Delta_r^{-1}\mathbb{Z}.$$

Moreover, for $h\in B_\ell^{(k)}\cap C_\ell^{(k)}$, if we assume that for all $j\in(r,k)$ $Y_j(h)\in[u_j,u_j+\Delta_j^{-1}]$, $S_k(h)\in[w,w+1]$ and $S_r(h)\in[v,v+\Delta_r^{-1}]$, then one must have

$$\begin{aligned}
&v+\sum_{r+1\le i\le k}u_i\le S_r(h)+\sum_{i=r+1}^{k}Y_i(h)\le w+1,\\
&v+\sum_{r+1\le i\le k}u_i\ge S_r(h)-\Delta_r^{-1}+\sum_{i=r+1}^{k}(Y_i(h)-\Delta_i^{-1})\ge w-2(\Delta_k^{-3/4}+\Delta_r^{-3/4}),
\end{aligned}\tag{34}$$

and under the same assumption for $j\in(r,k)$,

$$\begin{aligned}
&v+\sum_{r+1\le i\le j}u_i\le S_r(h)+\sum_{i=r+1}^{j}Y_i(h)\le m(j)+U_y(j),\\
&v+\sum_{r+1\le i\le j}u_i\ge S_r(h)-\Delta_r^{-1}+\sum_{i=r+1}^{j}(Y_i(h)-\Delta_i^{-1})\ge m(j)+L_y(j)-1.
\end{aligned}\tag{35}$$

These are the defining properties of the set $\mathcal{I}_{r,k}(v,w)$ in (27). These observations and the inequality (30) applied successively to every $Y_j(h)$ and to $S_r(h)$ yield

$$\begin{aligned}
&\mathbf{1}\Big(h\in B_\ell^{(k)}\cap C_\ell^{(k)}:S_k(h)\in[w,w+1]\Big)\\
&\le C\sum_{\substack{v\in\Delta_r^{-1}\mathbb{Z}\\ -L_y(r)\le v-m(r)\le U_y(r)\\ \mathbf{u}\in\mathcal{I}_{r,k}(v,w)}}|\mathcal{D}_{\Delta_r,A}(S_r(h)-v)|^2\prod_{j=r+1}^{k}\Big(|\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)|^2\,(1+Ce^{-\Delta_j^{A-1}})\Big).
\end{aligned}$$

Finally, we have $\prod_{j=r+1}^{k}(1+Ce^{-\Delta_j^{A-1}})\le C_0$ for some absolute constant $C_0>0$. This proves the lemma.

7.3. **Comparison with a random model.** Define the random variables

$$\mathcal{S}_k(h)=\sum_{e^{1000}\le\log p\le e^k}\mathrm{Re}\Big(Z_p\,p^{-(\frac12+\mathrm{i}h)}+\tfrac12\,Z_p^2\,p^{-(1+2\mathrm{i}h)}\Big),\qquad \mathcal{Y}_k(h)=\mathcal{S}_k(h)-\mathcal{S}_{k-1}(h),\tag{36}$$

where $(Z_p, p$ prime) are independent and identically distributed copies of a random variable uniformly distributed on the unit circle $|z|=1$. Notice that, since the increments

$\mathcal{Y}_k(h)$ are sums of independent variables, one expects that they are approximately Gaussian with mean zero and variance $\frac{1}{2}$. Moreover, denote

$$\mathcal{G}_k = \sum_{1000 \le \ell \le k} \mathcal{N}_\ell, \tag{37}$$

where the $\mathcal{N}_\ell$'s are centered, independent real Gaussian random variables, with variance $\frac{1}{2}$. Note that $\mathcal{G}$ does not depend on h.

The following lemma shows that one can replace the Dirichlet polynomial Y_j in expectation by the random variables $\mathcal{Y}_j, \mathcal{N}_j$ in the approximate indicators with a small error. This uses Lemma 13 and Lemma 14 in Appendix A.

Lemma 8. *Let $y > 4000$. Let $A > 10$ and $\ell \ge -1$ with $\exp(10^6(n-n_\ell)^{10A} e^{n_{\ell+1}}) \le \exp(\frac{1}{100}e^n)$ be given. Let $k \in (n_\ell, n_{\ell+1}]$. Let $L_y(r) \le v - m(r) \le U_y(r)$. One has for $h \in [-2,2]$,*

$$\mathbb{E}\Big[|\mathcal{D}_{\Delta_r,A}(S_r(h)-v)|^2 \prod_{j=r+1}^{k} |\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)|^2\Big]$$
$$\le (1+Ce^{-ce^n})\mathbb{E}[|\mathcal{D}_{\Delta_r,A}(\mathcal{S}_r(h)-v)|^2] \prod_{j=r+1}^{k} \mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2],$$

with $C, c > 0$ absolute constants. Furthermore, for $w - m(k) \in [L_y(k), U_y(k)]$, we have

$$\sum_{\substack{v \in \Delta_r^{-1}\mathbb{Z} \\ v-m(r)\in[L_y(r),U_y(r)] \\ \mathbf{u}\in\mathcal{I}_{r,k}(v,w)}} \mathbb{E}\Big[|\mathcal{D}_{\Delta_r,A}(\mathcal{S}_r(h)-v)|^2\Big] \prod_{j=r+1}^{k} \mathbb{E}\Big[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2\Big]$$
$$\le C \sum_{\substack{v \in \Delta_r^{-1}\mathbb{Z} \\ v-m(r)\in[L_y(r),U_y(r)] \\ \mathbf{u}\in\mathcal{I}_{r,k}(v,w)}} \mathbb{P}\big(\mathcal{G}_r \in [v, v+\Delta_r^{-1}] \text{ and } \mathcal{N}_j \in [u_j, u_j+\Delta_j^{-1}]\ \forall r < j \le k\big), \tag{38}$$

with $C > 0$ an absolute constant and $\mathcal{I}_{k,\ell}(v,w)$ defined in (34).

Proof. Note that $\mathcal{D}_{\Delta_r,A}(S_r(h)-v)\prod_{j=r+1}^{k}\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)$ is a Dirichlet polynomial of length at most

$$\exp\left(2\sum_{j=r}^{k} e^j \Delta_j^{10A}\right) \le \exp\left(10 e^{n_{\ell+1}} \Delta_{n_\ell}^{10A}\right).$$

The first claim then follows from Lemma 13 and Lemma 14, both in Appendix A. Note that the multiplicative error term from these lemmas is $1 + N/T$ with N the above degree of the Dirichlet polynomial; this error is bounded thanks to the assumption $\exp(10^6(n-n_\ell)^{10A}e^{n_{\ell+1}}) \le \exp(\frac{1}{100}e^n)$.

To prove the second assertion, it will suffice to show that for every $j \in (r,k]$ we have,

$$\mathbb{E}\Big[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2\Big] \le \mathbb{P}(\mathcal{N}_j \in [u_j, u_j+\Delta_j^{-1}]) \cdot (1+\mathrm{O}(\Delta_j^{-A/4})), \tag{39}$$

with an absolute implicit constant in $\mathrm{O}(\cdot)$, and moreover that,

$$\mathbb{E}\Big[|\mathcal{D}_{\Delta_r,A}(\mathcal{S}_r(h)-v)|^2\Big] \le C\mathbb{P}(\mathcal{G}_r \in [v, v+\Delta_r^{-1}]). \tag{40}$$

with $C > 0$ an absolute constant. Then taking the product of the above inequalities over all $j \in (r,k]$, we conclude that

$$\mathbb{E}\Big[|\mathcal{D}_{\Delta_r,A}(\mathcal{S}_r(h)-v)|^2 \prod_{j\in(r,k]} |\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2\Big]$$
$$\le C\mathbb{P}(\mathcal{S}_r \in [v, v+\Delta_r^{-1}]) \prod_{j\in(r,k]} \mathbb{P}(\mathcal{N}_j \in [u_j, u_j+\Delta_j^{-1}]) \cdot (1+\mathrm{O}(\Delta_j^{-A/4})).$$

This gives the claim since $\prod_{j=r}^{k}(1+\mathrm{O}(\Delta_j^{-A/4})) \le C$ with $C > 0$ an absolute constant.

It remains to prove (39) and (40). The first step is to replace $\mathcal{D}_{\Delta_j,A}$ by $G_{\Delta_j,A}$ with a good error using Equations (28) and (32) (with $\mathcal{Y}_j$ instead of Y_j and $\mathcal{S}_r$ instead of S_r). Note that on the event $|\mathcal{Y}_j(h)-u_j| \le \Delta_j^{6A}$, the estimate (33) still holds. Indeed we have, with $\nu = \Delta_j^{10A}$,

$$\frac{(2\pi)^\nu}{\nu!}|\mathcal{Y}_j(h)-u_j|^\nu \int_{\mathbb{R}} |\xi^\nu||\widehat{G}_{\Delta_j,A}(\xi)|\mathrm{d}\xi \le \frac{(10^6)^\nu}{\nu!}\Delta_j^{6A\nu}\cdot\Delta_j^{2A(\nu+1)} \le \frac{(10^6)^\nu}{\nu!}\Delta_j^{9A\nu}\cdot\Delta_j^{2\nu},$$

since $A > 10$. Moreover, since $\nu = \Delta_j^{10A}$, the above is $\le e^{-\Delta_j^{4A}}$. This implies

$$\begin{aligned}
&\mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2 \cdot \mathbf{1}(|\mathcal{Y}_j(h)-u_j| \le \Delta_j^{6A})] \\
&= \mathbb{E}[|G_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j) + \mathrm{O}(e^{-\Delta_j^{4A}})|^2 \cdot \mathbf{1}(|\mathcal{Y}_j(h)-u_j| \le \Delta_j^{6A})] \\
&\le \mathbb{E}[|G_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2] + \mathrm{O}(e^{-\Delta_j^{4A}}),
\end{aligned} \tag{41}$$

since by Lemma 6 we have $G_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j) \in [0,1]$. A quick computation shows that $\mathbb{E}[e^{K\mathcal{Y}_j(h)}] \ll_K 1$ for any given $K > 1$ and all $j \ge 1$ and $h \in [-2,2]$, see Lemma 15 in Appendix A. Therefore the contribution of the event $|\mathcal{Y}_j(h)-u_j| > \Delta_j^{6A}$ can be bounded by Chernoff's inequality:

$$\begin{aligned}
&\mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2 \cdot \mathbf{1}(|\mathcal{Y}_j(h)-u_j| > \Delta_j^{6A})] \\
&\qquad \le \mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^4]^{1/2}\, \mathbb{P}(|\mathcal{Y}_j(h)-u_j| > \Delta_j^{6A})^{1/2} \\
&\qquad \ll \mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^4]^{1/2}\, e^{-\frac{1}{4}\Delta_j^{6A}},
\end{aligned}$$

where we used $|u_j| \le 100\Delta_j^{1/4}$ in the Chernoff's inequality. The fourth moment is easily bounded using an estimate similar to (29):

$$\begin{aligned}
\mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^4] &\le \mathbb{E}\Big[\Big(\sum_{\ell \le \Delta_j^{10A}} \frac{(2\pi)^\ell}{\ell!} 2\Delta_j^{2A(\ell+1)}(|\mathcal{Y}_j(h)|+10^4\Delta_j^2)^\ell\Big)^4\Big] \\
&\ll \Delta_j^{2A}\,\mathbb{E}[\exp(9\pi\Delta_j^{2A}(|\mathcal{Y}_j(h)|+10^4\Delta_j^2))] \ll e^{\Delta_j^{5A}},
\end{aligned}$$

where we used Lemma 15 together with $e^{c|\mathcal{Y}|} \le e^{c\mathcal{Y}} + e^{-c\mathcal{Y}}$. Putting this together we get

$$\mathbb{E}[|\mathcal{D}_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2] \le \mathbb{E}[|G_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2] + \mathrm{O}(e^{-\frac{1}{8}\Delta_j^{6A}}). \tag{42}$$

Furthermore, by Lemma 6, we have

$$\mathbb{E}[|G_{\Delta_j,A}(\mathcal{Y}_j(h)-u_j)|^2] \le \mathbb{P}(\mathcal{Y}_j(h)\in[u_j-\Delta_j^{-A/2},u_j+\Delta_j^{-1}+\Delta_j^{-A/2}])+\mathrm{O}(e^{-\Delta_j^{A-1}}).$$

Since $|u_j|\le 100\min(j,n-j)$ and $j>y/4$, we obtain from Lemma 20 in Appendix B that for all h

$$\mathbb{P}(\mathcal{Y}_j(h)\in[u_j-\Delta_j^{-A/2},u_j+\Delta_j^{-1}+\Delta_j^{-A/2}])=\mathbb{P}(\mathcal{N}_j\in[u_j,u_j+\Delta_j^{-1}])\cdot(1+\mathrm{O}(\Delta_j^{-A/4})). \quad (43)$$

Note that the Gaussian distribution and the restriction on u_j and j are heavily used here to get the error term. This concludes the proof of (39).

The proof of (40) is similar, with the main difference being that we use Lemma 18 in order to show that

$$\mathbb{P}(\mathcal{S}_r\in[v-\Delta_r^{-A/2},v+\Delta_r^{-1}+\Delta_r^{-A/2}])+e^{-\Delta_r^{A-1}}\le C\mathbb{P}(\mathcal{G}_r\in[v,v+\Delta_r^{-1}]).$$

□

7.4. **Proof of Lemma 3.** Let $A=20$. By Lemma 7, we have

$$\begin{aligned}&\mathbf{1}\Big(h\in B_\ell^{(k)}\cap C_\ell^{(k)}\text{ and }S_k(h)\in(w,w+1]\Big)\\ &\qquad\le C\sum_{\substack{v\in\Delta_r^{-1}\mathbb{Z}\\ L_y(r)\le v-m(r)\le U_y(r)\\ \mathbf{u}\in\mathcal{I}_{r,k}(v,w)}}|\mathcal{D}_{\Delta_r,A}(S_r(h)-v)|^2\prod_{j\in(r,k]}|\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)|^2, \quad (44)\end{aligned}$$

$C>0$ an absolute constant. By the properties of $\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)$, we can write the right-hand side of (44) as

$$\sum_{i\in\mathcal{I}}|D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2 \quad (45)$$

a linear combination of squares of Dirichlet polynomials D_i, each of length

$$\le\exp\Big(2\sum_{0\le j\le k}e^j\Delta_j^{200}\Big)\le\exp(100e^k(n-k)^{800}).$$

Therefore multiplying (44) by an arbitrary Dirichlet polynomial $\mathcal{Q}$ of length $N\le\exp(\frac{1}{100}n)$ and applying the discretization in Lemma 27, we conclude that

$$\begin{aligned}&\mathbb{E}\Big[\max_{|h|\le2}|\mathcal{Q}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2\cdot\mathbf{1}(h\in B_\ell^{(k)}\cap C_\ell^{(k)}\text{ and }S_k(h)\in(w,w+1])\Big]\\ &\qquad\ll\Big(\log N+e^k(n-k)^{800}\Big)\sum_{i\in\mathcal{I}}\mathbb{E}\Big[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2\,|D_i(\tfrac12+\mathrm{i}\tau)|^2\Big].\end{aligned} \quad (46)$$

Here we use the fact that the expectations have the same values (up to negligible factors) for the $\mathrm{O}(\log N+e^k(n-k)^{800})$ relevant h's in Lemma 27, and the contribution of the remaining h's associated with very large j in Lemma 27 are bounded similarly to the paragraph after (8). All the following expressions are evaluated at $h=0$. The Dirichlet polynomials D_i are all of length $\le\exp(\frac{1}{100}n)$ and supported on integers n all of whose prime factors are in $\le\exp(e^k)$, while $\mathcal{Q}$ is supported on integers n all of whose prime factors are $>\exp(e^k)$. Therefore, Lemma 14 can be applied and yields

$$\mathbb{E}[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2\,|D_i(\tfrac12+\mathrm{i}\tau)|^2]\le2\mathbb{E}[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2]\,\mathbb{E}[|D_i(\tfrac12+\mathrm{i}\tau)|^2].$$

Finally, by the definition of D_i and $\mathcal{I}$ in (45) and Lemma 8, we have

$$\sum_{i\in\mathcal{I}}\mathbb{E}\Big[|D_i(\tfrac12+\mathrm{i}\tau)|^2\Big]\leq C\sum_{\substack{v\in\Delta_r^{-1}\mathbb{Z}\\ L_y(r)\leq v-m(r)\leq U_y(r)\\ \mathbf{u}\in\mathcal{I}_{r,k}(v,w)}}\mathbb{P}(\mathcal{G}_r\in[v,v+\Delta_r^{-1}]\text{ and }\mathcal{N}_j\in[u_j,u_j+\Delta_j^{-1}]\ \forall r<j\leq k),\tag{47}$$

with $C>0$ an absolute constant.

If for every $r<j\leq k$, we have $\mathcal{N}_j\in[u_j,u_j+\Delta_j^{-1}]$ and moreover $\mathcal{G}_r\in[v,v+\Delta_r^{-1}]$ and (27) holds, then we have

$$\begin{aligned}\forall j\in(r,k]:\mathcal{G}_j\leq m(j)+U_y(j)+1+\sum_{r<i\leq j}\Delta_i^{-1},\\ |\mathcal{G}_k-w|\leq 1+\sum_{r\leq j\leq k}\Delta_j^{-1},\\ \mathcal{G}_r\in[v,v+\Delta_r^{-1}].\end{aligned}\tag{48}$$

As a result after summing over $v\in\Delta_r^{-1}\mathbb{Z}$ we can bound (47) by

$$\leq C\,\mathbb{P}\Big(\mathcal{G}_j\leq m(j)+U_y(j)+2\text{ for all }r\leq j\leq k\text{ and }\mathcal{G}_k\in[w-2,w+2]\Big).$$

Consequently, plugging this into (46), we obtain for $h\in[-2,2]$,

$$\begin{aligned}&\mathbb{E}\Big[\max_{|h|\leq 2}|\mathcal{Q}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2\cdot\mathbf{1}(h\in B_\ell^{(k)}\cap C_\ell^{(k)}\text{ and }S_k(h)\in(w,w+1])\Big]\\ &\ll\Big(\log N+e^k(n-k)^{800}\Big)\mathbb{E}\Big[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2\Big]\\ &\times\mathbb{P}\Big(\mathcal{G}_j\leq m(j)+U_y(j)+2\text{ for all }r\leq j\leq k\text{ and }\mathcal{G}_k(0)\in[w-2,w+2]\Big).\end{aligned}$$

It remains to apply the version of the ballot theorem from Proposition 4 (with y replaced by $y+2$ and adding the bounds with w replaced by $w+i$, $i\in\{-2,-1,0,1\}$) to conclude that

$$\begin{aligned}&\mathbb{E}\Big[\max_{|h|\leq 2}|\mathcal{Q}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2\cdot\mathbf{1}\Big(h\in B_\ell^{(k)}\cap C_\ell^{(k)}:S_k(h)\in(w,w+1]\Big)\Big]\\ &\ll\mathbb{E}[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2]\Big(e^{-k}\log N+(n-k)^{800}\Big)\,y\,(U_y(k)-w+m(k)+2)\,e^{-2(w-m(k))}.\end{aligned}$$

This concludes the proof of Lemma 3.

8. Decoupling and Twisted Fourth Moment

We now prove Lemma 4. We will need the following class of "well-factorable" Dirichlet polynomials.

Definition 1. *Given $\ell\geq 0$ and $k\in[n_\ell,n_{\ell+1}]$, we will say that a Dirichlet polynomial $\mathcal{Q}$ is degree-k well-factorable if it can be written as*

$$\Big(\prod_{0\leq\lambda\leq\ell}\mathcal{Q}_\lambda(s)\Big)\mathcal{Q}_\ell^{(k)}(s),$$

where

$$\mathcal{Q}_\lambda(s) := \sum_{\substack{p|m \Rightarrow p\in(T_{\lambda-1},T_\lambda]\\ \Omega_\lambda(m)\leq 10(n_\lambda-n_{\lambda-1})^{10^4}}} \frac{\gamma(m)}{m^s} \quad \textit{and} \quad \mathcal{Q}_\ell^{(k)}(s) := \sum_{\substack{p|m\Rightarrow p\in(T_\ell,\exp(e^k)]\\ \Omega_\ell(m)\leq 10(n_{\ell+1}-n_\ell)^{10^4}}} \frac{\gamma(m)}{m^s},$$

and γ *are arbitrary coefficients such that* $|\gamma(m)| \ll \exp(\frac{1}{500}e^n)$ *for every* $m\geq 1$.

The proof of Lemma 4 will rely on the following result on the twisted fourth moment. We postpone the proof of this technical lemma to the next subsection.

Lemma 9. *Let* $\ell \geq 0$ *be such that* $\exp(10^6(n-n_\ell)^{10^5}e^{n_{\ell+1}}) \leq \exp(\frac{1}{100}e^n)$. *Let* $k \in [n_\ell, n_{\ell+1}]$. *Let* $\mathcal{Q}$ *be a degree-*k *well-factorable Dirichlet polynomial as in Definition 1. Then, we have*

$$\mathbb{E}\Big[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(\tfrac12+\mathrm{i}\tau)|^4\cdot|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2\Big] \ll e^{4(n-k)}\,\mathbb{E}\Big[|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2\Big].$$

We are now ready to prove Lemma 4.

Proof of Lemma 4. As in the proof of Lemma 3, by Lemma 7, we have for $A=20$

$$\mathbf{1}\Big(h\in B_\ell\cap C_\ell \text{ and } S_{n_\ell}(h)\in(u,u+1]\Big) \leq C \sum_{\substack{v\in\Delta_r^{-1}\mathbb{Z}\\ L_y(r)\leq v-m(r)\leq U_y(r)\\ \mathbf{u}\in\mathcal{I}_{r,n_\ell}(v,w)}} |\mathcal{D}_{\Delta_r,A}(S_r(h)-v)|^2 \prod_{j\in(r,n_\ell]} |\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)|^2, \tag{49}$$

with $C>0$ an absolute constant. By the properties of $\mathcal{D}_{\Delta_j,A}(Y_j(h)-u_j)$, we can write (49) as

$$\sum_{i\in\mathcal{I}} |D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2, \tag{50}$$

a linear combination of squares of Dirichlet polynomials of length

$$\leq \exp\Big(2\sum_{0\leq j\leq k} e^j\Delta_j^{200}\Big) \leq \exp(100e^k(n-k)^{800}).$$

We claim that, for every $i\in\mathcal{I}$, the Dirichlet polynomial

$$\mathcal{Q}_\ell^{(k)}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)$$

is degree-k well-factorable. This follows from the properties of $\mathcal{D}_{\Delta,A}$ listed after Equation (28). More precisely, each D_i has length

$$\leq \exp\Big(2\sum_{0\leq j\leq n_\ell} e^j\Delta_j^{200}\Big) \leq \exp(e^n/100).$$

Moreover, each D_i is supported on the set of integers m such that $p|m \Rightarrow p\leq e^{n_\ell}$, and for every $j\leq n_\ell$, $\Omega_j(m)\leq\Delta_j^{200}$. Furthermore, its coefficients are bounded by $\exp(e^n/500)$.

It then follows from Lemma 9 that

$$\sum_{i\in\mathcal{I}} \mathbb{E}[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(h)|^2\cdot|\mathcal{Q}_\ell^{(k)}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2\cdot|D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2]$$
$$\ll e^{4(n-k)}\sum_{i\in\mathcal{I}}\mathbb{E}[|\mathcal{Q}_\ell^{(k)}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2\cdot|D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2].$$

Moreover, since $\mathcal{Q}_\ell^{(k)}$ is supported on integers n having only prime factors in $(\exp(e^{n_\ell}),\exp(e^k)]$, while D_i is supported on integers n all of whose prime factors are $\le\exp(e^{n_\ell})$, and both Dirichlet polynomials have length $\le\exp(\frac{1}{100}n)$, we conclude from Lemma 14 that

$$\mathbb{E}[|\mathcal{Q}_\ell^{(k)}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2\,|D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2]\ll\mathbb{E}[|\mathcal{Q}_\ell^{(k)}(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2]\,\mathbb{E}[|D_i(\tfrac12+\mathrm{i}\tau+\mathrm{i}h)|^2].$$

Therefore, we obtain that

$$\mathbb{E}\Big[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(h)|^4\cdot|\mathcal{Q}_\ell^{(k)}(h)|^2\cdot\mathbf{1}\Big(h\in B_\ell\cap C_\ell\text{ and }S_{n_\ell}(h)\in[u,u+1]\Big)\Big]$$
$$\ll e^{4(n-k)}\,\mathbb{E}[|\mathcal{Q}_\ell^{(k)}(\tfrac12+\mathrm{i}\tau)|^2]\sum_{i\in\mathcal{I}}\mathbb{E}[|D_i(\tfrac12+\mathrm{i}\tau)|^2].$$

Now, proceeding exactly as in the proof of Lemma 3 starting from Equation (47) one gets

$$\sum_{i\in\mathcal{I}}\mathbb{E}[|D_i(\tfrac12+\mathrm{i}\tau)|^2]\ll y\,(U_y(n_\ell)-u+m(n_\ell)+2)e^{-2(u-m(n_\ell))}e^{-n_\ell}.$$

This concludes the proof. □

8.1. **Proof of Lemma 9.** We first need to introduce some notations. Define for $0\le i\le\ell+1$,

$$\beta_i(m):=\sum_{\substack{m=abc\\ \Omega_i(a),\Omega_i(b)\le(n_i-n_{i-1})^{10^5}\\ \Omega_i(c)\le 10(n_i-n_{i-1})^{10^4}}}\mu(a)\mu(b)\gamma(c),\tag{51}$$

where $\Omega_i(m)$ denotes as before the number of prime factors of m in the range $(T_{i-1},T_i]$. Given m, write $m=m_0\dots m_\ell m_\ell^{(k)}$ where m_j with $0\le j\le\ell$ has prime factors in $(T_{j-1},T_j]$, and $m_\ell^{(k)}$ has prime factors in the interval $(T_\ell,\exp(e^k)]$. Let $\beta(m)$ be defined by

$$\sum_{m\ge1}\frac{\beta(m)}{m^s}=\Big(\prod_{0\le i\le\ell}\mathcal{M}_i^2(s)\mathcal{Q}_i(s)\Big)(\mathcal{M}_\ell^{(k)}(s))^2\mathcal{Q}_\ell^{(k)}(s).$$

Note that,

$$\beta(m)=\prod_{0\le i\le\ell}\beta_i(m_i)\,\beta_{\ell+1}(m_\ell^{(k)}).\tag{52}$$

It will be convenient to redefine $T_{\ell+1}:=\exp(e^k)$ so that the above can be written as

$$\prod_{0\le i\le\ell+1}\beta_i(m_i),$$

with $m_{\ell+1}$ defined as the largest divisor of m all of whose prime factors belong to $(T_\ell, T_{\ell+1}]$ and where $T_{\ell+1} := \exp(e^k)$. Given complex numbers z_1, z_2, z_3, z_4 and $n \in \mathbb{N}$, set $\mathbf{z} := (z_1, z_2, z_3, z_4)$ and consider

$$B_{\mathbf{z}}(n) := B_{(z_1,z_2,z_3,z_4)}(n) = \prod_{p|n} \left(\sum_{j\geq 0} \frac{\sigma_{z_1,z_2}(p^{v_p(n)+j})\sigma_{z_3,z_4}(p^j)}{p^j} \right) \left(\sum_{j\geq 0} \frac{\sigma_{z_1,z_2}(p^j)\sigma_{z_3,z_4}(p^j)}{p^j} \right)^{-1},$$

with $\sigma_{z_1,z_2}(n) = \sum_{n_1 n_2 = n} n_1^{-z_1} n_2^{-z_2}$, and $v_p(n)$, the greatest integer k such that $p^k \mid n$. We are now ready to start the proof.

Proof of Lemma 9. As proved in [20, Section 6], the twisted fourth moment can be bounded by

$$\mathbb{E}[|(\zeta_\tau \mathcal{M}_{-1} \dots \mathcal{M}_\ell \mathcal{M}_\ell^{(k)})(0)|^4 \cdot |\mathcal{Q}(\tfrac{1}{2} + \mathrm{i}\tau)|^2] \ll e^{4n} \max_{\substack{j=1,2,3,4 \\ |z_j| = 3^j/e^n}} |G(z_1, z_2, z_3, z_4)|, \tag{53}$$

where

$$G(z_1, z_2, z_3, z_4) := \sum_{m_1, m_2} \frac{\beta(m_1)\overline{\beta(m_2)}}{[m_1, m_2]} B_{\mathbf{z}}\Big(\frac{m_1}{(m_1, m_2)}\Big) B_{\pi\mathbf{z}}\Big(\frac{m_2}{(m_1, m_2)}\Big), \tag{54}$$

and $\mathbf{z} = (z_1, z_2, z_3, z_4)$, $\pi\mathbf{z} = (z_3, z_4, z_1, z_2)$. Equation (54) relies on the reasoning from [20, Section 6]. It requires a slightly changed version of Proposition 4 in [20, Section 5], requiring a shorter Dirichlet polynomial with $\theta \leq \frac{1}{100}$ but allowing the coefficients to be as large as $T^{1/100}$. This change in the assumptions is possible by appealing to [21] instead of [6] in the argument, see the third remark after Theorem 1 in [21].

Allowing coefficients to be as large as $T^{1/100}$ is necessary because of our assumptions on the coefficients γ. We notice that by the definition of $\mathcal{M}_i$ and $\mathcal{M}_i^{(k)}$ the Dirichlet polynomial $\prod_{0\leq i\leq \ell} (\mathcal{M}_i \, \mathcal{M}_\ell^{(k)})^2$ is of length at most $\exp(2(n_{\ell+1} - n_\ell)^{10^5} e^{n_{\ell+1}})$. The assumptions of the lemma imply $\exp(2(n_{\ell+1} - n_\ell)^{10^5} e^{n_{\ell+1}}) \leq \exp(10^{-4} n)$. Furthermore by the definition of a degree-k well factorable Dirichlet polynomial, $\mathcal{Q}$ is of a length $\leq \exp(\frac{1}{500} n)$. Therefore, the total length of the Dirichlet polynomial $\prod_{0\leq i\leq \ell} (\mathcal{M}_i \, \mathcal{M}_\ell^{(k)})^2 \, \mathcal{Q}$ is $\leq \exp(\frac{1}{100} n)$ as needed.

From Equation (52), the function G can be written as the product

$$\prod_{i\leq \ell+1} \Big(\sum_{p|m_1,m_2 \Rightarrow T_{i-1} < p \leq T_i} \frac{\beta_i(m_1)\overline{\beta_i(m_2)}}{[m_1, m_2]} B_{\mathbf{z}}\Big(\frac{m_1}{(m_1, m_2)}\Big) B_{\pi\mathbf{z}}\Big(\frac{m_2}{(m_1, m_2)}\Big) \Big). \tag{55}$$

Applying the definition (51) with the decompositions $m_1 = a_1b_1c_1$ and $m_2 = a_2b_2c_2$, the inner sum in (55) at a given i can also be written as

$$\sum_{\substack{p|c_1,c_2\Rightarrow p\in(T_{i-1},T_i]\\ \Omega_i(c_1),\Omega_i(c_2)\leq 10(n_i-n_{i-1})^{10^4}}} \gamma(c_1)\overline{\gamma(c_2)}\times$$
$$\sum_{\substack{p|a_1,a_2\Rightarrow p\in(T_{i-1},T_i]\\ p|b_1,b_2\Rightarrow p\in(T_{i-1},T_i]\\ \Omega_i(a_1),\Omega_i(a_2)\leq(n_i-n_{i-1})^{10^5}\\ \Omega_i(b_1),\Omega_i(b_2)\leq(n_i-n_{i-1})^{10^5}}} \frac{\mu(a_1)\mu(a_2)\mu(b_1)\mu(b_2)}{[a_1b_1c_1,a_2b_2c_2]} B_{\mathbf{z}}\Big(\frac{a_1b_1c_1}{(a_1b_1c_1,a_2b_2c_2)}\Big) B_{\pi\mathbf{z}}\Big(\frac{a_2b_2c_2}{(a_1b_1c_1,a_2b_2c_2)}\Big). \tag{56}$$

Given an interval I, and integers $c_1, c_2 \geq 1$, we define the quantity

$$\mathfrak{S}_I(c_1,c_2) := \sum_{p|u,v\Rightarrow p\in I} \frac{f(u)f(v)}{[uc_1,vc_2]} B_{\mathbf{z}}\Big(\frac{uc_1}{(uc_1,vc_2)}\Big) B_{\pi\mathbf{z}}\Big(\frac{vc_2}{(uc_1,vc_2)}\Big), \tag{57}$$

where f is the multiplicative function such that $f(p) = -2$, $f(p^2) = 1$ and $f(p^\alpha) = 0$ for $\alpha \geq 3$. The rest of the argument relies on Lemma 10 and Lemma 11. Lemma 10 shows that the restriction on the number of factors for the a and b's can be dropped with a small error. Lemma 11 evaluates the sum of (57) without these restrictions.

Lemma 10. *For $0 \leq i \leq \ell+1$ the equation (56) is equal to*

$$\sum_{\substack{p|c_1,c_2\Rightarrow p\in(T_{i-1},T_i]\\ \Omega_i(c_1),\Omega_i(c_2)\leq 10(n_i-n_{i-1})^{10^4}}} \gamma(c_1)\overline{\gamma(c_2)}\mathfrak{S}_{(T_{i-1},T_i]}(c_1,c_2) + \mathrm{O}\Big(e^{-100(n_i-n_{i-1})} \sum_{p|c\Rightarrow p\in(T_{i-1},T_i]} \frac{|\gamma(c)|^2}{c}\Big),$$

with an absolute implicit constant in $\mathrm{O}(\cdot)$.

We now define

$$\mathcal{S}_I = \sum_{\substack{p|c_1,c_2\Rightarrow p\in I\\ \Omega_i(c_1),\Omega_i(c_2)\leq 10(n_i-n_{i-1})^{10^4}}} |\gamma(c_1)|\cdot|\gamma(c_2)|\cdot|\mathfrak{S}_I(c_1,c_2)|.$$

Lemma 11. *We have, for $0 \leq i \leq \ell+1$ and every interval $I \subset [T_{i-1}, T_i]$,*

$$\mathcal{S}_I \leq \exp\Big(e^{6000}(n_i-n_{i-1})^{4\cdot 10^4} e^{n_i-n}\Big) \exp\Big(-\sum_{p\in I}\frac{4}{p}\Big) \sum_{p|c\Rightarrow p\in I} \frac{|\gamma(c)|^2}{c}.$$

The proof of these lemmas is deferred to the next subsections. We first conclude the proof of Lemma 9.

It follows from (53), (55) and Lemma 10 that

$$\begin{aligned}&\mathbb{E}[|(\zeta_\tau\mathcal{M}_{-1}\dots\mathcal{M}_\ell\mathcal{M}_\ell^{(k)})(0)|^4\cdot|\mathcal{Q}(\tfrac12+\mathrm{i}\tau)|^2]\\ &\quad\ll e^{4n}\prod_{i=0}^{\ell+1}\Big(\mathcal{S}_{(T_{i-1},T_i]} + Ce^{-100(n_i-n_{i-1})}\sum_{p|c\Rightarrow p\in(T_i,T_{i+1}]}\frac{|\gamma(c)|^2}{c}\Big),\end{aligned} \tag{58}$$

with $C > 0$ an absolute constant. Combining (58) and Lemma 11, we conclude that (with $C > 0$ an absolute constant),

$$\begin{aligned}
&\mathbb{E}[|(\zeta_\tau \mathcal{M}_{-1}\dots\mathcal{M}_\ell \mathcal{M}_\ell^{(k)})(0)|^4 \cdot |\mathcal{Q}(\tfrac{1}{2}+\mathrm{i}\tau)|^2] \\
&\quad \ll e^{4n}\prod_{i=0}^{\ell+1}\Big(\exp(C(n_i-n_{i-1})^{10^5}e^{n_i-n}\,(1+Ce^{-(n_i-n_{i-1})})\Big) \\
&\qquad\qquad \times \prod_{i=0}^{\ell+1}\Big(\exp\Big(-\sum_{p\in(T_{i-1},T_i]}\frac{4}{p}\Big)\sum_{p|c\Rightarrow p\in(T_{i-1},T_i]}\frac{|\gamma(c)|^2}{c}\Big) \\
&\quad \ll e^{4(n-k)}\sum_{c\geq 1}\frac{|\gamma(c)|^2}{c} \ll e^{4(n-k)}\,\mathbb{E}[|\mathcal{Q}(\tfrac{1}{2}+\mathrm{i}\tau)|^2].
\end{aligned} \tag{59}$$

(Recall that $T_{\ell+1} = \exp(e^k)$.) In the last line, we used that

$$\prod_{i=0}^{\ell+1}\Big(\sum_{p|c\Rightarrow p\in(T_{-1},T_i]}\frac{|\gamma(c)|^2}{c}\Big) = \sum_{c\geq 1}\frac{|\gamma(c)|^2}{c},$$

which is a consequence of the assumption that the Dirichlet polynomial $\mathcal{Q}$ is degree-k well-factorable. We also used Lemma 13. □

8.2. **Proof of Lemma 10.** We bound the contribution from a_j's or b_j's, $j = 1, 2$, such that $\Omega_i(a_j) > (n_i - n_{i-1})^{10^5}$ or $\Omega_i(b_j) > (n_i - n_{i-1})^{10^5}$ for $j = 1$ or 2, using Chernoff's bound (also known in this setting as Rankin's trick). We write down the argument only for a_1 as the other cases are dealt with in an identical fashion. Note that since a_1 is square-free, we have $\Omega_i(a_1) = \omega_i(a_1)$, where ω_i denotes the number of distinct prime factors in $(T_{i-1}, T_i]$ counted without multiplicity. For any $\rho \in (0, 2000)$, the contribution of such a_1's is bounded by,

$$\begin{aligned}
&e^{-\rho(n_i-n_{i-1})^{10^5}}\sum_{\substack{p|c_1,c_2\Rightarrow p\in(T_{i-1},T_i]\\ \Omega_i(c_1),\Omega_i(c_2)\leq 10(n_i-n_{i-1})^{10^4}}}|\gamma(c_1)\gamma(c_2)|\sum_{\substack{p|a_1,b_1\Rightarrow p\in(T_{i-1},T_i]\\ p|a_2,b_2\Rightarrow p\in(T_{i-1},T_i]\\ a_1,b_1,a_2,b_2\leq T^{1/100}}} e^{\rho\omega_i(a_1)} \\
&\qquad \times \frac{\mu^2(a_1)\mu^2(a_2)\mu^2(b_1)\mu^2(b_2)}{[a_1b_1c_1, a_2b_2c_2]}\Big|B_{\mathbf{z}}\Big(\frac{a_1b_1c_1}{(a_1b_1c_1, a_2b_2c_2)}\Big)B_{\pi\mathbf{z}}\Big(\frac{a_2b_2c_2}{(a_1b_1c_2, a_2b_2c_2)}\Big)\Big|.
\end{aligned} \tag{60}$$

We now claim that $|B_{\mathbf{z}}(m)| \ll d_3(m)$ provided that $\mathbf{z} = (z_1, z_2, z_3, z_4)$ are such that $|z_j| = 3^j/\log T$ for all $1 \leq j \leq 4$ and $m \leq T$. Here $d_k(m)$ denotes the kth divisor function: $d_k(n) = \sum_{n=m_1\dots m_k} 1$. To prove this, from Lemma 24, for every $p^\alpha \leq T$ and integer $\alpha \geq 1$, we have

$$|B_{\mathbf{z}}(p^\alpha)| \leq d_2(p^\alpha)\Big(1 + \mathrm{O}\Big(\frac{\alpha\log p}{\log T} + \frac{1}{p}\Big)\Big).$$

Therefore, by taking the product over all $p|m$, we obtain

$$|B_{\mathbf{z}}(m)| \ll d_2(m)\prod_{p|m}\Big(1+\mathrm{O}\Big(\frac{\alpha\log p}{\log T}\Big)\Big)\prod_{p|m}\Big(1+\mathrm{O}\Big(\frac{1}{p}\Big)\Big) \ll d_2(m)\Big(\frac{3}{2}\Big)^{\omega(m)} \ll d_3(m),$$

since $\prod_{p|m}(1+\mathrm{O}(\alpha\log p/\log T)) \ll \exp(\log m/\log T) \ll 1$ for $m \leq T$ and $\prod_{p|m}(1+\mathrm{O}(1/p)) \ll (3/2)^{\omega(m)}$.

Furthermore, note that the factors μ^2 in (60) ensure that only the square-free a's and b's are counted. In particular, we have $v_p(a_jb_j) \leq 2$ for every j. Grouping a_1b_1 (resp. a_2b_2) as a single variable with $k_1 := v_p(a_1b_1)$ (resp. $k_2 := v_p(a_2b_2)$), we find that the sum over a_1, a_2, b_1, b_2 (for fixed c_1 and c_2) in (60) is bounded by the Euler product

$$C \prod_{p\in(T_{i-1},T_i]} \Big(\sum_{0\leq k_1,k_2\leq 2} \frac{e^{\rho\omega_i(p^{k_1})} d_2(p^{k_1}) d_3(p^{k_1+v_p(c_1)}) d_2(p^{k_2}) d_3(p^{k_2+v_p(c_2)})}{p^{\max(k_1+v_p(c_1),k_2+v_p(c_2))}} \Big), \tag{61}$$

where $d_2(p^{k_1})$ accounts for the number of choices of (a_1,b_1) giving the single variable a_1b_1, and the same for $d_2(p^{k_2})$. Note that we have not used the gcd factors and simply bounded $d_3(m/v) \leq d_3(m)$ for $v \mid m$.

The sum over the powers k_1 and k_2 in (61) can be bounded further by using the inequalities $d_3(p^{k_2+\alpha_2}) \leq d_3(p^{k_2}) d_3(p^{\alpha_2})$ and $d_2(p^{k_1}) = k_1+1$. This shows that the factor in (61) is

$$\begin{aligned} &\leq \frac{d_3(p^{v_p(c_1)}) d_3(p^{v_p(c_2)})}{p^{\max(v_p(c_1),v_p(c_2))}} \sum_{0\leq k_1,k_2\leq 2} \frac{e^{\rho\omega_i(p^{k_1})} d_2(p^{k_1}) d_3(p^{k_1}) d_2(p^{k_2}) d_3(p^{k_2})}{p^{\max(k_1+v_p(c_1),k_2+v_p(c_2))-\max(v_p(c_1),v_p(c_2))}} \\ &\leq \frac{d_3(p^{v_p(c_1)}) d_3(p^{v_p(c_2)})}{p^{\max(v_p(c_1),v_p(c_2))}} \cdot \begin{cases} 1+100e^{\rho}/p & \text{if } v_p(c_1) = v_p(c_2), \\ 100e^{\rho} & \text{if } v_p(c_1) \neq v_p(c_2). \end{cases} \end{aligned} \tag{62}$$

We notice that the contribution to (61) of primes $p \in (T_{i-1}, T_i]$ for which $v_p(c_1) = v_p(c_2) = 0$ is bounded by

$$\ll \Big(\frac{\log T_i}{\log T_{i-1}}\Big)^{100e^{\rho}} = e^{100e^{\rho}(n_i-n_{i-1})}.$$

As a result of the last two equations, the Euler product in (61) is bounded by

$$\ll d_3(c_1) d_3(c_2) \frac{f(c_1,c_2)}{[c_1,c_2]} \Big(\frac{\log T_i}{\log T_{i-1}}\Big)^{100e^{\rho}}, \tag{63}$$

where $f(c_1,c_2)$ is a multiplicative function of two variables such that $f(p^{\alpha},p^{\alpha}) = 1 + 100e^{\rho}/p$ for all $\alpha \geq 1$ and $f(p^{\alpha},p^{\beta}) = 100e^{\rho}$ for $\alpha \geq 0$ and $\beta \geq 0$ with $\alpha \neq \beta$. Going back to Equation (60), it remains to estimate the sum over c_1 and c_2 using (63):

$$\begin{aligned} &\sum_{\substack{p|c_1,c_2 \Rightarrow p\in(T_{i-1},T_i] \\ \Omega_i(c_1),\Omega_i(c_2)\leq 10(n_i-n_{i-1})^{10^4}}} \frac{|\gamma(c_1)\gamma(c_2)| f(c_1,c_2)}{[c_1,c_2]} d_3(c_1) d_3(c_2) \\ &\qquad \leq e^{1000(n_i-n_{i-1})^{10^4}} \sum_{\substack{p|c_1,c_2 \Rightarrow p\in(T_{i-1},T_i] \\ \Omega_i(c_1),\Omega_i(c_2)\leq 10(n_i-n_{i-1})^{10^4}}} \frac{|\gamma(c_1)\gamma(c_2)| f(c_1,c_2)}{[c_1,c_2]}, \end{aligned}$$

where the restriction on the number of prime factors of c_1 and c_2 is used to trivially bound d_3. Furthermore, using the inequality $|\gamma(c_1)\gamma(c_2)| \leq \frac{1}{2}|\gamma(c_1)|^2 + \frac{1}{2}|\gamma(c_2)|^2$ and dropping the restriction on $\Omega_i(c_1)$, the sum over c_1, c_2 above is less than

$$\sum_{\substack{p|c_2 \Rightarrow p\in(T_{i-1},T_i] \\ \Omega_i(c_2)\le 10(n_i-n_{i-1})^{10^4}}} |\gamma(c_2)|^2 \sum_{p|c_1\Rightarrow p\in(T_{i-1},T_i]} \frac{f(c_1,c_2)}{[c_1,c_2]}. \tag{64}$$

We note now that by the definition of f the sum over c_1 with c_2 fixed can be bounded by an Euler product

$$\sum_{p|c_1\Rightarrow p\in(T_{i-1},T_i]} \frac{f(c_1,c_2)}{[c_1,c_2]} = \prod_{p\in(T_{i-1},T_i]} \Big(\sum_{k\ge 0} \frac{f(p^k, p^{v_p(c_2)})}{p^{\max(k,v_p(c_2))}} \Big)$$
$$= \prod_{p\in(T_{i-1},T_i]} \Big(\sum_{0\le k<v_p(c_2)} \frac{f(p^k,p^{v_p(c_2)})}{p^{v_p(c_2)}} + \frac{f(p^{v_p(c_2)},p^{v_p(c_2)})}{p^{v_p(c_2)}} + \sum_{k>v_p(c_2)} \frac{f(p^k,p^{v_p(c_2)})}{p^k} \Big),$$

and using the definition of f we conclude that,

$$\sum_{p|c_1\Rightarrow p\in(T_{i-1},T_i]} \frac{f(c_1,c_2)}{[c_1,c_2]} \le \prod_{\substack{p\in(T_{i-1},T_i] \\ v_p(c_2)=0}} \Big(1+\frac{200e^{\rho}}{p}\Big) \prod_{\substack{p\in(T_{i-1},T_i] \\ v_p(c_2)>0}} \Big(\frac{200e^{\rho}\, v_p(c_2)}{p^{v_p(c_2)}}\Big).$$

Since c_2 has at most $10(n_i-n_{i-1})^{10^4}$ prime factors counted with multiplicity, this is

$$\le \frac{1}{c_2} \exp\Big(200e^{\rho}(n_i-n_{i-1}) + 10^4\rho(n_i-n_{i-1})^{10^4+1}\Big). \tag{65}$$

As a result, putting together equations (65), (64) and (63), we see that (60) is bounded by

$$\ll e^{-\rho(n_i-n_{i-1})^{10^5}+300e^{\rho}(n_i-n_{i-1})+10^4\rho(n_i-n_{i-1})^{10^4+1}} \sum_{p|c_2\Rightarrow p\in(T_{i-1},T_i]} \frac{|\gamma(c_2)|^2}{c_2}.$$

Here, the condition on $\Omega_i(c_2)$ was dropped. Choosing $\rho=1000$ we see that this is

$$\le e^{-100(n_i-n_{i-1})^{10^5}} \sum_{p|c\Rightarrow p\in(T_{i-1},T_i]} \frac{|\gamma(c)|^2}{c},$$

as needed.

8.3. **Proof of Lemma 11.** A crucial step in the proof of Lemma 11 will be the following estimate for $\mathfrak{S}_I(c_1,c_2)$ defined in Equation (57).

Lemma 12. *Let* $\ell\ge 0$ *be such that* $\exp(10^6(n_{\ell+1}-n_\ell)^{10^5}e^{n_{\ell+1}}) \le \exp(\frac{1}{100}e^n)$. *Let* $I\subset[\exp(e^{1000}),\exp(e^{n_{\ell+1}})]$ *be an interval. Let* $z_1,\dots,z_4$ *be complex number with* $|z_j|=3^j/e^n$ *for* $j=1,2,3,4$. *Let* $\mathbf{z}=(z_1,z_2,z_3,z_4)$ *and* $\pi\mathbf{z}=(z_3,z_4,z_1,z_2)$. *Given integers* $c_1,c_2\ge 1$ *with at most* $10(n_{\ell+1}-n_\ell)^{10^4}$ *prime factors, consider* $\mathfrak{S}_I(c_1,c_2)$ *as in Equation (57).*

Write $c_1=rc_1'$ *and* $c_2=rc_2'$ *with* $r:=(c_1,c_2)$. *Then, we have*

$$|\mathfrak{S}_I(c_1,c_2)| \le \prod_{p\in I}\Big(1-\frac{4}{p}+e^{4000}\frac{\log p}{pe^n}+\frac{e^{4000}}{p^2}\Big)\frac{h(c_1')h(c_2')}{rc_1'c_2'},$$

where h is a multiplicative function such that, for all prime $p \geq 2$ and integer $\alpha \geq 1$,

$$h(p^\alpha) = \frac{e^{5000}\alpha^2 \log p}{e^n}.$$

The lemma is proved in the next subsection. Assuming it, we prove Lemma 11. We start by writing $c_1 = rc_1'$, $c_2 = rc_2'$ and then we use the inequality $|\gamma(rc_1')\gamma(rc_2')| \leq \frac{1}{2}(|\gamma(rc_1')|^2 + |\gamma(rc_2')|^2)$. Lemma 12 then reduces the proof to evaluating

$$\sum_{\substack{p|r,c_1' \Rightarrow p\in I \\ \Omega_i(rc_1')\leq 10(n_i-n_{i-1})^{10^4}}} \frac{|\gamma(rc_1')|^2 h(c_1')}{rc_1'} \sum_{p|c_2'\Rightarrow p\in(T_{i-1},T_i]} \frac{h(c_2')}{c_2'}.$$

The definition of $h(p^\alpha)$ implies

$$\sum_{p|c_2'\Rightarrow p\in I} \frac{h(c_2')}{c_2'} = \prod_{p\in I}\Big(1 + \sum_{\alpha\geq 1} \frac{e^{5000}\alpha^2\log p}{p^\alpha e^n}\Big) \leq \exp(e^{6000}e^{n_i-n}),$$

using the fact that $I \subset [T_{i-1}, T_i]$ and the bound $\sum_{p\leq T_i} \frac{\log p}{p} \leq 2e^{n_i}$.

Finally, it remains to bound

$$\sum_{\substack{p|r\Rightarrow p\in I \\ \Omega_i(rc_1')\leq 10(n_i-n_{i-1})^{10^4}}} \frac{|\gamma(rc_1')|^2 h(c_1')}{rc_1'} = \sum_{\substack{p|m\Rightarrow p\in I \\ \Omega_i(m)\leq 10(n_i-n_{i-1})^{10^4}}} \frac{|\gamma(m)|^2}{m}\, g(m), \tag{66}$$

where g is a multiplicative function such that,

$$g(p^\alpha) := \sum_{rm=p^\alpha} h(m) \leq 1 + \frac{e^{5000}\alpha^3\log p}{e^n},$$

for every prime p and integer $\alpha \geq 1$. Since m has at most $10(n_i - n_{i-1})^{10^4}$ prime factors, all of which are less than $\exp(e^{n_i})$, we have

$$\begin{aligned} g(m) \leq \exp\Big(e^{5000}\sum_{p|m}\frac{v_p(m)^3\log p}{e^n}\Big) &\leq \exp\Big(e^{5000}\, 10000(n_i - n_{i-1})^{4\cdot 10^4}\, e^{n_i-n}\Big) \\ &\leq \exp\Big(e^{6000}(n_i - n_{i-1})^{4\cdot 10^4} e^{n_i-n}\Big). \end{aligned}$$

Therefore we obtain a final bound for (66)

$$\leq \exp\Big(e^{6000}(n_i - n_{i-1})^{4\cdot 10^4} e^{n_i-n}\Big) \sum_{p|c\Rightarrow p\in I} \frac{|\gamma(c)|^2}{c},$$

thereby concluding the proof of the lemma.

8.4. **Proof of Lemma 12.** Using multiplicativity, we can write $\mathfrak{S}_I(c_1, c_2)$ given in Equation (57) as a product

$$\mathfrak{S}_I(c_1, c_2) = \prod_{p\in I} \mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1, c_2, p),$$

where $\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1,c_2,p)$ is defined as

$$\sum_{0\le k_1,k_2\le 2}\frac{f(p^{k_1})f(p^{k_2})}{p^{\max(k_1+v_p(c_1),k_2+v_p(c_2))}}B_{\mathbf{z}}\Big(\frac{p^{k_1+v_p(c_1)}}{p^{\min(k_1+v_p(c_1),k_2+v_p(c_2))}}\Big)B_{\pi\mathbf{z}}\Big(\frac{p^{k_2+v_p(c_2)}}{p^{\min(k_1+v_p(c_1),k_2+v_p(c_2))}}\Big). \tag{67}$$

Writing $c_1=rc_1'$ and $c_2=rc_2'$ with $r=(c_1,c_2)$, we begin by noticing that

$$\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1,c_2,p)=\frac{1}{p^{v_p(r)}}\,\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1',c_2',p).$$

It therefore remains to understand $\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1',c_2',p)$. Since $(c_1',c_2')=1$ there are only two possibilities to consider: either $(p,c_1'c_2')=1$ or p divides only one of c_1', c_2'.

On one hand, if $(p,c_1'c_2')=1$, Lemma 24 in Appendix A yields

$$\begin{aligned}\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1',c_2',p)&=\sum_{0\le k_1,k_2\le 2}\frac{f(p^{k_1})f(p^{k_2})}{p^{\max(k_1,k_2)}}B_{\mathbf{z}}\Big(\frac{p^{k_1}}{p^{\min(k_1,k_2)}}\Big)B_{\pi\mathbf{z}}\Big(\frac{p^{k_2}}{p^{\min(k_1,k_2)}}\Big)\\&=\sum_{0\le k_1,k_2\le 2}\frac{f(p^{k_1})f(p^{k_2})}{p^{\max(k_1,k_2)}}B_{\mathbf{0}}\Big(\frac{p^{k_1}}{p^{\min(k_1,k_2)}}\Big)B_{\mathbf{0}}\Big(\frac{p^{k_2}}{p^{\min(k_1,k_2)}}\Big)+\mathrm{O}^\star\Big(e^{4000}\,\frac{\log p}{pe^n}\Big)\\&=1-\frac{4}{p}+\mathrm{O}^\star\Big(e^{4000}\,\frac{\log p}{pe^n}+\frac{e^{4000}}{p^2}\Big),\end{aligned} \tag{68}$$

where $\mathrm{O}^\star(\cdot)$ means that the implicit constant is ≤ 1. Note that we have used the simple bound $|f|\le 2$ and that either $k_1\ge 1$ or $k_2\ge 1$ if the above summands differ.

On the other hand, if $p|c_1'c_2'$ we can assume that $p|c_1'$ and $p\nmid c_2'$ as the case of $p|c_2'$ and $p\nmid c_1'$ is identical. Then we have $v_p(c_2')=0$ and hence,

$$\begin{aligned}\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1',c_2',p)=&\sum_{1\le k_1\le 2}\frac{f(p^{k_1})}{p^{k_1+v_p(c_1')}}\sum_{0\le k_2\le 2}f(p^{k_2})B_{\mathbf{z}}(p^{k_1+v_p(c_1')-k_2})+\frac{1}{p^{v_p(c_1')}}\sum_{0\le k_2\le 1}f(p^{k_2})B_{\mathbf{z}}(p^{v_p(c_1')-k_2})\\&+\begin{cases}\frac{1}{p^2}B_{\pi\mathbf{z}}(p) & \text{if } v_p(c_1')=1,\\ \frac{1}{p^{v_p(c_1')}}B_{\mathbf{z}}(p^{v_p(c_1')-2}) & \text{if } v_p(c_1')\ge 2.\end{cases}\end{aligned} \tag{69}$$

By Lemma 24, we have $|B_{\mathbf{z}}(p^j)-B_{\mathbf{0}}(p^j)|\le 100\,e^{3000}\,j^2\,(\log p)e^{-n}$ for all $0\le j\le v_p(c_1')+2$. Note that this uses that $p^{v_p(c_1')+2}\le\exp(100(n_{\ell+1}-n_\ell)^{10^4}e^{n_{\ell+1}})\le\exp(\frac{1}{100}e^n)$. Therefore, using this inequality and (69), we get

$$\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1',c_2',p)=\mathcal{P}_{\mathbf{0},\mathbf{0}}(c_1',c_2',p)+\mathrm{O}^\star\Big(e^{4000}\,\frac{v_p(c_1')^2\cdot\log p}{p^{v_p(c_1')}e^n}\Big), \tag{70}$$

where $\mathrm{O}^\star(\cdot)$ is a $\mathrm{O}(\cdot)$ with implicit constant ≤ 1. We claim that $\mathcal{P}_{\mathbf{0},\mathbf{0}}(c_1',c_2',p)=0$. For $v_p(c_1')\ge 2$, this follows from (69) and the identity

$$\sum_{0\le k\le 2}f(p^k)B_{\mathbf{0}}(p^{\ell-k})=0\quad\text{for } \ell\ge 2. \tag{71}$$

The above identity follows from Lemma 24 and the identities $\ell - 2(\ell-1) + (\ell-2) = 0$ and $1-2+1=0$. For $v_p(c_1') = 1$ by (69) and (71) it suffices to check that,

$$\frac{1}{p}\left(B_{\mathbf{0}}(p) - 2B_{\mathbf{0}}(1)\right) + \frac{1}{p^2}\, B_{\mathbf{0}}(p) = 0.$$

In particular upon factoring it is enough to check that $B_{\mathbf{0}}(p)\,(1+1/p) - 2 = 0$. This follows from $B_{\mathbf{0}}(p) = (1-p^{-2})^{-1}\,(2-2/p)$.

We conclude from (70) and $\mathcal{P}_{\mathbf{0},\mathbf{0}}(c_1', c_2', p) = 0$ that,

$$|\mathcal{P}_{\mathbf{z},\pi\mathbf{z}}(c_1', c_2', p)| \leq e^{4000}\, \frac{1}{p^{v_p(c_1')}}\, \frac{v_p(c_1')^2 \log p}{e^n}. \tag{72}$$

Equations (68) and (72) can then be used to get the bound

$$|\mathfrak{S}_I(c_1,c_2)| \leq \frac{1}{r c_1' c_2'} \prod_{\substack{p\in I \\ (p, c_1'c_2')=1}} \Big(1 - \frac{4}{p} + e^{4000}\, \frac{\log p}{p e^n} + \frac{e^{4000}}{p^2}\Big) \prod_{i=1}^{2} \prod_{\substack{p\in I \\ p | c_i'}} \Big(e^{4000}\, \frac{v_p(c_i')^2 \log p}{e^n}\Big)$$

The restriction $(p, c_1'c_2') = 1$ can be removed by multiplying and dividing by $(1-4/p)$ for primes p with $p | c_1'c_2'$. As a result, the above is bounded by

$$|\mathfrak{S}_I(c_1,c_2)| \leq \frac{1}{r c_1' c_2'} \prod_{p\in I} \Big(1 - \frac{4}{p} + e^{4000}\, \frac{\log p}{p e^n} + \frac{e^{4000}}{p^2}\Big) \times \prod_{i=1}^{2} \prod_{\substack{p\in I \\ p | c_i'}} \Big(2\cdot e^{4000}\, \frac{v_p(c_1')^2 \log p}{e^n}\Big).$$

This is the claimed bound.

Appendix A. Estimates on Sums over Primes

A.1 **Moments of Dirichlet Polynomials.** Let $(Z_p, p$ prime) a sequence of independent and identically distributed random variables, uniformly distributed on the unit circle $|z| = 1$. For an integer n with prime factorization $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ with $p_1, \dots, p_k$ all distinct, consider

$$Z_n := \prod_{i=1}^{k} Z_{p_i}^{\alpha_i}.$$

Then we have $\mathbb{E}[Z_n \overline{Z}_m] = \mathbf{1}_{n=m}$, and therefore, for an arbitrary sequence $a(n)$ of complex numbers, the following holds

$$\sum_{n\leq N} |a(n)|^2 = \mathbb{E}\Big[\Big|\sum_{n\leq N} a(n) Z_n\Big|^2\Big].$$

The next lemma shows that the mean value of Dirichlet polynomial is close to the one of the above random model. It follows directly from [25, Corollary 3].

Lemma 13 (Mean-value theorem for Dirichlet polynomials)**.** *We have,*

$$\mathbb{E}\Big[\Big|\sum_{n\leq N} a(n) n^{\mathrm{i}\tau}\Big|^2\Big] = \Big(1 + \mathrm{O}\Big(\frac{N}{T}\Big)\Big) \sum_{n\leq N} |a(n)|^2 = \Big(1 + \mathrm{O}\Big(\frac{N}{T}\Big)\Big) \mathbb{E}\Big[\Big|\sum_{n\leq N} a(n) Z_n\Big|^2\Big].$$

The above implies that Dirichlet polynomials that are supported on integers with prime factors in different ranges behave independently to some extent.

Lemma 14 (Splitting Lemma). *Let*

$$A(s) := \sum_{\substack{n\leq N\\ p|n\Rightarrow p\leq w}} \frac{a(n)}{n^s} \text{ and } B(s) := \sum_{\substack{n\leq N\\ p|n\Rightarrow p> w}} \frac{b(n)}{n^s}$$

be two Dirichlet polynomials with $N \leq T^{1/4}$. Then, we have

$$\mathbb{E}[|A(\tfrac12+\mathrm{i}\tau)|^2\,|B(\tfrac12+\mathrm{i}\tau)|^2] = (1+\mathrm{O}(T^{-1/2}))\mathbb{E}[|A(\tfrac12+\mathrm{i}\tau)|^2]\,\mathbb{E}[|B(\tfrac12+\mathrm{i}\tau)|^2].$$

Proof. Note that AB is a Dirichlet polynomial with length at most $T^{1/2}$, so Lemma 13 gives

$$\mathbb{E}[|A(\tfrac12+\mathrm{i}\tau)|^2\,|B(\tfrac12+\mathrm{i}\tau)|^2] = (1+\mathrm{O}(T^{-1/2}))\sum_n \frac{1}{n}\Big|\sum_{\substack{n=m_1m_2\\ m_1,m_2\leq T^{1/4}\\ p|m_1\Rightarrow p\leq w\\ p|m_2\Rightarrow p>w}} a(m_1)b(m_2)\Big|^2.$$

Expanding the square we find that the sum over n is equal to

$$\sum_{\substack{m_1m_2=m_1'm_2'\\ m_1,m_2,m_1',m_2'\leq T^{1/4}\\ p|m_1,m_1'\Rightarrow p\leq w\\ p|m_2,m_2'\Rightarrow p>w}} \frac{a(m_1)\overline{a(m_1')}b(m_2)\overline{b(m_2')}}{\sqrt{m_1m_1'm_2m_2'}} = \sum_{\substack{m_1\leq T^{1/4}\\ p|m_1\Rightarrow p\leq w}}\frac{|a(m_1)|^2}{m_1}\sum_{\substack{m_2\leq T^{1/4}\\ p|m_2\Rightarrow p>w}}\frac{|b(m_2)|^2}{m_2},$$

where we have used the condition on the prime factors of m_1, m_1', m_2, m_2', which implies that $m_1 = m_1'$ and $m_2 = m_2'$. By Lemma 13, the above right-hand side is equal to $(1+\mathrm{O}(T^{-3/4}))\mathbb{E}[|A(\frac12+\mathrm{i}\tau)|^2]\,\mathbb{E}[|B(\frac12+\mathrm{i}\tau)|^2]$, which concludes the proof. □

Moments of the Dirichlet polynomials S_k as defined in Equation (4) are very close to Gaussian ones provided that the moments are not too large compared to their length. This is the content of Lemma 16 below. For the proof of this, it is useful to consider the random variables

$$X_p(h) = \mathrm{Re}\Big(Z_p\,p^{-\frac12-\mathrm{i}h} + \tfrac12\,Z_p^2\,p^{-1-2\mathrm{i}h}\Big),\ p \text{ prime},\ h\in[-2,2], \tag{73}$$

where we remind that the variables Z_p are uniform on the unit circle. We also use a precise form of Mertens' theorem:

$$\sum_{a<p\leq b}\frac{1}{p} = \log_2 b - \log_2 a + \mathrm{O}(e^{-\kappa\sqrt{\log a}}), \tag{74}$$

for some $\kappa > 0$. Such estimates are given in [30, Corollary 2], for self-containedness we give a short proof below based on the following quantitative prime number theorem [26]: There exists $c>0$ such that uniformly in $x \geq 2$,

$$\pi(x) = |\{p\leq x\}| = \int_2^x \frac{\mathrm{d}t}{\log t} + \mathrm{O}\left(xe^{-c\sqrt{\log x}}\right). \tag{75}$$

This implies by integration by parts

$$\sum_{a<p\leq b}\frac{1}{p}=\int_a^b\frac{\mathrm{d}\pi(x)}{x}=\frac{\pi(b)}{b}-\frac{\pi(a)}{a}+\int_a^b\frac{\pi(x)}{x^2}\mathrm{d}x$$
$$=\int_2^b\frac{\mathrm{d}t}{b\log t}-\int_2^a\frac{\mathrm{d}t}{a\log t}+\int_a^b\frac{\mathrm{d}x}{x^2}\int_2^x\frac{\mathrm{d}t}{\log t}+\mathrm{O}(e^{-\kappa\sqrt{\log a}})=\log_2 b-\log_2 a+\mathrm{O}(e^{-\kappa\sqrt{\log a}}),$$

where we chose $\kappa=c/2$. For the proof of Lemma 16 below we will first need some control on the Laplace transform of the X_p's.

Lemma 15. *There exists an absolute $C>0$ such that for any $\lambda\in\mathbb{R}$ and $0\leq j\leq k$ we have*

$$\mathbb{E}\Big[\exp\big(\lambda\sum_{e^j<\log p<e^k}X_p\big)\Big]\leq\exp((k-j+C)\lambda^2/4). \tag{76}$$

Proof. As a preliminary, we consider the generating function of the increments X_p's. For any $B>0$ there exists $A>0$ such that $|e^w-(w+w^2/2)|\leq A|w|^3$ uniformly in $|w|<B$, so that uniformly in $p\geq 2$ and $|z|\leq C\sqrt{p}$ we have

$$\mathbb{E}\big[e^{zX_p}\big]=\int e^{z\left(\frac{e^{\mathrm{i}\theta}+e^{-\mathrm{i}\theta}}{2\sqrt{p}}+\frac{e^{2\mathrm{i}\theta}+e^{-2\mathrm{i}\theta}}{8p}\right)}\frac{\mathrm{d}\theta}{2\pi}$$
$$=1+z^2\int\left(\frac{e^{\mathrm{i}\theta}+e^{-\mathrm{i}\theta}}{2\sqrt{p}}+\frac{e^{2\mathrm{i}\theta}+e^{-2\mathrm{i}\theta}}{8p}\right)^2\frac{\mathrm{d}\theta}{2\pi}+\mathrm{O}\left(z^3p^{-\frac{3}{2}}\right)=1+\frac{z^2}{4p}+\mathrm{O}\left(z^3p^{-\frac{3}{2}}\right).$$

As a consequence, for any $C>0$ there exists $C'>0$ such that for any $p\geq 2$ and $|z|\leq C\sqrt{p}$, we have

$$\left|\mathbb{E}\big[e^{zX_p}\big]-e^{\frac{z^2}{4p}}\right|\leq C'\frac{|z^3|}{p^{3/2}}. \tag{77}$$

To prove (76), first note that for $1\leq p<\lambda^2/1000$, since the Z_p's are bounded, we trivially have

$$\mathbb{E}\big[e^{\lambda X_p}\big]\leq e^{\frac{|\lambda|}{\sqrt{p}}+\frac{|\lambda|}{2p}}\leq e^{\frac{\lambda^2}{4p}}.$$

Moreover, for $p>\lambda^2/1000$, from (77) there is an absolute $A>0$ such that

$$\mathbb{E}\big[e^{\lambda X_p}\big]\leq e^{\frac{\lambda^2}{4p}+A\frac{|\lambda|^3}{p^{3/2}}}.$$

We conclude that

$$\mathbb{E}\exp\Big(\lambda\sum_{e^j<\log p<e^k}X_p\Big)\leq\mathbb{E}\exp\Big(\lambda^2\sum_{e^j<\log p<e^k}\frac{1}{4p}+A|\lambda|^3\sum_{n>\frac{\lambda^2}{1000}}n^{-3/2}\Big)\leq\exp((k-j+C)\lambda^2/4),$$

where we have used (74). □

Lemma 16 (Gaussian moments of Dirichlet polynomials)**.** *For any $h \in [-2,2]$ and integers k, j, q satisfying $n_0 \leq j \leq k$, $2q \leq e^{n-k}$, and any constant $A > 0$ we have*

$$\mathbb{E}[|S_k(h) - S_j(h)|^{2q}] \ll \frac{(2q)!}{2^q q!} \Big(\frac{k-j}{2}\Big)^q, \tag{78}$$

$$\mathbb{E}[|S_k(h) - S_j(h) + A|^{2q}] \ll \frac{(2q)!}{q!} (k-j)^q + (2A)^{2q}. \tag{79}$$

Moreover, there exists $C > 0$ such that for any $0 \leq j \leq k$, $2q \leq e^{n-k}$, we have

$$\mathbb{E}[|S_k(h) - S_j(h)|^{2q}] \ll q^{1/2} \frac{(2q)!}{2^q q!} \Big(\frac{k-j+C}{2}\Big)^q. \tag{80}$$

Proof. Let Φ be a smooth function such that $\Phi \geq 0$ for all $x \in \mathbb{R}$, $\Phi(x) \gg 1$ for $x \in [-1,1]$ and $\widehat{\Phi}$ is compactly supported in $[-1,1]$, e.g,

$$\Phi(x) := \Big(\frac{\sin \pi(t-1)}{\pi(t-1)}\Big)^2 + \Big(\frac{\sin \pi t}{\pi t}\Big)^2 + \Big(\frac{\sin \pi(t+1)}{\pi(t+1)}\Big)^2.$$

For any two sets of primes $p_1, \dots, p_k$ and $q_1, \dots, q_\ell$ (with possible multiplicity) such that the products $p_1 \dots p_k$ and $q_1 \dots q_\ell$ are smaller than T, we have

$$\begin{aligned}\int_{\mathbb{R}} \Big(\frac{p_1 \dots p_k}{q_1 \dots q_\ell}\Big)^{\mathrm{i}t} \Phi\Big(\frac{t}{2T}\Big) = 2T\widehat{\Phi}\Big(2T \log \frac{p_1 \dots p_k}{q_1 \dots q_\ell}\Big) &= 2T\widehat{\Phi}(0)\mathbf{1}_{p_1 \dots p_k = q_1 \dots q_\ell} \\ &= 2T\widehat{\Phi}(0)\mathbb{E}\Big[Z_{p_1 \dots p_k} \overline{Z}_{q_1 \dots q_\ell}\Big].\end{aligned}$$

Therefore, for any $h \in [-2,2]$ and for primes $p_1, \dots, p_k$ such that $p_1 \dots p_k \leq T^{1/2}$, we have by developing the product

$$\int_{\mathbb{R}} \prod_{\ell=1}^{k} \Big(\mathrm{Re}\Big(\frac{1}{p_\ell^{1/2+\mathrm{i}t+\mathrm{i}h}} + \frac{1}{2}\frac{1}{p_\ell^{1+2\mathrm{i}t+2\mathrm{i}h}}\Big)\Big) \Phi\Big(\frac{t}{2T}\Big)\mathrm{d}t = 2T\widehat{\Phi}(0)\mathbb{E}\Big[\prod_{i=1}^{k} X_{p_i}(h)\Big].$$

We therefore find that, for any $h \in [-2,2]$

$$\begin{aligned}\mathbb{E}[|S_k(h) - S_j(h)|^{2q}] &\ll \frac{1}{2T} \int_{\mathbb{R}} \Big(\mathrm{Re} \sum_{e^j < \log p \leq e^k} \Big(\frac{1}{p^{1/2+\mathrm{i}t+\mathrm{i}h}} + \frac{1}{2}\frac{1}{p^{1+2\mathrm{i}t+2\mathrm{i}h}}\Big)\Big)^{2q} \Phi\Big(\frac{t}{2T}\Big)\mathrm{d}t \\ &= \widehat{\Phi}(0)\mathbb{E}\Big[\Big(\sum_{e^j < \log p \leq e^k} X_p(h)\Big)^{2q}\Big],\end{aligned}$$

where we have used $2q \leq e^{n-k}$ to ensure that all prime products in the expansion satisfy $p_1 \dots p_k \leq T^{1/2}$.

We now evaluate the above moment. Let Z be a uniform random variable on the unit circle, and Y a centered Gaussian random variable with variance $1/2$. For any integer m, we have $\mathbb{E}[(\mathrm{Re} Z)^{2m+1}] = \mathbb{E}[Y^{2m+1}] = 0$ and

$$\mathbb{E}[(\mathrm{Re} Z)^{2m}] = \frac{(2m)!}{2^{2m}(m!)^2} \leq \frac{(2m)!}{2^{2m} m!} = \mathbb{E}[Y^{2m}]. \tag{81}$$

Consider also $S_{jk}^{(1)}=\sum_{e^j<\log p\leq e^k}\mathrm{Re}\, Z_p\, p^{-\frac12}$, $S_{jk}^{(2)}=\frac12\sum_{e^j<\log p\leq e^k}\mathrm{Re}\, Z_p^2\, p^{-1}$, and $G_{jk}^{(1)}=\sum_{e^j<\log p\leq e^k}Y_p\, p^{-\frac12}$, $G_{jk}^{(2)}=\frac12\sum_{e^j<\log p\leq e^k}Y_p\, p^{-1}$, where $(Y_p)_p$ denote independent centered Gaussian random variables with variance $1/2$. With expansion through the binomial formula, (81) implies

$$\mathbb{E}[(S_{jk}^{(1)})^{2q}]\leq \mathbb{E}[(G_{jk}^{(1)})^{2q}],\quad \mathbb{E}[(S_{jk}^{(2)})^{2q}]\leq \mathbb{E}[(G_{jk}^{(2)})^{2q}].$$

Let $\sigma_1^2=\sum_{e^j<\log p\leq e^k}(2p)^{-1}$ and $\sigma_2^2=\sum_{e^j<\log p\leq e^k}(8p)^{-2}$. The above equation implies

$$\begin{aligned}\mathbb{E}[(S_{jk}^{(1)}+S_{jk}^{(2)})^{2q}]&\leq\left(\mathbb{E}[(S_{jk}^{(1)})^{2q}]^{\frac{1}{2q}}+\mathbb{E}[(S_{jk}^{(2)})^{2q}]^{\frac{1}{2q}}\right)^{2q}\\&\leq\left(\mathbb{E}[(G_{jk}^{(1)})^{2q}]^{\frac{1}{2q}}+\mathbb{E}[(G_{jk}^{(2)})^{2q}]^{\frac{1}{2q}}\right)^{2q}=(\sigma_1+\sigma_2)^{2q}\,\mathbb{E}(G^{2q})\end{aligned}$$

where G is a standard Gaussian random variable.

In the case $j\geq n_0$, the quantitative prime number theorem (75) implies $\sigma_1^2=\frac12(k-j)+\mathrm{O}(e^{-\kappa e^{n_0}})$, for some absolute $\kappa>0$. Moreover, we trivially have $\sigma_2^2\leq Ce^{-e^j}$, so that $q\sigma_2/\sigma_1\ll 1$ and (78) follows.

For (79), we have

$$\mathbb{E}[|S_k(h)-S_j(h)+A|^{2q}]^{1/2q}\leq(\mathbb{E}[|S_k(h)-S_j(h)|^{2q}])^{1/(2q)}+A,$$

so that

$$\mathbb{E}[|S_k(h)-S_j(h)+A]^{2q}]\leq 2^{2q}\cdot\mathbb{E}[|S_k(h)-S_j(h)|^{2q}]+(2A)^{2q}$$

and the claim follows from (78).

Finally, for (80), we rely on (76) and obtain, for any $\lambda>0$,

$$\mathbb{E}[(S_k(h)-S_j(h))^{2q}]\leq\frac{(2q)!}{\lambda^{2q}}\mathbb{E}[\cosh(\lambda(S_k(h)-S_j(h))]\leq\frac{(2q)!}{\lambda^{2q}}e^{(k-j+C)\lambda^2/4}.$$

The choice $\lambda^2=4q/(k-j+C)$ and Stirling's formula give the expected result. □

The Gaussian moments yield Gaussian tails for the probability. Indeed, if a random variable X is such that

$$\mathbb{E}[X^{2q}]\ll\frac{(2q)!}{2^q q!}\sigma^{2q},$$

then the Markov inequality together with Stirling's formula and optimization over q yield

$$\mathbb{P}(X>V)\ll\exp(-V^2/(2\sigma^2))\quad\text{for } 2q=\lceil V^2/\sigma^2\rceil.$$

This observation applied to S_k-S_j with $j\geq n_0$ yields for any $h\in[-2,2]$ and $V>0$

$$\mathbb{P}(S_k(h)-S_j(h)>V)\ll\exp(-V^2/(k-j+1)) \tag{82}$$

as long as $V^2\leq e^{n-k}\cdot(k-j+1)/2$. Note in particular that such large deviation estimates are harder to get as k gets closer to n.

We also recall the following analogous bound for the complex partial sums $\widetilde{S}(h)$.

Lemma 17. *For any $h\in[-2,2]$ and integers $n_0\leq j\leq k$ and $2q\leq e^{n-k}$ we have, We have,*

$$\mathbb{E}[|\widetilde{S}_k(h)-\widetilde{S}_j(h)|^{2q}]\ll q!(k-j+1)^q$$

Proof. This follows from [29, Lemma 3]. □

A simple consequence of Lemma 17 is that,

$$\mathbb{P}\Big(|\widetilde{S}_k(h) - \widetilde{S}_j(h)| > V\Big) \ll \frac{V+1}{(k-j+1)^{1/2}} \exp\Big(-\frac{V^2}{k-j+1}\Big). \tag{83}$$

A.2 **Gaussian approximation.** Recall the definition of the partial sums in (36) for $h = 0$:

$$\mathcal{S}_k = \sum_{e^{1000} < \log p \le e^k} X_p. \tag{84}$$

We have the following simple estimate for the probability density function of $\mathcal{S}_k$.

Lemma 18. *Let $|v| \le 100r$. Then, for $r > 1000$ and for all $\Delta \ge 1$, we have*

$$\mathbb{P}(\mathcal{S}_r \in [v, v + \Delta^{-1}]) \asymp \frac{1}{\Delta} \cdot \frac{1}{\sqrt{r}} \exp\Big(-\frac{v^2}{r}\Big).$$

Proof. We will merely sketch the proof of this standard result (see e.g [23, Theorem 1] or [11, Theorem 2.1] for more detailed accounts). The probability density function of $\mathcal{S}_r$ can be written (by inverse Fourier transform and contour deformation)

$$f_r(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbb{E}[\exp\Big((\sigma + \mathrm{i}t)\mathcal{S}_r\Big)\Big] \exp\Big(-(\sigma + \mathrm{i}t)x\Big) \mathrm{d}t \ , \ \sigma = \frac{2x}{r}.$$

It remains therefore to analyze the above integral using the saddle point method. First we notice that in the region $0 \le \mathrm{Re} z \le 200$ we have,

$$\mathbb{E}[e^{z\mathcal{S}_r}] = \exp\Big(\frac{z^2 r}{4}\Big) H(z),$$

with $H = H_r$ a function analytic in the strip $0 \le \mathrm{Re} z \le 200$ such that $\frac{1}{2} \le |H(z)| \le 10^3$ and $|H'(z)| \le 10^{-6}$ uniformly in the strip $0 \le \mathrm{Re} z \le 200$, and uniformly in r. (This uses that the X_p's appearing in $\mathcal{S}_r$ have $p > \exp(e^{1000})$). The rest of the proof now proceeds by a standard application of the saddle point method. The region $|t| > 100 \log r / \sqrt{r}$ gives a negligible contribution, while the region $|t| \le 100 \log r / \sqrt{r}$ contributes,

$$\frac{1}{2\pi} \int_{|t| \le 100 \log r / \sqrt{r}} H\Big(\frac{2x}{r} + \mathrm{i}t\Big) \exp\Big(\frac{r}{4}\Big(\frac{2x}{r} + \mathrm{i}t\Big)^2 - \frac{2x^2}{r} - \mathrm{i}tx\Big) \mathrm{d}t.$$

By a Taylor expansion, the above is equal to

$$\exp\Big(-\frac{x^2}{r}\Big) \cdot \frac{1}{2\pi} \int_{|t| \le 100 \log r / \sqrt{r}} \Big(H\Big(\frac{2x}{r}\Big) + \mathrm{O}^\star\Big(10^{-4} \frac{\log r}{\sqrt{r}}\Big)\Big) \exp\Big(-\frac{t^2 r}{4}\Big) \mathrm{d}t \asymp \frac{1}{\sqrt{r}} \exp\Big(-\frac{x^2}{r}\Big).$$

with $\mathrm{O}^\star$ denoting a O with implicit constant ≤ 1. Thus, uniformly in $|x| \le 100r$, we have $f_r(x) \asymp r^{-1/2} \exp(-x^2/r)$. The result follows upon integrating $x \in [v, v + \Delta^{-1}]$. □

We now remind the following version of the Berry-Esseen theorem, see for example Corollary 17.2 in [7]. The probability measure $\mathbb{P}$ below is arbitrary, and $\eta_{\mu,\sigma}$ denotes the Gaussian measure with mean μ and variance σ.

Lemma 19. *Let W_j be a sequence of independent random variables on $(\mathbb{R},\mathcal{B}(\mathbb{R}),\mathbb{P})$, with associated expectation denoted $\mathbb{E}$, and let $\mathbb{Q}_m$ be the distribution of $W_1+\cdots+W_m$. Let*

$$\mu_m=\sum_{j=1}^m \mathbb{E}[W_j],\quad \sigma_m=\sum_{j=1}^m \mathbb{E}[(W_j-\mathbb{E}(W_j))^2],$$

and $\mathcal{A}$ be the set of intervals in $\mathbb{R}$. There exists an absolute constant c such that

$$\sup_{A\in\mathcal{A}} |\mathbb{Q}_m(A)-\eta_{\mu_m,\sigma_m}(A)|\leq \frac{c}{\sigma_m^{3/2}}\sum_{j=1}^m \mathbb{E}[|W_j-\mathbb{E}(W_j)|^3].$$

The following consequence of Lemma 19 compares the probabilistic model $(\mathcal{S}_i)_{i\geq 1}$ defined previously with a natural Gaussian analogue. To state this comparison, remember the definitions (36) and (37). In the statement below we omit the argument h to mean $h=0$.

Lemma 20. *There exists a constant $c>0$ such that, for any interval A and $k\geq 1$,*

$$\mathbb{P}\Big(\mathcal{Y}_k\in A\Big)=\mathbb{P}\Big(\mathcal{N}_k\in A\Big)+\mathrm{O}(e^{-ce^{k/2}}).$$

Proof. Let

$$\mathcal{N}_k'=\sum_{e^{k-1}<\log p\leq e^k} X_p'$$

where the $(X_p', p \text{ prime})$ are centered, independent real Gaussian random variables, with variance $\frac{1}{2p}+\frac{1}{8p^2}$ matching exactly the variance of the summands X_p of $\mathcal{S}_k$. We apply Lemma 19: All random variables are centered with matching variances, $C^{-1}\leq \mathbb{E}[(\mathcal{Y}_k)^2)]\leq C$ and (we have $|X_p|<Cp^{-1/2}$ deterministically)

$$\sum_{e^{k-1}<\log p\leq e^k} \mathbb{E}[|X_p-\mathbb{E}(X_p)|^3]\leq C\sum_{e^{k-1}<\log p\leq e^k} p^{-3/2}\leq Ce^{-ce^k},$$

for some absolute constants $C,c>0$, so that

$$\mathbb{P}\Big(\mathcal{Y}_k\in A\Big)=\mathbb{P}\Big(\mathcal{N}_k'\in A\Big)+\mathrm{O}(e^{-ce^k}). \tag{85}$$

Moreover, denote $\beta_k=\sum_{e^{k-1}<\log p\leq e^k}\left(\frac{1}{2p}+\frac{1}{8p^2}\right)$. From Pinsker's inequality and (74), we have that the total variation between the distribution of $\mathcal{N}_k$ and $\mathcal{N}_k'$ is

$$2\mathrm{TV}(\mathcal{N}_k,\mathcal{N}_k')^2\leq \int\Big(\log\frac{\mathrm{d}\eta_{0,1/2}}{\mathrm{d}\eta_{0,\beta_k}}\Big)\mathrm{d}\eta_{0,1/2}=\mathrm{O}\Big(\big|\beta_k-\frac{1}{2}\big|\Big)=\mathrm{O}(e^{-ce^{k/2}}). \tag{86}$$

Equations (85) and (86) conclude the proof. □

A.3 **Moments of the Riemann zeta function.**

Lemma 21 (Second moment of the Riemann zeta function)**.** *For all $h\in[-2,2]$, we have*

$$\mathbb{E}[|\zeta_\tau(h)|^2]\ll e^n.$$

Proof. See [16, Theorem 2.41]. □

Lemma 22 (Fourth moment of the Riemann zeta function)**.** *For all $h \in [-2,2]$, we have*

$$\mathbb{E}[|\zeta_\tau(h)|^4] \ll e^{4n}.$$

More generally, for real $|\sigma - 1/2| \leq \frac{1}{100}$, we have

$$\mathbb{E}[|\zeta(\sigma + \mathrm{i}\tau + \mathrm{i}h)|^4] \ll \exp(1 + e^n(2-4\sigma))\, e^{4n}.$$

Proof. If $\sigma < \frac{1}{2}$ the functional equation yields

$$\mathbb{E}[|\zeta(\sigma + \mathrm{i}\tau + \mathrm{i}h)|^4] \ll \exp(e^n(2-4\sigma))\mathbb{E}[|\zeta(1-\sigma + \mathrm{i}\tau + \mathrm{i}h)|^4].$$

Now uniformly in $\frac{1}{2} \leq \sigma \leq \frac{3}{4}$ by [17, Theorem D] we have

$$\mathbb{E}[|\zeta(\sigma + \mathrm{i}\tau + \mathrm{i}h)|^4] \ll e^{4n}.$$

The result follows. □

A.4 **Some useful sums over primes.** The first lemma justifies the approximation of e^{-S_k} by mollifiers. Recall the definition of $\widetilde{S}_k$ in (5) and that $\mathrm{Re}\, \widetilde{S}_k = S_k$.

Lemma 23. *Let $\ell \geq 0$ and $k \in (n_{\ell-1}, n_\ell]$. Suppose that $|\widetilde{S}_k(h) - \widetilde{S}_{n_{\ell-1}}(h)| \leq 10^3(n_\ell - n_{\ell-1})$. We have,*

$$e^{-(S_k(h)-S_{n_{\ell-1}}(h))} \leq (1+e^{-n_{\ell-1}})\,|\mathcal{M}_{\ell-1}^{(k)}(h)| + e^{-10^5(n_\ell - n_{\ell-1})}.$$

Proof. Let

$$R_k(h) := \sum_{\substack{e^{n_{\ell-1}} < \log p \leq e^k \\ \alpha \geq 3}} \frac{1}{\alpha}\,\mathrm{Re}\, p^{-\alpha(\frac{1}{2}+\mathrm{i}\tau+\mathrm{i}h)}.$$

Notice that $|R_k(h)| \leq e^{-2n_{\ell-1}}$. As a result, we clearly have

$$e^{-(S_k(h)-S_{n_{\ell-1}}(h))} \leq (1+e^{-n_{\ell-1}})\, e^{-(S_k(h)-S_{n_{\ell-1}}(h))-R_k(h)}.$$

Set $s := \frac{1}{2} + \mathrm{i}\tau + \mathrm{i}h$. Notice that,

$$e^{-(S_k(h)-S_{n_{\ell-1}}(h))-R_k(h)} = \Big|\prod_{p \in (T_{\ell-1}, \exp(e^k)]} \Big(1 - \frac{1}{p^s}\Big)\Big|.$$

Furthermore, setting $V := (n_\ell - n_{\ell-1})^{10^5}$,

$$\prod_{p\in(T_{\ell-1},\exp(e^k)]}\Big(1-\frac{1}{p^s}\Big) = \sum_{p|n \Rightarrow p \in (T_{\ell-1},\exp(e^k)]} \frac{\mu(n)}{n^s}$$
$$= \mathcal{M}_{\ell-1}^{(k)}(h) + \sum_{\substack{p|n\Rightarrow p\in(T_{\ell-1},\exp(e^k)] \\ \Omega_{\ell-1}(n) > V}} \frac{\mu(n)}{n^s}$$

Therefore it remains to show that the second term is identically small. Notice that we can re-write the second term as

$$\sum_{\ell > V} (-1)^\ell \Big(\sum_{T_{\ell-1} < p_1 < \ldots < p_\ell \leq \exp(e^k)} \frac{1}{(p_1 \ldots p_\ell)^s}\Big). \tag{87}$$

Furthermore, using the Girard-Newton identities (see for example Equation 2.14' in [24]), we can re-write the inner sum as follows,

$$\sum_{T_{\ell-1}<p_1<\ldots<p_\ell<\exp(e^k)} \frac{1}{(p_1\ldots p_\ell)^s} = (-1)^\ell \sum_{\substack{m_1,\ldots,m_\ell,\ldots\geq 0\\ m_1+2m_2+\ldots+\ell m_\ell+(\ell+1)m_{\ell+1}+\ldots=\ell}} \prod_{1\leq j} \frac{(-\mathcal{P}(\ell s))^{m_j}}{m_j!\ j^{m_j}}$$

with

$$\mathcal{P}(s) := \sum_{T_{\ell-1}<p\leq\exp(e^k)} \frac{1}{p^s}.$$

Using this we can bound the absolute value of (87), for any $\alpha > 0$, by

$$\begin{aligned}\sum_{m_1,\ldots,m_\ell,\ldots\geq 0} &\exp\Big(-\alpha V + \alpha m_1 + 2\alpha m_2 + \ldots + \ell\alpha m_\ell + \ldots\Big) \prod_{1\leq j} \frac{|\mathcal{P}(js)|^{m_j}}{m_j!\ j^{m_j}}\\ &= e^{-\alpha V} \prod_{1\leq j}\Big(\sum_{m_j\geq 0} \frac{(e^{j\alpha}|\mathcal{P}(js)|/j)^{m_j}}{m_j!}\Big). \qquad (88)\end{aligned}$$

By assumption we have $|\mathcal{P}(s)| \leq 10^4(n_\ell - n_{\ell-1})$ and trivially we have $|\mathcal{P}(2s)| \leq n_\ell - n_{\ell-1} + 2$ and $|\mathcal{P}(\ell s)| \leq 10^{-\ell}$ for $\ell \geq 3$. As a result (88) is (for $\alpha = 1$) bounded by

$$\ll \exp(-(n_\ell - n_{\ell-1})^{10^5} + 10^5(n_\ell - n_{\ell-1}))$$

and the claim follows. □

Lemma 24. *Let $p > \exp(e^{1000})$ be a prime and $\alpha \geq 1$, an integer. Given $\pi\mathbf{z} = (z_3, z_4, z_1, z_2)$, define*

$$B_{\pi\mathbf{z}}(p^\alpha) := \frac{\sum_{j=0}^\infty \sigma_{z_1,z_2}(p^j)\sigma_{z_3,z_4}(p^{\alpha+j})p^{-j}}{\sum_{j=0}^\infty \sigma_{z_1,z_2}(p^j)\sigma_{z_3,z_4}(p^j)p^{-j}},$$

where $\sigma_{z,w}(p^\alpha) := \sum_{nm=p^\alpha} n^{-z}m^{-w}$. Then, uniformly in $|z_i| \leq 3^4/(\alpha\log p)$ for $i = 1,2,3,4$, we have

$$|B_{\mathbf{z}}(p^\alpha) - B_{\mathbf{0}}(p^\alpha)| \leq e^{3000}\,\alpha^2 \log p \sum_{i=1}^4 |z_i|,$$

where $\mathbf{0} := (0,0,0,0)$. Furthermore

$$B_{\mathbf{0}}(p^\alpha) = \Big(1 - \frac{1}{p^2}\Big)^{-1}\Big(1 + \alpha - \frac{2\alpha}{p} + \frac{\alpha-1}{p^2}\Big)$$

Proof. We start with the second claim, Lemma 6.9 of [21] (applied at $s = 0$) implies

$$B_{\mathbf{z}}(p^\alpha) = \frac{B_{\mathbf{z}}^{(0)}(p^\alpha) - p^{-1}B_{\mathbf{z}}^{(1)}(p^\alpha) + p^{-2}B_{\mathbf{z}}^{(2)}(p^\alpha)}{(p^{-z_3} - p^{-z_4})(1 - p^{-2-z_1-z_2-z_3-z_4})}, \qquad (89)$$

where

$$\begin{aligned} B_{\mathbf{z}}^{(0)}(p^\alpha) &= p^{-z_3(\alpha+1)} - p^{-z_4(\alpha+1)},\\ B_{\mathbf{z}}^{(1)}(p^\alpha) &= (p^{-z_1} + p^{-z_2})\,p^{-z_3-z_4}\,(p^{-z_3\alpha} - p^{-z_4\alpha}),\\ B_{\mathbf{z}}^{(2)}(p^\alpha) &= p^{-z_1-z_2-z_3-z_4}\,(p^{-z_4-z_3\alpha} - p^{-z_3-z_4\alpha}).\end{aligned}$$

The second claims follows by estimating this at $\mathbf{z} = \mathbf{0}$.

Note that, for $|w_i| \le 200/(\alpha \log p)$, (89) gives $|B_{\mathbf{w}}(p^\alpha)| \le e^{2000}\,\alpha$. Now, by Cauchy's theorem, we have

$$\begin{aligned}|B_{(z_1,z_2,z_3,z_4)}(p^\alpha) - B_{(0,z_2,z_3,z_4)}(p^\alpha)| &= \Big|\frac{1}{2\pi \mathrm{i}}\oint_{|w|=200/(\alpha\log p)} B_{(w,z_2,z_3,z_4)}(p^\alpha)\,\frac{z_1 \mathrm{d}w}{(w-z_1)w}\Big| \\ &\le |z_1|\, e^{10}\,\alpha\log p \max_{|\mathbf{w}|=200/(\alpha\log p)} |B_{\mathbf{w}}(p^\alpha)|,\end{aligned}$$

where $|\mathbf{w}| = C$ means that $|w_i| = C$ for $i = 1, 2, 3$, or 4. Note also that the last bound is true by the maximum modulus principle, since $|z_i| \le 200/(\alpha \log p)$. Now, iterating this on each variable z_2, z_3, z_4, using the bound $|B_{\mathbf{w}}(p^\alpha)| \le e^{2000}\,\alpha$, and adding the results, we conclude that

$$|B_{\mathbf{z}}(p^\alpha) - B_{\mathbf{0}}(p^\alpha)| \le e^{3000}\,\alpha^2 \log p \sum_{i=1}^{4} |z_i|.$$

This proves the first claim. □

Appendix B. Ballot Theorem

Recall the definition (37),

$$\mathcal{G}_k = \sum_{1000\le \ell\le k} \mathcal{N}_\ell,$$

where the $\mathcal{N}_\ell$'s are centered, independent real Gaussian random variables, with variance $\frac{1}{2}$. The main result in this section is the following Ballot theorem for the random walk $\mathcal{G}$, with Gaussian increments. It extends [31, Lemma 6.2] from a linear to a curved barrier, and our proof relies on this result.

Proposition 4. *Uniformly in $n \ge 1$, $1 \le y \le 2n$, $n/2 \le k \le n$ and $m(k) + L_y(k) - 4 \le w \le m(k) + U_y(k)$ (see Equation (6)), we have for $r := \lceil y/4 \rceil$,*

$$\begin{aligned}&\mathbb{P}(\{\mathcal{G}_k \in (w, w+1]\} \cap \{\mathcal{G}_r - m(r) \in [L_y(r), U_y(r)]\} \cap_{r<j\le k} \{\mathcal{G}_j < m(j) + U_y(j)\}) \\ &\qquad\qquad \ll (y+1)\,(U_y(k) + m(k) - w + 1)\, k^{-3/2}\, e^{-\frac{w^2}{k}}. \qquad (90)\end{aligned}$$

The above proposition is an immediate consequence of the following one.

Proposition 5. *For any fixed $c_1 > 0, 0 \le \theta < 1/2$, there exists C such that the following holds. Consider arbitrary $k \ge 1$,$|\alpha| < c_1^{-1}$ and g defined on $[0, k]$ satisfying $g(0) = g(k) = 0$,*

$$|g'(x)| < c_1^{-1} \min(x+1, k-x+1)^{\theta-1}, \quad 0 \le x \le k, \qquad (91)$$

$$-c_1 \min(x+1, k-x+1)^{\theta-2} < g''(x) \le 0, \quad 0 \le x \le k. \qquad (92)$$

Let $f_y(x) = g(x) + \alpha x + y$. Then for any such f_y and $0 < y < c_1^{-1} k$, $-c_1^{-1} k < w < f_y(k)$, we have

$$\mathbb{P}\Big(\bigcap_{1\le j\le k}\Big\{\sum_{1\le i\le j}\mathcal{N}_i \le f_y(j)\Big\} \cap \Big\{\sum_{1\le i\le k}\mathcal{N}_i \in (w, w+1]\Big\}\Big) \le C\frac{(y+1)\,(f_y(k)-w+1)}{k^{3/2}}\, e^{-\frac{w^2}{k}}. \qquad (93)$$

Proof. We abbreviate $W_j=\sum_{i\leq j}\mathcal{N}_i$. Let $\mathbb{P}_n^w$ denote the distribution of $(W_1,\dots,W_k)$ conditionally to $W_k=w$, and $\mathbb{E}_k^w$ the corresponding expectation. In our range of parameters for any $x\in[w,w+1)$ we have $e^{-x^2/k}\asymp e^{-w^2/k}$. It is therefore enough to prove that uniformly in the described f_y, y,w, we have

$$\mathbb{P}_k^x\Big(\bigcap_{j\leq k}\Big\{W_j\leq f_y(j)\Big\}\Big)\ll (y+1)\,(f_y(k)-x+1)\,k^{-1}. \tag{94}$$

By a linear change of variables eliminating α, we have

$$\mathbb{P}_k^x\Big(\bigcap_{j\leq k}\Big\{W_j\leq f_y(j)\Big\}\Big)=\mathbb{P}_k^{x-(f_y(k)-y)}\Big(\bigcap_{j\leq k}\Big\{W_j\leq g(j)+y\Big\}\Big). \tag{95}$$

We denote $\bar{x}=x-(f_y(k)-y)$. There is a constant $c(k)$ independent of all other parameters such that the above right-hand side is

$$c(k)\int_{u_j<y+g(j)} e^{-\sum_{i=1}^k(u_i-u_{i-1})^2}\prod_{j=1}^{k-1}\mathrm{d}u_j=c(k)\int_{v_j<y} e^{-\sum_{i=1}^k(v_i-v_{i-1}+g(i)-g(i-1))^2}\prod_{j=1}^{k-1}\mathrm{d}v_j, \tag{96}$$

where we use the conventions $u_0=v_0=0$, $u_k=v_k=\bar{x}$. From Equation (91), we have $|g(i)-g(i-1)|\leq c_1^{-1}\min(i,k-i+1)^{\theta-1}$, and Equation (92) gives $0\leq 2g(i)-g(i-1)-g(i+1)\leq 2c_1^{-1}\min(i,k-i+1)^{\theta-2}$. These bounds in the expansion of the Hamiltonian together with the assumption $0\leq\theta<1/2$ give

$$\begin{aligned}
&\sum_{i=1}^k(v_i-v_{i-1}+g(i)-g(i-1))^2\\
=&\sum_{i=1}^k(v_i-v_{i-1})^2+\mathrm{O}(1)+2\sum_{i=1}^k(v_i-i\frac{\bar{x}}{k}-v_{i-1}+(i-1)\frac{\bar{x}}{k})(g(i)-g(i-1))\\
\geq&\sum_{i=1}^k(v_i-v_{i-1})^2+\mathrm{O}(1)+\sum_{i=1}^k a_i\left(v_i-i\frac{\bar{x}}{k}\right)
\end{aligned} \tag{97}$$

where the constants a_i satisfy $0\leq a_i\leq 2c_1^{-1}\min(i,k-i+1)^{\theta-2}$. In the last line, we summed by parts to express the sum in the variables $v_i-i\frac{\bar{x}}{k}$ and in the difference $2g(i)-g(i-1)-g(i+1)$. Let $\overline{W}_j=W_j-j\frac{\bar{x}}{k}$. With equations (95), (96) and (97), Equation (94) follows once it is shown that

$$\mathbb{E}_k^{\bar{x}}\big[e^{-\sum_{j=1}^{k-1}a_j\overline{W}_j}\mathbf{1}_{\cap_{j\leq k}\{W_j\leq y)\}}\big]\ll\frac{(y+1)(f_y(k)-x+1)}{k}.$$

As $ab\leq(a^2+b^2)/2$, the above inequality will follow from

$$\mathbb{E}_k^{\bar{x}}\big[e^{-2\sum_{j\leq k/2}a_j\overline{W}_j}\mathbf{1}_{\cap_{j\leq k}\{W_j\leq y)\}}\big]\ll\frac{(y+1)(f_y(k)-x+1)}{k}, \tag{98}$$

$$\mathbb{E}_k^{\bar{x}}\big[e^{-2\sum_{j>k/2}a_j\overline{W}_j}\mathbf{1}_{\cap_{j\leq k}\{W_j\leq y)\}}\big]\ll\frac{(y+1)(f_y(k)-x+1)}{k}. \tag{99}$$

We start with (98). Suppose without loss of generality that $-2\sum_{j\leq k/2}a_j\overline{W}_j>1$. (On the event that this is <1, we can bound the exponential term by a constant, the estimate then follows by a standard ballot theorem with constant barrier as in (101).) Let $\varepsilon = (1/2-\theta)/2 > 0$. Note that there exists a constant $\kappa = \kappa(c_1) > 0$ such that for any $u > 1$, $-2\sum_{j\leq k/2} a_j \overline{W}_j \in [u, u+1]$ implies that there exists $1 \leq r \leq k/2$ such that $\overline{W}_r < -\kappa u r^{\frac12+\varepsilon}$. This observation together with the union bound gives

$$\begin{aligned}
&\mathbb{E}_k^{\bar x}\big[e^{-2\sum_{j\leq k/2} a_j \overline{W}_j}\mathbf{1}_{\cap_{j\leq k}\{W_j\leq y)\}}\big]\\
\ll &\sum_{u\geq 1, r\leq k/2, v\geq \kappa u r^{\frac12+\varepsilon}} e^u\, \mathbb{P}_k^{\bar x}\left(\{-\overline{W}_r\in[v,v+1]\}\cap_{j\leq k}\{W_j\leq y)\}\right)\\
\ll &\sum_{u\geq 1, r\leq k/2, v\geq \kappa u r^{\frac12+\varepsilon}} e^u\, \mathbb{P}_k^{\bar x}\left(-\overline{W}_r\in[v,v+1]\right)\sup_{a\in[v,v+1]}\mathbb{P}_{k-r}^{\bar x+a-r\frac{\bar x}{k}}\Big(\cap_{1\leq j\leq k-r}\{W_j\leq y-r\frac{\bar x}{k}+a)\}\Big)
\end{aligned} \tag{100}$$

where we used the Markov property for the second inequality. To bound the first probability above, note that under $\mathbb{P}_k^{\bar x}$, the random variable $\overline{W}_r$ is centered, Gaussian with variance $r - \frac{r^2}{k} \asymp r$. For the second probability, we will rely on [31, Lemma 6.2], which can be rephrased as follows: Uniformly in $m, z_1 \geq 1$, $z_2 \leq z_1$, we have

$$\mathbb{P}_m^{z_2}(\cap_{j\leq m}\{W_j\leq z_1)) \ll \frac{(z_1+1)(z_1-z_2+1)}{m}. \tag{101}$$

This allows to bound (100) with

$$\mathbb{E}_k^{\bar x}\big[e^{-2\sum_{j\leq k/2} a_j \overline{W}_j}\mathbf{1}_{\cap_{j\leq k}\{W_j\leq y)\}}\big] \ll \sum_{u\geq 1, 1\leq r\leq k/2, v>\kappa u r^{\frac12+\varepsilon}} e^{u-c\frac{v^2}{r}}\cdot\frac{(y-r\frac{\bar x}{k}+v+1)(y-\bar x+1)}{k}$$

for some absolute $c > 0$. The above sum over v and then u is $\ll e^{-c' r^{2\varepsilon}}$ for some $c' > 0$ depending on c_1. We conclude that uniformly in our parameters, (100) is bounded with

$$\sum_{1\leq r\leq k/2} e^{-cr^{2\varepsilon}}\frac{(y-r\frac{\bar x}{k}+1)(y-\bar x+1)}{k} \ll \frac{(y+\frac{|\bar x|}{k}+1)(y-\bar x+1)}{k}.$$

It follows from our hypotheses that $\bar x/k$ is uniformly bounded, so that the above equation gives (98). Equation (99) can be proved the same way, with r now chosen in $[k/2, k]$ and the barrier event between times 0 and r. This concludes the proof of (94), and the lemma. □

Appendix C. Discretization

The lemmas of this section allow to reduce the study the maximum of a Dirichlet polynomial of a given length on a typical interval to a finite set of points.

Lemma 25. *Let $\varepsilon > 0$ be given. Let V be a smooth function with $V(x) = 1$ for $0 \leq x \leq 1$ and compactly supported in $[-\varepsilon, 1+\varepsilon]$. Let $D(s)$ be a Dirichlet polynomial of length N. Then, for any $t, h_0 \in \mathbb{R}$,*

$$D(\tfrac12+\mathrm{i}t+\mathrm{i}h_0) = \frac{1}{2+\varepsilon}\sum_{h\in\frac{2\pi\mathbb{Z}}{(2+\varepsilon)\log N}} D\Big(\tfrac12+\mathrm{i}t+\mathrm{i}h\Big)\widehat V\Big(\frac{(h-h_0)\log N}{2\pi}\Big).$$

Proof. This proof is essentially a repetition of [5, Proposition 2.7] with slight differences. Let $G(x)=V(2\pi x/\log N)$, so that $\widehat{G}(x):=\frac{\log N}{2\pi}\widehat{V}(\frac{x\log N}{2\pi})$. By Poisson summation, for any fixed $0\leq n\leq N$, we have

$$\sum_{k\in\mathbb{Z}} n^{-\frac{2\pi \mathrm{i}k}{(2+\varepsilon)\log N}}\widehat{G}\Big(\frac{2\pi k}{(2+\varepsilon)\log N}-h_0\Big)=\sum_{\ell\in\mathbb{Z}}\int_{\mathbb{R}} n^{-\frac{2\pi \mathrm{i}x}{(2+\varepsilon)\log N}}\widehat{G}\Big(\frac{2\pi x}{(2+\varepsilon)\log N}-h_0\Big)e^{-2\pi\mathrm{i}\ell x}\mathrm{d}x. \tag{102}$$

For fixed ℓ, by inverse Fourier transform the above integral is

$$\frac{(2+\varepsilon)\log N}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}x(\log n+(2+\varepsilon)\ell\log N)}\widehat{G}(x-h_0)\mathrm{d}x$$
$$=\frac{(2+\varepsilon)\log N}{2\pi}e^{-\mathrm{i}h_0(\log n+(2+\varepsilon)\ell\log N)}G\left(\frac{\log n+(2+\varepsilon)\ell\log N}{2\pi}\right).$$

From the compact support assumption on V, for $0\leq n\leq N$ the above right-hand side is nonzero only for $\ell=0$. Equation (102) can therefore be written as

$$n^{-\mathrm{i}h_0}=\frac{1}{2+\varepsilon}\sum_{h\in\frac{2\pi\mathbb{Z}}{(2+\varepsilon)\log N}} n^{-\mathrm{i}h}\widehat{V}\Big(\frac{(h-h_0)\log N}{2\pi}\Big).$$

This conclude the proof by linearity. □

The following is a particular case of [5, Corollary 2.8].

Lemma 26. *Let $\mathcal{T}_n$ be a set of e^{-n-100} well-spaced points in $[-2,2]$ with $n=\log_2 T$. There exists an absolute constant $C>0$ such that for any $A>0$ and $v\geq 1$,*

$$\mathbb{P}(\max_{|h|\leq 1}|\zeta(\tfrac{1}{2}+\mathrm{i}\tau+\mathrm{i}h)|>v)\leq \mathbb{P}(\max_{h\in\mathcal{T}_n}|\zeta(\tfrac{1}{2}+\mathrm{i}\tau+\mathrm{i}h)|>v/C)+\mathrm{O}_A(e^{-An})$$

Proof. By [8, Proposition 2] for $t\in[T,2T]$, the zeta function is well-approximated by a Dirichlet polynomial of length T:

$$\zeta(\tfrac{1}{2}+\mathrm{i}t)=\sum_{n\leq T}\frac{1}{n^{1/2+\mathrm{i}t}}\left(1-\frac{\log n}{\log T}\right)^{100}+\mathrm{O}(T^{-100})=:D(t)+\mathrm{O}(T^{-100}). \tag{103}$$

Lemma 25 implies, for any $|h_0|\leq 1$,

$$|D(t+h_0)|\leq \sum_{h\in e^{-n-100}\mathbb{Z}}|D(t+h)|\cdot\Big|\widehat{V}\Big(\frac{(h-h_0)e^n}{2\pi}\Big)\Big|,$$

where V is a smooth compactly supported function such that $V(x)=1$ for $0\leq x\leq 1$. In particular, for any $|h_0|\leq 1$, we have

$$|D(t+h_0)|\leq C\max_{h\in\mathcal{T}_n}|D(t+h)|+\mathcal{E}(t)$$

with $C>0$ an absolute constant and where

$$\mathcal{E}(t):=\sum_{\substack{h\in e^{-n-100}\mathbb{Z}\\ |h|>2}}|D(t+h)|\cdot\Big|\widehat{V}\Big(\frac{(h-h_0)e^n}{2\pi}\Big)\Big|.$$

Since $|h_0| \leq 1$, the $\widehat{V}$ term decays faster than any polynomial of e^n. Lemma 13 and the Cauchy-Schwarz inequality therefore give

$$\mathbb{P}(\mathcal{E}(\tau) \geq 1) \leq \mathbb{E}[\mathcal{E}(\tau)] \ll_A e^{-An},$$

for any given $A > 0$. Putting it all together, we conclude that for $v \geq 1$, and all T sufficiently large,

$$\mathbb{P}(\max_{|h|\leq 1} |\zeta(\tfrac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)| > v) \leq \mathbb{P}(\max_{h \in \mathcal{T}_n} |\zeta(\tfrac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)| > v/(2C)) + \mathrm{O}_A(e^{-An})$$

for any given $A > 0$. □

Lemma 25 implies the following discretization for the maximum of Dirichlet polynomials.

Lemma 27. *Let $\mathcal{I}$ be a finite set of indices. Let D_i with $i \in \mathcal{I}$ be a sequence of Dirichlet polynomial of length $\leq N$. Then, for any $\ell \geq 1$, and any $A \geq 100$,*

$$\max_{|h|\leq 2} \Big(\sum_{i\in\mathcal{I}} |D_i(\tfrac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)|^2 \Big) \ll_A \sum_{|j|\leq 16\log N} \Big(\sum_{i\in\mathcal{I}} \Big| D_i\Big(\frac{1}{2} + \mathrm{i}\tau + \frac{2\pi \mathrm{i} j}{8\log N}\Big)\Big|^2 \Big) + \sum_{|j|>16\log N} \frac{1}{1+|j|^A} \Big(\sum_{i\in\mathcal{I}} \Big| D_i\Big(\frac{1}{2} + \mathrm{i}\tau + \frac{2\pi \mathrm{i} j}{8\log N}\Big)\Big|^2 \Big). \tag{104}$$

Proof. We can apply Lemma 25 to the Dirichlet polynomial D_i^2 (its proof for Dirichlet polynomials of length at most $2N$ only requires minor changes in constants) we get, for any $|h| \leq 2$,

$$|D_i(\tfrac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)|^2 \ll_A \sum_{|j|\leq 16\log N} \Big| D_i\Big(\frac{1}{2} + \mathrm{i}\tau + \frac{2\pi \mathrm{i} j}{8\log N}\Big)\Big|^2 + \sum_{|j|>16\log N} \frac{1}{1+|j|^A} \Big(\sum_{i\in\mathcal{I}} \Big| D_i\Big(\frac{1}{2} + \mathrm{i}\tau + \frac{2\pi \mathrm{i} j}{8\log N}\Big)\Big|^2 \Big),$$

using the decay bound $\widehat{V}(x) \ll_A (1+|x|)^{-A}$. Summing over $i \in \mathcal{I}$ and then taking the supremum over $|h| \leq 2$ yield the claim. □

Since the zeta function is well-approximated by a Dirichlet polynomial of length T as in (103), Lemma 27 can also be used to approximate the moments of the maximum of zeta on a short interval. We choose to prove this directly.

Lemma 28. *We have*

$$\mathbb{E}[\max_{|h|\leq 1} |\zeta(\tfrac{1}{2} + \mathrm{i}\tau + \mathrm{i}h)|^4] \ll e^{5n}.$$

Proof. As ζ is analytic, the function $|\zeta(\frac{1}{2}+\mathrm{i}t+\mathrm{i}z)|^4$ is subharmonic in the region $|z| < \frac{1}{100}$. Therefore for $|h| \leq e^{-n}$ we have,

$$|\zeta(\tfrac{1}{2} + \mathrm{i}t + \mathrm{i}h)|^4 \ll e^{2n} \int_{|x|,|y|\leq 2e^{-n}} |\zeta(\tfrac{1}{2} + \mathrm{i}t + x + \mathrm{i}y)|^4 \mathrm{d}x\mathrm{d}y.$$

Summing the above over a grid of e^{-n-100} well-spaced point we conclude that,

$$\max_{|h|\leq 1}|\zeta(\tfrac{1}{2}+\mathrm{i}\tau+\mathrm{i}h)|^4 \ll e^{2n}\sum_{h\in\mathcal{T}_n}\int_{|x|,|y|\leq 2e^{-n}}|\zeta(\tfrac{1}{2}+\mathrm{i}\tau+\mathrm{i}h+x+\mathrm{i}y)|^4\mathrm{d}x\mathrm{d}y.$$

Taking expectation on both sides and using Lemma 22 we obtain the desired result. □

References

[1] L.-P. Arguin. *Extrema of Log-correlated Random Variables: Principles and Examples*, page 166–204. Cambridge University Press, 2016.

[2] L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.*, 349(2):703–751, 2017.

[3] 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(3):500–535, 2019.

[4] L.-P. Arguin, A. Bovier, and N. Kistler. Genealogy of extremal particles of branching brownian motion. *Communications on Pure and Applied Mathematics*, 64(12):1647–1676, 2011.

[5] L-P. Arguin, F. Ouimet, and M. Radziwiłł. Moments of the Riemann zeta function on short intervals of the critical line. *arxiv:1901.04061*, 2019.

[6] S. Bettin, H. M. Bui, X. Li, and M. Radziwiłł. A quadratic divisor problem and moments of the Riemann zeta-function. *J. Eur. Math. Soc.*, to appear, 2020.

[7] R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions*. John Wiley & Sons, New York-London-Sydney, 1976. Wiley Series in Probability and Mathematical Statistics.

[8] E. Bombieri and J. B. Friedlander. Dirichlet polynomial approximations to zeta functions. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 22(3):517–544, 1995.

[9] M. Bramson. Maximal displacement of branching brownian motion. *Comm. Pure Appl. Math.*, 31(5):531–581, 1978.

[10] M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 69(1):62–123, 2016.

[11] N. R. Chaganty and J. Sethuraman. Large deviation local limit theorems for arbitrary sequences of random variables. *Ann. Probab.*, 13(1):97–114, 1985.

[12] R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Math. J.*, 167(12):2243–2345, 2018.

[13] J. Ding and O. Zeitouni. Extreme values for two-dimensional discrete Gaussian free field. *Ann. Probab.*, 42(4):1480–1515, 2014.

[14] Y. V. Fyodorov, G. A. Hiary, and J. P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta function. *Phys. Rev. Lett.*, 108:170601, Apr 2012.

[15] Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2007):20120503, 32, 2014.

[16] G. H. Hardy and J. E. Littlewood. Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. *Acta Math.*, 41(1):119–196, 1916.

[17] G. H. Hardy and J. E. Littlewood. The Approximate Functional Equation in the Theory of the Zeta-Function, with Applications to the Divisor-Problems of Dirichlet and Piltz. *Proc. London Math. Soc. (2)*, 21:39–74, 1923.

[18] A. J. Harper. On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture. *arXiv:1906.05783*, 2019.

[19] A. J. Harper. The Riemann zeta function in short intervals [after Najnudel, and Arguin, Belius, Bourgade, Radziwiłł, and Soundararajan]. *arxiv:1904.08204, Séminaire Bourbaki, to appear in Astérisque*, 2019.

[20] W. Heap, M. Radziwiłł, and K. Soundararajan. Sharp upper bounds for fractional moments of the Riemann zeta function. *Q. J. Math.*, 70(4):1387–1396, 2019.

[21] C. P. Hughes and M. P. Young. The twisted fourth moment of the Riemann zeta function. *J. Reine Angew. Math.*, 641:203–236, 2010.

[22] A. E. Ingham. A Note on Fourier Transforms. *J. London Math. Soc.*, 9(1):29–32, 1934.

[23] C. Joutard. Asymptotic approximation for the probability density function of an arbitrary sequence of random variables. *Statist. Probab. Lett.*, 90:100–107, 2014.

[24] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[25] H. L. Montgomery and R. C. Vaughan. Hilbert's inequality. *J. London Math. Soc. (2)*, 8:73–82, 1974.

[26] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[27] J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018.

[28] E. Paquette and O. Zeitouni. The maximum of the CUE field. *Int. Math. Res. Not. IMRN*, pages 5028–5119, 2018.

[29] Kannan Soundararajan. Moments of the Riemann zeta function. *Ann. of Math. (2)*, 170(2):981–993, 2009.

[30] R. Vanlalngaia. Explicit Mertens sums. *Integers*, 17:Paper No. A11, 18, 2017.

[31] C. Webb. Exact asymptotics of the freezing transition of a logarithmically correlated random energy model. *J. Stat. Phys.*, 145(6):1595–1619, 2011.

Department of Mathematics, Baruch College and Graduate Center, City University of New York, USA
Email address: louis-pierre.arguin@baruch.cuny.edu

Courant Institute, New York University, USA
Email address: bourgade@cims.nyu.edu

Department of Mathematics, Caltech, USA
Email address: maksym.radziwill@gmail.com