MAXIMUM OF THE RIEMANN ZETA FUNCTION ON A SHORT INTERVAL OF THE CRITICAL LINE

L.-P. ARGUIN, D. BELIUS, P. BOURGADE, M. RADZIWIŁŁ, AND K. SOUNDARARAJAN

Abstract. We prove the leading order of a conjecture by Fyodorov, Hiary and Keating, about the maximum of the Riemann zeta function on random intervals along the critical line. More precisely, if t is uniformly distributed in $[T, 2T]$, then

$$\max_{|t-u|\leq 1} \log \left|\zeta\left(\frac{1}{2}+\mathrm{i}u\right)\right| = (1+\mathrm{o}(1))\log\log T$$

with probability converging to 1 as $T \to \infty$.

1. Introduction

1.1. Maximum of the Riemann ζ function on large and short intervals. Bounding the maximum of the Riemann zeta function on the critical line has been the source of many investigations since Lindelöf, who conjectured that for any $\varepsilon > 0$, we have

$$\zeta\left(\frac{1}{2}+\mathrm{i}T\right) = \mathrm{O}(T^{\varepsilon}) \qquad \text{as } T \to \infty.$$

Among the many arithmetic consequences of the Lindelöf hypothesis we highlight the existence of primes in all intervals $[x, x+x^{1/2+\varepsilon}]$ for all x large enough, and in almost all intervals of the form $[x, x+x^{\varepsilon}]$. The current best bound towards the Lindelöf hypothesis states that $|\zeta(\frac{1}{2}+it)| \ll 1+|t|^{13/84+\varepsilon}$, see [7].

Conditionally on the Riemann hypothesis, Littlewood [24] showed that

$$\zeta\left(\frac{1}{2}+\mathrm{i}T\right) = \mathrm{O}\left(\exp\left(C\frac{\log T}{\log\log T}\right)\right) , \tag{1}$$

for some constant $C > 0$. Apart from the value of the constant C [32, 36, 10], this bound remains the best that is known.

There has been more progress on lower bounds for the maximal size of the zeta function. The first result is due to Titchmarsh, who proved that for any $\alpha < \frac{1}{2}$, and large enough T,

$$\max_{t\in[T,2T]} |\zeta(1/2+\mathrm{i}t)| \geq \exp(\log^{\alpha} T) .$$

Date: February 18, 2017.

L.-P. A. is supported by NSF CAREER 1653602, NSF grant DMS-1513441, and a Eugene M. Lang Junior Faculty Research Fellowship.

D. B. is grateful for the hospitality of the Courant Institute during visits when part of this work was carried out.

P. B. is supported by NSF grant DMS-1513587.

M. R. is grateful to NSERC for partially funding this research.

K.S. is partly supported by a grant from the NSF, and a Simons Investigator grant from the Simons Foundation.

This result was improved to

$$\max_{t\in[T,2T]} |\zeta(1/2+\mathrm{i}t)| \geq \exp\left(c\sqrt{\frac{\log T}{\log\log T}}\right)$$

in [26] under the Riemann hypothesis, then the constant c was improved in [4] and, unconditionally, in [35]. The best lower bounds on the maximum were recently obtained in [6]: for any $c < 1/\sqrt{2}$,

$$\max_{t\in[T,2T]} \left|\zeta\left(\frac{1}{2}+\mathrm{i}t\right)\right| \geq \exp\left(c\sqrt{\frac{\log T \log\log\log T}{\log\log T}}\right). \tag{2}$$

Previous results (with differing methods) were established in [26], [4], [35].

From (1) and (2), the actual size order of this maximum is unclear, and its asymptotics are not known to follow from conjectures in number theory. However, interesting probabilistic arguments [13] suggest that

$$\max_{t\in[T,2T]} \log|\zeta(1/2+\mathrm{i}t)| \sim \sqrt{\frac{1}{2}\log T \log\log T}.$$

There are however dissenting views, advocating that (1) is closer to the maximal size of the Riemann zeta function (see end of [13]). These arguments are motivated by Waldspurger's theorem and an analogy between integral weight and half-integral weight modular forms.

Although the global maximum up to height T can hardly be tested numerically, the maximum along random short intervals was very precisely conjectured by Fyodorov, Hiary and Keating, and their prediction is supported by numerics [15, 16]. This conjecture states that if t is random, uniform in $[T, 2T]$, then

$$\max_{|t-u|\leq 1} \log\left|\zeta\left(\frac{1}{2}+\mathrm{i}u\right)\right| = \log\log T - \frac{3}{4}\log\log\log T + X_T\ , \tag{3}$$

where the random variable X_T converges weakly as $T\to\infty$, to an explicit distribution[1]. Our main result is a proof for the first order asymptotics in (3).

Theorem 1.1. *For any $\varepsilon > 0$, as $T\to\infty$ we have*

$$\frac{1}{T}\text{meas}\left\{T\leq t\leq 2T : (1-\varepsilon)\log\log T < \max_{|t-u|\leq 1}\log\left|\zeta\left(\frac{1}{2}+\mathrm{i}u\right)\right| < (1+\varepsilon)\log\log T\right\} \to 1.$$

While completing this work, we learned that the above result (as well as the analogue for $\operatorname{Im}\log\zeta$) was independently proved in [28] under the assumption of the Riemann hypothesis.

1.2. **Extrema of log-correlated fields.** Fyodorov, Hiary and Keating's conjecture was motivated by a connection with random matrices. This analogy has been a source of many investigations, for example at the microscopic level since Montgomery's pair correlation conjecture [25], and the Keating–Snaith conjecture about the moments of the Riemann zeta function [21]. The prediction (3) relies on the analogy at a different, mesoscopic scale. Indeed, Selberg proved that $\log|\zeta|$ is normally distributed on the critical axis, with variance

[1] For convenience we state the conjecture (3) with a random interval of length 2, without loss of generality: only the parameters of the limiting distribution of X_T are supposed to depend on this fixed length.

of order $\frac{1}{2}\log\log T$ [34]. His central limit theorem has been extended to evaluation points at close distance in [8]: for example, if t is uniform on $[T, 2T]$ and $0 < h < 1$ depends on t, the covariance between $\log|\zeta(\frac{1}{2}+\mathrm{i}t)|$ and $\log|\zeta(\frac{1}{2}+\mathrm{i}(t+h))|$ is of order

$$-\frac{1}{2}\log\max\left(h, \frac{1}{\log T}\right), \tag{4}$$

where $(\log T)^{-1}$ is a natural threshold corresponding to the typical spacing between ζ zeros. A parallel story holds for the logarithm of the characteristic polynomial of $N \times N$ Haar-distributed unitary matrices, $\log|\mathrm{P}_N(z)|$: for $|z| = 1$, it is asymptotically Gaussian with variance $\frac{1}{2}\log N$ [21], and for two points on the unit circle within distance $|z_1 - z_2| = h$, the covariance between $\log|\mathrm{P}_N(z_1)|$ and $\log|\mathrm{P}_N(z_2)|$ is of order $-\frac{1}{2}\log\max\left(h, \frac{1}{N}\right)$, analogously to (4), see [8]. Fyodorov, Hiary and Keating gave a very precise conjecture for the maximum of $\{\log|P_N(z)|, |z| = 1\}$ by relying on the replica method, and techniques from statistical mechanics predicting extreme values in disordered systems [14, 17, 18]. Finally, assuming the logarithmic covariance form characterizes the extrema, they proposed the asymptotics (3).

The above Fyodorov-Hiary-Keating analogy, about extreme value theory, has recently been proved in a variety of cases. For a probabilistic model of the Riemann zeta function the leading order of the maximum on short intervals was obtained in [20], and the second order in [3]. For the characteristic polynomial of random unitary matrices, the asymptotics of the maximum at first order [2] and then second order [29] are known, together with tightness of the third order [11] in the more general context of circular beta ensembles. In the context of Hermitian invariant ensembles, the first order of the maximum of the characteristic polynomial was proved in [23] and precise conjectures can be found in [19]. Theorem 1.1 and its conditional analogue in [28] are the first results about the maxima of ζ itself, with the only source of randomness being the choice of the interval. Moreover, in connection with the prediction from [15, 16] that $\log|\zeta|$ behaves like a real log-correlated random field, we note that [33] recently proved that ζ converges to a complex Gaussian multiplicative chaos.

Finally, this work builds on, and adds to, the efforts to develop extreme value theory of correlated systems. Such statistics are expected to lie on the same universality class for any covariance of type (4). This class includes the two-dimensional Gaussian free field, branching random walks, cover times of random walks, Gaussian multiplicative chaos, random matrices and the Riemann zeta function. We do not give here a list of the many rigorous works on this topic in recent years, pointing instead to [1] and the references therein.

1.3. **About the proof.** The short proof of the upper bound in Theorem 1.1 is given in section 2. We only rely on a Sobolev inequality and classical second moment estimates for ζ and ζ'. The proved upper bound is actually stronger than as stated in Theorem 1.1: for any function f diverging to $+\infty$, as $T \to \infty$, we have

$$\frac{1}{T}\mathrm{meas}\left\{\max_{|t-u|\leq 1}\log\left|\zeta\left(\frac{1}{2}+\mathrm{i}u\right)\right| < (\log\log T) + f(T)\right\} \to 1.$$

This result is obtained in a different way in [28], also unconditionally.

The lower bound in Theorem 1.1 is proved in section 3. As a first step, we show that we only need to prove the lower bound for a sufficiently long but finite length Dirichlet sum, which corresponds to the first $\exp((\log T)^{1-\varepsilon})$ terms in the expansion of $\log|\zeta|$ over primes, for some $\varepsilon > 0$. This step makes use of ideas developed to give an alternative proof of Selberg's central limit theorem in [31]. Specifically we need to develop a maximal analogue of the argument in [31]. The first observation is that it suffices to prove the lower bound for the local maxima of $\zeta(\sigma + it)$ with σ slightly off the half-line, as large values off the half-line propagate to the half-line. The second point is the construction of a special mollifier $M(s)$ which is shown to have two properties: On the one hand for almost all t we show that $|\zeta(\sigma + iu)M(\sigma + iu)| \approx 1$ for all $|t - u| \leq 1$. On the other hand we show that for almost all t we have $M(\sigma + iu) \approx \prod_{p\leq X}(1 - p^{-\sigma+iu})^{-1}$ for all $|t - u| \leq 1$ (and some suitably chosen cut-off X). The combination of these two properties implies that for almost all t we have $|\zeta(\sigma + iu)| \approx \prod_{p\leq X}(1 - p^{-\sigma+iu})^{-1}$ for all $|t - u| \leq 1$. Thus the problem is reduced to understanding the local maximum of $\sum_{p\leq X} p^{-\sigma+iu}$, which is a short sum over the primes.

A key idea in the proof of the lower bound for the local maximum of $\sum_{p\leq X} p^{-\sigma+iu}$ is the identification of an approximate branching random walk in this Dirichlet polynomial. This idea was used in [3] to study the extrema of a random model of the zeta function, and in the subsequent works regarding the extremes of characteristic polynomials [2, 11, 23, 29] and the zeta function [28]. The conceptual picture is explained in detail in [3, 2]. Once the branching structure appears, we follow a second moment method which goes back to Bramson's study of branching Brownian motion [9]; specifically we use Kistler's robust multiscale refinement from [22], in a manner close to [2]. This method requires large deviation estimates and speed of convergence to the normal distribution for our Dirichlet sums, when evaluated at distinct points. For that purpose, the Fourier and Laplace transforms with high arguments are evaluated through an expansion over moments. Large moments are typically hard to approximate, a problem we overcome here following a decomposition from [30]: over a good subset B of $[T, 2T]$ the first $\log\log T$ moments provide an accurate approximation of the characteristic function, and the complement of B has small enough measure.

It should be possible to obtain more refined results by pushing the parameter K to infinity, but we have not attempted to carry this out.

For the rest of the paper, we will think of t has a uniform random variable on $[T, 2T]$. We will write accordingly $\mathbb{P}$ for $\frac{1}{T}$meas and $\mathbb{E}$ for $\frac{1}{T}\int_T^{2T}$. We will also write $f(T) = \mathrm{O}(g(T))$ if $\frac{|f(T)|}{|g(T)|}$ is bounded and $f(T) = \mathrm{o}(g(T))$ if $\frac{|f(T)|}{|g(T)|} \to 0$. Finally, we will sometimes write for short $f(T) \ll g(T)$ when $f(T) = \mathrm{O}(g(T))$.

2. Proof of the upper bound

The upper bound is a direct consequence of unconditional moment bounds and of the Sobolev-type inequality in Lemma A.1.

Proposition 2.1. *If $V = V(T)$ tends to infinity as $T \to \infty$, then*

$$\mathbb{P}\Big(\max_{|t-u|\leq 1} |\zeta(1/2 + \mathrm{i}u)| > V \log T\Big) = \mathrm{O}(1/V^2) = \mathrm{o}(1)\ .$$

Proof. Chebyshev's inequality implies

$$\mathbb{P}\Big(\max_{|t-u|\leq 1} |\zeta(1/2 + \mathrm{i}u)| > V \log T\Big) \leq \frac{1}{V^2(\log T)^2}\mathbb{E}\left[\max_{|t-u|\leq 1} |\zeta(1/2 + \mathrm{i}u)|^2\right]\ . \tag{5}$$

Lemma A.1 bounds the second moment of the maximum by

$$\mathbb{E}\left[\max_{|t-u|\le 1} |\zeta(1/2+\mathrm{i}u)|^2\right] \ll \left(\mathbb{E}\left[|\zeta(1/2+\mathrm{i}u)|^2\right]\right)^{1/2}\cdot\left(\mathbb{E}\left[|\zeta'(1/2+\mathrm{i}u)|^2\right]\right)^{1/2}+\mathbb{E}\left[|\zeta(1/2+\mathrm{i}u)|^2\right] .$$

The following bounds of the second moment of ζ and its derivative are known unconditionally (see, e.g., [37] and [12]):

$$\begin{aligned}\mathbb{E}\big[|\zeta(1/2+\mathrm{i}t)|^2\big] &\ll \log T\\ \mathbb{E}\big[|\zeta'(1/2+\mathrm{i}t)|^2\big] &\ll (\log T)^3 .\end{aligned} \tag{6}$$

We conclude that

$$\mathbb{E}\left[\max_{|t-u|\le 1} |\zeta(1/2+\mathrm{i}u)|^2\right] \ll (\log T)^2 .$$

The proposition follows from this and (5). □

3. Proof of the lower bound

The lower bound of Theorem 1.1 is proved in two main steps. First, it is shown that the maximum on a short interval of $\log|\zeta|$ is close to the maximum of a Dirichlet polynomial slightly off the critical axis. This is the content of Proposition 3.1. Second, a lower bound for the maximum of Dirichlet polynomials on an interval is proved using the robust approach of [22] in Proposition 3.2.

The following notation will be used throughout the section. Following [22], we will fix a large integer $K = K(\varepsilon)$. For this K, we take

$$\sigma_0 = \frac{1}{2} + \frac{(\log T)^{\frac{3}{2K}}}{\log T} . \tag{7}$$

The primes will be divided into ranges as follows

$$J_j = [\exp((\log T)^{\frac{j}{K}}), \exp((\log T)^{\frac{j+1}{K}})], \qquad j = 1,\dots,K-2, \tag{8}$$

and $J_0 = [2, \exp((\log T)^{\frac{1}{K}})]$. We also denote

$$X = \exp((\log T)^{1-\frac{1}{K}}) . \tag{9}$$

Finally, the relevant Dirichlet polynomials are

$$P_j(u) = \mathrm{Re}\sum_{p\in J_j} \frac{1}{p^{\sigma_0+\mathrm{i}u}}, \qquad j = 0,\dots,K-2 . \tag{10}$$

The lower bound in the theorem follows from the two main propositions stated below.

Proposition 3.1. *For any $\varepsilon > 0$ and $K = K(\varepsilon)$ large enough,*

$$\mathbb{P}\left(\max_{|t-u|\le 1} \log|\zeta(1/2+\mathrm{i}u)| > (1-2\varepsilon)\log\log T\right) \ge \mathbb{P}\left(\max_{|t-u|\le \frac{1}{4}} \sum_{j=1}^{K-3} P_j(u) > (1-\varepsilon)\log\log T\right) + \mathrm{o}(1).$$

Proposition 3.2. *For any $K > 3$ and $0 < \lambda < 1$,*

$$\mathbb{P}\left(\exists u :\ |t-u| \le \frac{1}{4}, P_j(u) > \frac{\lambda}{K}\log\log T \text{ for all } 1 \le j \le K-3\right) = 1 + \mathrm{o}(1). \tag{11}$$

Proof of Theorem 1.1. On the event of Proposition 3.2, we have

$$\sum_{j=1}^{K-3} P_j(u) > \lambda\left(1-\frac{3}{K}\right)\log\log T \text{ for some } u \text{ with } |t-u| \leq \frac{1}{4}.$$

Therefore it suffices to take K large and λ close to 1 in terms of ε to get the lower bound of the theorem using Proposition 3.1. □

3.1. **Proof of Proposition 3.1.** The proof of the proposition is divided into three lemmas. In the first, we bound the maximum on the critical axis by the ones off axis.

Lemma 3.3. *Let $\varepsilon > 0$, $V > 2$ and $\frac{1}{2} \leq \sigma \leq \frac{1}{2} + (\log T)^{-1/2-\varepsilon}$. Then*

$$\mathbb{P}\left(\max_{|t-u|\leq 1} |\zeta\left(1/2+\mathrm{i}t\right)| > V\right) \geq \mathbb{P}\left(\max_{|t-u|\leq \frac{1}{4}} |\zeta\left(\sigma+\mathrm{i}t\right)| > 2V\right) + \mathrm{o}(1).$$

Proof. Recall that, for any $t \in \mathbb{R}$, we can write $\zeta(\sigma+\mathrm{i}t)$ as an average on the critical line using the Poisson kernel:

$$\zeta(\sigma+\mathrm{i}t) = \int_{-\infty}^{\infty} \zeta\left(1/2+\mathrm{i}u\right)\cdot\frac{1}{\pi}\frac{\sigma-1/2}{(u-t)^2+(\sigma-1/2)^2}\mathrm{d}u \ + O(T^{-\delta}) \tag{12}$$

for some $\delta > 0$. This is proven by recalling that for $t \in [T, 2T]$ and $\sigma \geq \frac{1}{2}$ we have $\zeta(\sigma+\mathrm{i}t) = \sum_{n\leq T} n^{-\sigma-\mathrm{i}t} + O(T^{-1/2})$ and by observing that for any finite Dirichlet polynomial $D(s) = \sum_{n\leq N} a(n)n^{-s}$ we have,

$$D(\sigma+\mathrm{i}t) = \int_{-\infty}^{\infty} D(\tfrac{1}{2}+\mathrm{i}u)\cdot\frac{1}{\pi}\cdot\frac{(\sigma-\frac{1}{2})\mathrm{d}u}{(\sigma-\frac{1}{2})^2+(u-t)^2}.$$

We denote the above Poisson kernel by $f_t(u)$.

Consider now $t \in [T, 2T]$ such that $\max_{|v|\leq\frac{1}{4}} |\zeta\left(\sigma+\mathrm{i}(t+v)\right)| > 2V$, with $V > 2$. We write $v^\star = v^\star(t)$ for the $v \in [-\frac{1}{4}, \frac{1}{4}]$ that achieves the maximum for this t. Equation (12) implies

$$2V < |\zeta(\sigma+\mathrm{i}(t+v^\star))| \leq \int_{-\infty}^{\infty} |\zeta\left(1/2+\mathrm{i}(u+v^\star)\right)| \ f_t(u)\mathrm{d}u \ .$$

In particular, for this t, it must be that

$$\int_{t-1/2}^{t+1/2} |\zeta\left(1/2+\mathrm{i}(u+v^\star)\right)| \ f_t(u)\mathrm{d}u > V \ ,$$

or

$$\int_{[t-\frac{1}{2},t+\frac{1}{2}]^c} |\zeta\left(1/2+\mathrm{i}(u+v^\star)\right)| \ f_t(u)\mathrm{d}u > V-1 \ .$$

Since $|v^\star| \leq \frac{1}{4}$ and $f_t(u)$ is a probability density centered at t, the first case implies that $\max_{|t-u|\leq 1} |\zeta\left(1/2+\mathrm{i}t\right)| > V$. Therefore it remains to prove that the set

$$\{t : \int_{[t-\frac{1}{2},t+\frac{1}{2}]^c} |\zeta\left(1/2+\mathrm{i}(u+v^\star)\right)| \ f_t(u)\mathrm{d}u > V-1\} \ \text{ has small probability.}$$

Again, since $|v^\star| \le \frac{1}{4}$, it suffices to show that the probability of

$$\int_{[t-\frac{1}{4},t+\frac{1}{4}]^c} |\zeta\left(1/2+\mathrm{i}u\right)|\ f_t(u)\mathrm{d}u > V-1 \text{ is small }.$$

By Chebyshev's inequality, the probability is smaller than

$$\begin{aligned}&\frac{(\sigma-1/2)^2}{\pi^2V^2}\mathbb{E}\left[\left(\int_{|t-u|>1/4}\frac{|\zeta\left(1/2+\mathrm{i}u\right)|}{(u-t)^2+(\sigma-1/2)^2}\mathrm{d}u\right)^2\right]\\ &\ll (\log T)^{-1-2\varepsilon}\mathbb{E}\left[\int_{[-\frac{1}{4},\frac{1}{4}]^c}|\zeta\left(1/2+\mathrm{i}(t+v)\right)|^2\ \frac{\mathrm{d}v}{v^2}\right]\ ,\end{aligned}$$

where the second inequality follows by definition of σ and Jensen's inequality applied with the probability measure $\mathbf{1}_{[-\frac{1}{4},\frac{1}{4}]^c}\frac{1}{8}\frac{\mathrm{d}v}{v^2}$. The last expression is $\mathrm{O}((\log T)^{-2\varepsilon})$. To see this, note that $\mathbb{E}[|\zeta\left(1/2+\mathrm{i}(t+v)\right)|^2]$ is $\mathrm{O}(\log T)$ for $|v|\le T/2$, say, by (6). For $|v|>T/2$, one can use the fact that $\zeta(1/2+it)=\mathrm{O}(t^{1/4})$ (see, e.g. Theorem 5.12 in [37]) to conclude that the integral over this range of v is $\mathrm{O}(1)$. □

For the second lemma, we take the following approximation for the inverse of ζ:

$$M(s)=\sum_n\frac{\mu(n)a(n)}{n^s}\ , \tag{13}$$

where for n square-free $\mu(n)=(-1)^{\omega(n)}$ and $\omega(n)$ is the number of distinct prime factors, and for n non-square-free $\mu(n)=0$. The factor $a(n)$ equals 1 if all primes factors of n are smaller than X and $\Omega(n)\le 100K\log\log T=:\nu$, where $\Omega(n)$ is the number of primes factors of n. The following shows that the maxima of $|\zeta|$ and M^{-1} are close.

Lemma 3.4. *For any $\varepsilon>0$, we have*

$$\mathbb{P}\left(\max_{|t-u|\le 1}|M(\sigma_0+\mathrm{i}u)\zeta(\sigma_0+\mathrm{i}u)-1|>\varepsilon\right)=\mathrm{o}(1)\ .$$

Proof. We closely follow [31, Section 4], adapting it to different scalings. Assume we can prove that, uniformly in $\sigma>\frac{1}{2}+(\log T)^{-1+\frac{1}{K}+\varepsilon}$,

$$\int_T^{2T}|M(\sigma+\mathrm{i}t)\zeta(\sigma+\mathrm{i}t)-1|^2\,\mathrm{d}t=\mathrm{o}(T). \tag{14}$$

Then Lemma (3.4) follows by Chebyshev's inequality and Lemma A.2.

To prove (14), we first consider the cross term in the square expansion. We integrate along the rectangular contour with vertices $\sigma+\mathrm{i}T,\sigma+\mathrm{i}2T,2+\mathrm{i}T,2+\mathrm{i}2T$, and use the estimates $\zeta=\mathrm{O}(T^{1/4})$ [37] and $M=\mathrm{O}(T^\varepsilon)$ on the horizontal parts. This gives

$$\int_T^{2T}\zeta(\sigma+\mathrm{i}t)M(\sigma+\mathrm{i}t)\mathrm{d}t=\int_T^{2T}\zeta(2+\mathrm{i}t)M(2+\mathrm{i}t)\mathrm{d}t+\mathrm{O}(T^{1/4+\varepsilon})=T+\mathrm{O}(T^{1/4+\varepsilon}),$$

where we used the simple expansion

$$\int_T^{2T}\zeta(2+\mathrm{i}t)M(2+\mathrm{i}t)\mathrm{d}t=T+\sum_{m,n:mn\neq 1}\frac{a(m)\mu(m)}{(mn)^2}\int_T^{2T}(mn)^{-\mathrm{i}t}\mathrm{d}t=T+\mathrm{O}(1).$$

We therefore proved

$$\int_T^{2T} |M(\sigma+\mathrm{i}t)\zeta(\sigma+\mathrm{i}t)-1|^2\,\mathrm{d}t = \int_T^{2T} |M(\sigma+\mathrm{i}t)\zeta(\sigma+\mathrm{i}t)|^2\,\mathrm{d}t - T + \mathrm{O}(T^{1/4+\varepsilon}), \tag{15}$$

and we now turn to the evaluation of the above second moment:

$$\int_T^{2T} |M(\sigma+\mathrm{i}t)\zeta(\sigma+\mathrm{i}t)|^2\,\mathrm{d}t = \sum_{h,k}\frac{\mu(h)\mu(k)a(h)a(k)}{(hk)^\sigma}\int_T^{2T}\left(\frac{h}{k}\right)^{\mathrm{i}t}|\zeta(\sigma+\mathrm{i}t)|^2\,\mathrm{d}t\ .$$

It can be estimated based on the following result (see [31, Lemma 4]): for any $h,k\leq T$ and $1/2<\sigma\leq 1$, we have

$$\int_T^{2T}\left(\frac{h}{k}\right)^{\mathrm{i}t}|\zeta(\sigma+\mathrm{i}t)|^2\,\mathrm{d}t = \int_T^{2T}\left(\zeta(2\sigma)\left(\frac{(h,k)^2}{hk}\right)^{\sigma}+\left(\frac{t}{2\pi}\right)^{1-2\sigma}\zeta(2-2\sigma)\left(\frac{(h,k)^2}{hk}\right)^{1-\sigma}\right)\mathrm{d}t + \mathrm{O}(T^{1-\sigma+\varepsilon}\min(h,k)). \tag{16}$$

The contribution of the above error term is bounded by

$$T^{1-\sigma+\varepsilon}\sum_{h,k}a(h)a(k)\frac{\min(h,k)}{(hk)^\sigma}\leq T^{1-\sigma+\varepsilon}\left(\sum_h a(h)\right)^2\leq T^{1-\sigma+\varepsilon}\left(\sum_{p\leq X}1\right)^{2\nu}=\mathrm{O}\left(T^{1-\sigma+2\varepsilon}\right).$$

The contribution of the first term in (16) is

$$T\zeta(2\sigma)\sum_{h,k}\frac{a(h)a(k)\mu(h)\mu(k)}{(hk)^{2\sigma}}(h,k)^{2\sigma}\ . \tag{17}$$

If we omit the constraint $\Omega(h),\Omega(k)\leq\nu$ in the above sum, we obtain (for A a subset of $\mathcal{P}_X$, the set of primes smaller than X, we write $\pi_A=\prod_{p\in A}p$):

$$\sum_{h,k:p|hk\Rightarrow p\leq X}\frac{\mu(h)\mu(k)}{(hk)^{2\sigma}}(h,k)^{2\sigma} = \sum_{A,B\subset\mathcal{P}_X}\frac{(-1)^{|A|+|B|}}{(\pi_A\pi_B)^{2\sigma}}\pi_{A\cap B}^{2\sigma}$$
$$= \sum_{A\subset\mathcal{P}_X,C\subset A,D\subset(\mathcal{P}_X\setminus A)}\frac{(-1)^{|A|+|C|+|D|}}{(\pi_A\pi_D)^{2\sigma}}=\prod_{p\leq X}\left(1-\frac{1}{p^{2\sigma}}\right)$$

where we decomposed $B=C\cup D$ and used $\sum_{C\subset A}(-1)^{|C|}=0$ unless $A=\emptyset$. The error consisting in terms such that $\Omega(h)>\nu$ or $\Omega(k)>\nu$ is bounded using (58) by

$$\sum_{\substack{h,k:p|hk\Rightarrow p\leq X\\ \Omega(h)>\nu}}\frac{|\mu(h)\mu(k)|(h,k)^{2\sigma}}{(hk)^{2\sigma}}\leq e^{-\nu}\sum_{h,k:p|hk\Rightarrow p\leq X}\frac{|\mu(h)\mu(k)|(h,k)^{2\sigma}}{(hk)^{2\sigma}}e^{\Omega(h)}$$

$$=e^{-\nu}\sum_{A\subset\mathcal{P}_X,D\subset(\mathcal{P}_X-A)}\frac{e^{|A|}}{(\pi_A\pi_D)^{2\sigma}}\leq e^{-\nu}\left(\sum_{A\subset\mathcal{P}_X}\frac{e^{|A|}}{\pi_A^{2\sigma}}\right)^2=e^{-\nu}\prod_{p\in\mathcal{P}_X}\left(1+\frac{e}{p^{2\sigma}}\right)^2=\mathrm{O}\left(\frac{1}{(\log T)^{98}}\right)$$

where we used (54). We therefore proved that

$$(17) = T\zeta(2\sigma)\left(\prod_{p\le X}\left(1-\frac{1}{p^{2\sigma}}\right)+\mathrm{O}\left(\frac{1}{(\log T)^{98}}\right)\right)$$
$$= T\left(\prod_{p>X}\left(1-\frac{1}{p^{2\sigma}}\right)^{-1}+\mathrm{O}\left(\frac{1}{(2\sigma-1)(\log T)^{98}}\right)\right) = T+\mathrm{O}\left(\frac{T}{(\log T)^{96}}\right),$$

where we used (53) to estimate $\log\prod_{p>X}\left(1-\frac{1}{p^{2\sigma}}\right)^{-1} = \sum_{p>X}p^{-2\sigma}+\mathrm{O}(1/X)$. This is negligible whenever $(2\sigma-1)\log X\to\infty$ (in the case of interest, $(2\sigma_0-1)\log X = (\log T)^{1/(2K)}$). Finally, the contribution from the second term in (16) is shown to be negligible as in [31]. □

For the last reduction, define the Dirichlet polynomials

$$\mathcal{P}_j(s) = \sum_{n\in J_j}\frac{\Lambda(n)}{n^s\log n}, \qquad j=0,\dots,K-2\ . \tag{18}$$

Then again, the maxima of M and $\exp(-\sum_j\mathcal{P}_j)$ are close.

Lemma 3.5. *We have*

$$\mathbb{P}\left(\max_{|t-u|\le 1}\left|M(\sigma_0+\mathrm{i}u)-\exp\left(\sum_{j=0}^{K-2}-\mathcal{P}_j(\sigma_0+\mathrm{i}u)\right)\right| > (\log T)^{-2}\right) = \mathrm{o}(1)\ .$$

Proof. We defined the truncated exponential

$$\mathcal{M}(s) = \sum_{k\le\nu}\frac{(-1)^k}{k!}\Big(\sum_j\mathcal{P}_j(s)\Big)^k\ . \tag{19}$$

Consider the good set

$$\mathcal{U} = \left\{t:\max_{|t-u|\le 1}|\mathcal{P}_j(\sigma_0+\mathrm{i}u)|\le\frac{10}{K^{1/2}}\log\log T,\ \forall 0\le j\le K-3\right\}\ . \tag{20}$$

(The choice of upper bound is motivated by the logarithmic correlations (4). These imply that, for each j, the values of $\mathcal{P}_j$ are almost perfectly correlated for points at a distance less than $(\log T)^{-\frac{j+1}{K}}$, and close to independent for points farther than $(\log T)^{-\frac{j}{K}}$ apart. Therefore, at a heuristic level, one expects that the maximum of $\mathcal{P}_j$ corresponds to the one of $(\log T)^{\frac{j+1}{K}}$ independent Gaussian variables of variance $\frac{1}{K}\log\log T$, which is of the order of $\frac{\sqrt{j+1}}{K}\log\log T$. Here we pick a bound that holds simultaneously for all j.) Note that on the set $\mathcal{U}$, $\mathcal{M}$ is arbitrarily close to $\exp(-\sum_j\mathcal{P}_j)$ since

$$\sum_{k>\nu}\frac{(-1)^k}{k!}\Big(\sum_j\mathcal{P}_j(s)\Big)^k \ll \sum_{k>\nu}\frac{1}{k!}(\log\log T)^k \ll (\log T)^{-100}\ .$$

Moreover, the complement has small probability, since by Chebyshev's inequality applied with $\ell = \log\log T$ and the Sobolev's inequality (A.2) with $L = \log T$, we get

$$\mathbb{P}(\mathcal{U}^c) \leq \sum_j \frac{1}{(\frac{10}{K^{1/2}}\log\log T)^{2\ell}} \mathbb{E}\left[\max_{|t-u|\leq 1} |\mathcal{P}_j(\sigma_0 + \mathrm{i}u)|^{2\ell}\right]$$
$$\ll \sum_j \frac{L}{(\frac{10}{K^{1/2}}\log\log T)^{2\ell}} \sup_{|\sigma-\sigma_0|\leq L^{-1}} \mathbb{E}\left[|\mathcal{P}_j(\sigma + \mathrm{i}u)|^{2\ell}\right] + \mathrm{o}(1) \ .$$

The bound on the moments in Lemma B.2 yields $\mathbb{E}\left[|\mathcal{P}_j(\sigma+\mathrm{i}u)|^{2\ell}\right] \ll \ell!(\frac{1}{K}\log\log T)^\ell$ for all j uniformly for $|\sigma - \sigma_0| \leq L^{-1}$. We conclude that

$$\mathbb{P}(\mathcal{U}^c) \ll \log T \ \ell^\ell e^{-\ell}\ell^{1/2} \frac{(\frac{1}{K}\log\log T)^\ell}{(\frac{10}{K^{1/2}}\log\log T)^{2\ell}} \ll (\log T)^{-4} \ .$$

With the above observations, the proof of the lemma is reduced to showing that

$$\mathbb{P}\left(\max_{|t-u|\leq 1} |M(\sigma_0+\mathrm{i}u) - \mathcal{M}(\sigma_0+\mathrm{i}u)| > (\log T)^{-3}\right) = \mathrm{o}(1) \ .$$

Again, this follows by Chebyshev's inequality and the Sobolev's inequality (A.2) if

$$\mathbb{E}\left[|M(\sigma+\mathrm{i}u) - \mathcal{M}(\sigma+\mathrm{i}u)|^2\right] = o(L^{-4}) \text{ uniformly for } |\sigma-\sigma_0| \leq L^{-1}.$$

To see this, note that $\mathcal{M}$ can be written as a Dirichlet sum $\sum_n b(n)n^{-s}$ where $|b(n)| \leq 1$, by expanding the powers of $\mathcal{P}_j$. Moreover $b(n) = a(n)\mu(n)$ whenever $\Omega(n) \leq \nu$, and $b(n) = 0$ if $n \leq X^\nu$. (This is because the difference between M and $\mathcal{M}$ only comes from the prime powers in $\mathcal{P}_j$.) Thus if we write $c(n) = b(n) - \mu(n)a(n)$, the L^2-difference becomes

$$\ll \sum_{m,n} \frac{|c(n)c(m)|}{(mn)^\sigma} \mathbb{E}[(m/n)^{it}] \ .$$

If $m \neq n$, since $\log(m/n) \gg (mn)^{-1/2}$, the sum is smaller than $X^{2\nu}/T$ (by an upper bound on the number of terms). If $m = n$, then we have that the sum is

$$\ll \sum_{\substack{n \\ \Omega(n)>\nu \\ p\backslash n \Rightarrow p\leq X}} n^{-1} \ll r^{-\nu} \prod_{p\leq X}\left(1 + \frac{r}{p} + \sum_{j>1}\left(\frac{r}{p}\right)^j\right) \ll (\log T)^{-99r}$$

where we used (58) with $1 < r < 2$. This concludes the proof of the lemma. □

Proof of Proposition 3.1. Note that the difference between the Dirichlet polynomials P_j and $\mathcal{P}_j$ is given by the prime powers

$$Q(\sigma_0+\mathrm{i}u) := \sum_{j=0}^{K-2} \mathcal{P}_j(u) - \sum_{p\leq X} \frac{1}{p^{\sigma_0+\mathrm{i}u}} = \frac{1}{2}\sum_{p\leq X} \frac{1}{p^{2\sigma_0+2\mathrm{i}u}} + \mathrm{O}(1) \ ,$$

since the higher powers are summable by (53). Therefore by Chebyshev's inequality and the Sobolev inequality A.1 we have

$$\mathbb{P}(\max_{|t-u|\leq\frac{1}{4}}|Q(u)|>\log\log\log T)$$
$$\ll\frac{1}{(\log\log\log T)^2}\left(\mathbb{E}[|Q(2\sigma_0+\mathrm{i}t)|^2]+(\mathbb{E}[|Q(2\sigma_0+\mathrm{i}t)|^2])^{1/2}(\mathbb{E}[|Q(2\sigma_0+\mathrm{i}t)|^2])^{1/2}\right)\ .$$

A quick calculation with Lemma B.2 shows that the parenthesis is $\mathrm{O}(1)$, so that the probability goes to 0. Furthermore, we have on the set $\mathcal{U}$ in (20) that $\max_{|t-u|\leq 1/4}|\mathcal{P}_j|\leq\frac{10}{K^{1/2}}\log\log T$ for $j=0$ and $K-2$ with probability going to 1. By taking $K=K(\varepsilon)$ large enough, these two observations imply

$$\mathbb{P}\left(\max_{|t-u|\leq\frac{1}{4}}\mathrm{Re}\sum_{j=0}^{K-2}\mathcal{P}_j(u)\geq(1-2\varepsilon)\log\log T\right)\geq\mathbb{P}\left(\max_{|t-u|\leq\frac{1}{4}}\sum_{j=1}^{K-3}P_j(u)\geq(1-\varepsilon)\log\log T\right)+\mathrm{o}(1)\ .$$

The complements of the two events appearing in Lemma 3.4 and 3.5 have probability $1+\mathrm{o}(1)$, and so does their intersection. (The two lemmas were proved for $|t-u|\leq 1$ but holds the same way for a smaller interval.) If u is such that $\mathrm{Re}\sum_{j=0}^{K-2}\mathcal{P}_j(u)\geq(1-3\varepsilon/2)\log\log T$, we must have on the intersection of the complements

$$|\zeta(\sigma_0+\mathrm{i}u)|\geq|M^{-1}(\sigma_0+\mathrm{i}u)|\geq|(\log T)^{-2}+e^{-\sum_j\mathcal{P}_j(u)}|^{-1}\geq(1+\mathrm{o}(1))(\log T)^{1-2\varepsilon})\ .$$

Altogether, we have so far shown

$$\mathbb{P}\left(\max_{|t-u|\leq\frac{1}{4}}|\zeta(\sigma_0+\mathrm{i}u)|\geq(\log T)^{1-2\varepsilon}\right)\geq\mathbb{P}\left(\max_{|t-u|\leq\frac{1}{4}}\sum_{j=1}^{K-3}P_j(u)\geq(1-\varepsilon)\log\log T\right)+\mathrm{o}(1)\ .$$

To conclude, it suffices to redefine ε and apply Lemma 3.3 with σ_0 and $V=(\log T)^{1-2\varepsilon}$. □

3.2. **Proof of Proposition 3.2.** The proof of the proposition is based on large deviation estimates of the variables $P_j(u)$ as defined in (10) for σ_0 given in (7), see Propositions 3.9 and 3.10. Proposition 3.2 is stated for the interval $|t-u|\leq 1/4$ as needed for the proof of the theorem. Since the size of the interval does not matter in the proof, we take $|t-u|\leq 1$ for simplicity. The large deviation estimates are derived using large moments of Dirichlet polynomials given in Lemma B.1, and by estimating the Laplace-Fourier transform of the polynomials using the moments, see Proposition 3.8.

The first step is to show that the moments of sums of P_j's are very close to Gaussian moments.

Proposition 3.6. *Let* $(\xi_j,1\leq j\leq K-3)$ *and* $(\xi'_j,1\leq j\leq K-3)$ *in* $\mathbb{C}$ *such that* $|\xi_j|,|\xi'_j|\leq(\log T)^{\frac{1}{16K}}$. *Then for any* $n\leq(\log T)^{\frac{1}{2K}}$, *we have for* $|\tau|,|\tau'|\leq 1$,

$$\begin{aligned}&\mathbb{E}\left[\left(\sum_{j=1}^{K-3}\{\xi_jP_j(t+\tau)+\xi'_jP_j(t+\tau')\}\right)^{2n}\right]\\ &=(2n-1)!!\ \left(\sum_{j=1}^{K-3}\{s_j^2(\xi_j^2+\xi'^2_j)+2\rho_j(\tau,\tau')\xi_j\xi'_j\}\right)^n+\mathrm{O}(e^{-(\log T)^{\frac{1}{2K}}})\end{aligned}\tag{21}$$

for

$$s_j^2 = \frac{1}{2}\sum_{p\in J_j} p^{-2\sigma_0} \text{ and } \rho_j(\tau,\tau') = \frac{1}{2}\sum_{p\in J_j} p^{-2\sigma_0}\cos(|\tau-\tau'|\log p). \tag{22}$$

The odd moments are $\mathrm{O}(e^{-(\log T)^{\frac{1}{2K}}})$ *under the same condition on* n.

Note that without the error term the right-hand side of (21) is precisely what the n-th moment would be if $P_j(t+\tau)$, $P_j(t+\tau')$ were jointly Gaussian with variance s_j^2 and covariance $\rho_j(\tau,\tau')$, and uncorrelated for different j. Using (53), one gets the bounds for $j\geq 1$

$$2s_j^2 = \frac{\log\log T}{K} + \mathrm{O}((\log T)^{-\frac{1}{2K}}),\ \rho_j(\tau,\tau') = \mathrm{O}\left((\log T)^{-\frac{1}{2K}}\right) \text{ if } |\tau-\tau'|\geq (\log T)^{-\frac{j+1}{m}}. \tag{23}$$

Of course $-s_j^2\leq \rho_j(\tau,\tau')\leq s_j^2$ and in fact $\rho_j(\tau,\tau')$ is close to s_j^2 if $|\tau-\tau'|\leq (\log T)^{-\frac{j-1}{m}}$, since then the cosine in the sum is close to 1.

Proof of Proposition 3.6. Let $a(p)=a^\star(p)=0$ for $p\notin\cup_j J_j$ and

$$a(p) = (\xi_j p^{-i\tau}+\xi_j' p^{-i\tau'})p^{-\sigma_0}\ , \qquad a^\star(p) = (\xi_j p^{i\tau}+\xi_j' p^{i\tau'})p^{-\sigma_0} \qquad \text{for } p\in J_j.$$

Then we have

$$\sum_{j=1}^{K-3}\{\xi_j P_j(t+\tau)+\xi_j' P_j(t+\tau')\} = \frac{1}{2}\sum_p\{a(p)p^{-it}+a^\star(p)p^{it}\}$$

$$a(p)a^\star(p) = (\xi_j^2+{\xi_j'}^2+2\xi_j\xi_j'\cos(|\tau-\tau'|\log p))p^{-2\sigma_0} \text{ if } p\in J_j.$$

In this case, the function $\mathcal{J}$ appearing in Lemma B.1 takes the form

$$\prod_p \mathcal{J}(a(p)a^\star(p)z^2) = \prod_p\left(1+\frac{a(p)a^\star(p)z^2}{4}+\mathrm{O}(|a(p)a^\star(p)|^2z^4)\right)$$

$$= \exp\left(\frac{z^2}{2}\left(\frac{1}{2}\sum_j\sum_{p\in J_j} p^{-2\sigma_0}(\xi_j^2+{\xi_j'}^2+2\xi_j\xi_j'\cos(|\tau-\tau'|\log p))\right)\right)F_X(z)\ ,$$

for $F_X(z)$ a function which is analytic in a neighborhood of 0, satisfies $F_X(0)=1$, and whose derivatives at 0 are uniformly bounded by

$$\sum_{j=1}^{K-3}\sum_{p\in J_j}|a(p)a^\star(p)|^2 \ll (\log T)^{\frac{1}{8K}}\sum_{j=1}^{K-3}\sum_{p\in J_j}p^{-2} \ll (\log T)^{\frac{1}{8K}}e^{-(\log T)^{\frac{1}{K}}}\ .$$

The claim (21) follows from Lemma B.1 by taking the n-th derivative (note that the exponential term is exactly the moment generating function of a Gaussian) and noting that the terms involving a derivative of $F_X(z)$ contribute at most $\mathrm{O}(e^{-(\log T)^{\frac{1}{2K}}})$. Also the error term from Lemma B.1 is $X^{2n}/T\ll e^{-(\log T)^{\frac{1}{2K}}}$ under the assumption $n\leq(\log T)^{\frac{1}{2K}}$.

□

It is necessary to introduce a cutoff to obtain a precise comparison of the Dirichlet polynomials with Gaussians. With this in mind, we introduce the set

$$B(\tau) = \{T\leq t\leq 2T : P_j(t+\tau)\leq(\log T)^{\frac{1}{4K}},\ \forall 1\leq j\leq K-3\}\ . \tag{24}$$

Note that if the moments of a random variable Y grow no faster than Gaussian moments for some $\sigma^2 > 0$ up to $\ell \le L$, then

$$\mathbb{P}(Y > x) \ll \exp(-\frac{x^2}{2\sigma^2}) \qquad \text{whenever } \frac{x^2}{2\sigma^2} \le L. \tag{25}$$

This is proved using Chebyshev's inequality and optimizing over ℓ. Together with Proposition 3.6 this observation yields the following:

Lemma 3.7. *Let $|\tau| \le 1$. We have*

$$\mathbb{P}(B(\tau)^c) \ll \exp\left(-\frac{(\log T)^{\frac{1}{2K}}}{\log\log T}\right) . \tag{26}$$

Proof. A union bound gives $\mathbb{P}(B(\tau)^c) \le \sum_{j=1}^{K-3} \mathbb{P}(P_j(t+\tau) > (\log T)^{\frac{1}{4K}})$. Recall that K is fixed. We use the Gaussian bound (25) and (23). Note that $\frac{\left((\log T)^{1/4K}\right)^2}{2s_j^2} \le (\log T)^{1/2K}$. This yields the claimed bound. □

On $B(\tau)$, we can derive precise bounds for the Fourier-Laplace transforms of the P_j's.

Proposition 3.8. *Let $\xi_j \in \mathbb{C}$, $1 \le j \le K-3$, such that $|\xi_j| \le (\log T)^{\frac{1}{16K}}$. Then for $|\tau| \le 1$,*

$$\mathbb{E}\left[\exp\left(\sum_{j=1}^{K-3} \xi_j P_j(t+\tau)\right) \ \mathbf{1}_{B(\tau)}\right] = (1 + \mathrm{O}(e^{-(\log T)^{\frac{1}{8K}}})) \ \exp\left(\frac{1}{2}\sum_{j=1}^{K-3} \xi_j^2 s_j^2\right) . \tag{27}$$

Moreover, if $\tau \ne \tau'$, then for $\xi'_j \in \mathbb{C}$, $1 \le j \le K-3$, with $|\xi'_j| \le (\log T)^{\frac{1}{16K}}$, we also have

$$\begin{aligned} &\mathbb{E}\left[\exp\left(\textstyle\sum_{j=1}^{K-3} \xi_j P_j(t+\tau) + \xi'_j P_j(t+\tau')\right)\right) \ \mathbf{1}_{B(\tau)\cap B(\tau')}\Big] \\ &= (1 + \mathrm{O}(e^{-(\log T)^{\frac{1}{8K}}})) \exp\left(\tfrac{1}{2}\textstyle\sum_{j=1}^{K-3} \{s_j^2(\xi_j^2 + {\xi'}_j^2) + 2\rho_j(\tau,\tau')\xi_j\xi'_j\}\right) . \end{aligned} \tag{28}$$

Proof. The one-point bound (27) follows from the two-point bound (28) with $\xi'_j = 0$. The proof of the latter consists in developing the exponential and use the Gaussian moments in Proposition 3.6. These hold up at least to moments $(\log T)^{\frac{1}{2K}}$. The restriction to the set $B(\tau)$ takes care of the higher moments. The difference between the restricted moments and the ones calculated in the corollary induce some error when computing the moments. This error gets worse as the moments grow. With this in mind, we choose to cut the development at

$$N = 10\ (\log T)^{\frac{1}{3K}} \ .$$

Write for simplicity P_j for $P_j(t+\tau)$ and P'_j for $P_j(t+\tau')$, and similarly for $B(\tau)$ and $B(\tau')$. The exponential can be written as

$$\begin{aligned} &\mathbb{E}\left[\exp\left(\sum_j \xi_j P_j + \xi'_j P'_j\right) \ \mathbf{1}_{B\cap B'}\right] = \\ &\sum_{n\le N} \frac{1}{n!}\mathbb{E}\left[\Big(\sum_j \xi_j P_j + \xi'_j P'_j\Big)^n \ \mathbf{1}_{B\cap B'}\right] + \sum_{n>N} \frac{1}{n!}\mathbb{E}\left[\Big(\sum_j \xi_j P_j + \xi'_j P'_j\Big)^n \ \mathbf{1}_{B\cap B'}\right] . \end{aligned} \tag{29}$$

By the definition of $B(\tau)$, the bound on ξ_j and ξ'_j, and the choice of N, the second term is bounded by

(30)
$$\sum_{n>N}\frac{1}{n!}((\sum_j|\xi_j|+|\xi'_j|)(\log T)^{\frac{1}{4K}})^n \ll \frac{(2K(\log T)^{\frac{1}{16K}+\frac{1}{4K}})^N}{N!}\exp\left(2K(\log T)^{\frac{1}{16K}+\frac{1}{4K}}\right) \ll e^{-(\log T)^{\frac{1}{3K}}}\ ,$$

since $2K(\log T)^{\frac{1}{16K}+\frac{1}{4K}} \le N/10 = (\log T)^{\frac{1}{3K}}$. For the first sum in (29), the unrestricted even moments are given by (21) in Proposition 3.6:

(31)
$$\mathbb{E}\left[\Big(\sum_j \xi_j P_j + \xi'_j P'_j\Big)^{2n}\right] = (2n-1)!!\left(\sum_j \left\{s_j^2(\xi_j^2+\xi'^2_j)+2\rho_j(\tau,\tau')\xi_j\xi'_j\right\}\right)^n + \mathrm{O}(e^{-(\log T)^{\frac{1}{2K}}}).$$

For the odd moments, we simply have by Proposition 3.6:

$$\mathbb{E}\left[\Big(\sum_j \xi_j P_j + \xi'_j P'_j\Big)^{2n+1}\right] \ll e^{-(\log T)^{\frac{1}{2K}}}\ . \tag{32}$$

We claim that for any $n \le N$, the unrestricted moments are close to the unrestricted ones:

$$\mathbb{E}[(\sum_j \xi_j P_j + \xi'_j P'_j)^n\ \mathbf{1}_{B\cap B'}] = \mathbb{E}[(\sum_j \xi_j P_j + \xi'_j P'_j)^n] + \mathrm{O}(e^{-(\log T)^{\frac{1}{4K}}})\ . \tag{33}$$

Indeed, Cauchy-Schwartz inequality implies

$$\mathbb{E}[(\sum_j \xi_j P_j + \xi'_j P'_j)^n\ \mathbf{1}_{(B\cap B')^c}] \le \mathbb{E}[(\sum_j \xi_j P_j + \xi'_j P'_j)^{2n}]^{1/2}\cdot(2\mathbb{P}(B^c))^{1/2}\ .$$

From (31), the bounds on ξ_j and (23), we see that the first term is bounded by

$$\left(\frac{(2n)!}{2^n n!}\right)^{1/2}((\log T)^{\frac{1}{4K}})^n \ll N^N((\log T)^{\frac{1}{4K}})^N \quad, n\le N.$$

Equation (33) then follows from the choice of N and Lemma 3.7. Using (33) together with (31) and (32) as well as (30), we get

$$\begin{aligned}&\mathbb{E}\left[\exp\left(\sum_j \xi_j P_j + \xi'_j P'_j\right)\ \mathbf{1}_{B\cap B'}\right]\\ &= \sum_{n\le N/2}\frac{1}{2^n n!}\left(\sum_j\left\{s_j^2(\xi_j^2+\xi'^2_j)+2\rho_j(\tau,\tau')\xi_j\xi'_j\right\}\right)^n + \mathrm{O}(e^{-(\log T)^{\frac{1}{4K}}})\\ &= \exp\left(\frac12\sum_j\left\{s_j^2(\xi_j^2+\xi'^2_j)+2\rho_j(\tau,\tau')\xi_j\xi'_j\right\}\right) + \mathrm{O}(e^{-(\log T)^{\frac{1}{4K}}})\ .\end{aligned}\tag{34}$$

The second equality comes from the fact that for $|\xi_j|,|\xi'_j| \le (\log T)^{\frac{1}{16K}}$,

$$\sum_{n>N/2}\frac{1}{n!}\left(\frac12\sum_j\left\{s_j^2(\xi_j^2+\xi'^2_j)+2\rho_j(\tau,\tau')\xi_j\xi'_j\right\}\right)^n \ll e^{-(\log T)^{\frac{1}{4K}}}\ .$$

The additive error in (34) can be absorbed in the main term since

$$e^{-(\log T)^{\frac{1}{4K}}} \ll e^{-(\log T)^{\frac{1}{8K}}} \exp\left(\frac{1}{2}\sum_j \left\{s_j^2(\xi_j^2+\xi'^2_j)+2\rho_j(\tau,\tau')\xi_j\xi'_j\right\}\right).$$

□

The statements of Proposition 3.8 are used to get precise large deviation estimates on the variables P_j. For x_j to be fixed, define the event

$$\mathcal{T}(\tau)=\{t: P_j(t+\tau)>x_j,\ \forall 1\le j\le K-3\}. \tag{35}$$

We first prove a two point bound for τ and τ' that are at mesoscopic distance.

Proposition 3.9. *Let τ and τ' such that $(\log T)^{-(m+1)/K}\le|\tau-\tau'|<(\log T)^{-m/K}$ for some $0\le m\le K-3$. We have for $0<x_j\ll\log\log T$*

$$\mathbb{P}(\mathcal{T}(\tau)\cap\mathcal{T}(\tau'))\ll\exp\left(\sum_{j=1}^{m}-\frac{x_j^2}{2s_j^2}+\sum_{j=m+1}^{K-3}-\frac{x_j^2}{s_j^2}\right). \tag{36}$$

Proof. Since the right-hand side of (26) is of lower order than the right-hand side of (36), it suffices to prove a bound on $\mathbb{P}(\mathcal{T}(\tau)\cap\mathcal{T}(\tau')\cap B(\tau)\cap B(\tau'))$. For simplicity, we write $P_j=P_j(t+\tau)$, $P'_j=P_j(t+\tau')$, $B=B(\tau)$ and $B'=B(\tau')$. Note that by definition of the set $\mathcal{T}$, the probability $\mathbb{P}(\mathcal{T}\cap\mathcal{T}'\cap B\cap B')$ is bounded by

$$\mathbb{E}\left(\exp\left(\sum_{j=1}^{K-3}\beta_j(P_j+P'_j)\right)\mathbf{1}_{B\cap B'}\right)\exp\left(-2\sum_{j=1}^{K-3}\beta_jx_j\right)\quad\text{, for any }\beta_j>0.$$

Using (28) we get that this is

$$\ll\exp\left(\frac{1}{2}\sum_{j=1}^{K-3}2\beta_j^2(s_j^2+\rho_j(\tau,\tau')^2)-2\sum_{j=1}^{K-3}\beta_jx_j\right)\quad\text{, for any }0<\beta_j<(\log T)^{\frac{1}{16K}}.$$

For $j\le m$ we have the trivial bound $\rho_j(\tau,\tau')\le s_j^2$ and for $j\ge m+1$ we have by (23) that $\rho_j(\tau,\tau')=\mathrm{O}(1)$, so that the probability is

$$\ll\exp\left(\frac{1}{2}\sum_{j=1}^{m}4\beta_j^2s_j^2+\frac{1}{2}\sum_{j=m+1}^{K-3}2\beta_j^2s_j^2-2\sum_{j=1}^{K-3}\beta_jx_j+\mathrm{O}\left(\sum_{j=m+1}^{K-3}\beta_j^2\right)\right).$$

By setting $\beta_j=x_j/s_j^2$ for $j\ge m+1$ and $\beta_j=x_j/(2s_j^2)$ for $j\le m$ we obtain (36). □

We also need precise large deviation bounds for one point τ and two points τ and τ' that are at near macroscopic distance.

Proposition 3.10. *For $|\tau|\le1$ we have for $0<x_j\ll\log\log T$*

$$\mathbb{P}(\mathcal{T}(\tau))\ge(1+\mathrm{o}(1))\prod_{j=1}^{K-3}\int_0^\infty\frac{1}{\sqrt{2\pi s_j^2}}e^{-\frac{(x_j+y)^2}{2s_j^2}}\,\mathrm{d}y\,. \tag{37}$$

Moreover, if $|\tau-\tau'|\ge(\log T)^{-\frac{1}{2K}}$, then

$$\mathbb{P}(\mathcal{T}(\tau)\cap\mathcal{T}(\tau'))\le(1+\mathrm{o}(1))\ \mathbb{P}(\mathcal{T}(\tau))\ \mathbb{P}(\mathcal{T}(\tau'))\,. \tag{38}$$

The proof consists of inverting the Fourier-Laplace transform of Proposition 3.8. We will do this by directly using the following lemma used in [2] and based on [5, Corollary 11.5].

Lemma 3.11. *Let $d \geq 1$. Let μ and ν be two probability measures on $\mathbb{R}^d$ with Fourier transform $\hat{\mu}$ and $\hat{\nu}$. There exists constant $c > 0$ such that for any function $f : \mathbb{R}^d \to \mathbb{R}$ with Lipschitz constant C and for any $R, N > 0$,*

$$\left| \int_{\mathbb{R}^d} f \mathrm{d}\mu - \int_{\mathbb{R}^d} f \mathrm{d}\nu \right| \leq \frac{C}{N} + \|f\|_\infty \left\{ (RN)^d \|(\hat{\mu} - \hat{\nu})\ \mathbf{1}_{(-N,N)^d}\|_\infty + \mu(((-R,R)^d)^c) + \nu(((-R,R)^d)^c) \right\} .$$

There will be a small error in the inversion if there is a broad range of parameters where the Fourier transforms are close. In the present setting, the errors on the probability need to be $\mathrm{o}((\log T)^{-1}))$ to compensate the number of points in the discretization. The range of the Fourier transform in Proposition 3.8 is thus *a priori* insufficient. To improve the error, we make a change of measure for the value x_j to be typical.

Proof of Proposition 3.10. As in the proof of Proposition 3.9, we drop the dependence on τ and τ' for simplicity. We start with the one-point bound (37). It suffices to bound $\mathbb{P}(\mathcal{T} \cap B)$ below. Define the measure $\mathbb{Q}$ by the Radon-Nikodym derivative

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{\exp(\sum_j \beta_j P_j)}{\mathbb{E}[\exp(\sum_j \beta_j P_j)]} \mathbf{1}_B \quad , \quad \beta_j = \frac{x_j}{s_j^2} .$$

The probability of $\mathcal{T} \cap B$ under $\mathbb{P}$ can be written as

$$\begin{aligned} \mathbb{P}(\mathcal{T} \cap B) &= \mathbb{E}\left[e^{\sum_j \beta_j P_j}\ \mathbf{1}_B \right] e^{-\sum_j \beta_j x_j}\ \mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j (\beta_j P_j - x_j)} \mathbf{1}_{\mathcal{T}} \right] \\ &= (1 + \mathrm{o}(1)) e^{-\sum_j \frac{x_j^2}{2 s_j^2}}\ \mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j (\beta_j P_j - x_j)} \mathbf{1}_{\mathcal{T}} \right] \end{aligned}$$

by Proposition 3.8 and the choice of β. It remains to prove that

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j (\beta_j P_j - x_j)} \mathbf{1}_{\mathcal{T}} \right] \geq (1 + \mathrm{o}(1)) \prod_j \int_0^\infty e^{-\frac{x_j y}{s_j^2}}\ \frac{e^{-\frac{y^2}{2 s_j^2}}}{\sqrt{2\pi s_j^2}}\ \mathrm{d}y . \tag{39}$$

Consider the functions g_j defined by

$$g_j(y) = \begin{cases} 0 & \text{if } y \leq 0 \\ e^{-\beta_j y} & \text{if } y \geq (\log T)^{-\frac{1}{64K^2}} \end{cases} ,$$

with a linear interpolation on the interval $[0, (\log T)^{-\frac{1}{64K^2}}]$. By construction, g_j is Lipschitz with constant $(\log T)^{\frac{1}{64K^2}}$. Moreover,

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j \beta_j (P_j - x_j)} \mathbf{1}_{\mathcal{T}} \right] \geq \mathbb{E}_{\mathbb{Q}}\left[\prod_j g_j(P_j - x_j) \mathbf{1}_{\mathcal{T}} \right] .$$

The expectation on the right can now be compared to the expectation over independent Gaussians variables with variance s_j^2 via Fourier inversion. The Fourier transform of the

distribution of $(P_j - x_j, 1 \leq j \leq K-3)$ under $\mathbb{Q}$ is
(40)

$$\mathbb{E}_{\mathbb{Q}}\left[e^{\mathrm{i}\sum_j t_j(P_j - x_j)}\right] = e^{-\mathrm{i}\sum_j t_j x_j}\frac{\mathbb{E}\left[e^{\sum_j(\beta_j+\mathrm{i}t_j)P_j}\ \mathbf{1}_B\right]}{\mathbb{E}\left[e^{\sum_j \beta_j P_j}\ \mathbf{1}_B\right]} = (1+\mathrm{O}(e^{-(\log T)^{\frac{1}{8K}}}))\exp\left(-\frac{1}{2}\sum_j t_j^2 s_j^2\right) ,$$

for $|t_j| \leq (\log T)^{\frac{1}{16K}}$ by Proposition 3.8. The right side corresponds to the Fourier transform of independent Gaussian variables of variance s_j^2. The difference between the Fourier transforms on that range is $\mathrm{O}(e^{-(\log T)^{\frac{1}{8K}}})$. We apply Lemma 3.11 with $d = K-3$, $N = R = (\log T)^{\frac{1}{32K^2}}$ and $f = \prod_j g_j$. This yields

$$\begin{aligned}&\left|\mathbb{E}_{\mathbb{Q}}\left[\prod_j g(P_j - x_j)\mathbf{1}_{\mathcal{T}}\right] - \prod_j \int_0^\infty g_j(y)\frac{e^{-\frac{y^2}{2s_j^2}}}{\sqrt{2\pi s_j^2}}\mathrm{d}y\right| \\ &\ll \frac{(\log T)^{\frac{1}{64K^2}}}{(\log T)^{\frac{1}{32K^2}}} + \frac{(\log T)^{\frac{K-3}{16K^2}}}{e^{(\log T)^{\frac{1}{8K}}}} + e^{-(\log T)^{\frac{1}{32K^2}}} + e^{-(\log T)^{\frac{1}{32K^2}}} \ll (\log T)^{-\frac{1}{64K^2}} ,\end{aligned} \tag{41}$$

where the last two terms in the first inequality come from the measure of $((-R,R)^d)^c$ under the Gaussian and $\mathbb{Q}$ measure; they are bounded respectively by a standard Gaussian tail bound and the corresponding bound under $\mathbb{Q}$, which can be obtained from the exponential Chebyshev inequality and the exponential moment bound (40) (with t_j purely complex). The claim (39) follows from (41) and the fact that

$$\left|\prod_j \int_0^\infty g_j(y)\frac{e^{-\frac{y^2}{2s_j^2}}}{\sqrt{2\pi s_j^2}}\mathrm{d}y - \prod_j \int_0^\infty e^{-\frac{x_j y}{s_j^2}}\ \frac{e^{-\frac{y^2}{2s_j^2}}}{\sqrt{2\pi s_j^2}}\ \mathrm{d}y\right| \ll (\log T)^{-\frac{1}{64K^2}} , \tag{42}$$

by definition of g_j. The error term $(\log T)^{-\frac{1}{64K^2}}$ is of smaller order than the integral since

$$\int_0^\infty e^{-\frac{x_j y}{s_j^2}}\ \frac{e^{-\frac{y^2}{2s_j^2}}}{\sqrt{2\pi s_j^2}}\mathrm{d}y \geq e^{\frac{x_j^2}{2s_j^2}}\int_{x_j/s_j}^\infty \frac{e^{-u^2/2}}{\sqrt{2\pi}}\mathrm{d}u \gg \frac{s_j}{x_j}, \tag{43}$$

and $(\log T)^{-\frac{1}{64K^2}} = \mathrm{o}(s_j/x_j)$. This concludes the proof of (39).

The proof of the two-point bound (38) is done similarly. As in the proof of Proposition 3.9, it suffices to prove a bound on $\mathbb{P}(\mathcal{T} \cap \mathcal{T} \cap B \cap B)$. The right change of measure is now

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{\exp(\sum_j \beta_j(P_j + P_j'))}{\mathbb{E}[\exp(\sum_j \beta_j(P_j + P_j'))]}\mathbf{1}_{B\cap B'} \quad , \quad \beta_j = \frac{x_j}{s_j^2} .$$

The probability $\mathbb{P}(\mathcal{T} \cap \mathcal{T}' \cap B \cap B')$ becomes

$$\mathbb{E}\left[e^{\sum_j \beta_j(P_j+P_j')}\ \mathbf{1}_{B\cap B'}\right] e^{-2\sum_j \beta_j x_j}\ \mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j \beta_j(P_j + P_j' - 2x_j)}\mathbf{1}_{\mathcal{T}\cap\mathcal{T}'}\right]. \tag{44}$$

It follows from (28) that

$$\mathbb{E}\left[\exp\left(\sum_{j=1}^{K-3}\xi_j P_j + \xi_j' P_j'\right)\ \mathbf{1}_{B\cap B'}\right] = (1 + \mathrm{O}((\log T)^{-\frac{1}{4K}})\exp\left(\frac{1}{2}\sum_{j=1}^{K-3}(\xi_j^2 + {\xi_j'}^2)s_j^2\right) , \tag{45}$$

for $|\xi_j|,|\xi'_j| \leq \log T^{\frac{1}{16K}}$, where we have used that $|\sum_j \xi_j\xi'_j\rho_j(\tau,\tau')| \ll (\log T)^{\frac{1}{8K}}(\log T)^{-\frac{1}{2K}} = \mathrm{O}((\log T)^{-\frac{1}{4K}})$, by (23). We thus expect the variables $((P_j,P'_j), 1 \leq j \leq K-3)$ to be approximatively independent Gaussians of variance s_j^2. By (45) for the choice $\xi_j = \xi'_j = \beta_j$, the quantity in (44) equals

$$(1+\mathrm{o}(1))e^{-\sum_j \frac{x_j^2}{s_j^2}}\ \mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j \beta_j(P_j+P_j-2x_j)}\mathbf{1}_{\mathcal{T}\cap\mathcal{T}'}\right] .$$

We now show that

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j \beta_j(P_j+P_j-2x_j)}\mathbf{1}_{\mathcal{T}\cap\mathcal{T}'}\right] \leq (1+\mathrm{o}(1))\prod_j\left(\int_0^\infty e^{-\frac{x_jy}{s_j^2}}\frac{e^{-\frac{y^2}{2s_j^2}}}{\sqrt{2\pi s_j^2}}\mathrm{d}y\right)^2 . \tag{46}$$

This is proved by using Lemma 3.11 with $d = 2(K-3)$, $N = R = (\log T)^{\frac{1}{32K^2}}$ and $f = \prod_j g_j$ where g_j is now

$$g_j(y) = \begin{cases} 0 & \text{if } y \leq -(\log T)^{-\frac{1}{64K^2}} \\ e^{-\beta_j y} & \text{if } y > 0 \ , \end{cases}$$

with linear interpolation in-between. This function has Lipschitz constant $(\log T)^{\frac{1}{64K^2}}$, and by definition,

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\sum_j \beta_j(P_j+P'_j-2x_j)}\mathbf{1}_{\mathcal{T}\cap\mathcal{T}'}\right] \leq \mathbb{E}_{\mathbb{Q}}\left[\prod_j g_j(P_j-x_j)g_j(P'_j-x_j)\mathbf{1}_{\mathcal{T}\cap\mathcal{T}'}\right] . \tag{47}$$

The Fourier transform under the $\mathbb{Q}$-measure is

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[e^{\mathrm{i}\sum_j(t_jP_j+t'_jP'_j-2x_j)}\right] &= e^{-\mathrm{i}\sum_j(t_j+t'_j)x_j}\ \frac{\mathbb{E}\left[e^{\sum_j(\beta_j+\mathrm{i}t_j)P_j+(\beta_j+\mathrm{i}t'_j)P'_j}\ \mathbf{1}_{B\cap B'}\right]}{\mathbb{E}\left[e^{\sum_j\beta_j(P_j+P'_j)}\ \mathbf{1}_{B\cap B'}\right]} \\ &= (1+\mathrm{O}((\log T)^{-\frac{1}{4K}}))\ \exp\left(-\frac{1}{2}\sum_j(t_j^2+{t'_j}^2)s_j^2\right) ,\end{aligned} \tag{48}$$

for $|t_j|,|t'_j| \leq (\log T)^{\frac{1}{16K}}$ by (45). Again, the right side corresponds to the Fourier transform of pairs of independent Gaussian variables of variance s_j^2. The difference between the two Fourier transform is $\mathrm{O}((\log T)^{-\frac{1}{4K}})$. Lemma 3.11 directly implies

$$\begin{aligned}&\left|\mathbb{E}_{\mathbb{Q}}\left[\prod_j g_j(P_j-x_j)g_j(P'_j-x_j)\mathbf{1}_{\mathcal{T}\cap\mathcal{T}'}\right] - \prod_j\left(\int_0^\infty g_j(y)e^{-\frac{y^2}{2s_j^2}}\mathrm{d}y\right)^2\right| \\ &\ll \frac{(\log T)^{\frac{1}{64K^2}}}{(\log T)^{\frac{1}{32K^2}}} + \frac{(\log T)^{\frac{2(K-3)}{16K^2}}}{(\log T)^{\frac{1}{4K}}} + e^{-(\log T)^{\frac{1}{32K^2}}} + e^{-(\log T)^{\frac{1}{32K^2}}} \ll (\log T)^{-\frac{1}{64K^2}} ,\end{aligned} \tag{49}$$

similarly to (41). The claim (46) follows from (49) and (47). The error coming from g_j is controlled as in (42). This concludes the proof of (46) and of the proposition. □

It is now possible to finish the proof of Proposition 3.2.

Proof of Proposition 3.2. Without loss of generality, we suppose that $\log T$ is an integer. Consider a finite set $(\tau_\ell, 1 \leq \ell \leq \log T)$ of $\log T$ points that are equidistant on $[-1,1]$. As mentioned in (23), we have that $2s_j^2 = \frac{1}{K}\log\log T + \mathrm{o}(1)$. Therefore, by picking $x_j = \lambda' 2s_j^2$ for $1 > \lambda' > \lambda$ in the definition of $\mathcal{T}(\tau)$, the probability in the statement is greater than $\mathbb{P}\left(\bigcup_\ell \mathcal{T}(\tau_\ell)\right)$. Moreover, by a direct application of the Cauchy-Schwarz inequality (the Paley-Zygmund inequality), this is bounded below by

$$\frac{\left(\mathbb{E}\left[\sum_\ell \mathbf{1}_{\mathcal{T}(\tau_\ell)}\right]\right)^2}{\mathbb{E}\left[\left(\sum_\ell \mathbf{1}_{\mathcal{T}(\tau_\ell)}\right)^2\right]} = \frac{\left(\sum_\ell \ \mathbb{P}(\mathcal{T}(\tau_\ell))\right)^2}{\sum_{\ell,\ell'} \mathbb{P}(\mathcal{T}(\tau_\ell) \bigcap \mathcal{T}(\tau_{\ell'}))} . \tag{50}$$

On one hand, by (37) and (43), the sum in the numerator is bounded below by

$$\begin{aligned}\sum_\ell \mathbb{P}(\mathcal{T}(\tau_\ell)) &\geq (1+\mathrm{o}(1)) \log T \ \prod_j \int_0^\infty \frac{1}{\sqrt{2\pi s_j^2}} e^{-\frac{(x_j+y)^2}{2s_j^2}} \mathrm{d}y \\ &\gg (\log T)\ e^{\sum_j -\frac{x_j^2}{2s_j^2}} \ (\log\log T)^{-1/2}.\end{aligned} \tag{51}$$

On the other hand, the sum in the denominator can be split into three parts. The first is over pairs with $|\tau_\ell - \tau_{\ell'}| \geq (\log T)^{-\frac{1}{2K}}$. There are at most $(\log T)^2$ such pairs, therefore by (38),

$$\sum_{|\tau_\ell - \tau_{\ell'}| \leq (\log T)^{-\frac{1}{2K}}} \mathbb{P}(\mathcal{T}(\tau_\ell) \cap \mathcal{T}(\tau_{\ell'})) \leq (1+\mathrm{o}(1)) \left(\sum_\ell \mathbb{P}(\mathcal{T}(\tau_\ell))\right)^2 . \tag{52}$$

In view of (50) and (51), it remains to show that the sum over pairs such that $|\tau_\ell - \tau_{\ell'}| < (\log T)^{-\frac{1}{2K}}$ are negligible with respect to the numerator. There are at most $(\log T)^{2-\frac{1}{2K}}$ pairs with $(\log T)^{-\frac{1}{K}} < |\tau_\ell - \tau_{\ell'}| \leq (\log T)^{-\frac{1}{2K}}$. Proposition 3.9 (with $m=0$) implies that the sum over these pairs is

$$\ll (\log T)^{2-\frac{1}{2K}} \ e^{\sum_j -\frac{x_j^2}{s_j^2}} \ll (\log T)^{-\frac{1}{3K}} \left(\sum_\ell \mathbb{P}(\mathcal{T}(\tau_\ell))\right)^2 = \mathrm{o}\left(\left(\sum_\ell \mathbb{P}(\mathcal{T}(\tau_\ell))\right)^2\right),$$

where the second inequality is from the bound (51). Finally, for $m = 1, \dots, K-3$, there are at most $(\log T)^{2-\frac{m}{K}}$ pairs with $(\log T)^{-\frac{(m+1)}{K}} < |\tau_\ell - \tau_{\ell'}| \leq (\log T)^{-\frac{m}{K}}$. By Proposition 3.9, the sum of the probabilities over these pairs is

$$\begin{aligned}\ll (\log T)^{2-\frac{m}{K}} \exp\left(\sum_{j=1}^m -\frac{x_j^2}{2s_j^2} + \sum_{j=m+1}^{K-3} -\frac{x_j^2}{s_j^2}\right) &= (\log T)^{2-\frac{m}{K}} \exp\left(-\sum_{j=1}^{K-3} \frac{x_j^2}{s_j^2}\right) \exp\left(\sum_{j=1}^m \frac{x_j^2}{2s_j^2}\right) \\ &\ll (\log T)^2 \exp\left(-\sum_{j=1}^{K-3} \frac{x_j^2}{s_j^2}\right) (\log T)^{-\frac{m}{K}(1-\lambda'^2)} .\end{aligned}$$

This implies that the sum over m from 1 to $K-3$ is $\mathrm{O}\left((\log T)^{2-\frac{1-\lambda'^2}{K}}\exp\left(-\sum_{j=1}^{K-3}\frac{x_j^2}{s_j^2}\right)\right)$, which is also negligible with respect to the numerator by (51) for $\lambda'<1$. This concludes the proof of the proposition. □

Appendix A. Sobolev-type inequalities

The following Sobolev-type inequalities are used several times throughout the paper.

Lemma A.1. *Let f be a differentiable function on $[0,\infty)$ such that $f(t)=\mathrm{o}(t^{1/2})$ and $f'(t)=\mathrm{o}(t^{1/2})$. Then*

$$\mathbb{E}\left[\max_{u\in[-1,1]}|f(t+u)|^2\right]\ll\left(\mathbb{E}[|f(t)|^2]\right)^{1/2}\cdot\left(\mathbb{E}[|f'(t)|^2]\right)^{1/2}+\mathbb{E}[|f(t)|^2]\ .$$

Proof. For any $t\in[T,2T]$ and any $u\in[-1,1]$, integration yields the inequality

$$\begin{aligned}|f(t+u)|^2&\le\int_{t-1}^{t+u}|f'(s)\cdot f(s)|\mathrm{d}s+\int_{t+u}^{t+1}|f'(s)\cdot f(s)|\mathrm{d}s+\frac{1}{2}|f(t-1)|^2+\frac{1}{2}|f(t+1)|^2\\&=\int_{t-1}^{t+1}|f'(s)\cdot f(s)|\mathrm{d}s+\frac{1}{2}|f(t-1)|^2+\frac{1}{2}|f(t+1)|^2\ .\end{aligned}$$

Therefore, by taking the maximum over u and averaging on t, we get

$$\mathbb{E}\left[\max_{u\in[-1,1]}|f(t+u)|^2\right]\le\int_{-1}^{1}\mathbb{E}\left[|f'(t+s)\cdot f(t+s)|\right]\mathrm{d}s+\frac{1}{2}\mathbb{E}[|f(t-1)|^2]+\frac{1}{2}\mathbb{E}[|f(t+1)|^2]\ .$$

By assumption on f and f', we have for any $u\in[-1,1]$, $\mathbb{E}\left[|f'(t+u)|^2\right]\ll\mathbb{E}\left[|f'(t)|^2\right]$ and $\mathbb{E}\left[|f(t+u)|^2\right]\ll\mathbb{E}\left[|f(t)|^2\right]$. The claim follows from this observation and Cauchy-Schwarz inequality applied on the first term. □

Lemma A.2. *Let $p>0$ and f be an analytic function in the strip $\sigma_0<\sigma<\sigma_1$. Then for any $L>1$, and σ with $\sigma_0+1/L\le\sigma\le\sigma_1-1/L$,*

$$\int_T^{2T}\max_{|t-u|\le1}|f(\sigma+\mathrm{i}u)|^p\mathrm{d}t\ll L\sup_{\sigma^\star\in[\sigma-1/L,\sigma+1/L]}\int_{T-1}^{2T+1}|f(\sigma^\star+\mathrm{i}t)|^p\mathrm{d}t.$$

Proof. To prove the lemma it suffices to notice that

$$\max_{|t-u|\le1}|f(\sigma+\mathrm{i}u)|^p\le\sum_{|j|\le L}\max_{|t-u-j/L|\le1/L}|f(\sigma+\mathrm{i}u)|^p\ .$$

By subharmonicity of $|f|^p$, the above is

$$\ll L^2\int_{|x|\le1/L}\sum_{|j|\le L}\int_{|t-y-j/L|\le2/L}|f(\sigma+x+\mathrm{i}y)|^p\mathrm{d}y\mathrm{d}x\ll\ L^2\int_{-1/L}^{1/L}\int_{t-1}^{t+1}|f(\sigma+x+iy)|^p\mathrm{d}y\mathrm{d}x\ .$$

The claim follows by integrating over t. □

Appendix B. Elementary Number Theory Estimates

The Prime Number Theorem states that, cf. [27],

$$\#\{p \le x : p \text{ prime}\} = \int_2^x \frac{1}{\log y}\mathrm{d}y + \mathrm{O}(xe^{-c\sqrt{\log x}}) \ .$$

In particular, this implies

$$\sum_{P\le p\le Q} \frac{1}{p^{2\sigma}} = \int_P^Q \frac{1}{x^{2\sigma}\log x}\mathrm{d}x + \mathrm{O}(e^{-c\sqrt{\log P}}) \ . \tag{53}$$

When $\sigma = 1/2+\delta$ and $\delta \log Q < 1$, it is easy to check that this yields

$$\sum_{P\le p\le Q} \frac{1}{p^{2\sigma}} = \log\log Q - \log\log P + \mathrm{O}(\delta \log Q) \ . \tag{54}$$

This is the case in particular for the choice of σ_0 in (7) where $\delta = (\log T)^{-1+\frac{3}{2K}}$ and the largest Q is the upper bound of the set J_{K-3} given in (8).

The next lemma confirms the heuristics that the Dirichlet polynomials are very close to polynomials where (p^{-it}) is replaced by IID uniform variables on the circle. This result appears elsewhere, in particular in [8, proof of Proposition 3.1]. We prove it for completeness.

Lemma B.1. *Let $X \ge 1$ and let $(a(p), p$ primes$)$ and $(a^\star(p), p$ primes$)$ be two sequences in $\mathbb{C}$ with $|a(p)|, |a^\star(p)| \le 1$ for all p and $a(p) = a^\star(p) = 0$ for $p > X$. Then for $k \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\left(\frac{1}{2}\sum_p a(p)p^{-it} + a^\star(p)p^{it}\right)^k\right] = \partial_z^k\left(\prod_p \mathcal{J}(a(p)a^\star(p)z^2)\right)\Big|_{z=0} + \mathrm{O}\left(\frac{X^{2k}}{T}\right)$$

where $\mathcal{J}(z) = \sum_{n\ge 0} z^n/(2^{2n}(n!)^2)$. In particular, the expression is $\mathrm{O}\left(X^{2k}/T\right)$ for odd k.

Proof. The expectation can be written as

$$\sum_{p_1,\dots,p_k} \mathbb{E}\left[\prod_{l=1}^k \left(\frac{1}{2}a(p_l)p_l^{-\mathrm{i}t} + a^\star(p_l)p_l^{\mathrm{i}t}\right)\right].$$

Denote the number of prime factors of n (with multiplicities) by $\Omega(n)$. We define the multiplicative function $g(q^\alpha) = 1/\alpha!$ if q is prime. If we write $n = q_1^{\alpha_1}\dots q_r^{\alpha_r}$ for the prime decomposition of n, the above sum becomes

$$\sum_{\Omega(n)=k} k!g(n)\ \mathbb{E}\left[\prod_{l=1}^k \left(\frac{1}{2}a(q_l)q_l^{-\mathrm{i}t} + \frac{1}{2}a^\star(q_l)q_l^{\mathrm{i}t}\right)^{\alpha_l}\right] \tag{55}$$

Let $\mathcal{N}_l$ be independent Binomial random variable with parameters $(\alpha_l, 1/2)$. Denote the expectation over these by E. Note that by expanding the power α_l each term can be written as

$$\left(\frac{1}{2}a(q_l)q_l^{-\mathrm{i}t} + \frac{1}{2}a^\star(q_l)q_l^{\mathrm{i}t}\right)^{\alpha_l} = E[a(q_l)^{\mathcal{N}_l}a^\star(q_l)^{\alpha_l-\mathcal{N}_l}e^{-\mathrm{i}t\log q_l^{2\mathcal{N}_l-\alpha_l}}].$$

The expectations $\mathbb{E}$ and E can be interchanged. Moreover, unless $\mathcal{N}_l = \alpha_l/2$ for all l we have that $|\mathrm{i}\log\prod_l q_l^{2\mathcal{N}_l-\alpha_l}|\gg \frac{1}{X^k}$. In this case, the factor coming from the integration of the exponential in t is bounded by X^k/T. These observations imply altogether the following:

$$E\left[\mathbb{E}\left[\prod_l a(q_l)^{\mathcal{N}_l} a^\star(q_l)^{\alpha_l-\mathcal{N}_l} e^{\mathrm{i}t\log q_i^{2\mathcal{N}_l-\alpha_l}}\right]\right] = \prod_l \frac{(a(q_l)a^\star(q_l))^{\alpha_l/2}}{2^{\alpha_l}}\binom{\alpha_l}{\alpha_l/2} + \mathrm{O}\left(\frac{X^k}{T}\right) ,$$

where we use the bound on $a(p), a^\star(p)$ and the convention that $\binom{\alpha}{\alpha/2} = 0$ if $\alpha/2$ is not an integer. Consider the multiplicative function f defined by

$$f(p^\alpha) = \frac{(a(p)a^\star(p))^{\alpha/2}}{2^\alpha}\binom{\alpha}{\alpha/2} .$$

With these observations, (55) becomes

$$\sum_{\Omega(n)=k} k! f(n)g(n) + \mathrm{O}\left(\frac{X^{2k}}{T}\right) , \tag{56}$$

where for the error term we used the fact that there are at most X^k terms in the sum. The condition $\Omega(n) = k$ can be expressed using a contour integral

$$\sum_{\Omega(n)=k} k! f(n)g(n) = \frac{k!}{2\pi\mathrm{i}}\oint \sum_{n\geq 1} f(n)g(n) z^{\Omega(n)} \frac{\mathrm{d}z}{z^{k+1}} . \tag{57}$$

Since the summand is a multiplicative function, the finite sum may be expressed as an Euler product

$$\sum_{n\geq 1} f(n)g(n)z^{\Omega(n)} = \prod_p \left(1 + \sum_{\ell\geq 1}\frac{(a(p)a^\star(p)z^2)^\ell}{2^{2\ell}(2\ell)!}\binom{2\ell}{\ell}\right) = \prod_p \mathcal{J}(a(p)a^\star(p)z^2) .$$

The lemma follows by putting this back in (57) and using Cauchy's formula. □

A similar argument can be used to prove an upper bound on the moment, see Lemma 3 in [36].

Lemma B.2. *Let $a(p)$ be a sequence of complex coefficients and $\sigma \geq 1/2$. Then for X and $k \in \mathbb{N}$ such that $X^k \leq T(\log T)^{-1}$, we have*

$$\mathbb{E}\left[\Big|\sum_{p\leq T^{1/k}}\frac{a(p)}{p^{\sigma+it}}\Big|^{2k}\right] \ll k!\cdot\Big(\sum_{p\leq T^{1/k}}\frac{|a(p)|^2}{p^{2\sigma}}\Big)^k.$$

Note that when the sum does not involve only primes but also prime powers (like in the case of $\mathcal{P}_i$) one can apply the Lemma with the choice of $a(p) = \alpha(p) + \beta(p^2)/p^{1/2}$ for example with $\alpha(p)$ the coefficient at the primes and $\beta(p^2)$ the coefficients at squares of primes.

The following observation called Rankin's trick will be used often:

$$\sum_{n>x} a(n) \leq x^{-\alpha}\sum_{n>x} a(n)n^\alpha \leq x^{-\alpha}\sum_{n=1}^{\infty} a(n)n^\alpha \qquad a(n)\geq 0, \alpha > 0. \tag{58}$$

References

[1] L.-P. Arguin, *Extrema of Log-correlated Random Variables: Principles and Examples*, in Advances in Disordered Systems, Random Processes and Some Applications, Cambridge Univ. Press, Cambridge (2016), 166–204.

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

[3] L.-P. Arguin, D. Belius, and A.J. Harper, *Maxima of a randomized Riemann zeta function, and branching random walks*, to appear in *Ann. Appl. Prob.* (2015).

[4] R. Balasubramanian and K. Ramachandra, *On the frequency of Titchmarchs phenomenon for $\zeta(s)$-III*, Proc. Indian Acad. Sci. **86** (1977), 341–351.

[5] 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.

[6] A. Bondarenko and K. Seip, *Large GCD sums and extreme values of the Riemann zeta function*, Preprint arXiv:1507.05840 (2015).

[7] J. Bourgain, *Decoupling, exponential sums and the Riemann zeta function*, Journal Amer. Math. Soc. (2016).

[8] P. Bourgade, *Mesoscopic fluctuations of the zeta zeros*, Probab. Theory Related Fields **148** (2010), no. 3-4, 479–500.

[9] M. D. Bramson, *Maximal displacement of branching Brownian motion*, Comm. Pure Appl. Math. **31** (1978), no. 5, 531–581.

[10] V. Chandee and K. Soundararajan, *Bounding $|\zeta(1/2 + \mathrm{i}t)|$ on the Riemann hypothesis*, Bull. London Math. Soc. (2011).

[11] R. Chhaibi, T. Madaule, and J. Najnudel, *On the maximum of the CβE field*, Preprint arXiv:1607.00243 (2016).

[12] J. B. Conrey, *The fourth moment of derivatives of the Riemann zeta-function*, Quart. J. Math. Oxford Ser. (2) **39** (1988), no. 153, 21–36.

[13] D. Farmer, S.M. Gonek, and C. Hughes, *The maximum size of L-functions*, J. Reine Angew. Math. **609** (2007), 215–236.

[14] Y. V. Fyodorov and J.-P. Bouchaud, *Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential*, J. Phys. A **41** (2008), no. 37, 372001, 12 pp.

[15] 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** (2012), 170601, 5pp.

[16] Y. V. Fyodorov and J. P. Keating, *Freezing transitions and extreme values: random matrix theory, and disordered landscapes*, Phil. Trans. R. Soc. A **372** (2014), 20120503, 32 pp.

[17] Y. V. Fyodorov, P. Le Doussal, and A. Rosso, *Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields*, J. Stat. Mech. Theory Exp. **10** (2009), P10005, 32 pp.

[18] ———, *Counting function fluctuations and extreme value threshold in multifractal patterns: the case study of an ideal $1/f$ noise*, J. Stat. Phys. **149** (2012), no. 5, 898–920.

[19] Y. V. Fyodorov and N. J. Simm, *On the distribution of maximum value of the characteristic polynomial of GUE random matrices*, preprint, arXiv:1503.07110 (2015).

[20] A. J. Harper, *A note on the maximum of the Riemann zeta function, and log-correlated random variables*, Preprint arXiv:1304.0677 (2013).

[21] J. Keating and N. Snaith, *Random Matrix Theory and $\zeta(1/2+\mathrm{i}t)$*, Comm. Math. Phys. **1** (2000), 57–89.

[22] N. Kistler, *Derrida's Random Energy Models*, Lecture Notes in Math., vol. 2143, Springer, Berlin, 2015.

[23] G. Lambert and E. Paquette, *The law of large numbers for the maximum of almost Gaussian log-correlated fields coming from random matrices*, preprint, arXiv:1611.08885 (2016).

[24] J.E. Littlewood, *On the zeros of the Riemann zeta-function*, Proc. Camb. Philos. Soc. **22** (1924), 295–318.

[25] H. Montgomery, *The pair correlation of zeros of the zeta function*, Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), Amer. Math. Soc., Providence, R.I., 1973, pp. 181–193.

[26] ———, *Extreme values of the Riemann zeta function*, Comment. Math. Helv. **52** (1977), 511–518.

[27] H. Montgomery and R. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[28] J. Najnudel, *On the extreme values of the Riemann zeta function on random intervals of the critical line*, Preprint arXiv:1611.05562 (2016).

[29] E. Paquette and O. Zeitouni, *The maximum of the CUE field*, preprint, arXiv:1602.08875 (2016).

[30] M. Radziwill, *Large deviations in Selberg's central limit theorem*, Preprint arxiv:1108.5092 (2011).

[31] M. Radziwill and K. Soundararajan, *Selberg's central limit theorem for* $\log|\zeta(\frac{1}{2}+it)|$, Preprint arxiv:1509.06827 (2015).

[32] K. Ramachandra and A. Sankaranarayanan, *On some theorems of Littlewood and Selberg*, J. Number Theory **44** (1993), 281–291.

[33] E. Saksman and C. Webb, *The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line*, preprint, arXiv:1609.00027 (2016).

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

[35] K. Soundararajan, *Extreme values of zeta and L-functions*, Mathematische Annalen **342** (2008), 467–486.

[36] ———, *Moments of the Riemann zeta-function*, Annals of Math. **170** (2009), 981–993.

[37] E.C. Titchmarsh, *The theory of the Riemann zeta-function*, Second Edition, Oxford Univ. Press, New York (1986).

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

Institute of Mathematics, University of Zurich, Switzerland
E-mail address: `david.belius@cantab.net`

Courant Institute, New York University, USA
E-mail address: `bourgade@cims.nyu.edu`

Department of Mathematics and Statistics, McGill University, Canada
E-mail address: `maksym.radziwill@mcgill.ca`

Department of Mathematics, Stanford University, USA
E-mail address: `ksound@stanford.edu`