Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 151 (2022) 174–190

stochastic processes and their applications

www.elsevier.com/locate/spa

High points of a random model of the Riemann-zeta function and Gaussian multiplicative chaos[☆]

Louis-Pierre Arguin[a], Lisa Hartung[b,*], Nicola Kistler[c]

[a] *Department of Mathematics Baruch College and Graduate Center City University of New York, New York, NY 10010, USA*

[b] *Institut für Mathematik, Johannes Gutenberg-Universität Mainz, Staudingerweg 9, 55099 Mainz, Germany*

[c] *Institut für Mathematik, Goethe-Universität Frankfurt, Robert-Mayer-Str. 10, 60325 Frankfurt, Germany*

Received 12 October 2019; received in revised form 28 September 2021; accepted 29 April 2022
Available online 7 May 2022

Abstract

We study the total mass of high points in a random model for the Riemann-zeta function. We consider the same model as in Harper (2013) and Arguin et al. (2017), and build on the convergence to Gaussian multiplicative chaos proved in Saksman and Webb (2016). We show that the total mass of points which are a linear order below the maximum, divided by their expectation, converges almost surely to the Gaussian multiplicative chaos of the approximating Gaussian process times a random function. We use the second moment method together with a branching approximation to establish this convergence.

MSC: 60J80; 60G70; 82B44

Keywords: Riemann-Zeta function; High points; Gaussian multiplicative chaos; Extreme values

1. Introduction

1.1. The model

Let $\mathcal{P}$ denote the set of all prime numbers. Let $(\theta_p)_{p\in\mathcal{P}}$ be independent identically distributed random variables, being uniformly distributed on $[0, 2\pi]$. For $N \in \mathbb{N}$, a good model for the

☆ The research of L.-P. A. is supported in part by NSF CAREER, USA DMS-1653602.
* Corresponding author.
E-mail addresses: louis-pierre.arguin@baruch.cuny.edu (L.-P. Arguin), lhartung@uni-mainz.de (L. Hartung), kistler@math.uni-frankfurt.de (N. Kistler).

https://doi.org/10.1016/j.spa.2022.04.017
0304-4149/

large values of the logarithm of the Riemann-zeta function on a typical interval of length 1 of the critical line as proposed in [10] is

$$X_N(x) = \sum_{p\in\mathcal{P}\cap[0,N]} \frac{1}{\sqrt{p}} \left(\cos(x \ln p)\cos(\theta_p) + \sin(x\ln p)\sin(\theta_p)\right), \quad x \in [0,1] . \tag{1.1}$$

We denote by $\mathbb{E}$ the expectation with respect to the θ_p's.

The maximum of the process on a small interval was studied in [2]. There it was shown that with high probability, depending on ϵ,

$$\max_{x\in[0,1]} X_N(x) = \ln\ln N - (\frac{3}{4}+\epsilon)\ln\ln\ln N. \tag{1.2}$$

In this paper, we are interested in the values of the process of the order of $\frac{\alpha}{2}\ln\ln N$ with $\alpha < 2$. Some of the behavior of the large values of the process $X_N(x)$, $x \in [0,1]$, is captured by the random measure

$$M_{\alpha,N}(\mathrm{d}x) = \frac{e^{\alpha X_N(x)}}{\mathbb{E}(e^{\alpha X_N(x)})}\mathrm{d}x. \tag{1.3}$$

By the independence of the θ_p's, it is not hard to see that $M_{\alpha,N}$ converges almost surely as $N \to \infty$. By Theorem 4 in [16], the almost sure weak limit of $M_{\alpha,N}(\mathrm{d}x)$ is non-trivial for $0 < \alpha < 2$. We denote the limit of the total mass by M_α

$$M_\alpha = \lim_{N\to\infty} \int_0^1 M_{\alpha,N}(\mathrm{d}x) \quad a.s. \tag{1.4}$$

For log-correlated Gaussian field the analogous limiting measure is called Gaussian multiplicative chaos and M_α corresponds to the total mass of the limiting measure. For Gaussian multiplicative chaos it was first proved in [11] that the limit is nontrivial for small α and was recently revisited (see for example [15,14]). Note that in our case the limit of $M_{\alpha,N}(\mathrm{d}x)$ is almost a Gaussian multiplicative measure (see [16]). The connection between the Riemann-zeta function and Gaussian multiplicative chaos has been further analyzed in [17].

The fact that the Riemann-zeta function (or a random model of it) can be well approximated by a log-correlated field have recently been used to study the extremes on a random interval [5, 13,2].

1.2. Main result

Consider the Lebesgue measure of α-high points:

$$W_{\alpha,N} = \mathrm{Leb}\left\{x \in [0,1] : X_N(x) > \tfrac{\alpha}{2}\ln\ln N\right\}. \tag{1.5}$$

The main result of this note is to relate the limit M_α to the Lebesgue measure of high points building on the ideas of [8]:

Theorem 1.1. *For any $0 < \alpha < 2$ and M_α as in (1.4), we have*

$$\frac{W_{\alpha,N}}{\mathbb{E}\left(W_{\alpha,N}\right)} \to M_\alpha \quad \text{in probability as } N \to \infty. \tag{1.6}$$

In view of Eq. (1.2) and of Theorem 1.1, it is not surprising to see that the M_α is non-trivial for $\alpha < 2$. The critical case where $\alpha \to 2$ is interesting as it is related to the fluctuations of the maximum of X_N. It is reasonable to expect that our approach can be adapted to the method of [6] to prove the critical case. Another upshot of the proof is that it highlights the fact that M_α depends on small primes, cf. Lemma 3.1. In a branching random walk the corresponding martingale limit encodes the effect of the first few generations. The effect of larger primes is somehow averaged out, which might also seem natural as the structure of primes become more regular.

The problem for the Riemann-zeta function is trickier. We expect that the equivalent of Theorem 1.1 still holds:

Conjecture 1.2. *Let τ be a uniform random variable on $[T, 2T]$. Let $W_{\alpha,T} = \mathrm{Leb}\{h \in [0,1] : \ln|\zeta(1/2 + i(\tau + h))| > \frac{\alpha}{2}\ln\ln T\}$. Then we have for $\alpha < 2$*

$$\lim_{T\to\infty} \frac{W_{\alpha,T}}{\mathbb{E}(W_{\alpha,T})} = \lim_{T\to\infty} \frac{\int_0^1 |\zeta(1/2 + i(\tau + h))|^\alpha \mathrm{d}h}{\mathbb{E}(|\zeta(1/2 + i\tau)|^\alpha)} \quad \textit{in probability.}$$

This would be consistent with the conjecture of Fyodorov & Keating for the Lebesgue measure of high points, see Section 2.5 in [7]. There might be hope to prove this as the proof of Theorem 1.1 relies on a Gaussian comparison for one point and two points. This is accessible to some extent for the zeta function, see [1].

The analog of Theorem 1.1 was proved for the two-dimensional discrete Gaussian free field in [4] (see Corollary 2.2). There, the result is proved as a consequence of a much more detailed result on the (joint) point measure of the value of high points and their location (see Theorem 2.1). Note that the convergence of the measure level sets there is in distribution, but their method should also yield convergence in probability, as in Eq. (1.5). Another notable difference is that Theorem 1.1 holds for a process that is *a priori* non-Gaussian. In fact, the main novelty of the present paper is to concretely enlarge the universality class where multiplicative chaos phenomena can be found.

1.3. Outline of the proof

The proof of Theorem 1.1 is based on a first and a second moment estimate and follow the global strategy proposed in [8] for branching Brownian motion. First, we prove convergence of a conditional first moment to the desired limiting object in Lemma 3.1. The proof relies on an explicit Gaussian comparison, cf. Proposition 2.3. Next, a localization result is established in Lemma 3.3. Finally, we turn to the proof of Proposition 4.1 which is based on a second moment computation. We use a branching approximation similar to the one employed in [2]. Using the obtained first and second moment estimates we are finally in the position to prove Theorem 1.1.

Notations. To lighten some computations, we will sometimes use Vinogradov's notation where $f(N) \ll g(N)$ stands for $f(N)/g(N) = \mathrm{O}(1)$. The notations O and o will always be meant for the limit $N \to \infty$ with implicit dependence on the fixed parameter α. In some proofs, it is convenient to use a loglog-scale for the parameters $R < K < N$, in which case we will use lower case letters and write $r = \ln\ln R$, $k = \ln\ln K$ and $n = \ln\ln N$. To keep the computations as clear as possible we assume r, n are natural numbers. The general case follows in the same way by considering the first and last summands of X_N separately. The desired estimates carry over with minor adjustments, but would require a more involved notation.

2. Comparison with a Gaussian process

It turns out that the process X_N is well approximated by a log-correlated Gaussian field $G_N(x), x \in [0, 1]$. A precise result in this direction is the following result of [17].

Theorem 2.1 (*Theorem 1.7 in [17]*). *For $N \geq 2$, the field $X_N(x)$, $x \in [0, 1]$, can be decomposed as $G_N(x) + E_N(x)$ where*

$$G_N(x) = \sum_{p\in\mathcal{P}\cap[0,N]} \frac{1}{\sqrt{2p}} \left(Z_p^{(1)} \cos(x \ln p) + Z_p^{(2)} \sin(x \ln p)\right), \tag{2.1}$$

for $(Z_p^{(i)})_{p\in\mathcal{P},i\in\{1,2\}}$ i.i.d. standard normal random variables. The error E_N is such that

$$\lim_{N\to\infty} \max_{x\in[0,1]} |E_N(x) - E(x)| = 0 \quad a.s., \tag{2.2}$$

where the limit $E(x)$ is a smooth (random) function. Moreover, the error $E_N(x)$ has uniform exponential moments

$$\mathbb{E}\left(\exp\Big(\lambda \sup_{N\geq 1, x\in[0,1]} E_N(x)\Big)\right) < \infty. \tag{2.3}$$

The statement of Theorem 1.3 in [17] is in terms of a random Euler product. The difference between the Euler product formulation and the Dirichlet polynomial formulation in (1.1) is small and is controlled by Lemma 3.2 of [17].

It might be tempting to prove Theorem 1.1 by simply proving it for G_N and control the error E_N using Eq. (2.3). However, there is a major difficulty in taking this approach as one might lose the independence between the small primes in G_N and the error E_N.[1] Instead, we rely on the following Berry–Esseen approximation as in [2] which allows for precise first and second moment estimates.

Lemma 2.2 (*Corollary 17.2 in [3], see also Theorem 1.3 in [9]*). *Let $(\boldsymbol{Y}_j, j \geq 1)$ be a sequence of independent random vectors on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P)$ with mean $E(\boldsymbol{Y}_j)$ and covariance matrix $\mathrm{Cov}(\boldsymbol{Y}_j)$. Define*

$$\boldsymbol{\mu}_m = \sum_{j=1}^{m} E(\boldsymbol{Y}_j) \text{ and } \Sigma_m = \sum_{j=1}^{m} \mathrm{Cov}(\boldsymbol{Y}_j) .$$

Let λ_m be the smallest eigenvalue of Σ_m and Q_m be the law of $\boldsymbol{Y}_1 + \cdots + \boldsymbol{Y}_m$.

There exists an absolute constant c depending only on the dimension d such that

$$\sup_{A\in\mathcal{A}} \left| Q_m(A) - \eta_{\boldsymbol{\mu}_m,\Sigma_m}(A)\right| \leq c\lambda_m^{-3/2} \sum_{j=1}^{m} E(\|\boldsymbol{Y}_j - E[\boldsymbol{Y}_j]\|^3), \tag{2.4}$$

where $\eta_{\boldsymbol{\mu}_m,\Sigma_m}$ is the Gaussian measure of mean μ_m and covariance matrix Σ_m, and $\mathcal{A}$ is the collection of Borel measurable convex subsets of $\mathbb{R}^d$.

[1] We thank the referee for pointing out this in the first version of the manuscript.

In the context of Eq. (1.1), we take as increments for $k > 1$

$$Y_k(x) = \sum_{e^{k-1} < \ln p \le e^k} \frac{1}{p^{1/2}} \big(\cos(x \ln p)\cos(\theta_p) + \sin(x \ln p)\sin(\theta_p)\big)$$
$$= \sum_{e^{k-1} < \ln p \le e^k} \frac{\cos(x \ln p - \theta_p)}{p^{1/2}}, \quad x \in [0, 1].$$

For $k = 1$, the sum is the same with the primes ranging from 2 to e^e. By definition, we then have

$$X_N(x) = \sum_{k=1}^{n} Y_k(x), \tag{2.5}$$

where we set $n = \ln \ln N$. Since the uniform random variables are bounded, the error in approximating Y_k by Gaussian random variables in Eq. (2.4) is the sum over $p^{-3/2}$. To ensure this error is small, it is necessary to truncate the small primes.

With this in mind, consider $R \le N$. Define $\mathcal{F}_R$ to be the σ-algebra generated by $(\theta_p)_{p \le R}$. We will often condition on $\mathcal{F}_R$ to fix the dependence on the small primes. To shorten notation, we also write

$$X_{R,N} = X_N - X_R.$$

The variance of $X_{R,N}(x)$, $x \in [0, 1]$, is by definition

$$\sigma^2_{R,N} \equiv \mathrm{Var}(X_{R,N}(x)) = \frac{1}{2} \sum_{R < p \le N} p^{-1}. \tag{2.6}$$

The prime number theorem, see e.g. [12], implies that the density of the primes goes like $(\ln p)^{-1}$. More precisely, we have by Merten's second theorem

$$\left| \sigma^2_{R,N} - \frac{1}{2}(\ln \ln N - \ln \ln R) \right| = \mathrm{o}(1) \quad \text{as } N \to \infty \text{ and } R \to \infty. \tag{2.7}$$

In the next two sections, we state the results from [2] derived from Lemma 2.2. The reader is referred to [2] for more details of the proofs.

2.1. One-point Gaussian comparison

For one point, the Gaussian comparison is simple.

Proposition 2.3. *For any $a \in \mathbb{R}$ and $R > (\ln N)^{100}$, we have*

$$\mathbb{P}(X_{R,N}(x) > a) = (1 + \mathrm{o}(1)) \int_a^\infty \frac{e^{-y^2/(2\sigma^2_{R,N})}}{\sqrt{2\pi}\sigma_{R,N}} \mathrm{d}y + \mathrm{O}((\ln N)^{-2}).$$

where the error is uniform in a.

Proof. This is Proposition 2.11 in [2] (with $\lambda = 0$). This is a direct consequence of Lemma 2.2. Note that the error is $\ll \sum_{p>R} p^{-3/2} = \mathrm{O}(R^{-1/2}) \ll (\ln N)^{-2}$, by the choice of R. □

For some estimates, Proposition 2.3 is too precise. A plain Chernoff bound is often enough. For this reason, it is useful to compute the moment generating function.

Lemma 2.4. *Let $\lambda \in \mathbb{R}$. Then for any $x \in [0, 1]$ and $R < N$, we have*

$$\mathbb{E}(\exp(\lambda X_{R,N}(x))) = (1 + \mathrm{O}(\lambda^4 R^{-1})) \cdot \exp(\lambda^2 \sigma_{R,N}^2/2),$$

and for an absolute constant $c > 1$,

$$c^{-1}\exp(\lambda^2 \ln\ln R/4 + c\lambda^4) \le \mathbb{E}(\exp(\lambda X_R(x))) \le c\exp(\lambda^2 \ln\ln R/4 + c\lambda^4). \tag{2.8}$$

Proof. Without loss of generality, we can assume $x = 0$. Note that by independence of the θ_p's, we have

$$\mathbb{E}(\exp(\lambda X_{R,N}(0))) = \prod_{R<p\le N} \frac{1}{2\pi}\int_0^{2\pi} \exp\Big(\frac{\lambda}{p^{1/2}}\cos\theta\Big)\mathrm{d}\theta.$$

An expansion of the exponential and integration over θ yields

$$\mathbb{E}(\exp(\lambda X_{R,N}(0))) = \prod_{R<p\le N}\left(1 + \frac{\lambda^2}{4p} + \mathrm{O}(\lambda^4 p^{-2})\right). \tag{2.9}$$

The result then follows from (2.6) by taking the logarithm and by noticing that $\sum_{p>R} p^{-2} = \mathrm{O}(R^{-1})$. The claim (2.8) is obtained the same way using (2.7) by considering the sum over all $p \le R$. □

We stress that Lemma 2.4 implies a Gaussian-like behavior for $X_{R,N}$ and X_R only if λ is small compared to R. However, this will always be the case in the forthcoming estimates. In fact, we will take λ to be fixed as N and R go to infinity. Of course, a Chernoff bound for the large deviation of $X_{R,N}$ and X_R can be upgraded to a Gaussian tail by optimizing over λ. More precisely, one gets for $\lambda = V/\sigma^2$,

$$\mathbb{P}(X_{R,N}(0) > V) \ll \exp(-V^2/(2\sigma_{R,N}^2)). \tag{2.10}$$

The estimate will be used only for V of the order of the variance ensuring that λ is of order one.

2.2. *Two-point comparison*

As in the case of one-point estimates, it will often be enough to use a Chernoff bound for two points. For this purpose, we compute the two-point moment generating function.

Lemma 2.5. *Let $\lambda \in \mathbb{R}$. Then for any $x, x' \in [0, 1]$ and $R < P < Q < N$, we have*

$$\begin{aligned}&\mathbb{E}(\exp(\lambda X_{P,Q}(x) + \lambda' X_{P,Q}(x')))\\ &= (1 + \mathrm{O}(\lambda^4 R^{-1}))\cdot\exp\left(\frac{\lambda^2\sigma_{P,Q}^2}{2} + \frac{\lambda'^2\sigma_{P,Q}^2}{2} + \frac{\lambda\lambda'}{2}\sum_{P<p\le Q}\frac{\cos(|x-x'|\ln p)}{p}\right).\end{aligned}$$

Proof. This is done as in the one-point case. Without loss of generality, we can assume $x' = 0$. The independence of the θ_p's gives

$$\begin{aligned}
&\mathbb{E}(\exp(\lambda X_{P,Q}(x) + \lambda' X_{P,Q}(0)))\\
&= \prod_{P<p\leq Q} \frac{1}{2\pi}\int_0^{2\pi} \exp\Big(\frac{\lambda}{p^{1/2}}\cos(x\ln p - \theta) + \frac{\lambda'}{p^{1/2}}\cos\theta\Big)\mathrm{d}\theta\\
&= \prod_{P<p\leq Q}\left(1 + \frac{\lambda^2}{4p} + \frac{\lambda'^2}{4p} + \frac{\lambda\lambda'}{2p}\cos(x\ln p) + \mathrm{O}(\lambda^{-4}p^2)\right),
\end{aligned}$$

where the second line follows by expanding the exponential and integrating. The claim follows from the observation that $1 + x = e^x + \mathrm{O}(x^2)$. □

As for the one-point estimate, it is possible to get a two-dimensional Chernoff bound

$$\mathbb{P}(X_{P,Q}(x) > u, X_{P,Q}(x') > v) \ll \exp\Big(-\tfrac{1}{2}(u,v)\cdot C_{P,Q}^{-1}(u,v)\Big), \quad u,v>0. \tag{2.11}$$

Here, $C_{P,Q}$ is the covariance matrix of $(X_{P,Q}(x), X_{P,Q}(x'))$

$$C_{P,Q} = \begin{pmatrix}\sigma^2_{P,Q} & \rho_{P,Q}\\ \rho_{P,Q} & \sigma^2_{P,Q}\end{pmatrix} \qquad \rho_{P,Q} = \frac{1}{2}\sum_{P<p\leq Q}\frac{\cos(|x-x'|\ln p)}{p}.$$

Eq. (2.11) is achieved by optimizing (λ,λ'), i.e., $(\lambda,\lambda') = C_{P,Q}^{-1}(u,v)$. The sum of cosines in Lemma 2.5 has very different behavior depending on the distance $|x-x'|$. On one hand, if $|x-x'|\ln Q < 1$, then a Taylor expansion of the cosine yields

$$\begin{aligned}
\frac{1}{2}\sum_{P<p\leq Q}\frac{\cos(|x-x'|\ln p)}{p} &= \sigma^2_{P,Q} + \sum_{P<p\leq Q}\mathrm{O}((|x-x'|\ln p)^2/p)\\
&= \sigma^2_{P,Q} + \mathrm{O}((|x-x'|\ln Q)^2),
\end{aligned}$$

where we use the fact that $\sum_{P<p\leq Q}(\ln p)^2/p \ll (\ln Q)^2$. Roughly speaking, this shows that $X_{P,Q}(x)$ and $X_{P,Q}(x')$ are essentially perfectly correlated whenever $|x-x'| < (\ln Q)^{-1}$. On the other hand, if $|x-x'|\ln P > 1$, the prime number theorem and integration by parts yield

$$\sum_{P<p\leq Q}\frac{\cos(|x-x'|\ln p)}{p} = \mathrm{O}((|x-x'|\ln P)^{-1}),$$

see Lemma 2.1 in [2]. Since the error is typically small, this suggests that $X_{P,Q}(x)$ and $X_{P,Q}(x')$ are essentially independent whenever $|x-x'|\ln P > 1$.

When x, x' are far away, the Chernoff bound is not precise enough. We then resort to the following precise Gaussian comparison. The Gaussian comparison is quite powerful and applies not only for $X_{R,N}$ but for the whole random walk as defined in (2.5). More precisely, consider the loglog-scale notation: $n = \ln\ln N$, $r = \ln\ln R$, $k = \ln\ln K$. We restrict the events below to the discrete set of integers $k \in [0,n]$. Consider the Gaussian random walk

$$S_{R,N} = \sum_{r<k\leq n} G_k\ ,$$

where G_k, $k \leq n$, are IID centered Gaussian random variables of variance $1/2$.

Proposition 2.6. *For $R > (\ln N)^{100}$, we have for $|x - x'| > (\ln R)^{-1/2}$ and the notation as above*

$$\begin{aligned}&\mathbb{P}(X_{R,K}(x) \in A_K, X_{R,K}(x') \in A'_K,\ \forall k \in [r,n])\\&= (1 + o(1))\ \mathbb{P}(S_{R,K} \in A_K,\ \forall k \in [r,n]) \cdot \mathbb{P}(S_{R,K} \in A'_K,\ \forall k \in [r,n]) + O((\ln N)^{-2}).\end{aligned}$$

where A_K and A'_K are intervals of $\mathbb{R}$ and the error term is uniform in the choice of these intervals.

Proof. This is Proposition 2.9 in [2] (with $\lambda = 0$, $m = r$, $\Delta = r/2$). This is a direct consequence of Lemma 2.2. □

3. First moment estimates

The next lemma highlights the fact that the non-trivial contribution to Theorem 1.1 comes from small primes. As before, we write $n = \ln\ln N$ and $r = \ln\ln R$.

Lemma 3.1. *For $W_{\alpha,N}$ as in (1.5), we have for $0 < \alpha < 2$ and $R > (\ln N)^{100}$*

$$\lim_{N\to\infty} \frac{\mathbb{E}\left(W_{\alpha,N}|\mathcal{F}_R\right)}{\mathbb{E}\left(W_{\alpha,N}\right)} = M_\alpha \quad \textit{in probability.} \tag{3.1}$$

Proof. We compute $\mathbb{E}(W_{\alpha,N}|\mathcal{F}_R)$ and $\mathbb{E}(W_{\alpha,N})$ simultaneously. Define the (random) subsets

$$\begin{aligned}B_1 &= \left\{x \in [0,1] : |X_R(x)| > \frac{\alpha}{4}n\right\},\\B_2 &= \left\{x \in [0,1] : |X_R(x)| \in [n^{1/4}, \tfrac{\alpha}{4}n]\right\},\\B_3 &= \left\{x \in [0,1] : |X_R(x)| \le n^{1/4}\right\}.\end{aligned}$$

In view of Eq. (1.2), there are no points in B_2 and B_1 with high probability. However, since we are dealing with the expectation of $W_{\alpha,N}$, these events could still have an effect and need to be controlled. Note that the functions $\mathbb{1}_{B_j}(x)$, $j = 1, 2, 3$ are Borel measurable as subsets of $[0, 1]$ with $\mathbb{P}$-probability one. We split the integral on B_1, B_2 and B_3. As expected, the dominant contribution is from B_3 in expectation and on an event of high probability. The set B_1 is useful since on its complement, the quantity $\frac{\alpha}{2}n - X_R(x)$ is much larger than 1, so a Gaussian estimate will be possible. The contribution of the set B_2 is handled with more care, as one needs a joint control of X_R and $X_{N,R}$.

The set B_3 has large measure in expectation and in probability. Indeed, one has by a Chernoff bound using (2.8)

$$\mathbb{E}(\mathrm{Leb}(B_3)) = \mathbb{P}\Big(|X_R(0)| \le n^{1/4}\Big) \ge 1 - e^{-n^{1/4}}.$$

In particular, this implies by a Markov inequality that B_3 has Lebesgue measure greater than $1/2$ with high probability:

$$\mathbb{P}(\mathrm{Leb}(B_3) \le 1/2) = \mathbb{P}(\mathrm{Leb}(B_3^c) > 1/2) \le 2 \cdot \mathbb{P}\Big(|X_R(0)| > n^{1/4}\Big) \le e^{-n^{1/4}}.$$

One has, using the independence of $X_{R,N}$ and X_R as well as Proposition 2.3,

$$\int_{B_3} \mathbb{P}\Big(X_{R,N}(x) > \frac{\alpha}{2}n - X_R(x)\Big|\mathcal{F}_R\Big)\,\mathrm{d}x = \int_{B_3} \mathrm{d}x \int_{\frac{\alpha}{2}n - X_R(x)}^{\infty} \frac{e^{-y^2/(2\sigma_{R,N}^2)}}{\sqrt{2\pi}\sigma_{R,N}}\mathrm{d}y + O((\ln N)^{-2}).$$

(Note that the error in Proposition 2.3 is uniform in $X_R(x)$). The integral in y can be evaluated using the Gaussian estimate

$$\mathbb{P}(Y > V) = (1+\mathrm{o}(1))\frac{\sigma/V}{\sqrt{2\pi}}\, e^{-V^2/(2\sigma^2)}, \quad V > 1,$$

for Y a Gaussian random variable of mean 0 and variance σ^2. With this, the above equals

$$\begin{aligned} &= (1+\mathrm{o}(1))\int_{B_3}\frac{\sigma_{R,N}}{\sqrt{2\pi n}}e^{-(\frac{\alpha}{2}n - X_R(x))^2/(2\sigma^2_{R,N})} + \mathrm{O}((\ln N)^{-2}) \\ &= (1+\mathrm{o}(1))\frac{e^{-\frac{\alpha^2}{4}(n+r)}}{\alpha\sqrt{\pi n}}\int_{B_3} e^{\alpha X_R(x)}\mathrm{d}x + \mathrm{O}((\ln N)^{-2}), \end{aligned}$$

where the estimate (2.7) and the bound on $X_R(x)$ for $x \in B_3$ are used. Note that on the event $\{\mathrm{Leb}(B_3) > 1/2\}$, the first term is at least

$$\frac{e^{-\frac{\alpha^2}{4}(n+r)}}{\alpha\sqrt{\pi n}}\int_{B_3} e^{\alpha X_R(x)}\mathrm{d}x \geq \frac{e^{-\frac{\alpha^2}{4}(n+r)}}{\alpha\sqrt{\pi n}}\cdot\frac{1}{2}e^{-\alpha n^{1/4}}, \tag{3.2}$$

which is much larger than $(\ln N)^{-2}$ since $\alpha < 2$. Therefore, the error term can be absorbed as a multiplicative error:

$$\int_{B_3}\mathbb{P}\Big(X_{R,N}(x) > \frac{\alpha}{2}n - X_R(x)\Big|\mathcal{F}_R\Big)\,\mathrm{d}x = (1+\mathrm{o}(1))\frac{e^{-\frac{\alpha^2}{4}n+r}}{\alpha\sqrt{\pi n}}\int_{B_3} e^{\alpha X_R(x)}\,\mathrm{d}x. \tag{3.3}$$

Furthermore, note that, by a Chernoff bound (with $\lambda = 100$ say) and (2.8),

$$\mathbb{E}\left(\int_{B_3^c} e^{\alpha X_R(x)}\mathrm{d}x\right) \leq e^{-100n^{1/4}}\cdot\mathbb{E}(e^{\alpha X_R(0)+100|X_R(0)|}) \ll e^{-99n^{1/4}}.$$

We deduce from this that the integral on B_3 can be extended to the whole $[0, 1]$ in the expectation

$$\mathbb{E}\left(\int_{B_3}\mathbb{P}\Big(X_{R,N}(x) > \frac{\alpha}{2}n - X_R(x)\Big|\mathcal{F}_R\Big)\,\mathrm{d}x\right) = (1+\mathrm{o}(1))\frac{e^{-\frac{\alpha^2}{4}(n+r)}}{\alpha\sqrt{\pi n}}\int_0^1\mathbb{E}(e^{\alpha X_R(x)})\mathrm{d}x. \tag{3.4}$$

Similarly, the integral on B_3 can be extended to $[0, 1]$ on a $\mathcal{F}_R$-measurable event of large probability, since $\mathbb{P}\left(\int_{B_3^c} e^{\alpha X_R(x)}\mathrm{d}x > e^{-10n^{1/4}}\right) = \mathrm{o}(1)$, so that on this event

$$\int_{B_3}\mathbb{P}\Big(X_{R,N}(x) > \frac{\alpha}{2}n - X_R(x)\Big|\mathcal{F}_R\Big)\,\mathrm{d}x = (1+\mathrm{o}(1))\frac{e^{-\frac{\alpha^2}{4}(n+r)}}{\alpha\sqrt{\pi n}}\int_0^1 e^{\alpha X_R(x)}\mathrm{d}x\ . \tag{3.5}$$

The conclusion of the lemma follows by considering the ratio of the right-hand side of (3.5) and Eq. (3.4), and by taking the limit $N \to \infty$ as in (1.4). It remains to show that the contributions of B_1 and B_2 are small compared to the one of B_3 in expectation and on an event of high probability.

For B_1, one brutally bounds the probability by 1 to get

$$\int_{B_1}\mathbb{P}\Big(X_{R,N}(x) > \frac{\alpha}{2}n - X_R(x)\Big|\mathcal{F}_R\Big)\,\mathrm{d}x \leq \mathrm{Leb}\Big(B_1\Big).$$

But this measure is small in expectation and with high probability, since by a Chernoff bound again,

$$\mathbb{E}\Big(\text{Leb}(B_1)\Big) = \mathbb{P}(|X_R(x)| > \tfrac{\alpha}{4}n) \leq (\ln N)^{-100}. \tag{3.6}$$

The estimate is a bit more subtle for B_2. We divide the range $[n^{1/4}, \frac{\alpha}{4}n]$ into intervals of length 1. Proceeding as for B_3 and using a Gaussian estimate, one has

$$\begin{aligned}
&\int_{B_2} \mathbb{P}\Big(X_{R,N}(x) > \frac{\alpha}{2}n - X_R(x)\Big|\mathcal{F}_R\Big)\,\mathrm{d}x \\
&\leq \sum_{u=\lfloor n^{1/4}\rfloor}^{\lceil \frac{\alpha}{4}n\rceil} e^{-(\frac{\alpha}{2}n-u)^2/(2\sigma_{R,N}^2)} \cdot \int_{\{x:|X_R(x)|\in[u,u+1]\}} e^{\alpha X_R(x)}\mathrm{d}x + \mathrm{O}((\ln N)^{-2}) \\
&\leq e^{-\frac{\alpha^2}{4}(n+r)} \sum_{u>n^{1/4}} e^{\alpha u} \cdot \int_{\{x:|X_R(x)|\in[u,u+1]\}} e^{\alpha X_R(x)}\mathrm{d}x + \mathrm{O}((\ln N)^{-2}).
\end{aligned} \tag{3.7}$$

This is much smaller than the integral on B_3 in expectation since

$$\mathbb{E}\left(\sum_{u>n^{1/4}} e^{\alpha u} \cdot \int_{\{x:|X_R(x)|\in[u,u+1]\}} e^{\alpha X_R(x)}\mathrm{d}x\right) \leq \sum_{u>n^{1/4}} e^{(\alpha-100)u} \cdot \mathbb{E}(e^{\alpha X_R(0)+100|X_R(0)|}). \tag{3.8}$$

This is $\mathrm{O}(e^{-98(\ln\ln N)^{1/4}})$ by (2.8). This also implies by Markov's inequality that on an event of large probability, this is negligible. This concludes the proof of the lemma. □

The proof of the last lemma also yields a precise estimate for the average measure of high points.

Corollary 3.2. *For $W_{\alpha,N}$ as in (1.5), we have for $0 < \alpha < 2$ and $R > (\ln N)^{100}$,*

$$\mathbb{E}(W_{\alpha,N}|\mathcal{F}_R) = (1+\mathrm{o}(1))\frac{(\ln N)^{-\alpha^2/4}}{\alpha\sqrt{\pi \ln\ln N}} \int_0^1 e^{\alpha X_R(x) - \frac{\alpha^2}{4}\ln\ln R}\,\mathrm{d}x\ . \tag{3.9}$$

In particular, this implies

$$\mathbb{E}(W_{\alpha,N}) \gg \frac{(\ln N)^{-\alpha^2/4}}{\sqrt{\ln\ln N}}.$$

Proof. The equality is a direct consequence of Eq. (3.4) which is the dominant contribution, and of Eqs. (3.6), (3.7), (3.8) which show that the contribution of B_2 and B_1 to the expectation is negligible. The inequality follows from Eq. (2.8) that gives a bound on the moment generating function of X_R. □

We now want to show a barrier-type estimate: the points x such that $X_N(x) > \frac{\alpha}{2}\ln\ln N$ must be such that $X_K(x)$ is close to $\frac{\alpha}{2}\ln\ln K$ for most K's in $[R, N]$. For conciseness, we turn again to a loglog-scale notation: $n = \ln\ln N$, $r = \ln\ln R$, $k = \ln\ln K$. With analogy with random walks, consider the discrete set of integers $k \in [0, n]$. Define for fixed $\epsilon > 0$ and $0 < \delta < 1$

$$\begin{aligned}
W_{\alpha,N}^{>} &= \text{Leb}\left\{x \in [0,1] : X_N(x) \geq \tfrac{\alpha}{2}n;\, \exists k \in [r/2, n(1-\delta)] :\ X_K(x) > (\tfrac{\alpha}{2}+\epsilon)k\right\} \\
W_{\alpha,N}^{<} &= \text{Leb}\left\{x \in [0,1] : X_N(x) \geq \tfrac{\alpha}{2}n;\, \exists k \in [r/2, n(1-\delta)] :\ X_K(x) < (\tfrac{\alpha}{2}-\epsilon)k\right\}.
\end{aligned} \tag{3.10}$$

We need to pick R such that $r = n^{1/100}$, as it needs to be much smaller than n, yet not too small. Picking $r = \ln n$ for example would lead to errors too big in what follows. Note that we have $R > (\ln N)^{100}$ for the choice $r = n^{1/100}$, thereby fulfilling the assumptions of the previous results. The restriction on the range of k for the barrier is necessary as the behavior for small and large primes is not as regular.

Lemma 3.3. *For R such that $r = n^{1/100}$, we have for $0 < \delta < 1$ and $0 < \epsilon < 1 \wedge \frac{\alpha}{4}\delta$,*

$$\frac{\mathbb{E}(W^{>}_{\alpha,N})}{\mathbb{E}(W_{\alpha,N})} = \mathrm{o}(1).$$

In particular, for all $c > 0$, we have

$$\mathbb{P}\left(W^{>}_{\alpha,N} > c\ \mathbb{E}(W_{\alpha,N})\right) = \mathrm{o}(1) \qquad \mathbb{P}\left(\mathbb{E}(W^{>}_{\alpha,N}|\mathcal{F}_R) > c\ \mathbb{E}(W_{\alpha,N})\right) = \mathrm{o}(1), \tag{3.11}$$

where the o*-term depends on c. The same estimates hold for $W^{<}_{\alpha,N}$.*

Proof. We prove the lemma for $W^{>}_{\alpha,N}$ as the proof is very similar for $W^{<}_{\alpha,N}$. Eq. (3.11) is a direct consequence of the first claim by Markov's inequality.

We bound the expectation of $W^{>}_{\alpha,N}$ from above:

$$\begin{aligned}\mathbb{E}\left(W^{>}_{\alpha,N}\right) &\le \int_0^1 \sum_{k=r/2}^{n(1-\delta)} \mathbb{P}\Big(X_N(x) > \tfrac{\alpha}{2}n,\, X_K(x) > (\tfrac{\alpha}{2}+\epsilon)k\Big)\,\mathrm{d}x \\ &\le \int_0^1 \sum_{k=r/2}^{n(1-\delta)} \sum_{v>(\frac{\alpha}{2}+\epsilon)k} \mathbb{P}\Big(X_{K,N}(x) > \tfrac{\alpha}{2}n - v,\, X_K(x) > v\Big)\,\mathrm{d}x,\end{aligned}$$

where the second inequality is obtained by partitioning the range of X_K. Recall that X_K and $X_{K,N}$ are independent. We estimate the probability depending on the value of k and v. Note that we always have $(\frac{\alpha}{2}+\epsilon)k < 2k$, since $\alpha < 2$ and $\epsilon < 1$. We first consider the range $(\frac{\alpha}{2}+\epsilon)k < v < 2k$ and $v < \frac{\alpha}{2}n$. This is the sharpest case since X_K is expected to lie close to $\frac{\alpha}{2}k$ if X_N is around $\frac{\alpha}{2}n$. In this case, the estimate (2.10) can be applied to X_K since v is of the order of the variance. It can also be applied to $X_{K,N}$ since $0 < \frac{\alpha}{2}n - v < \frac{\alpha}{2}(n-k) - \epsilon k$, which is of the order of the variance. This yields

$$\sum_{k=r/2}^{n(1-\delta)} \sum_{v=(\frac{\alpha}{2}+\epsilon)k}^{2k\wedge\frac{\alpha}{2}n} \exp\left(-\frac{(\frac{\alpha}{2}n - v)^2}{n-k} - \frac{v^2}{k}\right).$$

A direct computation shows that the exponential term is maximized at $v = \frac{\alpha}{2}k$. This is much smaller than $v = (\frac{\alpha}{2}+\epsilon)k$. This is therefore the dominant v, and we conclude that the above is

$$\ll \sum_{k=r/2}^{n(1-\delta)} \exp\left(\frac{-(\frac{\alpha}{2}n - (\frac{\alpha}{2}+\epsilon)k)^2}{n-k} + \frac{-\left((\frac{\alpha}{2}+\epsilon)k\right)^2}{k}\right) \ll e^{-\frac{\alpha^2}{4}n} \cdot \sum_{k=r/2}^{n(1-\delta)} e^{-\epsilon^2 k} \ll e^{-\frac{\alpha^2}{4}n - \frac{\epsilon^2}{2}r}.$$

This is $\mathrm{o}(\mathbb{E}(W_{\alpha,N}))$ by Corollary 3.2 for any fixed $\epsilon > 0$ by the choice $r = n^{1/100}$.

The case where $(\frac{\alpha}{2}+\epsilon)k < v < 2k$ and $v > \frac{\alpha}{2}n$ can only occur when $k > \frac{\alpha}{4}n$. For this, the restriction to $X_{K,N}$ can be dropped. This yields

$$\sum_{k=\alpha n/4}^{n(1-\delta)} \sum_{v=\frac{\alpha}{2}n}^{2k} \mathbb{P}\Big(X_{K,N}(0) > \tfrac{\alpha}{2}n - v, X_K(0) > v\Big) \ll \sum_{k=\alpha n/4}^{n(1-\delta)} \sum_{v=\frac{\alpha}{2}n}^{2k} e^{-\frac{v^2}{k}}. \tag{3.12}$$

This is $\ll \exp(\frac{-\alpha^2}{4}\frac{n^2}{n(1-\delta)})$, which is $o(\mathbb{E}[W_{\alpha,N}])$ by Corollary 3.2 for any $0 < \delta < 1$.

It remains to handle the case $v > 2k$. Again, we split into the cases when $v > \frac{\alpha}{2}n$ and $v \leq \frac{\alpha}{2}n$. The latter can only occur for $k \leq \frac{\alpha}{4}n$. In this case, we can apply the Gaussian estimate for $X_{K,N}$. However, it might not hold for X_K for large v. Instead, we rely on a plain exponential Chernoff bound (with parameter a) to get

$$\sum_{k=r/2}^{\frac{\alpha}{4}n} \sum_{v=2k}^{\frac{\alpha}{2}n} \exp\left(-\frac{(\frac{\alpha}{2}n - v)^2}{n-k} - av + \frac{a^2}{4}k\right).$$

The summand is maximized at $v = \frac{\alpha}{2}n - \frac{a}{2}(n-k)$. We pick $a = 2 > \alpha$ so that the maximizer is simply $v = k + n(\frac{\alpha}{2}-1) < k$. The maximizer is outside the range of v, hence the maximizing value of v is $2k$ as expected. Writing $2k = \frac{\alpha}{2}k + (2-\frac{\alpha}{2})k$, the above is

$$\ll \sum_{k=r/2}^{\frac{\alpha}{4}n} n \cdot e^{-\frac{\alpha^2}{4}n} \cdot e^{-\frac{k}{4}(\alpha^2-8\alpha+12)} \ll \sum_{k=r/2}^{\frac{\alpha}{4}n} n \cdot e^{-\frac{\alpha^2}{4}n} \cdot e^{-(2-\alpha)k} \ll n \cdot e^{-\frac{\alpha^2}{4}n} e^{-(1-\frac{\alpha}{2})r}.$$

This is $o(\mathbb{E}(W_{\alpha,N}))$ by Corollary 3.2 for any fixed $\epsilon > 0$ by the choice $r = n^{1/100}$. In the case $v > 2k$ and $v > \frac{\alpha}{2}n$, the idea is again to drop the probability of $X_{K,N}$. The probability $\mathbb{P}(X_K(x) > v)$ is estimated for $k \leq \frac{\alpha}{4}n$ and $k > \frac{\alpha}{4}n$ separately. In the latter case, we can use a Gaussian estimate as the value of v is of the order of the variance. This gives an estimate as in (3.12), which is small. In the case, $k \leq \frac{\alpha}{4}n$, we can use an exponential Chernoff bound (with $a = 2$) to get that

$$\sum_{k=r/2}^{\frac{\alpha}{4}n} \sum_{v>2k\vee\frac{\alpha}{2}n} \mathbb{P}(X_K(x) > v) \ll \sum_{k=r/2}^{\frac{\alpha}{4}n} \sum_{v>2k\vee\frac{\alpha}{2}n} e^{-2v+k} \ll n \cdot e^{-\alpha n + \frac{\alpha}{4}n},$$

which is smaller than $\frac{e^{-\frac{\alpha^2}{4}n}}{\sqrt{n}}$, since $\alpha < 2$. This completes the proof of the lemma. □

4. Second moment estimates

The main result of this section is:

Proposition 4.1. *For R such that $r = n^{1/100}$, we have for any fixed $\alpha \in (0,2)$*

$$\mathbb{P}\left(\left|\frac{W_{\alpha,N} - \mathbb{E}\left(W_{\alpha,N}|\mathcal{F}_R\right)}{\mathbb{E}\left(W_{\alpha,N}\right)}\right| > c\right) = o(1), \quad \textit{as } N \to \infty, \tag{4.1}$$

where the o*-term depends on* c.

4.1. Proof of Proposition 4.1

To prove the proposition, we consider the following reduction. Let

$$W^{=}_{\alpha,N} = \mathrm{Leb}\Big\{x \in [0,1] : X_N(x) \geq \tfrac{\alpha}{2}n; \forall k \in [r/2, n(1-\delta)] : \ (\tfrac{\alpha}{2}-\epsilon)k \leq X_K(x) \leq (\tfrac{\alpha}{2}+\epsilon)k\Big\}.$$

Note that $W^{=}_{\alpha,N} = W_{\alpha,N} - W^{>}_{\alpha,N} - W^{<}_{\alpha,N}$, as defined in (3.10). In particular, we have the following decomposition

$$\begin{aligned}\frac{W_{\alpha,N} - \mathbb{E}(W_{\alpha,N}|\mathcal{F}_R)}{\mathbb{E}(W_{\alpha,N})} &= \frac{W^{=}_{\alpha,N} - \mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R)}{\mathbb{E}(W_{\alpha,N})} + \frac{W^{>}_{\alpha,N}}{\mathbb{E}(W_{\alpha,N})} - \frac{\mathbb{E}(W^{>}_{\alpha,N}|\mathcal{F}_R)}{\mathbb{E}(W_{\alpha,N})} \\ &\quad + \frac{W^{<}_{\alpha,N}}{\mathbb{E}(W_{\alpha,N})} - \frac{\mathbb{E}(W^{<}_{\alpha,N}|\mathcal{F}_R)}{\mathbb{E}(W_{\alpha,N})}.\end{aligned} \tag{4.2}$$

The last four terms are small in probability by Lemma 3.3. Therefore, the proof of the proposition is reduced to show that the first term is also small in probability. For $\eta > 0$, consider the $\mathcal{F}_R$-measurable event

$$\mathcal{A}_{\eta,N} = \{\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R) \leq \eta^{-1}\mathbb{E}(W_{\alpha,N})\}. \tag{4.3}$$

Note that the complement has small probability, uniformly in N, since by Markov's inequality,

$$\mathbb{P}(\mathcal{A}^c_{\eta,N}) \leq \eta,$$

Therefore, the proof of the proposition can be reduced to showing that for fixed $\eta > 0$,

$$\lim_{N\to\infty} \mathbb{E}\left(\left(\frac{W^{=}_{\alpha,N} - \mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R)}{\mathbb{E}(W_{\alpha,N})}\right)^2 \mathbb{1}_{\mathcal{A}_{\eta,N}}\right) = 0\ .$$

This will follow once it is shown that for fixed η

$$\begin{aligned}\mathbb{E}\Big((W^{=}_{\alpha,N} - \mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2\ \mathbb{1}_{\mathcal{A}_{\eta,N}}\Big) &= \mathbb{E}\Big((W^{=}_{\alpha,N})^2\mathbb{1}_{\mathcal{A}_{\eta,N}} - (\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2\ \mathbb{1}_{\mathcal{A}_{\eta,N}}\Big) \\ &= \mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2).\end{aligned} \tag{4.4}$$

Clearly, we have

$$\begin{aligned}&(W^{=}_{\alpha,N})^2 \\ &= \mathrm{Leb}^{\times 2}\{(x,x') \in [0,1]^2 : \forall_{y\in\{x,x'\}}\ X_N(y) > \tfrac{\alpha}{2}n, \\ &\quad \forall_{k\in[r/2,n(1-\delta)]}\ (\tfrac{\alpha}{2}-\epsilon)k \leq X_K(y) \leq (\tfrac{\alpha}{2}+\epsilon)k\}.\end{aligned}$$

Let $0 < \delta < 1-\alpha^2/4$. We divide $(W^{=}_{\alpha,N})^2$ into three terms depending on the distance between x and x':

$$(I) : |x-x'| < e^{-n(1-\delta)} \qquad (II) : e^{-n(1-\delta)} \leq |x-x'| < e^{-r/2}$$
$$(III) : |x-x'| > e^{-r/2}\ .$$

With this notation, Eq. (4.4) gives

$$0 \leq \mathbb{E}(I) + \mathbb{E}(II) + \mathbb{E}\Big((III)\mathbb{1}_{\mathcal{A}_{\eta,N}} - (\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2\mathbb{1}_{\mathcal{A}_{\eta,N}}\Big),$$

where we dropped the event in the first two terms. Lemmas 4.2 and 4.3, and 4.4 prove that this is $\mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2)$, and thereby finish the proof of the proposition. Note that it is sufficient to show that the positive part of $\mathbb{E}\Big((III)\mathbb{1}_{\mathcal{A}_{\eta,N}} - (\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2\mathbb{1}_{\mathcal{A}_{\eta,N}}\Big)$ is small.

Lemma 4.2. *Let* $0 < \alpha < 2$*. For* R *such that* $r = n^{1/100}$*, we have as* $N \to \infty$*,*

$$\mathbb{E}(I) = \mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2). \tag{4.5}$$

Proof. By noting that for x fixed we have $\mathrm{Leb}\{x' : |x - x'| < e^{-n(1-\delta)}\} \le e^{-n(1-\delta)}$, we bound $\mathbb{E}(I)$ from above by simply dropping all the restrictions on x' and keeping only the endpoint restriction for x:

$$\mathbb{E}(I) \ll e^{-n(1-\delta)} \int_0^1 \mathbb{P}\left(X_N(x) > \frac{\alpha}{2} n\right) \mathrm{d}x = e^{-n(1-\delta)} \mathbb{E}(W_{\alpha,N}). \tag{4.6}$$

Hence, we have

$$\mathbb{E}(I) \ll \mathbb{E}(W_{\alpha,N})^2 \cdot \frac{e^{-n(1-\delta)}}{\mathbb{E}(W_{\alpha,N})} = \mathbb{E}(W_{\alpha,N})^2 \cdot \mathrm{o}(1), \tag{4.7}$$

by Corollary 3.2 and the choice $\delta < 1 - \alpha^2/4$. □

The estimate of (II) is where the restriction established in Lemma 3.3 comes in handy. We will only use the restriction at each K.

Lemma 4.3. *Let* $0 < \alpha < 2$*. For* R *such that* $r = n^{1/100}$*, we have as* $N \to \infty$

$$\mathbb{E}(II) = \mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2). \tag{4.8}$$

Proof. Write $(II)_k$ for the contribution of the pairs of points with $e^{-(k+1)} \le |x - x'| \le e^{-k}$. For conciseness, we write $\mathbf{X}_K(\mathbf{x}) = (X_K(x), X_K(x'))$ with $\mathbf{x} = (x, x')$ and $\mathbf{u} = (u, u')$. By decomposing the values of $X_K(x)$ and $X_K(x')$, we get that the contribution of a fixed k is

$$\mathbb{E}((II)_k) \ll \iint_{|x-x'| \le e^{-k}} \sum_{(\frac{\alpha}{2}-\epsilon)k \le u,u' \le (\frac{\alpha}{2}+\epsilon)k} \mathbb{P}\Big(\mathbf{X}_K(\mathbf{x}) > \mathbf{u}\Big) \cdot \mathbb{P}\Big(\mathbf{X}_{K,N}(\mathbf{x}) > \tfrac{\alpha}{2}(n,n) - \mathbf{u}\Big)\, \mathrm{d}x\mathrm{d}x'. \tag{4.9}$$

Note that we can assume that $u, u' < \frac{\alpha}{2} n$ provided we choose $\epsilon < \delta\alpha/4$ from the equation $(\frac{\alpha}{2} + \epsilon)k < \frac{\alpha}{2} n$, $k \le n(1-\delta)$. The two probabilities can then be estimated by a Chernoff bound as in Eq. (2.11). We evaluate the first probability in (4.9). We expect that $X_K(x)$ and $X_K(x')$ are almost perfectly correlated. The covariance matrix of $(X_K(x), X_K(x'))$ is by Lemma 2.5

$$C_K = \begin{pmatrix} \sigma_K^2 & \rho_K \\ \rho_K & \sigma_K^2 \end{pmatrix}, \qquad \rho_K = \frac{1}{2} \sum_{1 < p \le K} \frac{\cos(|x-x'| \ln p)}{p} = \sigma_K^2 + \mathrm{O}(|x - x'|^2 e^{2k}).$$

Note that the error term $\mathrm{O}(|x - x'|^2 e^{2k})$ is order one for the range of $x - x'$ considered. Denote this error term by $-c^{-1}$. The inverse of this matrix is then

$$(1 + \mathrm{o}(1)) \left\{ \frac{c}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{1}{k} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

The Chernoff bound (2.11) then yields

$$\mathbb{P}\Big(\mathbf{X}_K(\mathbf{x}) > \mathbf{u}\Big) \ll \exp\Big(-\tfrac{c}{4}(u - u')^2\Big) \cdot \exp\Big(-\tfrac{u \cdot u'}{k}\Big). \tag{4.10}$$

The first term is an effective delta function for $u = u'$, whereas the second term provides the Gaussian decay for a single X_K (as $u \approx u'$ effectively from the first term). The second probability in (4.9) is evaluated the same way. The covariance matrix is now

$$C_{K,N} = \begin{pmatrix} \sigma^2_{K,N} & \rho_{K,N} \\ \rho_{K,N} & \sigma^2_{K,N} \end{pmatrix}, \qquad \rho_{K,N} = \frac{1}{2} \sum_{K<p\le N} \frac{\cos(|x-x'|\ln p)}{p} = \mathrm{O}(1).$$

Therefore, the Chernoff bound (2.11) yields

$$\mathbb{P}(\mathbf{X}_{K,N}(\mathbf{x}) > \tfrac{\alpha}{2}(n,n) - \mathbf{u}) \ll \exp\Big(-\frac{(\frac{\alpha}{2}n-u)^2}{n-k} - \frac{(\frac{\alpha}{2}n-u')^2}{n-k}\Big). \tag{4.11}$$

The dominant term in Eq. (4.9) is obtained by optimizing (4.10) and (4.11) over u, u'. The solution is

$$u, u' = \tfrac{\alpha}{2}k \cdot \tfrac{2n}{n+k}.$$

This is larger than $(\frac{\alpha}{2}+\epsilon)k$ for the choice $\epsilon < \alpha\delta/4$ when $k < n(1-\delta)$. The upshot is that the dominant term in the range of $\mathbf{u}$ of interest is simply $u = u' = (\frac{\alpha}{2}+\epsilon)k$. Putting this back in (4.9) with the estimates of (4.10) and (4.11) yields

$$\begin{aligned}\mathbb{E}((II)_k) &\ll e^{-k} \cdot \exp\Big(-(\tfrac{\alpha}{2}+\epsilon)^2 k - \tfrac{2}{n-k}(\tfrac{\alpha}{2}n - (\tfrac{\alpha}{2}+\epsilon)k)^2\Big) \\ &\ll e^{-\frac{\alpha^2}{2}n} \cdot \exp(-k(1-\tfrac{\alpha^2}{4}-\alpha\epsilon)) .\end{aligned}$$

Summing this over $k \ge r/2$ is $\mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2)$ by Corollary 3.2 and the choice of r for ϵ small enough, since $\alpha < 2$. □

Lemma 4.4. *Let $0 < \alpha < 2$ and $\mathcal{A}_{\eta,N}$ as in Eq. (4.3) for $\eta > 0$ fixed. For R such that $r = n^{1/100}$, we have as $N \to \infty$*

$$\mathbb{E}\Big((III)\mathbb{1}_{\mathcal{A}_{\eta,N}} - (\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2\mathbb{1}_{\mathcal{A}_{\eta,N}}\Big)_+ = \mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2). \tag{4.12}$$

Proof. Recall that the event $\mathcal{A}_{\eta,N}$ is $\mathcal{F}_R$-measurable. First, note that the indicator function of the barrier event $\{\forall k \in [r/2, n(1-\delta)] : (\frac{\alpha}{2}-\epsilon)k \le X_K(x) \le (\frac{\alpha}{2}+\epsilon)k\}$ can be written as $\mathbb{1}_{\mathcal{B}(x)} = \mathbb{1}_{\mathcal{B}_R(x)} \cdot \mathbb{1}_{\mathcal{B}_{R,K}(x)}$ where

$$\begin{aligned}\mathbb{1}_{\mathcal{B}_R(x)} &= \prod_{r/2\le k\le r} \mathbb{1}_{\{(\frac{\alpha}{2}-\epsilon)k \le X_K(x) \le (\frac{\alpha}{2}+\epsilon)k\}} \\ \mathbb{1}_{\mathcal{B}_{R,K}(x)} &= \prod_{r\le k\le n(1-\delta)} \mathbb{1}_{\{(\frac{\alpha}{2}-\epsilon)k - X_R(x) \le X_{R,K}(x) \le (\frac{\alpha}{2}+\epsilon)k - X_R(x)\}}.\end{aligned}$$

Note also that the event $\mathcal{B}_R(x)$ is $\mathcal{F}_R$-measurable for every x. By linearity of conditional expectation, we have

$$\begin{aligned}&\mathbb{E}((III)|\mathcal{F}_R) \\ &= \iint_{|x-x'|>e^{-r/2}} \mathbb{P}\Big(\{X_N(x) > \frac{\alpha}{2}n\} \cap \mathcal{B}(x) \cap \{X_N(x') > \frac{\alpha}{2}n\} \cap \mathcal{B}(x') \Big| \mathcal{F}_R\Big)\, \mathrm{d}x\mathrm{d}x' \\ &= \iint_{|x-x'|>e^{-r/2}} \mathbb{P}\Big(\forall y \in \{x,x'\}\ \{X_{R,N}(y) > \frac{\alpha}{2}n - X_R(y)\} \cap \mathcal{B}_{R,N}(y) \Big| X_R\Big) \\ &\quad \times \mathbb{1}_{\mathcal{B}_R(x)\cap\mathcal{B}_R(x')} \mathrm{d}x\mathrm{d}x'.\end{aligned}$$

The conditional probability is of the form of Proposition 2.6 for the process $X_{R,K}$, $R < K \le N$. Hence, we have that $\mathbb{E}((III)|\mathcal{F}_R)$ equals

$$(1+\mathrm{o}(1)) \times \iint_{|x-x'|>e^{-r/2}} \prod_{y\in\{x,x'\}} \mathbb{1}_{\mathcal{B}_R(y)} \cdot \mathbb{P}\Big(\{X_{R,N}(y) > \frac{\alpha}{2}n - X_R(y)\} \cap \mathcal{B}_{R,N}(y) \Big| X_R\Big) \,\mathrm{d}x\mathrm{d}x' + \mathrm{O}((\ln N)^{-2}).$$

Note that by definition

$$\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R)^2 = \iint_{[0,1]^2} \prod_{y\in\{x,x'\}} \mathbb{1}_{\mathcal{B}_R(y)} \cdot \mathbb{P}\Big(\{X_{R,N}(y) > \frac{\alpha}{2}n - X_R(y)\} \cap \mathcal{B}_{R,N}(y) \Big| X_R\Big) \,\mathrm{d}x\mathrm{d}x'.$$

By dropping the contribution of $|x-x'| \le e^{-r/2}$ in the double integral, we get

$$\Big(\mathbb{E}((III)|\mathcal{F}_R) - (\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2\Big)_+ = \mathrm{o}(1)\cdot(\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R))^2 + \mathrm{O}((\ln N)^{-2}).$$

On the event $\mathcal{A}_{\eta,N}$, the first term is $\mathrm{o}((\mathbb{E}(W_{\alpha,N}))^2)$. Therefore, it remains to show that the error term $\mathrm{O}((\ln N)^{-2})$ is small compared to the double integral. Note that by Jensen's inequality

$$\mathbb{E}\Big(\mathbb{E}(W^{=}_{\alpha,N}|\mathcal{F}_R)^2\Big) \gg \Big(\mathbb{E}(W^{=}_{\alpha,N})\Big)^2 \gg \frac{e^{-\frac{\alpha^2}{2}n}}{n},$$

by Corollary 3.2. This is much larger than $(\ln N)^{-2}$ for $\alpha < 2$. This proves the lemma. □

5. Proof of Theorem 1.1

We are in the position to prove Theorem 1.1 using Lemma 3.1 and Proposition 4.1. First, by picking R such that $r = n^{1/100}$, we write

$$\frac{W_{\alpha,N}}{\mathbb{E}\left(W_{\alpha,N}\right)} = \frac{\mathbb{E}\left(W_{\alpha,N}|\mathcal{F}_R\right)}{\mathbb{E}\left(W_{\alpha,N}\right)} + \frac{W_{\alpha,N} - \mathbb{E}\left(W_{\alpha,N}|\mathcal{F}_R\right)}{\mathbb{E}\left(W_{\alpha,N}\right)}\,. \tag{5.1}$$

By Proposition 4.1 the second summand converges to zero in probability as $N \to \infty$. By Lemma 3.1, the first summand converges almost surely to M_α defined in (1.4). This completes the proof of Theorem 1.1.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.spa.2022.04.017.

Acknowledgments

Lisa Hartung and Nicola Kistler thank the Rhine-Main Stochastic group for creating an interactive research environment leading to this article. We are indebted to the referee who found a major mistake in a first version of the proof that eventually led to a simplification of the proof of the theorem.

References

[1] L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwill, K. Soundararajan, Maximum of the Riemann zeta function on a short interval of the critical line, Comm. Pure Appl. Math. 72 (3) (2019) 500–535.

[2] L.-P. Arguin, D. Belius, A.J. Harper, Maxima of a randomized Riemann zeta function, and branching random walks, Ann. Appl. Probab. 27 (1) (2017) 178–215.

[3] R.N. Bhattacharya, R. Ranga Rao, Normal approximation and asymptotic expansions, in: Wiley Series in Probability and Mathematical Statistics, 1976, p. xiv+274.

[4] M. Biskup, O. Louidor, On intermediate level sets of two-dimensional discrete Gaussian free field, Ann. Inst. Henri Poincare 55 (4) (2019) 1948–1987.

[5] R. Chhaibi, J. Najnudel, A. Nikeghbali, The circular unitary ensemble and the Riemann zeta function: The microscopic landscape and a new approach to ratios, Invent. Math. 207 (1) (2017) 23–113.

[6] B. Duplantier, R. Rhodes, S. Sheffield, V. Vargas, Critical Gaussian multiplicative chaos: Convergence of the derivative martingale, Ann. Probab. 42 (5) (2014) 1769–1808.

[7] Y.V. Fyodorov, J.P. Keating, Freezing transitions and extreme values: Random matrix theory, $\zeta(\frac{1}{2}+it)$ and disordered landscapes, Phil. Trans. R. Soc. A 372 (20120503) (2014) 1–32.

[8] C. Glenz, N. Kistler, M.A. Schmidt, High points of branching Brownian motion and McKean's martingale in the Bovier-Hartung extremal process, Electron. Commun. Probab. 23 (2018) 12.

[9] F. Götze, On the rate of convergence in the multivariate CLT, Ann. Probab. 19 (2) (1991) 724–739.

[10] A.J. Harper, A note on the maximum of the Riemann zeta function, and log-correlated random variables, 2013, pp. 1–26, Preprint, arXiv:1304.0677.

[11] J.-P. Kahane, Sur le chaos multiplicatif, Ann. Sci. Math. Québec 9 (2) (1985) 105–150.

[12] H.L. Montgomery, R.C. Vaughan, Multiplicative number theory, in: I. Classical Theory, in: Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[13] 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) (2018) 387–452.

[14] R. Rhodes, V. Vargas, Gaussian multiplicative chaos and applications: A review, Probab. Surv. 11 (2014) 315–392.

[15] R. Robert, V. Vargas, Gaussian multiplicative chaos revisited, Ann. Probab. 38 (2) (2010) 605–631.

[16] E. Saksman, C. Webb, Multiplicative chaos measures for a random model of the Riemann zeta function, 2016, arXiv e-prints, arXiv:1604.08378.

[17] E. Saksman, C. Webb, The Riemann zeta function and Gaussian multiplicative chaos: Statistics on the critical line, Ann. Probab. 48 (6) (2020) 2680–2754.