

# Maxima of a random model of the Riemann zeta function over intervals of varying length

Louis-Pierre Arguin<sup>1,a</sup>, Guillaume Dubach<sup>2,b</sup> and Lisa Hartung<sup>3,c</sup>

<sup>1</sup>*CUNY, Baruch College & Graduate Center, New York, NY, USA, <sup>a</sup>[louis-pierre.arguin@baruch.cuny.edu](mailto:louis-pierre.arguin@baruch.cuny.edu)*

<sup>2</sup>*CMLS, École Polytechnique, Palaiseau, France, <sup>b</sup>[guillaume.dubach@polytechnique.edu](mailto:guillaume.dubach@polytechnique.edu)*

<sup>3</sup>*Johannes Gutenberg Universität, Mainz, Germany, <sup>c</sup>[lhartung@uni-mainz.de](mailto:lhartung@uni-mainz.de)*

Received 30 April 2021; revised 7 September 2022; accepted 23 September 2022

**Abstract.** We consider a model of the Riemann zeta function on the critical axis and study its maximum over intervals of length  $(\log T)^\theta$ , where  $\theta$  is either fixed or tends to zero at a suitable rate. It is shown that the deterministic level of the maximum interpolates smoothly between the ones of log-correlated variables and of i.i.d. random variables, exhibiting a smooth transition ‘from  $\frac{3}{4}$  to  $\frac{1}{4}$ ’ in the second order. This provides a natural context where extreme value statistics of log-correlated variables with *time-dependent variance and rate* occur. A key ingredient of the proof is a precise upper tail tightness estimate for the maximum of the model on intervals of size one, that includes a Gaussian correction. This correction is expected to be present for the Riemann zeta function and pertains to the question of the correct order of the maximum of the zeta function in large intervals.

**Résumé.** Nous considérons un modèle aléatoire des valeurs de la fonction zêta de Riemann sur sa droite critique pour en étudier le maximum sur des intervalles de longueur  $(\log T)^\theta$ , où  $\theta$  est soit fixé, soit tend vers zéro à une vitesse spécifiquement calibrée. Nous établissons que la valeur déterministe attendue de ce maximum passe continûment de celle obtenue pour des variables log-corrélées à celle de variables i.i.d., effectuant une transition ‘de  $3/4$  à  $1/4$ ’ au second ordre. Il s’agit d’un contexte naturel pour l’étude des valeurs extrêmes de variables log-corrélées avec un taux et une variance variables en temps. Un ingrédient-clé de notre preuve est une estimation fine de la queue de distribution du maximum sur des intervalles de longueur unité, qui inclut une correction de type gaussien. Une telle correction est également escomptée dans les statistiques du maximum de la véritable fonction zêta sur des intervalles aléatoires de taille unité, ce qui est lié à la question analogue du maximum de la fonction zêta sur des intervalles plus longs de la droite critique.

*MSC2020 subject classifications:* Primary 11M06; 60G70; secondary 60J80

*Keywords:* Extreme value theory; Riemann zeta function; Branching random walk

## 1. Introduction

The Riemann zeta function is defined for  $\operatorname{Re} s > 1$  as

$$(1.1) \quad \zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} = \prod_{p \geq 1} \frac{1}{1 - p^{-s}},$$

where the Euler product on the left is over all primes. It extends uniquely to an analytic function over  $\mathbb{C} \setminus \{1\}$ . The Riemann hypothesis states that, except for the negative even integers (the so-called trivial zeros), the zeros of the function lie on the critical line  $\operatorname{Re} s = 1/2$ . The large values of the function on the critical line also plays a major role in number theory. The Riemann hypothesis implies that the maximum over a large interval  $[T, 2T]$  is

$$(1.2) \quad \max_{t \in [T, 2T]} |\zeta(1/2 + it)| = O\left(\exp\left(C \frac{\log T}{\log \log T}\right)\right),$$

for some  $C > 0$ , as first shown in [31]. The question of whether or not the right-hand side is the true order of the maximum is still subject of debates, see for example [20] for probabilistic arguments motivating

$$(1.3) \quad \max_{t \in [T, 2T]} |\zeta(1/2 + it)| \approx \exp\left(\frac{1}{2} \sqrt{\log T \cdot \log \log T}\right).$$

The question of the order of the maximum in shorter intervals is much better understood. Fyodorov, Hiary & Keating [24] and Fyodorov & Keating [25] conjectured using techniques of random matrix theory and log-correlated processes that, if  $\tau$  is chosen uniformly on  $[T, 2T]$ ,

$$(1.4) \quad \max_{|h| \leq 1} |\zeta(1/2 + i(\tau + h))| = \frac{\log T}{(\log \log T)^{3/4}} e^{\mathcal{M}(T)},$$

where  $(\mathcal{M}(T), T > 1)$  is a tight sequence of random variables converging as  $T \rightarrow \infty$  to a random variable  $\mathcal{M}$  with right tail  $\mathbb{P}(\mathcal{M} > y) \sim Cye^{-2y}$ . The leading order  $\log T$  was proved in [35] (on the Riemann hypothesis) and in [2] unconditionally. An upper bound for the leading order, and subleading order was proved in [28] with an error of  $\log \log \log T$ . The upper tail tightness with the correct decay  $ye^{-2y}$  was recently established in [4].

The techniques developed so far can be also used to investigate large values in intervals with length varying with  $T$ . It was conjectured in [6] that, for intervals of size  $(\log T)^\theta$  where  $\theta > 0$  is fixed, the maximum is

$$(1.5) \quad \max_{|h| \leq (\log T)^\theta} |\zeta(1/2 + i(\tau + h))| = \frac{(\log T)^{\sqrt{1+\theta}}}{(\log \log T)^{\frac{1}{4\sqrt{1+\theta}}}} e^{\mathcal{M}_\theta(T)},$$

where  $(\mathcal{M}_\theta(T), T > 1)$  is a tight sequence of random variables. The leading order  $(\log T)^{\sqrt{1+\theta}}$  was also proved there. This would suggest an interesting jump discontinuity in the subleading order in (1.4) and (1.5) where the exponent  $\frac{1}{4\sqrt{1+\theta}}$  does not approach  $\frac{3}{4}$  as  $\theta \rightarrow 0$ . The main goal of this paper is to shed light on this discontinuity.

Using a random model of the Riemann zeta function, we motivate the following conjecture that identifies the proper scale to smoothe the discontinuity:

**Conjecture.** Let  $0 < \alpha < 1$  and  $\theta \sim (\log \log T)^{-\alpha}$ . Consider an interval of size  $(\log T)^\theta = \exp(\log \log T)^{1-\alpha}$ . If  $\tau$  is chosen uniformly on  $[T, 2T]$ , we have

$$(1.6) \quad \max_{|h| \leq \exp(\log \log T)^{1-\alpha}} |\zeta(1/2 + i(\tau + h))| = \frac{(\log T)^{\sqrt{1+\theta}}}{(\log \log T)^{\frac{1+2\alpha}{4}}} e^{\mathcal{M}_\alpha(T)},$$

for some tight sequence of random variables  $(\mathcal{M}_\alpha(T), T > 1)$ . In other words, the relation between Equations (1.4) for  $\theta > 0$  and (1.5) for  $\theta = 0$  is provided by taking  $\theta \downarrow 0$  like  $(\log \log T)^{-\alpha}$ ,  $0 < \alpha < 1$ .

Note that the non-trivial exponent

$$\frac{1+2\alpha}{4}$$

interpolates between the case  $\theta > 0$  corresponding to  $\alpha = 0$ , and  $\theta = 0$  for  $\alpha = 1$ . This transition is proved for a random Euler product defined in the next section, see Theorems 1.1 and 1.2. Along the way, a correction for the decay  $ye^{-2y}$  for the right tail of the maximum will be proved in the case  $\theta = 0$ , see Theorem 1.3. This is relevant for making sense of the two competing scenarios (1.2) and (1.3) for the global maximum of the zeta function as explained in the next section.

There has been much interest recently in the study of log-correlated processes with *time-inhomogeneous variances and rates*, see for example [5,7,10,13,14,19,21–23,32,33,36]. In particular, such models can be designed in such a way that extreme value statistics interpolate between those of log-correlated variables and those of i.i.d. random variables [9,15,30]. The above conjecture provides a natural context, the Riemann zeta function, where such interpolated extreme value statistics would occur.

### 1.1. Main results

The following random model has been proposed in [27] to study the large values of  $\log |\zeta|$  in a short interval  $I$ :

$$(1.7) \quad W_T(h) = \sum_{p \leq T} \frac{\operatorname{Re}(U_p p^{-ih})}{p^{1/2}}, \quad h \in I,$$

where the sum is over primes less than  $T$  and where  $(U_p, p \text{ primes})$  are i.i.d. random variables distributed uniformly on the unit circle. The model morally corresponds to a formal expansion of the logarithm of the Euler product in (1.1) around  $s = 1/2 + i(\tau + h)$ . The identification of the random phases  $p^{-i\tau} \leftrightarrow U_p$  is then made with the extra assumption that the  $U_p$ 's are independent, which is not exactly the case for the  $p^{-i\tau}$ 's. The resulting model (1.7), when exponentiated, can thus be seen as a random Euler product.<sup>1</sup>

It was shown in [3] that with high probability

$$(1.8) \quad \lim_{T \rightarrow \infty} \max_{|h| \leq 1} \frac{W_T(h) - \log \log T}{\log \log \log T} = -\frac{3}{4} \quad \text{in probability,}$$

thereby providing evidence for the conjecture (1.4) for the log of the actual zeta function.

This work is concerned with a Gaussian version of this model where the  $U_p$ 's are replaced by standard complex Gaussian

$$(1.9) \quad X_T(h) = \sum_{p \leq T} \frac{\operatorname{Re}(G_p p^{-ih})}{p^{1/2}}, \quad h \in I,$$

where the  $G_p$ 's are i.i.d. standard complex Gaussian variables. As explained in Section 2, the process  $(X_T(h), |h| \leq 1)$  is a Gaussian log-correlated process with variance  $\frac{1}{2} \sum_{p \leq T} p^{-1} = \frac{1}{2} \log \log T + O(1)$ .

We choose to work directly with the Gaussian model (1.9) to highlight the new ideas of the proof. A standard approach to studying model (1.7) is to resort to Gaussian comparison; so that working with a Gaussian process from the start makes the outline of the proof much clearer. This is done at virtually no cost to the strength of the result, as the two processes are known to be close, as far as the relevant statistics are concerned, see for example [3] for a pointwise comparison relying on Berry-Esseen bounds, and Theorem 1.7 in [37] for a global approximation. It was shown in [37] that the exponential of the process  $(W_T(h), |h| \leq 1)$ , suitably normalized, converges to Gaussian multiplicative chaos. Any of these methods could be similarly applied on top of the present work, in order to extend its results to other models such as (1.7).

The first result is concerned with intervals of length  $(\log T)^\theta$  for fixed  $\theta > 0$ , where i.i.d. extreme value statistics prevail.

**Theorem 1.1.** *For  $\theta > 0$  fixed, we have*

$$(1.10) \quad \lim_{T \rightarrow \infty} \max_{|h| \leq (\log T)^\theta} \frac{X_T(h) - \sqrt{1+\theta} \log \log T}{\log \log \log T} = \frac{-1}{4\sqrt{1+\theta}} \quad \text{in probability.}$$

This is consistent with Equation (1.5), after exponentiation. The result is easier to prove than Equation (1.8), as correlations play a lesser role. The second and main result is an explicit interpolation between Equations (1.8) and (1.10), when  $\theta$  goes to zero with  $T$  at a suitable rate.

**Theorem 1.2.** *For  $\alpha \in (0, 1)$  and  $\theta = (\log \log T)^{-\alpha}$ , we have*

$$(1.11) \quad \lim_{T \rightarrow \infty} \max_{|h| \leq (\log T)^\theta} \frac{X_T(h) - \sqrt{1+\theta} \log \log T}{\log \log \log T} = \frac{-(1+2\alpha)}{4} \quad \text{in probability.}$$

This is the basis for the conjecture (1.6). This result is also relevant for the pseudomoments of the Riemann zeta function where it provides evidence for the asymptotic behavior of such moments in certain regimes, see Conjecture 5.4 in [26]. To our knowledge, these interpolated statistics are also new in the context of many weakly correlated branching random walks. This connection is explained in Section 1.2.

A key ingredient of the proof, which is of independent interest, is a precise upper bound for the right tail of the maximum in an interval of order one ( $\theta = 0$ ). It gives an alternative and direct approach to proving the upper tail tightness for log-correlated processes, in particular for the 2D Gaussian free field, that avoids the use of Gaussian comparison [11,17].

<sup>1</sup> Another possible approach to model the large values of the zeta function would be to randomize the Dirichlet sum in (1.1) instead of the product. This model is investigated in [8].

**Theorem 1.3.** *Let  $y > 0$  and  $y = o(\frac{\log \log T}{\log \log \log T})$ . Then we have*

$$\mathbb{P}\left(\max_{|h| \leq 1} X_T(h) > \log \log T - \frac{3}{4} \log \log \log T + y\right) \leq C y e^{-2y} e^{-y^2 / \log \log T},$$

for some constant  $C > 0$ .

We expect that the result holds up to  $y = o(\log \log T)$  by analogy with branching Brownian motion, see for example Proposition 2.1 in [15]. The presence of both the exponential and Gaussian decays is relevant to the question of the maximum of the zeta function in large intervals quantified by Equation (1.2) and (1.3), at least at the heuristic level. Indeed, if the Gaussian correction were not present, a back-of-the-envelope calculation for the maximum of independent maxima in intervals of order one shows that Equation (1.2) should yield the true order of the maximum. However, Theorem 1.3 gives evidence that the Gaussian behavior prevails for large  $y$ , in which case Equation (1.3) should be closer to the truth.

## 1.2. Structure of the proof

Conjecture (1.4) is based on the assumption that the statistics of the zeta function on an interval of size one of the critical line resemble the ones of the characteristic polynomial of a random unitary matrix (CUE). It has been known since Bourgade [12] that the finite-dimensional distributions of the logarithm of  $\zeta$  and of the characteristic polynomial of CUE are Gaussian, when suitably normalized, with correlations that decay logarithmically with the distance. In the context of the model  $(X_T(h), h \in I)$ , this is also verified as we have

$$(1.12) \quad \mathbb{E}[X_T(h)X_T(h')] = \begin{cases} \frac{1}{2} \log|h - h'|^{-1} + O(1) & \text{if } |h - h'| \leq 1, \\ O(|h - h'|^{-1}) & \text{if } |h - h'| > 1. \end{cases}$$

This is explained in Section 2. The dichotomy here is important: for intervals of size one, the variables  $X_T(h)$ ,  $X_T(h')$  are log-correlated, whereas for large intervals, most pairs are weakly correlated. Therefore, the extreme value statistics of the process  $(X_T(h), |h| \leq (\log T)^\theta)$  should interpolate between log-correlated and IID for  $\theta \geq 0$ . We are mainly concerned with the intermediate regime where hybrid statistics should appear. Note that even though the process  $(X_T(h), |h| \leq 1)$  is continuous, the analysis can be reduced to a finite number of points: this is made precise in the proof of Lemma 3.6. The distance  $(\log T)^{-1}$  between these points is consistent with Equation (1.12), as in the case  $|h - h'| \approx (\log T)^{-1}$  the covariance is essentially equal to the variance  $1/2 \log \log T$ .

The following analogous problem is useful to keep in mind throughout the proof. An important example of log-correlated process is branching Brownian motion (BBM), constructed as follows. At time 0, there is a single Brownian motion at 0. It evolves in time with diffusion constant  $1/2$  so the variance at time  $t$  is  $t/2$ . We choose  $1/2$  to match the variance of  $X_T(h)$ . After an exponentially-distributed time, the Brownian motion splits into two independent Brownian motions, with the same branching property. It is not hard to check that after a time  $t$  this procedure produces an average of  $e^t$  Brownian motions. It is known since Bramson [18] that the maximum at time  $t$  over all Brownian particles is

$$(1.13) \quad t - \frac{3}{4} \log t + \mathcal{M}_t,$$

where the fluctuations  $\mathcal{M}_t$  are known in the limit  $t \rightarrow \infty$ . In view of Equation (1.4), we see that the extreme statistics of a BBM at time  $t$  should be the same as the ones of  $(X_T(h), |h| \leq 1)$ , after the identification  $\log \log T \longleftrightarrow t$ .

What about longer intervals? It is not hard to check that the maximum of two independent BBM's (or any finite number for that matter) also satisfy (1.13). Indeed, this is the same as the situation of a single BBM right after the first splitting, which occurs in a finite time. Hence, the statistics of the maximum are not affected for  $t$  large. However, if one takes many (depending on  $t$ ) independent BBM's, Equation (1.13) should eventually morph to IID statistics. Equation (1.12) suggests that the process  $(X_T(h), |h| \leq (\log T)^\theta)$  should behave like  $(\log T)^\theta$  independent copies of  $(X_T(h), |h| \leq 1)$ , because of the strong decoupling for  $h$ 's at a distance more than 1. In particular, the right number of independent BBM's needed to match to the subleading order the extreme value statistics of  $(X_T(h), |h| \leq (\log T)^\theta)$  would be  $e^{t\theta}$ . To sum up, the proofs of Theorems 1.1 and 1.2 are guided by the analysis of the extreme values of  $e^{t\theta}$  independent BBM's. We stress that the precise decay  $\theta \sim t^{-\alpha}$  needed to capture the intermediate regime is a non-trivial consequence of Theorem 1.3.

The upper bounds of the theorems are derived in Section 3. For  $\theta > 0$ , the proof is not hard and is done in Section 3.3. It follows by a simple union bound on the number of points, which is  $(\log T)^\theta \cdot \log T = (\log T)^{1+\theta}$ , since there are

$e^t = \log T$  points in intervals of size one, and there are essentially  $e^{t\theta} = (\log T)^\theta$  such intervals. The case  $\theta \downarrow 0$  is much more subtle. It is established by taking a union bound over  $e^{t\theta}$  intervals of size one, each of which have a precise right tail given in Theorem 1.3. The calculation involves every factor in the estimate of the right tail. This is done in Section 3.2. The proof of the right tail of the maximum for  $\theta = 0$  is given in Section 3.1.

The proof of the lower bound in Section 4 is based on the multiscale refinement of the second moment method given in [29]. In the case  $\theta = 0$  analyzed in [3], it is necessary to truncate the second moment by introducing a linear barrier that prevents the partial sums of  $X_T(h)$  to be too high, an idea going back to [18] for a single BBM. The linear barrier can be justified as follows: since there are on average  $e^k$  Brownian motions at time  $k$ , a Gaussian estimate yields that the maximum at time  $k$  of a single BBM should be  $\approx \sqrt{k \cdot \log e^k} = k$ . There are some important variations in the intermediate regime where  $\theta \sim t^{-\alpha}$ . Following the above heuristics for  $e^{t\theta}$  independent BBM's, the maximum at time  $0 \leq k \leq t$  of  $e^{t\theta}$  independent BBM's is  $\sqrt{k \cdot \log e^{k+t\theta}} = \sqrt{k(k+t\theta)}$ . In particular, this is much larger than  $k$  for small  $k$ 's! This means that the barrier should not be too small at the beginning to capture the typical path of the maximizer. Since the barrier is fairly high at the beginning, it does not affect the trajectories of a single path very much. As a matter of fact, it will be clear from the proof that the barrier is only effective for times larger than  $t - t^\alpha$ .

It would be interesting to describe the finer asymptotics of the extreme values of  $(X_T(h), |h| \leq (\log T)^\theta)$  down to the fluctuations, a level of precision still elusive for the Riemann zeta function. It seems feasible to prove tightness of the recentered maximum by carrying a finer analysis of the random walk avoiding the specific barrier used for the lower bound. However, information on the specifics of the fluctuations seem much harder to obtain as one expects non-trivial arithmetic and random-matrix-type corrections of order one whenever  $\theta > 0$ . Numerical evidence of this phenomenon appeared in [1].

*Notations.* We write  $f(T) = O(g(T))$  whenever  $\limsup_{T \rightarrow \infty} |f(T)/g(T)| < \infty$  with possibly other parameters fixed, such as  $\varepsilon$  and  $\theta$  which will be clear from context. We often use Vinogradov's notation  $f(T) \ll g(T)$  whenever  $f(T) = O(g(T))$ . We also use  $f(T) = o(g(T))$  whenever  $\lim_{T \rightarrow \infty} |f(T)/g(T)| = 0$ . Some expressions in the proofs are more palatable in a log log scale. The parameters at that scale will be denoted by the corresponding lower case letter. For example, we write respectively

$$(1.14) \quad k = \log \log K \quad \ell = \log \log L \quad t = \log \log T, \quad \text{for } K, L, T.$$

Finally, we use the double-bracket notation  $\llbracket k, t \rrbracket$  for  $[k, t] \cap \mathbb{Z}$ , the integers in the interval  $[k, t]$ .

## 2. Probability estimates for the model

This preliminary section is devoted to giving more details about the correlation structure of the process  $X_T$ . It is similar to Section 2 of [3]. The results here are more straightforward, as the process is exactly Gaussian. We include them for completeness.

In some cases, we need to study the process in a smaller range of primes. For this purpose, we define, for  $1 \leq K < L \leq T$ ,

$$(2.1) \quad X_{K,L}(h) = \sum_{K < p \leq L} \frac{\operatorname{Re}(G_p p^{-ih})}{p^{1/2}}, \quad |h| \leq (\log T)^\theta.$$

We simply write  $X_K(h)$  for  $X_{1,K}(h)$ . The Laplace transform of the process is easy to compute:

**Lemma 2.1.** *Let  $1 \leq K < L \leq T$ . For  $\lambda, \lambda' \in \mathbb{R}$  and  $|h|, |h'| \in \mathbb{R}$ , we have*

$$(2.2) \quad \mathbb{E}[e^{\lambda X_{K,L}(h) + \lambda' X_{K,L}(h')}] = \exp\left(\sum_{K < p \leq L} \frac{1}{4p} (\lambda^2 + \lambda'^2 + 2\lambda\lambda' \cos(|h - h'| \log p))\right).$$

**Proof.** This is obvious from the fact that

$$(2.3) \quad |\lambda p^{-ih} + \lambda' p^{-ih'}|^2 = \lambda^2 + \lambda'^2 + 2\lambda\lambda' \cos(|h - h'| \log p).$$

Therefore the Gaussian random variable  $\lambda X_{K,L}(h) + \lambda' X_{K,L}(h')$  has variance

$$(2.4) \quad \sum_{K < p \leq L} \frac{1}{4p} (\lambda^2 + \lambda'^2 + 2\lambda\lambda' \cos(|h - h'| \log p))$$

and the claim follows. □

Setting  $\lambda' = 0$  in the above yields the variance of  $X_{K,L}(h)$

$$(2.5) \quad \frac{1}{2} \sum_{K < p \leq L} \frac{1}{p} = \frac{1}{2}(\ell - k) + O(e^{-c\sqrt{e^k}}), \quad K > 2,$$

where we use the log log-notation (1.14). The equality is a strong form of Mertens's theorem, and can be proved using the Prime Number Theorem with classical error, see for example [34, Theorem 6.9],

$$(2.6) \quad \#\{p \leq x : p \text{ prime}\} = \int_2^x \frac{1}{\log u} du + O(xe^{-c\sqrt{\log x}}).$$

It is not hard to evaluate the same way the cosine sum appearing in the covariance of  $X_{K,L}(h)$  and  $X_{K,L}(h')$ .

**Lemma 2.2.** *Let  $1 \leq K < L \leq T$ . We have, using the notation (1.14),*

$$\sum_{K < p \leq L} \frac{1}{p} \cos(\delta \log p) = \begin{cases} \ell - k + O(\delta^2 e^{2\ell}) + O(e^{-c\sqrt{e^k}}) & \text{if } \delta e^\ell \leq 1, \\ O(\delta^{-1} e^{-k}) + O(e^{-c\sqrt{e^k}}) & \text{if } \delta e^k > 1. \end{cases}$$

**Proof.** This is done in Lemma 2.1 of [3]. We reproduce the details for completeness. The sum can be expressed using integration by parts and Equation (2.6) as

$$\int_K^L \frac{\cos(\delta \log u)}{u \log u} du + O(e^{-c\sqrt{\log K}}).$$

The claim for  $\delta e^\ell \leq 1$  follows from  $\cos(\delta \log u) = 1 + O(\delta^2 (\log u)^2)$ . When  $\delta e^k > 1$ , we do the change of variable  $v = \log u$  and integrate by parts. The integral then becomes

$$\frac{\sin(\delta v)}{\delta v} \Big|_{\log K}^{\log L} + \int_{\log K}^{\log L} \frac{\sin(\delta v)}{\delta v^2} dv.$$

Both terms are  $O(\delta^{-1} e^{-k})$ . □

Based on Lemma 2.2, we see that there are two very different regimes for the covariance based on the value of  $|h - h'|$ : for  $1 \leq K < L \leq T$ ,

(i) If  $|h - h'| < e^{-\ell}$ , we get

$$(2.7) \quad \mathbb{E}[X_{K,L}(h)X_{K,L}(h')] = \frac{1}{2}(\ell - k) + O(e^{2\ell}|h - h'|^2).$$

In this case the two variables are almost perfectly correlated.

(ii) In the case  $|h - h'| > e^{-k}$ , the oscillatory behavior of the sum causes the variables to decorrelate polynomially

$$(2.8) \quad \mathbb{E}[X_{K,L}(h)X_{K,L}(h')] = O(e^{-k}|h - h'|^{-1}).$$

It is useful to think of  $X_{K,L}(h)$  as the sum of increments of a random walk between time  $k$  and  $\ell$ . With this in mind, the above dichotomy says that the random walks at  $h$  and  $h'$  are almost the same until time  $-\log|h - h'|$ , at which point they decorrelate rapidly.

We make the decorrelation of the random walks more precise in the following lemma. The decorrelation will be needed under a tilted measure  $\overline{\mathbb{P}}$ : for  $\lambda, \lambda' \geq 0$ , take

$$(2.9) \quad \frac{d\overline{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{\lambda X_{K,L}(h) + \lambda' X_{K,L}(h')}}{\mathbb{E}[e^{\lambda X_{K,L}(h) + \lambda' X_{K,L}(h')}]}, \quad h, h' \in \mathbb{R}.$$

It is straightforward to check by differentiating the logarithm of Equation (2.2) that the covariance is unchanged under  $\overline{\mathbb{P}}$ , and that the mean is

$$(2.10) \quad \overline{\mathbb{E}}[X_{K,L}(h)] = \sum_{K < p \leq L} \frac{\lambda + \lambda' \cos(|h - h'| \log p)}{2p}.$$

**Lemma 2.3.** *Let  $1 \leq K < T$ . Consider  $A_\ell, A'_\ell$ , a collection of intervals in  $\mathbb{R}$ , indexed by  $\ell \in \llbracket k+1, t \rrbracket$ . Then for any  $h, h'$  with  $|h - h'| > e^{-k}$ , we have under the measure  $\overline{\mathbb{P}}$  defined by Equation (2.9):*

$$(2.11) \quad \begin{aligned} & \overline{\mathbb{P}}(X_{K,L}(h) \in A_\ell, X_{K,L}(h') \in A'_\ell, \forall \ell \in \llbracket k+1, t \rrbracket) \\ &= (1 + O((e^k |h - h'|)^{-\frac{1}{2}})) \overline{\mathbb{P}}(X_{K,L}(h) \in A_\ell, \forall \ell) \overline{\mathbb{P}}(X_{K,L}(h') \in A'_\ell, \forall \ell) \\ & \quad + O(e^{-c\sqrt{e^k |h - h'|}}), \end{aligned}$$

where the error does not depend on the choice of  $A$ 's and  $c$  denotes an absolute positive constant.

**Proof.** As the process  $X$  has independent increments, the probability can be decomposed in terms of the variables

$$(2.12) \quad Y_\ell(h) = \sum_{e^{\ell-1} < \log p \leq e^\ell} \frac{\operatorname{Re}(G_p p^{-ih})}{p^{1/2}}.$$

This is  $X_{L^{1/e}, L}(h)$  in the notation of Equation (2.10). We prove that for every  $\ell \in \llbracket k+1, t \rrbracket$

$$(2.13) \quad \begin{aligned} & \overline{\mathbb{P}}(Y_\ell(h) \in B_\ell, Y_\ell(h') \in B'_\ell) \\ &= (1 + O((e^\ell |h - h'|)^{-1/2})) \overline{\mathbb{P}}(Y_\ell(h) \in B_\ell) \overline{\mathbb{P}}(Y_\ell(h') \in B'_\ell) + O(e^{-c\sqrt{e^\ell |h - h'|}}), \end{aligned}$$

for some intervals  $B_\ell$  and  $B'_\ell$ , where the error terms are uniform in  $\ell$  and in the choice of  $B$ 's. Equation (2.11) then follows by writing the probability increment by increment using a Markov-type decomposition and taking the product.

To prove Equation (2.13), consider first the subset

$$E_\ell = \left\{ \left\| (Y_\ell(h), Y_\ell(h')) - (\mu_\ell, \mu_\ell) \right\|_{\mathbb{R}^2} \leq (e^\ell |h - h'|)^{1/4} \right\},$$

where

$$(2.14) \quad \mu_\ell = \overline{\mathbb{E}}[Y_\ell] = \sum_{e^{\ell-1} < \log p \leq e^\ell} \frac{\lambda + \lambda' \cos(|h - h'| \log p)}{2p}$$

by Equation (2.10). A straightforward Gaussian estimate (or a Chernoff bound using Lemma 2.1) yields  $\overline{\mathbb{P}}(E_\ell^c) \leq \exp(-c\sqrt{e^\ell |h - h'|})$ . This accounts for the error term in Equation (2.13).

It remains to estimate the Gaussian density on the event  $E_\ell$ . The covariance matrix  $\Sigma_\ell$  of  $(Y_\ell(h), Y_\ell(h'))$  is

$$(2.15) \quad \Sigma_\ell = \begin{pmatrix} \sigma_\ell^2 & \rho_\ell \\ \rho_\ell & \sigma_\ell^2 \end{pmatrix}, \quad \rho_\ell = O(e^{-\ell} |h - h'|^{-1}), \quad \sigma_\ell^2 = \frac{1}{2} + O(e^{-c\sqrt{e^\ell}}),$$

by Equation (2.8). The inverse covariance admits the following expansion around the identity.

$$(2.16) \quad \Sigma_\ell^{-1} = \sigma_\ell^{-2} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \varepsilon_\ell \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + O(\varepsilon_\ell^2) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}, \quad \varepsilon_\ell = \frac{\rho_\ell}{\sigma_\ell} = O(e^{-\ell} |h - h'|^{-1}).$$

For  $(y, y') \in E_\ell$ , the above implies the following estimate:

$$(y - \mu_\ell, y' - \mu_\ell)^T \Sigma_\ell^{-1} (y - \mu_\ell, y' - \mu_\ell) = \frac{(y - \mu_\ell)^2 + (y' - \mu_\ell)^2}{\sigma_\ell^2} + O((e^\ell |h - h'|)^{-1/2}).$$

The estimate for the determinant of  $\Sigma^{-1}$  is immediate from Equation (2.16). Therefore, the density of the Gaussian vector  $(Y_\ell(h), Y_\ell(h'))$  under  $\overline{\mathbb{P}}$  is

$$(1 + O((e^\ell |h - h'|)^{-1/2})) \frac{1}{2\pi\sigma_\ell^2} \exp\left(-\frac{(y - \mu_\ell)^2 + (y' - \mu_\ell)^2}{2\sigma_\ell^2}\right).$$

This proves Equation (2.13) and concludes the proof of the lemma.  $\square$



The above proof can be also used to compare the probability to the one of a random walk with increments of variance exactly equal to  $1/2$ . We will only need this in the case of one point  $h$ . To use the standard version of the ballot theorem, we will need to compare the random walk at  $h$  with a random walk with increments of variance exactly equal to  $1/2$ .

**Corollary 2.4.** *Let  $1 \leq K < T$ . Consider  $A_\ell$ , a collection of intervals in  $\mathbb{R}$ , indexed by  $\ell \in \llbracket k+1, t \rrbracket$ . Consider  $\mathcal{Y}_\ell$ ,  $\ell \geq 1$ , i.i.d. Gaussian random variables of mean  $\mu_\ell$  as in Equation (2.14), and variance  $1/2$  under  $\overline{\mathbb{P}}$ . Then for any  $h \in \mathbb{R}$  and  $\ell > k$ , we have under the measure  $\overline{\mathbb{P}}$  defined by Equation (2.9):*

$$(2.17) \quad \begin{aligned} & \overline{\mathbb{P}}(X_{K,L}(h) \in A_\ell, \forall \ell \in \llbracket k+1, t \rrbracket) \\ &= (1 + O(e^{-c\sqrt{e^k}})) \overline{\mathbb{P}}\left(\sum_{j=k+1}^{\ell} \mathcal{Y}_j \in A_\ell, \forall \ell \in \llbracket k+1, t \rrbracket\right) + O(e^{-ce^{2k}}). \end{aligned}$$

**Proof.** The proof follows the same line as the one of Lemma 2.3, setting  $\lambda' = 0$  in (2.9). Consider the increments  $Y_\ell$  and the event

$$E_\ell = \{|Y_\ell(h) - \mu_\ell| \leq e^\ell\}.$$

A Gaussian estimate shows that  $\mathbb{P}(E_\ell^c) \ll e^{-ce^{2\ell}}$ . For  $y \in E_\ell$ , we have that

$$\frac{1}{\sigma_\ell} e^{-(y-\mu_\ell)^2/2\sigma_\ell^2} = (1 + O(e^{-c\sqrt{e^\ell}})) e^{-(y-\mu_\ell)^2}.$$

Putting both estimates together yield for any interval  $B \subset \mathbb{R}$ ,

$$\overline{\mathbb{P}}(Y_\ell(h) \in B) = (1 + O(e^{-c\sqrt{e^\ell}})) \overline{\mathbb{P}}(\mathcal{Y}_\ell \in B \cap E_\ell) + O(e^{-ce^{2\ell}}).$$

The probability of  $X_{K,L}$  is then obtained by conditioning successively on the  $Y_\ell$ 's and computing the product.  $\square$

We remark that if the  $A_\ell$ 's do not grow too fast in such a way that the event  $\{X_{K,L}(h) \in A_\ell, \forall \ell \in \llbracket k+1, t \rrbracket\}$  is included in the events  $E_\ell$  for all  $\ell \in \llbracket k+1, t \rrbracket$ , the additive error term is not present and we may conclude that

$$(2.18) \quad \begin{aligned} & \overline{\mathbb{P}}(X_{K,L}(h) \in A_\ell, \forall \ell \in \llbracket k+1, t \rrbracket) \\ &= (1 + O(e^{-c\sqrt{e^k}})) \overline{\mathbb{P}}\left(\sum_{j=k+1}^{\ell} \mathcal{Y}_j \in A_\ell, \forall \ell \in \llbracket k+1, t \rrbracket\right). \end{aligned}$$

### 3. Upper bounds

#### 3.1. Upper tail tightness for $\theta = 0$

The purpose of this section is to prove Theorem 1.3. The starting point is in the spirit of Lemma 2.4 in Bramson, Ding, and Zeitouni [16] for branching random walks; however, their approach has to be generalized to account for the lack of perfect branching in the model of the zeta function. Our method is applicable to other log-correlated processes that do not exhibit an exact branching structure.

As in Equation (2.12), we define

$$(3.1) \quad Y_k(h) = \sum_{e^{k-1} < \log p \leq e^k} \frac{1}{\sqrt{p}} \operatorname{Re}(p^{-ih} G_p), \quad k \leq t,$$

where  $t = \log \log T$ . Each  $Y_k(h)$  is a Gaussian variable of variance roughly  $1/2$ . We also define the partial sums

$$(3.2) \quad S_j(h) = \sum_{k=1}^j Y_k(h), \quad j \leq t,$$



and their recentered version,

$$(3.3) \quad \bar{S}_j = S_j - m(j), \quad m(j) = j \left( 1 - \frac{3 \log t}{4t} \right).$$

The method requires to introduce a logarithmic barrier:

$$(3.4) \quad \psi_j = \log(\min(j, t - j)), \quad 1 \leq j \leq t - 1.$$

so that we have the symmetry  $\psi_j = \psi_{t-j}$ . We also set  $\psi_0 = \psi_t = 0$ .

The idea of the proof is as follows. The goal is to bound the probability  $p_{\text{cross}}$  that the value at some  $h$  is above the level  $m(t) + y$ :

$$(3.5) \quad p_{\text{cross}} = \mathbb{P} \left( \max_{|h| \leq 1} X_T(h) > m(t) + y \right).$$

We start by bounding  $p_{\text{cross}}$  by a more amenable expression. The event can be decomposed according to the first time  $j$  after some fixed time  $r$  when the barrier  $y + \psi$  is crossed by the maximum over  $h$ . It is convenient to take  $r$  away from 1 for two reasons. First, the increments for small primes have a variance, which is not close to  $1/2$ , which makes the estimate such as Proposition A.3 more cumbersome. The choice

$$r = r(y) = y \vee 1$$

is good. Second, it is then possible to get an *a priori bound* on  $\max_{|h| \leq 1} |S_j(h)|$  for all  $r < j \leq t$ . Indeed, a union bound gives

$$\begin{aligned} \mathbb{P} \left( \exists r < j \leq t : \max_{|h| \leq 1} |S_j(h)| > 2j + y \right) &\ll \sum_{j=r+1}^t e^j \mathbb{P}(|S_\ell(h)| > 2j + y) \\ &\ll \sum_{j=r+1}^t e^j \cdot e^{-(2j+y)^2/2\sigma_j^2} \\ &\ll e^{-3r} \cdot e^{-4y} e^{-y^2/t}. \end{aligned}$$

The first inequality is obtained by a discretization argument (see Corollary 2.6 in [3] which is proved again in Lemma 3.6 below). The third inequality comes from the fact that  $\sigma_j^2 = j/2 + O(1)$  and  $y$  is much smaller than  $t$ . The above is much smaller than the claimed bound in Theorem 1.3. Therefore, we can assume without loss of generality in the proof that

$$(3.6) \quad \max_{|h| \leq 1} |S_j(h)| \leq 2j + y, \quad \text{for all } r < j \leq t.$$

We now estimate Equation (3.5) with an union bound on  $r < j \leq t$ :

$$(3.7) \quad p_{\text{cross}} \leq \sum_{r < j \leq t-1} \mathbb{P} \left( \forall r < \ell \leq j \max_{|h| \leq 1} \bar{S}_\ell(h) \leq y + \psi_\ell, \max_{|h| \leq 1} \bar{S}_{j+1}(h) > y + \psi_{j+1} \right).$$

Secondly, the interval  $[-1, 1]$  can be partitioned into  $e^j$  intervals of width  $2e^{-j}$  and centers  $h_j^{(i)}$  for  $i \leq e^j$ . Therefore, the probability can be bounded by a union bound over these intervals:

$$(3.8) \quad p_{\text{cross}} \ll \sum_{r < j \leq t-1} \sum_{i=1}^{e^j} \mathbb{P} \left( \forall r < \ell \leq j \max_{|h| \leq 1} \bar{S}_\ell(h) \leq y + \psi_\ell, \max_{|h - h_j^{(i)}| \leq e^{-j}} \bar{S}_{j+1}(h) > y + \psi_{j+1} \right).$$

Finally, the constraint on the maximum over  $[-1, 1]$  is limited to  $h = h_j$ . Using the symmetry of the distribution of the model under translation, the above expression then becomes

$$(3.9) \quad p_{\text{cross}} \ll \sum_{r < j \leq t-1} e^j \mathbb{P} \left( \forall r < \ell \leq j \bar{S}_\ell(0) \leq y + \psi_\ell, \max_{|h| \leq e^{-j}} \bar{S}_{j+1}(h) > y + \psi_{j+1} \right).$$

We now make a simple but important observation: the overshoot of  $\max_{|h| \leq e^{-j}} \bar{S}_{j+1}(h)$  is either due to a large value of the increment  $\bar{S}_{j+1} - \bar{S}_j$  at  $h = 0$  or to a large value of the difference  $S_{j+1}(h) - S_{j+1}(0)$  for some  $|h| \leq e^{-j}$ . The first case is exactly the one that is obtained for an exact branching random walk. The second case is due to the fact the correlation structure of the model is not exactly a tree. However, these fluctuations can be controlled very precisely by a chaining argument because of the log-correlations. Specifically, define the following event that includes the barrier and the value of the endpoint  $S_j$ :

$$(3.10) \quad \mathcal{B}_j(x) = \{\forall r < \ell \leq j, \bar{S}_\ell(0) \leq y + \psi_\ell \text{ and } \bar{S}_j(0) \in I_x\}, \quad j \leq t, x \in \mathbb{Z},$$

where  $I_x = (x, x + 1]$  for  $x > 0$ ,  $I_x = (x - 1, x]$  if  $x < 0$ , and  $I_x = [-1, 1]$  if  $x = 0$ . Performing a union bound on the range of values of  $\bar{S}_j(0)$ , we prove the following two estimates:

**Lemma 3.1.** *With the notation above, we have for  $y > 0$  and  $y = o(\frac{\log \log T}{\log \log \log T})$ ,*

$$(3.11) \quad \sum_{r < j \leq t-1} \sum_{x \leq y + \psi_j} e^j \mathbb{P}(\mathcal{B}_j(x), \bar{S}_{j+1}(0) - \bar{S}_j(0) > y + \psi_{j+1} - x) \ll (y + 1) e^{-2y} e^{-y^2/t}.$$

This is the same estimate one would prove for a branching random walk. The new input is an estimate for the overshoot due to a large value in the difference in a small neighborhood:

**Lemma 3.2.** *With the notation above, we have for  $y > 0$  and  $y = o(\frac{\log \log T}{\log \log \log T})$ ,*

$$(3.12) \quad \sum_{r < j \leq t-1} \sum_{x \leq y + \psi_{j+1}} e^j \mathbb{P}(\mathcal{B}_{j+1}(x), \max_{|h| \leq e^{-j}} \bar{S}_{j+1}(h) - \bar{S}_{j+1}(0) > y + \psi_{j+1} - x) \ll (y + 1) e^{-2y} e^{-y^2/t}.$$

**Proof of Theorem 1.3.** From the previous discussion, it suffices to bound the right-hand side of Equation (3.9). As explained above, this splits into the two cases given by Equations (3.11) and (3.12), thereby proving the claim.  $\square$

We now prove the two lemmas. The first proof is along the lines of Lemma 2.4 of [16].

**Proof of Lemma 3.1.** Consider the change of measure, from  $\mathbb{P}$  to  $\bar{\mathbb{P}}$ , as in Equation (2.9) with  $\lambda' = 0$ ,

$$(3.13) \quad \frac{d\bar{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{\lambda S_j(0)}}{\mathbb{E}[e^{\lambda S_j(0)}]} = e^{-\lambda^2 \sigma_j^2 / 2} e^{\lambda S_j(0)},$$

where  $\sigma_j^2 = \frac{1}{2} \sum_{\log p \leq e^j} 1/p = \frac{1}{2} j + O(1)$  as in Equation (2.5), and  $\lambda$  is chosen in such a way that  $\bar{\mathbb{E}}[\bar{S}_j] = 0$  as in Equation (2.10). The condition for this centering to hold is  $m(j) = \lambda \sigma_j^2 = \frac{1}{2} \lambda j (1 + O(1/j))$ , which yields the following identities:

$$(3.14) \quad \left\{ \begin{aligned} \lambda &= \frac{m(j)}{\sigma_j^2} = (2 - \frac{3 \log t}{2t})(1 + O(1/j)) = 2 - \frac{3 \log t}{2t} + O(1/j), \\ \frac{\lambda^2 \sigma_j^2}{2} &= (j - \frac{3}{2} \log t + \frac{9 (\log t)^2}{16t})(1 + O(1/j)) = j - \frac{3 \log t}{2t} j + O(1), \\ \lambda m(j) &= (2j - 3 \log t + \frac{9 (\log t)^2}{8t})(1 + O(1/j)) = 2j - 3 \frac{\log t}{t} j + O(1). \end{aligned} \right.$$

$$(3.15) \quad \left\{ \begin{aligned} \lambda^2 \sigma_j^2 &= (j - \frac{3}{2} \log t + \frac{9 (\log t)^2}{16t})(1 + O(1/j)) = j - \frac{3 \log t}{2t} j + O(1), \\ \lambda m(j) &= (2j - 3 \log t + \frac{9 (\log t)^2}{8t})(1 + O(1/j)) = 2j - 3 \frac{\log t}{t} j + O(1). \end{aligned} \right.$$

$$(3.16) \quad \left\{ \begin{aligned} \lambda^2 \sigma_j^2 &= (j - \frac{3}{2} \log t + \frac{9 (\log t)^2}{16t})(1 + O(1/j)) = j - \frac{3 \log t}{2t} j + O(1), \\ \lambda m(j) &= (2j - 3 \log t + \frac{9 (\log t)^2}{8t})(1 + O(1/j)) = 2j - 3 \frac{\log t}{t} j + O(1). \end{aligned} \right.$$

The probability in the sum can be bounded using  $\bar{\mathbb{P}}$  as follows, for  $w = y + \psi_{j+1} - x$ ,

$$(3.17) \quad \begin{aligned} \mathbb{P}(\mathcal{B}_j(x), \bar{S}_{j+1}(0) - \bar{S}_j(0) > w) &\ll e^{j - \frac{3}{2} \frac{\log t}{t} j} \cdot \bar{\mathbb{E}}[e^{-\lambda S_j(0)} \mathbf{1}_{\mathcal{B}_j(x)} \mathbf{1}_{\bar{S}_{j+1}(0) - \bar{S}_j(0) > w}] \\ &\ll e^{-j + \frac{3}{2} \frac{\log t}{t} j} \cdot \bar{\mathbb{E}}[e^{-\lambda \bar{S}_j(0)} \mathbf{1}_{\mathcal{B}_j(x)} \mathbf{1}_{\bar{S}_{j+1}(0) - \bar{S}_j(0) > w}] \\ &\ll e^{-j} e^{\frac{3}{2} \frac{\log t}{t} j} e^{-\lambda x} \cdot \bar{\mathbb{P}}(\mathcal{B}_j(x)) e^{-202w} \end{aligned}$$

where (3.15) is used in the first inequality, (3.16) in the second, as well as the values of  $\bar{S}_j(0)$  on the event  $\mathcal{B}_j(x)$  and a Chernoff bound on  $\bar{S}_{j+1}(0) - \bar{S}_j(0)$  in the third. We could have used a Gaussian estimate for  $\bar{S}_{j+1}(0) - \bar{S}_j(0)$ . However, as it will be clear below, the convergence of the sum is ensured as long as the probability for the increment is  $\ll e^{-(2+\delta)w}$  for some  $\delta > 0$ . We choose  $\delta = 200$ . This highlights the similarities with the proof of Lemma 3.2, where Gaussian estimates are not available.

We split the estimate of Equation (3.17) into two cases. The estimate for  $r < j \leq t/\log t$  is simpler, as no ballot theorem is needed. We trivially have

$$(3.18) \quad \mathbb{P}(\mathcal{B}_j(x)) \ll e^{-x^2/2\sigma_j^2} \ll e^{-x^2/j},$$

since  $\sigma_j^2 = \frac{1}{2}j + O(1)$  and  $|x| \ll j$ ,  $j > r$ , by the *a priori* bound (3.6). Therefore, the following estimate holds

$$(3.19) \quad \mathbb{P}(\mathcal{B}_j(x), \bar{S}_{j+1}(0) - \bar{S}_j(0) > w) \ll e^{-j} \cdot e^{-2x} \cdot e^{-x^2/j} \cdot e^{-202w}.$$

The assumption  $y = o(t/\log t)$  is needed in estimating  $\lambda x$ , so that  $e^{\frac{3}{2}\frac{\log t}{t}x} \leq e^{\frac{3}{2}\frac{\log t}{t}(y+\psi_j)} \ll 1$ . In the case  $t/\log t < j < t$ , we first note that  $\frac{3}{2}\frac{\log t}{t}j \leq j^{3/2}$ , since the function  $\log x/x$  is decreasing for  $x \geq 1$ . Moreover, Proposition A.3 can be applied thanks to the *a priori* bounds (3.6) and Equation (2.17). All this implies

$$(3.20) \quad \mathbb{P}(\mathcal{B}_j(x), \bar{S}_{j+1}(0) - \bar{S}_j(0) > w) \ll e^{-j} e^{-2x} (y+1)(y+\psi_j-x+1) e^{-x^2/j} e^{-202w}.$$

Note that the right-hand side is also an upper bound to (3.19). Thus, the sum of the lemma becomes

$$\begin{aligned} & \sum_{r < j \leq t-1} \sum_{x \leq y+\psi_j} e^j \mathbb{P}(\mathcal{B}_j(x), \bar{S}_{j+1}(0) - \bar{S}_j(0) > y + \psi_{j+1} - x) \\ & \ll \sum_{r < j \leq t-1} \sum_{x \leq y+\psi_j} e^{-2x} (y+1)(y+\psi_j-x+1) e^{-x^2/j} e^{-202(y+\psi_{j+1}-x)}. \end{aligned}$$

It is convenient to do a change of variable  $v = y + \psi_j - x$ , expressing the distance between the barrier and  $S_j(0)$ . With this notation and the fact that  $|\psi_{j+1} - \psi_j| < 1$ , the above is

$$\begin{aligned} & \ll (y+1) \sum_{r < j \leq t-1} \sum_{v \geq 0} e^{-2(y+\psi_j-v)} (v+1) e^{-(y+\psi_j-v)^2/j} e^{-202(v-1)} \\ (3.21) \quad & \ll (y+1) e^{-2y} \sum_{j \leq t-1} e^{-2\psi_j} \sum_{v \geq 0} (v+1) e^{-(y+\psi_j-v)^2/t} e^{-200v}, \end{aligned}$$

where we bounded  $e^{-(y+\psi_j-v)^2/j}$  by  $e^{-(y+\psi_j-v)^2/t}$  and extended the sum to  $j \leq r$ . This allows us to use the symmetry  $\psi_j = \psi_{t-j}$  to pair the terms  $j$  and  $t-j$  together. Indeed, for a given  $v$ ,  $y + \psi_j - v$  has the same value for both values of  $j$ , but  $(y + \psi_j - v)^2/t$ , so that it is enough to bound the second half of the sum:

$$(3.22) \quad (y+1) e^{-2y} \sum_{j=\lfloor t/2 \rfloor}^{t-1} e^{-2\psi_j} \sum_{v \geq 0} (v+1) e^{-100v} \cdot e^{-(y+\psi_j-v)^2/t} e^{-100v}.$$

Note that the term  $e^{-(y+\psi_j-v)^2/t} e^{-100v}$  is maximal at  $v = y + \psi_j - 50t$ , which is negative since  $y = o(t/\log t)$ . In particular, this implies that  $v = 0$  is the maximizer in the range considered, hence,  $e^{-(y+\psi_j-v)^2/t} e^{-100v} < e^{-(y+\psi_j)^2/t} \leq e^{-y^2/t}$  since  $y > 0$ . With this observation, we conclude that the above is

$$\ll (y+1) e^{-2y} e^{-y^2/t} \sum_{j=\lfloor t/2 \rfloor}^{t-1} e^{-2\psi_j} \sum_{v \geq 0} (v+1) e^{-100v} \ll (y+1) e^{-2y} e^{-y^2/t},$$

by the definition of  $\psi_j$ , which ensures summability. This proves the claim.  $\square$

**Proof of Lemma 3.2.** We now decompose according to the value of  $\bar{S}_{j+1}(0)$ . We need to bound:

$$(3.23) \quad \sum_{r < j \leq t-1} \sum_{x \leq y+\psi_{j+1}} e^j \mathbb{P}(\mathcal{B}_{j+1}(x), \max_{|h| \leq e^{-j}} \bar{S}_{j+1}(h) - \bar{S}_{j+1}(0) > y + \psi_{j+1} - x).$$

The term  $x = y + \psi_{j+1}$  must be dealt with separately. This is to ensure that we consider a jump of size at least 1. To bound this term, the restriction on the maximum is simply dropped. More precisely, proceeding exactly as in Lemma 3.1 with a change of measure and using the ballot theorem, but dropping the event on the increment  $S_{j+1}(0) - S_j(0)$ , one gets

$$\sum_{r < j \leq t-1} e^j \mathbb{P}(\mathcal{B}_{j+1}(y + \psi_{j+1})) \ll (y+1) \sum_{j \leq t} e^{-2(y+\psi_j)} e^{-(y+\psi_j)^2/j} \ll (y+1) e^{-2y} e^{-y^2/t},$$

as desired.

Let  $v = y + \psi_{j+1} - x$ . We can now restrict the proof to the case  $v \geq 1$ . On the event considered, there is an  $h \in [-e^{-j}, e^{-j}]$  such that

$$\bar{S}_{j+1}(h) - \bar{S}_{j+1}(0) > v \geq 1.$$

We decompose this increment using a chaining argument. For each  $h$ , we consider a sequence  $h_k \rightarrow h$ , within the following dyadic sets

$$(3.24) \quad h_k \in \mathcal{H}_k \quad \text{where } \mathcal{H}_k = \left\{ e^{-j} \frac{l}{2^k}, l \in [-2^k, 2^k] \cap \mathbb{Z} \right\} \subset [-e^{-j}, e^{-j}], \quad k \geq 0.$$

In particular, we can choose it such that  $h_0 = 0$ , and  $|h_{k+1} - h_k| \leq \frac{e^{-j}}{2^k}$  (by taking  $h_k$  the closest to  $h$  in  $\mathcal{H}_k$  when  $k > 0$ ). By continuity, we have

$$\bar{S}_{j+1}(h) - \bar{S}_{j+1}(0) = \sum_{k=0}^{\infty} (\bar{S}_{j+1}(h_{k+1}) - \bar{S}_{j+1}(h_k)).$$

The left-hand side being larger than  $v \geq 1$ , there is a  $k^* \geq 0$  such that

$$\bar{S}_{j+1}(h_{k^*+1}) - \bar{S}_{j+1}(h_{k^*}) > \frac{v}{2(1+k)^2}.$$

We perform a union bound over every possibility of neighboring elements of  $\mathcal{H}_{k+1}$ . Let  $(h_i)_i$  be an increasing enumeration of all elements of  $\mathcal{H}_{k+1}$ . With these observations, the contribution to Equation (3.23) coming from  $v \geq 1$  is bounded by

$$(3.25) \quad \ll \sum_{j \leq t-1} \sum_{v \geq 1} \sum_{k \geq 0} e^j \sum_{h_i \in \mathcal{H}_{k+1}} \mathbb{P}\left(\mathcal{B}_{j+1}(y + \psi_{j+1} - v), \bar{S}_j(h_{i+1}) - \bar{S}_j(h_i) > \frac{v}{2(1+k)^2}\right).$$

We use Chernoff's inequality with one real parameters  $\lambda = \lambda_k$  (to be specified later), and obtain the bound

$$(3.26) \quad \sum_{j \leq t} \sum_{v \geq 1} \sum_{k \geq 0} e^j \sum_{h_i \in \mathcal{H}_{k+1}} e^{-\lambda \frac{v}{2(1+k)^2}} \mathbb{E}\left[\mathbf{1}_{\mathcal{B}_{j+1}(y+\psi_{j+1}-v)} e^{\lambda(\bar{S}_{j+1}(h_{i+1}) - \bar{S}_{j+1}(h_i))}\right].$$

The key step is to quantify the effect of the exponential factor on the probability of the event  $\mathcal{B}_{j+1}$ . For this, note that for each  $h_i \in \mathcal{H}_{k+1}$ ,

$$\bar{S}_{j+1}(h_{i+1}) - \bar{S}_{j+1}(h_i) = S_{j+1}(h_{i+1}) - S_{j+1}(h_i) = \sum_{\ell=1}^{j+1} (Y_{\ell}(h_{i+1}) - Y_{\ell}(h_i)).$$

In order to make the correlation with the path at  $h = 0$  explicit, we write the orthogonal decomposition of these increments, i.e., we consider the orthogonal  $L^2$ -projection of these increments  $Y_{\ell}(h_{i+1}) - Y_{\ell}(h_i)$  on  $Y_{\ell}(0)$ :

$$(3.27) \quad Y_{\ell}(h_{i+1}) - Y_{\ell}(h_i) = \pi_{\ell}^{(h_i)} Y_{\ell}(0) + \Delta_{\ell}^{(h_i)},$$

where  $\Delta_\ell^{(h_i)}$  is a Gaussian variable independent of  $Y_\ell(0)$  and of all the other increments  $Y_m$ ,  $m \neq \ell$ . We now estimate  $\Delta_\ell^{(h_i)}$  and  $\pi_\ell^{(h_i)}$ . The bounds we obtain are the same for all  $h_i \in \mathcal{H}_{k+1}$ . We first compute the variance of the difference

$$\begin{aligned} \text{Var}(Y_\ell(h_{i+1}) - Y_\ell(h_i)) &= \sum_{e^{\ell-1} < \log p \leq e^\ell} \frac{1}{p} (1 - \cos(|h_{i+1} - h_i| \log p)) \\ (3.28) \quad &\leq \sum_{e^{\ell-1} < \log p \leq e^\ell} \frac{(\log p)^2}{p} (h_{i+1} - h_i)^2. \end{aligned}$$

Note that  $|h_{i+1} - h_i| \leq e^{-j} 2^{-k}$  for all  $h_i \in \mathcal{H}_{k+1}$  by construction, and  $\log p \leq e^\ell$  in this sum, so that

$$\text{Var}(Y_\ell(h_{i+1}) - Y_\ell(h_i)) \leq \sum_{e^{\ell-1} < \log p \leq e^\ell} \frac{1}{p} (e^\ell)^2 e^{-2j} 2^{-2k} \ll 2^{-2k} e^{2(\ell-j)}.$$

The Cauchy-Schwarz inequality then implies

$$|\mathbb{E}[(Y_\ell(h_{i+1}) - Y_\ell(h_i)) Y_\ell(0)]| \leq \mathbb{E}[(Y_\ell(h_{i+1}) - Y_\ell(h_i))^2]^{1/2} \cdot \mathbb{E}[Y_\ell(0)^2]^{1/2} \ll 2^{-k} e^{-(j-\ell)},$$

for all  $h_i \in \mathcal{H}_{k+1}$ . The bound for  $\pi_\ell^{(h_i)}$  is then straightforward

$$(3.29) \quad \pi_\ell^{(h_i)} = \frac{\mathbb{E}[(Y_\ell(h_{i+1}) - Y_\ell(h_i)) Y_\ell(0)]}{\mathbb{E}[Y_\ell(0)^2]} \ll 2^{-k} e^{-(j-\ell)}.$$

The orthogonality in (3.27) implies

$$\text{Var}(Y_\ell(h_{i+1}) - Y_\ell(h_i)) = \mathbb{E}[(\Delta_\ell^{(h_i)})^2] + (\pi_\ell^{(h_i)})^2 \mathbb{E}[Y_\ell(0)^2].$$

Together with (3.29), this gives the bound:

$$(3.30) \quad \mathbb{E}[(\Delta_\ell^{(h_i)})^2] \ll 2^{-2k} e^{-2(j-\ell)}.$$

We now insert the decomposition (3.27) into the bound (3.26). Equation (3.30) yields the following exponential bound, for some absolute  $c > 0$ ,

$$(3.31) \quad \mathbb{E}[e^{\lambda \sum_{\ell \leq j+1} \Delta_\ell^{(h_i)}}] \ll e^{c \sum_{\ell \leq j+1} \lambda^2 e^{-2(j-\ell)} 2^{-2k}} = e^{O(\lambda^2 2^{-2k})},$$

since  $\ell \leq j+1$ . Note that the bounds in (3.29), (3.30) and (3.31) are the same for all  $h_i \in \mathcal{H}_{k+1}$ . Provided that  $\lambda = \lambda_k \ll 2^k$  (our choice of  $\lambda$  in the end will meet this criterion), the bound (3.26) becomes

$$(3.32) \quad \ll \sum_{r < j \leq t-1} \sum_{v \geq 1} \sum_{k \geq 0} e^j \sum_{h_i \in \mathcal{H}_{k+1}} e^{-\lambda \frac{v}{2(1+k)^2}} \mathbb{E}[e^{\lambda \sum_{\ell \leq j+1} \pi_\ell^{(h_i)} Y_\ell(0)} \mathbf{1}_{\mathcal{B}_{j+1}(y+\psi_{j+1}-v)}].$$

The exponential factor is now amenable to analysis, since its effect is to put a small drift on  $Y_\ell(0)$ . More precisely, we combine the change of measure (3.13) from  $\mathbb{P}$  to  $\tilde{\mathbb{P}}$  with a second change of measure from  $\tilde{\mathbb{P}}$  to  $\bar{\mathbb{P}}$  such that

$$(3.33) \quad \frac{d\tilde{\mathbb{P}}}{d\bar{\mathbb{P}}} = \frac{e^{\lambda \sum_{\ell \leq j+1} \pi_\ell^{(h_i)} Y_\ell(0)}}{\mathbb{E}[e^{\lambda \sum_{\ell \leq j+1} \pi_\ell^{(h_i)} Y_\ell(0)}]}.$$

Under  $\tilde{\mathbb{P}}$ , the increments  $Y_\ell(0)$ ,  $\ell \leq j+1$ , are independent Gaussians of variance  $\mathbb{E}[Y_\ell(0)^2] \ll 1$  and shifted mean  $\mu_\ell + \lambda \pi_\ell^{(h_i)} \mathbb{E}[Y_\ell(0)^2]$ . It is straightforward to check that the cumulative drift is of order 1 for all  $h_i \in \mathcal{H}_{k+1}$  using (3.29):

$$\lambda \sum_{\ell \leq j+1} \pi_\ell^{(h_i)} \leq \lambda \sum_{\ell \leq j+1} e^{-(j-\ell)} 2^{-k} \ll \lambda 2^{-k} \ll 1,$$

again, provided that  $\lambda \ll 2^k$ . By interpreting the small upward drift at each  $\ell$  by a variation of the barrier, we have from Equation (2.17) and Proposition A.3 the estimate

$$(3.34) \quad \tilde{\mathbb{P}}(\mathcal{B}_{j+1}(y + \psi_{j+1} - v)) \ll \frac{(y+1)(v+1)}{j^{3/2}} e^{-(y+\psi_{j+1}-v+c)^2/j}, \quad j \geq t/\log t,$$

by simply raising the original barrier by a constant  $c$ . The endpoint value is also affected by this, but the constant  $c$  will not matter in the end. In the case,  $j \leq t/\log t$ , we can simply use that

$$\tilde{\mathbb{P}}(\mathcal{B}_{j+1}(y + \psi_{j+1} - v)) \ll e^{-(y+\psi_{j+1}-v+c)^2/j}, \quad j \leq t/\log t,$$

as in Equation (3.18). Therefore, bounding Equation (3.32) first by introducing  $\bar{\mathbb{P}}$  as in Equation (3.17), then by introducing  $\tilde{\mathbb{P}}$  and using (3.34), we are left with the bound

$$(3.35) \quad (y+1) \sum_{r < j \leq t-1} \sum_{v \geq 1} \sum_{k \geq 0} 2^k e^{-\lambda \frac{v}{2(1+k)^2}} e^{-2(y+\psi_{j+1}-v)} (v+1) e^{-(y+\psi_{j+1}-v+c)^2/j},$$

as  $\mathcal{H}_{k+1}$  has of order  $2^k$  elements. The only constraint on  $\lambda$  is  $\lambda \ll 2^k$ . Choosing  $\lambda = 404(1+k)^3$  will suffice. Summing first according to  $k$  gives the bound

$$(y+1) e^{-2y} \sum_{r < j \leq t-1} e^{-2\psi_{j+1}} \sum_{v \geq 1} e^{-200v} (v+1) e^{-(y+\psi_{j+1}-v+c)^2/j}.$$

This corresponds to Equation (3.21) up to the constant  $c$ . It is bounded the same way. The constant  $c$  does not matter in the end, because  $e^{-(y+\psi_{j+1}+c)^2/t} \ll e^{-y^2/t}$  since  $y = o(t)$ . This proves the lemma.  $\square$

### 3.2. Upper bound for Theorem 1.2

In this section, we prove the upper bound to Theorem 1.2. Note that the convergence is slightly better than the one stated as the  $\varepsilon \log \log \log T$  term from the definition of the convergence in probability can be replaced at no cost by a function  $g(T)$  going to infinity arbitrarily slowly.

**Proposition 3.3.** *Let  $g(T)$  be a function with  $g(T) \rightarrow +\infty$  arbitrarily slowly, and  $\theta = (\log \log T)^{-\alpha}$  with  $0 < \alpha < 1$ . We have*

$$(3.36) \quad \mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > \sqrt{1+\theta} \log \log T - \frac{1+2\alpha}{4\sqrt{1+\theta}} \log \log \log T + g(T)\right) = o(1).$$

**Proof.** We use the short-hand notation  $t = \log \log T$ , and write, with abuse of notation,  $g(t)$  for  $g(T)$ . For  $\theta = t^{-\alpha}$ , define

$$(3.37) \quad m_\alpha(t) = \sqrt{1+\theta} t - \frac{1+2\alpha}{4\sqrt{1+\theta}} \log t + g(t).$$

The probability of the maximum on the interval  $[-(\log T)^\theta, (\log T)^\theta]$  is bounded above by a union bound on  $(\log T)^\theta$  intervals of length 2. The translation invariance of the distribution of the model then gives

$$(3.38) \quad \mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > m_\alpha(t)\right) \leq (\log T)^\theta \cdot \mathbb{P}\left(\max_{|h| \leq 1} X_T(h) > m_\alpha(t)\right).$$

The recentering  $m_\alpha$  can be expressed as  $m_\alpha(t) = m_1(t) + y$  with  $m_1(t) = t - \frac{3}{4} \log t$  and

$$(3.39) \quad y = (\sqrt{1+\theta} - 1)t - \left(\frac{1+2\alpha}{4\sqrt{1+\theta}} - 3/4\right) \log t + g(t).$$

Note that  $y > 0$  and  $y = o(t)$  if  $0 < \alpha < 1$ , so that Theorem 1.3 can be applied. It gives

$$(3.40) \quad \mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > m_\alpha(t)\right) \leq e^{t\theta} \cdot y e^{-2y} e^{\frac{-y^2}{t}}.$$

From Equation (3.39) and the fact that  $\theta = t^{-\alpha}$ , we deduce that

$$\begin{aligned}\frac{y^2}{t} &= (2 + \theta - 2\sqrt{1 + \theta})t + O(t^{-\alpha} \log t) \\ 2y &= 2(\sqrt{1 + \theta} - 1)t - (\alpha - 1) \log t + 2g(t) + O(t^{-\alpha} \log t),\end{aligned}$$

and

$$(3.41) \quad \frac{y^2}{t} + 2y = \theta t + (1 - \alpha) \log t + 2g(t) + O(t^{-\alpha} \log t).$$

We also have by a Taylor expansion

$$(3.42) \quad \log y = \log \theta + \log t - \log 2 + O(t^{-\alpha} \log t) = (1 - \alpha) \log t + O(1).$$

With these estimates, the right-hand side of Equation (3.40) becomes

$$\begin{aligned}&\ll \exp(\theta t + \log y - \theta t + (\alpha - 1) \log t - 2g(t)) \\ &\ll \exp(-2g(t) + O(1)) = o(1).\end{aligned}$$

This proves the claim for any  $\alpha \in (0, 1)$  and concludes the proof.  $\square$

**Remark 3.4.** In the case  $2\alpha \geq 1$ , the above proof can be made more direct. It is not hard to check that the weaker estimate

$$(3.43) \quad \mathbb{P}\left(\max_{|h| \leq 1} X_T(h) > m_1(t) + y\right) \ll ye^{-2y}$$

is sufficient for the proof.

### 3.3. Upper bound of Theorem 1.1

In this section, we prove the upper bound to Theorem 1.1.

**Proposition 3.5.** *Let  $g(T)$  be a function with  $g(T) \rightarrow +\infty$  arbitrarily slowly, and  $\theta > 0$  fixed. We have*

$$(3.44) \quad \mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > \sqrt{1 + \theta} \log \log T - \frac{1}{4\sqrt{1 + \theta}} \log \log \log T + g(T)\right) = o(1).$$

We will need the following discretization lemma which states that the maximum of  $S_j$  over a neighborhood of size  $e^{-j}$  behaves like a single Gaussian random variable of variance  $j$ .

**Lemma 3.6.** *Let  $C > 0$  and  $1 \leq j \leq t$ . For any  $1 < y \leq Cj$ , we have*

$$\mathbb{P}\left(\max_{|h| \leq e^{-j}} S_j(h) > y\right) \ll_C \frac{e^{-y^2/j}}{j^{1/2}},$$

where  $\ll_C$  means that the implicit constant depends on  $C$ .

The proof of the proposition is a straightforward union bound using the lemma.

**Proof of Proposition 3.5.** Let

$$m_0(t) = \sqrt{1 + \theta} t - \frac{1}{4\sqrt{1 + \theta}} \log t + g(t).$$

The interval  $[-(\log T)^\theta, (\log T)^\theta]$  can be split into  $(\log T)^{1+\theta}$  intervals each of length  $2e^{-t}$ . Therefore, a union bound gives

$$\mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > m_0(t)\right) \leq (\log T)^{1+\theta} \cdot \mathbb{P}\left(\max_{|h| \leq e^{-t}} X_T(h) > m_0(t)\right).$$



Lemma 3.6 can be applied with  $j = t$  (recall that  $S_t = X_T$ ), and  $C = \sqrt{1 + \theta}$ , yielding

$$\mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > m_0(t)\right) \ll e^{t(1+\theta)} \cdot \frac{e^{-m_0(t)^2/t}}{t^{1/2}}.$$

This is  $o(1)$ , thereby concluding the proof.  $\square$

Lemma 3.6 is proved by a chaining argument similarly to Lemma 3.2.

**Proof of Lemma 3.6.** The probability can be divided into

$$\mathbb{P}\left(\max_{|h| \leq e^{-j}} S_j(h) > y, S_j(0) > y - 1\right) + \mathbb{P}\left(\max_{|h| \leq e^{-j}} S_j(h) > y, S_j(0) \leq y - 1\right).$$

The first term is  $\leq \mathbb{P}(S_j(0) > y - 1) \ll \frac{e^{-y^2/j}}{j^{1/2}}$  for  $y > 1$ . It remains to bound the second term. Summing over the values of  $S_j(0)$  gives

$$(3.45) \quad \begin{aligned} & \mathbb{P}\left(\max_{|h| \leq e^{-j}} S_j(h) > y, S_j(0) \leq y - 1\right) \\ & \ll \sum_{x \leq y-1} \mathbb{P}\left(\max_{|h| \leq e^{-j}} S_j(h) - S_j(0) > y - x, S_j(0) \in I_x\right), \end{aligned}$$

where the intervals  $I_x$  are defined below (3.10). This is of the same form as the expression in Equation (3.12) of Lemma 3.2 with  $j + 1$  instead of  $j$ , without the sum over  $j$ , with  $\psi_{j+1} = 0$ , and with the event  $\mathcal{B}_{j+1}(x)$  being simply  $\{S_j(0) \in I_x\}$  (without the recentring). Proceeding the same way using a chaining argument up to Equation (3.35) leads to the bound for  $v = y - x$

$$\sum_{v \geq 1} \sum_{k \geq 0} 2^k e^{-\lambda \frac{v}{2(1+k)^2}} \cdot \frac{e^{-(y-v)^2/j}}{j^{1/2}}.$$

Note that the full ballot estimate (3.34) is not needed here, as only the endpoint is involved. We choose  $\lambda = 200C(1+k)^3$ , so that the above becomes

$$\sum_{v \geq 1} \sum_{k \geq 0} 2^k e^{-100C(1+k)v} e^{-(y-v)^2/j} \ll \sum_{v \geq 1} e^{-100Cv} \cdot \frac{e^{-(y-v)^2/j}}{j^{1/2}}.$$

The summand is maximal at  $v = y - 100Cj$ . This is negative for  $y \leq Cj$ . We conclude that the maximizer is at  $v = 1$  so that

$$\mathbb{P}\left(\max_{|h| \leq e^{-j}} S_j(h) > y, S_j(0) \leq y - 1\right) \ll \frac{e^{-y^2/j}}{j^{1/2}}$$

from which the claim follows.  $\square$

#### 4. Lower bounds

The lower bounds to Theorem 1.1 and Theorem 1.2 are stated in the next two propositions.

**Proposition 4.1.** *For any  $\varepsilon > 0$ , we have for  $\theta \sim (\log \log T)^{-\alpha}$*

$$\mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > \sqrt{1 + \theta} \log \log T - \frac{1 + 2\alpha}{4\sqrt{1 + \theta}} \log \log \log T - \varepsilon \log \log \log T\right) = 1 - o(1).$$

**Proposition 4.2.** *For any  $\varepsilon > 0$ , we have for  $\theta > 0$  fixed*

$$\mathbb{P}\left(\max_{|h| \leq (\log T)^\theta} X_T(h) > \sqrt{1 + \theta} \log \log T - \frac{1}{4\sqrt{1 + \theta}} \log \log \log T - \varepsilon \log \log \log T\right) = 1 - o(1).$$

Proposition 4.1 is proved in the following subsection. The proof of Proposition 4.2 is very similar, and the differences are discussed in Section 4.2.

#### 4.1. Proof of Proposition 4.1

Recall the log log notation in Equation (1.14). Define also

$$(4.1) \quad \mu = \sqrt{1 + \theta} - \frac{1 + 2\alpha}{4\sqrt{1 + \theta}} \frac{\log t}{t} - \varepsilon \frac{\log t}{t}.$$

Consider the set of points in the interval  $[-e^{\theta t}, e^{\theta t}]$  which lie at a distance  $e^{-t}$  of each other

$$\mathcal{H} = [-e^{\theta t}, e^{\theta t}] \cap e^{-t} \mathbb{Z}.$$

Note that  $\#\mathcal{H} = 2e^{(1+\theta)t}$ . Clearly, it is sufficient for the purpose of the lower bound to bound the maximum on  $\mathcal{H}$ .

$$\mathbb{P}\left(\max_{h \in \mathcal{H}} X_T(h) > \mu t\right) = 1 - o(1).$$

We need to introduce some notations to state the lemmas needed for the proof. Define for a fixed  $\delta > 0$

$$(4.2) \quad Z_\delta = \{h \in \mathcal{H} : X_T(h) - \mu t \in [0, \delta], X_K(h) \leq k\mu + b(k), \forall k \leq t\},$$

where the *linear barrier*  $b(k)$  is

$$(4.3) \quad b(k) = \frac{k}{t} + \frac{t^{1-\alpha}}{10} \left(1 - \frac{k}{t}\right), \quad k \leq t.$$

The barrier  $b(k)$  is such that  $b(0) = \frac{t^{1-\alpha}}{10}$  and  $b(t) = 1$ . (The factor  $1/10$  could be replaced by any positive number smaller than  $1/2$ .) Consider also  $\bar{\mathbb{P}}$  defined in Equation (2.9) with  $K = 1$  and  $\lambda' = 0$ :

$$(4.4) \quad \frac{d\bar{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{\lambda X_T(0)}}{\mathbb{E}[e^{\lambda X_T(0)}]}.$$

From Lemma 2.1, we have

$$(4.5) \quad \mathbb{E}[e^{\lambda X_T(0)}] = e^{\lambda^2 \sigma_t^2 / 2}, \quad \sigma_t^2 = \frac{1}{2} \sum_{p \leq T} \frac{1}{p} = \frac{1}{2} t + O(1),$$

where Equation (2.5) is used to estimate the variance. The mean is by Equation (2.10)

$$(4.6) \quad \bar{\mathbb{E}}[X_T(0)] = \lambda \sigma_t^2.$$

Therefore, for the tilted mean to be  $\mu t$ , we take

$$(4.7) \quad \lambda = \frac{\mu t}{\sigma_t^2} = 2\mu + O(1/t).$$

With this in mind, define the recentered partial sums  $\bar{X}_K(0) = X_K(0) - \bar{\mathbb{E}}[X_T(0)]$ ,  $k \leq t$ . It is also convenient to define the event

$$(4.8) \quad J(h) = \{\bar{X}_T(h) \in [0, \delta], \bar{X}_K(h) \leq b(k), \forall k \in \llbracket 1, t \rrbracket\}, \quad h \in \mathcal{H}.$$

**Lemma 4.3.** *Let  $Z_\delta$  be as in Equation (4.2) for  $\delta > 0$  and  $\mu$  as in Equation (4.1). We have*

$$\mathbb{E}[Z_\delta] = 2e^{(1+\theta)t} \cdot e^{-\mu^2 t^2 / (2\sigma_t^2)} \cdot e^{-\delta \mu t / \sigma_t^2} \cdot \bar{\mathbb{P}}(J(0)).$$

*In particular, this gives the lower bound*

$$(4.9) \quad \mathbb{E}[Z_\delta] \gg t^\varepsilon \cdot \delta e^{-2\mu\delta},$$

*which shows that  $\mathbb{E}[Z_\delta] \rightarrow \infty$  as  $t \rightarrow \infty$ , for fixed  $\delta$ .*

**Lemma 4.4.**

$$(4.10) \quad \mathbb{E}[Z_\delta^2] \leq \{(2 + o(1)) e^{(1+\theta)t} \cdot e^{-\mu^2 t^2/(2\sigma_t^2)} \bar{\mathbb{P}}(J(0))\}^2.$$

**Proof of Proposition 4.1.** The Paley-Zygmund inequality implies

$$\mathbb{P}(Z_\delta \geq 1) \geq \frac{(\mathbb{E}[Z_\delta])^2}{\mathbb{E}[Z_\delta^2]} = (e^{-4\mu\delta} + o(1)),$$

by Lemmas 4.3 and 4.4, and the fact that  $\mu t/\sigma_t^2 \rightarrow 2\mu$ . The claim follows by taking the limit  $T \rightarrow \infty$  and then  $\delta \rightarrow 0$ .  $\square$

**Proof of Lemma 4.3.** By linearity of expectation and the change of measure (4.4), the expectation can be written, with the choice  $\lambda = \mu t/\sigma_t^2$ , as

$$(4.11) \quad \begin{aligned} \mathbb{E}[Z_\delta] &= (\#\mathcal{H}) \cdot \mathbb{E}[e^{\lambda X_T(0)}] \cdot e^{-\lambda \bar{\mathbb{E}}[X_T(0)]} \cdot \bar{\mathbb{E}}[e^{-\lambda \bar{X}_T(0)} \mathbb{1}_{J(0)}] \\ &= 2e^{(1+\theta)t} \cdot e^{\mu^2 t^2/(2\sigma_t^2)} \cdot e^{-\mu^2 t^2/\sigma_t^2} \cdot e^{-\delta \mu t/\sigma_t^2} \cdot \bar{\mathbb{P}}(J(0)) \\ &= 2e^{(1+\theta)t} \cdot e^{-\mu^2 t^2/(2\sigma_t^2)} \cdot e^{-\delta \mu t/\sigma_t^2} \cdot \bar{\mathbb{P}}(J(0)) \end{aligned}$$

where we used Equations (4.5), (4.6) and (4.7). This proves the first claim. For the second claim, the definition of  $\mu$  gives

$$e^{-\mu^2 t^2/(2\sigma_t^2)} \ll e^{-(1+\theta)t} \cdot t^{\alpha+1/2+\varepsilon}.$$

It remains to apply Proposition A.1 to  $\bar{\mathbb{P}}(J(0))$ :

$$(4.12) \quad \bar{\mathbb{P}}(J(0)) \gg \delta \frac{b(t) \cdot b(0)}{t^{3/2}} \gg \delta t^{-1/2-\alpha}. \quad \square$$

**Proof of Lemma 4.4.** The second moment of  $Z_\delta$  can be written as a sum over pairs of  $h$ 's

$$\mathbb{E}[Z_\delta^2] = \sum_{h, h'} \mathbb{P}(J(h) \cap J(h')) \leq (\#\mathcal{H}) \cdot \sum_{|h| \leq (\log T)^\theta} \mathbb{P}(J(0) \cap J(h)),$$

where the last equality is by the symmetry in  $h$ 's. The sum is split into three terms:  $|h| \leq 1$ ,  $|h| > e^{t\theta/2}$ , and  $1 < |h| \leq e^{t\theta/2}$ . The dominant term is the latter.

For the sum over  $|h| > e^{t\theta/2}$ , we use the biased measure  $\bar{\mathbb{P}}$  in Equation (2.9) with  $\lambda = \lambda'$  and  $K = 1$

$$(4.13) \quad \frac{d\bar{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{\lambda(X_T(0) + X_T(h))}}{\mathbb{E}[e^{\lambda(X_T(0) + X_T(h))}]}.$$

Lemma 2.1 gives for the Laplace transform

$$(4.14) \quad \begin{aligned} \mathbb{E}[e^{\lambda(X_T(0) + X_T(h))}] &= \exp\left(\frac{\lambda^2}{2} \sum_{p \leq T} \frac{1 + \cos(|h| \log p)}{p}\right) \\ &= \exp(\lambda^2 \sigma_T^2 + O(|h|^{-1})). \end{aligned}$$

In particular, the tilted mean is by Equation (2.10)

$$(4.15) \quad \bar{\mathbb{E}}[X_T(0)] = \lambda \sigma_T^2 + O(|h|^{-1}).$$

We choose  $\bar{\mathbb{E}}[X_T(0)] = \mu t$  so that

$$(4.16) \quad \lambda = \frac{\mu t}{\sigma_T^2} + O(e^{-t\theta/2}) = 2\mu + O(1/t).$$

Lemma 2.3 gives

$$(4.17) \quad \begin{aligned} \bar{\mathbb{P}}(J(0) \cap J(h)) &= (1 + O(e^{-t\theta/2})) \bar{\mathbb{P}}(J(0)) \cdot \bar{\mathbb{P}}(J(h)) + O(\exp(-ce^{t\theta/4})) \\ &= (1 + o(1)) \bar{\mathbb{P}}(J(0)) \cdot \bar{\mathbb{P}}(J(h)), \end{aligned}$$

where Equation (4.12) is used to get the second equality. As before, define the recentered process  $\bar{X}_K = X_K - \bar{\mathbb{E}}[X_K]$ ,  $k \in \llbracket 1, t \rrbracket$ . With this notation, the contribution of  $|h| > e^{t\theta/2}$  to  $\mathbb{E}[Z_\delta]$  is

$$(4.18) \quad \begin{aligned} (\#\mathcal{H}) \cdot \sum_{|h| > e^{t\theta/2}} \mathbb{P}(J(0) \cap J(h)) &= e^{\lambda^2 \sigma_t^2} \cdot e^{-2\lambda \bar{\mathbb{E}}[X_T(0)]} \cdot (\#\mathcal{H}) \cdot \sum_{|h| > e^{t\theta/2}} \bar{\mathbb{P}}(J(0) \cap J(h)) \\ &= (2 + o(1))^2 e^{-\mu^2 t^2 / \sigma_t^2} \cdot e^{2t(1+\theta)} \cdot \bar{\mathbb{P}}(J(0))^2, \end{aligned}$$

where in the first inequality we bounded the factor  $e^{-\lambda(\bar{X}_T(0) + \bar{X}_T(h))}$  in the expectation with respect to  $\bar{\mathbb{P}}$  from above by 1 (on the event  $J(0) \cap J(h)$ ). In the second equality, we used Equations (4.15), (4.17), and the fact that there are  $(2 + o(1))e^{(1+\theta)t}$  terms in the sum. This is (to leading order) equal to the right-hand side of Equation (4.10). It remains to prove that the other contributions are negligible compared to  $(\mathbb{E}[Z_\delta])^2$ .

For the case  $1 < |h| \leq e^{t\theta/2}$ , we proceed as in Equation (4.18). The number of  $h'$  in the sum is now  $\ll e^{t+t\theta/2}$  instead of being  $\gg e^{t(1+\theta)}$ . Moreover, the additive error term in (4.17) is  $O(1)$ . Since  $\bar{\mathbb{P}}(J(0)) \gg e^{-t\theta/4}$  by (4.12), we get

$$(4.19) \quad (\#\mathcal{H}) \cdot \sum_{1 < |h| \leq e^{t\theta/2}} \mathbb{P}(J(0) \cap J(h)) \ll e^{-t\theta/4} (e^{-\mu^2 t^2 / \sigma_t^2} \cdot e^{2t(1+\theta)} \cdot \bar{\mathbb{P}}(J(0))^2).$$

This is  $o((\mathbb{E}[Z_\delta])^2)$  by Lemma 4.3.

It remains to bound the contribution of  $|h| \leq 1$ . It is convenient to split into the intervals  $e^{-k-1} < |h| \leq e^{-k}$  for  $0 \leq k \leq t$ . For a given  $k$ , the contribution to  $\mathbb{E}[Z_\delta^2]$ , that we denote by  $(\mathbb{E}[Z_\delta^2])_k$  is

$$(\mathbb{E}[Z_\delta^2])_k \ll (\#\mathcal{H}) \cdot \sum_{e^{-k-1} < |h| \leq e^{-k}} \mathbb{P}(J(0) \cap J(h)).$$

We handle two ranges of  $k$ 's differently:

$$(I): \quad 0 \leq k \leq t - t^{\alpha+\varepsilon/100}, \quad (II): \quad t - t^{\alpha+\varepsilon/100} < k \leq t.$$

In the first range, most of the barrier is not needed to show that the contribution is negligible compared to  $(\mathbb{E}[Z_\delta])^2 \gg t^{2\varepsilon}$ . Only the value at the “branching time  $k$ ” is necessary. This is because the barrier is too high at these times to affect the probability of the walks. For technical reason, we need the boundary case  $k = t - t^{\alpha+\varepsilon/100}$  to be slightly smaller than  $t - t^\alpha$ , where the barrier becomes of order one.

To handle the additive error term coming out from the decoupling in Lemma 2.3, it is necessary to split the random walk  $X_T$  into  $X_{K'}$  and  $X_{K',T}$  where  $K'$  is slightly larger than  $K$ . We take  $k' = k + r$  where  $r = 10 \log \log t$ . We consider the biased measure  $\bar{\mathbb{P}}$  now defined separately for  $X_{K'}$  and  $X_{K',T}$

$$(4.20) \quad \frac{d\bar{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{\lambda_1(X_{K'}(0) + X_{K'}(h))}}{\mathbb{E}[e^{\lambda_1(X_{K'}(0) + X_{K'}(h))}]} \cdot \frac{e^{\lambda_2(X_{K',T}(0) + X_{K',T}(h))}}{\mathbb{E}[e^{\lambda_2(X_{K',T}(0) + X_{K',T}(h))}]}$$

We pick  $\lambda_1 = \mu$  and  $\lambda_2 = 2\mu$  to reflect the fact that  $X_{K'}(0)$  and  $X_{K'}(h)$  are almost perfectly correlated for  $|h| \approx e^{-k}$ . The Laplace transforms are easily computed using Lemma 2.1 and Lemma 2.2:

$$(4.21) \quad \mathbb{E}[e^{\mu(X_{K'}(0) + X_{K'}(h))}] \ll e^{\mu^2 k'} = e^{\mu^2(k+r)} \quad \mathbb{E}[e^{2\mu(X_{K',T}(0) + X_{K',T}(h))}] \ll e^{2\mu^2(t-k')} = e^{2\mu^2(t-k-r)}.$$

The tilted mean are given by Equation (2.10) together with Lemma 2.2:

$$\bar{\mathbb{E}}[X_{K'}(0)] = \bar{\mathbb{E}}[X_{K'}(h)] = \mu k' + O(1) \quad \bar{\mathbb{E}}[X_{K',T}(0)] = \bar{\mathbb{E}}[X_{K',T}(h)] = \mu(t - k') + O(e^{-e^{k'/2}}).$$

Again, the recentered sums are  $\bar{X}_{K'} = X_{K'} - \bar{\mathbb{E}}[X_{K'}]$  and  $\bar{X}_{K',T} = X_{K',T} - \bar{\mathbb{E}}[X_{K',T}]$ .

Case (I). The probability is split using the value of  $\bar{X}_{K'}$  at 0 and  $h$ . Note that these values are bounded by  $b(k')$ . Since there are approximately  $e^{t-k}$   $h$ 's with  $|h| \approx e^{-k}$  in  $\mathcal{H}$ , we get

$$\begin{aligned}
 (\mathbb{E}[Z_\delta^2])_k &\ll (\#\mathcal{H}) \cdot e^{t-k} \cdot \sum_{q_1, q_2 \leq b(k')} \mathbb{P}(\bar{X}_{K'}(0) \in (q_1 - 1, q_1], \bar{X}_{K'}(h) \in (q_2 - 1, q_2]) \\
 &\quad \cdot \mathbb{P}(\bar{X}_{K',T}(0) + q_1 \in [0, \delta + 1], \bar{X}_{K',T}(h) + q_2 \in [0, \delta + 1]) \\
 (4.22) \quad &\ll e^{t(1+\theta)+(t-k)} \cdot \sum_{q_1, q_2 \leq b(k+r)} e^{-\mu(q_1+q_2)} e^{-\mu^2(k+r)} \cdot e^{2\mu(q_1+q_2)} e^{-2\mu^2(t-k-r)} \\
 &\ll e^r \cdot e^{-\theta(t-k)} \cdot t^{(\frac{k}{t} + 2\frac{t-k}{t})(\alpha+1/2+\varepsilon)} \sum_{q_1, q_2 \leq b(k+r)} e^{\mu(q_1+q_2)}.
 \end{aligned}$$

In the second inequality, we changed measure to  $\bar{\mathbb{P}}$  using (4.21), and simply dropped the probabilities of the events under  $\bar{\mathbb{P}}$  keeping only the normalizing constants. In the third inequality, we simply replaced the definition of  $\mu$  given in (4.1). Note that  $e^{-\theta(t-k)} \cdot t^{(\frac{k}{t} + 2\frac{t-k}{t})(\alpha+1/2+\varepsilon)} \ll e^{-\frac{9}{10}\theta(t-k)}$  for  $k \leq t - t^{\alpha+\varepsilon/100}$ . Moreover, by definition of  $b(k)$ , we have  $b(j) \leq \theta \frac{(t-j)}{10} + 1$  for all  $j$ , therefore the above is

$$(4.23) \quad (\mathbb{E}[Z_\delta^2])_k \ll e^r \cdot e^{-\frac{9}{10}\theta(t-k)} \cdot e^{2\mu b(k+r)} \ll e^r \cdot e^{-\frac{\theta}{2}(t-k)}.$$

The sum over the range of  $k$  considered then yields

$$\sum_{1 \leq k \leq t - t^{\alpha+\varepsilon/100}} (\mathbb{E}[Z_\delta^2])_k \ll e^r \cdot \theta e^{-\frac{\theta t^{\alpha+\varepsilon/100}}{2}} \ll e^r.$$

This is much smaller than  $(\mathbb{E}[Z_\delta])^2 \gg t^{2\varepsilon} \cdot \delta^2 e^{-4\mu\delta}$ , by Equation (4.9) and the choice of  $r$ .

Case (II). It remains to estimate the contribution of  $t - t^{\alpha+\varepsilon/100} < k \leq t$ . This is done as in Equation (4.22) keeping now the probabilities after the change of measure. This gives

$$\begin{aligned}
 (\mathbb{E}[Z_\delta^2])_k &\ll e^r \cdot e^{-\theta(t-k)} \cdot t^{(\frac{k}{t} + 2\frac{t-k}{t})(\alpha+1/2+\varepsilon)} \\
 &\quad \cdot \sum_{q_1, q_2 \leq b(k)} e^{\mu(q_1+q_2)} \bar{\mathbb{P}}(A_0(q_1) \cap A_h(q_2)) \cdot \bar{\mathbb{P}}(B_0(q_1) \cap B_h(q_2)),
 \end{aligned}$$

where

$$\begin{aligned}
 (4.24) \quad A_h(q) &= \{\bar{X}_{K'}(h) \in (q - 1, q], \bar{X}_L(h) \leq b(\ell), \forall \ell \in \llbracket 1, k' \rrbracket\} \\
 B_h(q) &= \{\bar{X}_{K',T}(h) + q \in [0, \delta + 1], \bar{X}_{K',L}(h) + q \leq b(\ell), \forall \ell \in \llbracket k' + 1, t \rrbracket\}.
 \end{aligned}$$

The idea is that  $X_L(0)$  and  $X_L(h)$  should be essentially equal for  $\ell \leq k'$ , whereas  $X_{L,T}(0)$  and  $X_{L,T}(h)$  are essentially independent for  $\ell > k'$ . To obtain an upper bound we can simply drop  $A_h(q_2)$  in the inequality. This is expected to be almost sharp since  $X_L(0)$  and  $X_L(h)$  should be essentially equal for  $\ell \leq k$ . For the decoupling, since  $k' = k + r$  slightly larger than  $k$ , the additive error term in Lemma 2.3 will be absorbed. More precisely, in order to estimate  $\bar{\mathbb{P}}(B_0(q_1) \cap B_h(q_2))$ , we apply Lemma 2.3, Corollary 2.4 and Equation (A.2) to get

$$\begin{aligned}
 (4.25) \quad \bar{\mathbb{P}}(B_0(q_1) \cap B_h(q_2)) &\ll \bar{\mathbb{P}}(B_0(q_1)) \cdot \bar{\mathbb{P}}(B_h(q_2)) + O(e^{-c\sqrt{e^r}}) \\
 &\ll \frac{1}{(t-k-r)^3} \prod_{i=1,2} (b(k+r) - q_i + 1),
 \end{aligned}$$

by the choice of  $r$  and the fact that  $b(t) = 1$ . The probability  $\bar{\mathbb{P}}(A_0(q_1))$  is bounded using Corollary 2.4 and Proposition A.2 to give

$$(4.26) \quad \bar{\mathbb{P}}(A_0(q_1)) \ll \frac{(b(k+r) - q_1 + 1)(b(0) + 1)}{(k+r)^{3/2}} \ll \frac{(b(k+r) - q_1 + 1) \cdot t^{1-\alpha}}{k^{3/2}},$$

where we use the fact that  $k$  is much larger than  $r$  in the range considered. (The increments here do not have variance exactly  $1/2$ , but can be simply fixed by considering the slightly larger event where the barrier start at  $r$ , and by conditioning on the position at time  $r$ . The error term in Corollary 2.4 can then be easily absorbed.)

Putting Equations (4.26) and (4.25) together yield

$$\begin{aligned} (\mathbb{E}[Z_\delta^2])_k &\ll e^r \cdot e^{-\theta(t-k)} \cdot \frac{t^{(\frac{k}{t}+2\frac{t-k}{t})(\alpha+1/2+\varepsilon)+(1-\alpha)}}{k^{3/2}(t-k)^3} \\ &\cdot \sum_{q_1, q_2 \leq b(k+r)} e^{\mu(q_1+q_2)} (b(k+r) - q_1 + 1) \cdot (b(k+r) - q_1 + 1) (b(k+r) - q_2 + 1). \end{aligned}$$

The three factors in parentheses are smaller than  $t^{\varepsilon/2}$  since  $b(k+r) \ll t^{\varepsilon/100}$ . Moreover, we have similarly as in Case (I):  $e^{-\theta(t-k)} e^{2\mu b(k+r)} \ll e^{-\frac{9}{10}\theta(t-k)} \ll 1$  for all  $k$  in the range. The above becomes

$$(\mathbb{E}[Z_\delta^2])_k \ll e^r \cdot t^{\varepsilon/2} \cdot \frac{t^{(\frac{k}{t}+2\frac{t-k}{t})(\alpha+1/2+\varepsilon)+(1-\alpha)}}{k^{3/2}(t-k)^3}.$$

The numerator can also be simplified:

$$t^{(\frac{k}{t}+2\frac{t-k}{t})(\alpha+1/2+\varepsilon)+(1-\alpha)} = t^{3/2+\varepsilon} \cdot t^{(1-\frac{k}{t})(\alpha+1/2+\varepsilon)}.$$

We are left with

$$(\mathbb{E}[Z_\delta^2])_k \ll \frac{e^r \cdot t^{3/2} \cdot t^{3\varepsilon/2} \cdot t^{(1-\frac{k}{t})(\alpha+1/2)}}{k^{3/2}(t-k)^3} \ll \frac{e^r \cdot t^{(1-\frac{k}{t})(\alpha+1/2+\varepsilon)}}{(t-k)^3},$$

since  $k^{3/2} > (t - t^{\alpha+\varepsilon/100})^{3/2} \gg t^{3/2}$ . It remains to sum over  $k$ . After the change of index  $\hat{k} = t - k$ , the contribution of the range is

$$\sum_{k=t-t^{\alpha+\varepsilon/100}+1}^t (\mathbb{E}[Z_\delta^2])_k \ll e^r \cdot t^{3\varepsilon/2} \cdot \sum_{\hat{k}=1}^{t^{\alpha+\varepsilon/100}} \frac{t^{\frac{\hat{k}}{t}(\alpha+1/2+\varepsilon)}}{\hat{k}^3} \ll e^r \cdot t^{3\varepsilon/2} \cdot \sum_{\hat{k}=1}^{\infty} \frac{1}{\hat{k}^3} \ll t^{7\varepsilon/4}.$$

This proves that this contribution is  $o(\mathbb{E}[Z_\delta]^2)$  by Equation (4.9) and concludes the proof of the lemma.  $\square$

#### 4.2. Proof of Proposition 4.2

We now explain which modifications in the proof of Proposition 4.1 are necessary to prove Proposition 4.2. Note that the case  $\theta$  greater than zero is actually simpler as we do not need the ballot-type estimates used in Case (II) in the proof of Lemma 4.4.

**Proof of Proposition 4.2.** As seen in the proof of Proposition 4.1 it suffices to proof a lower bound on the first moment and an upper bound on the second moment of  $Z_\delta$ . Adapting the barrier in (4.3) to the case  $\theta > 0$  we choose

$$(4.27) \quad b(k) = \frac{k}{t} + \frac{\theta t}{10} \left(1 - \frac{k}{t}\right), \quad k \leq t.$$

Moreover, in the definition of  $\mu$  we set  $\alpha = 0$ .

We start by showing a lower bound on  $\mathbb{E}[Z_\delta]$ . The first bound in Lemma 4.3 (and its proof) still hold, with  $\alpha = 0$ , without any modification. Noting that the barrier in (4.27) is linear and of order  $t$  at the beginning we have

$$\overline{\mathbb{P}}(J(0)) \gg t^{-1/2}.$$

The bound in (4.9) follows. Next, we turn to the upper bound of  $\mathbb{E}[Z_\delta^2]$ . Equations (4.18) and (4.19) hold as is.

It remains to bound the terms with  $e^{-k-1} \leq |h| \leq e^{-k}$  for  $k \geq 0$ . We proceed in a more straightforward way than the proof of Lemma 4.4. We split the walk  $X_T$  into  $X_K$  and  $X_{K,T}$ . (There is no need for an extra spacing  $K'$  to allow for a

better error in the decoupling here.) We have

$$\begin{aligned}
 (\mathbb{E}[Z_\delta^2])_k &\ll (\#\mathcal{H}) \cdot e^{t-k} \cdot \sum_{q_1, q_2 \leq b(k)} \mathbb{P}(\bar{X}_K(0) \in (q_1 - 1, q_1], \bar{X}_K(h) \in (q_2 - 1, q_2]) \\
 &\quad \cdot \mathbb{P}(\bar{X}_{K,T}(0) + q_1 \in [0, \delta + 1], \bar{X}_{K,T}(h) + q_2 \in [0, \delta + 1]) \\
 (4.28) \quad &\ll e^{t(1+\theta)+(t-k)} \cdot \sum_{q_1, q_2 \leq b(k)} \frac{1}{\sqrt{k}} e^{-\mu(q_1+q_2)} e^{-\mu^2 k} \cdot e^{2\mu(q_1+q_2)} e^{-2\mu^2(t-k)} \\
 &\ll e^{-\theta(t-k)} \cdot t^{(\frac{k}{t} + 2\frac{t-k}{t})(1/2+\varepsilon)} \sum_{q_1, q_2 \leq b(k)} \frac{1}{\sqrt{k}} e^{\mu(q_1+q_2)}.
 \end{aligned}$$

In the second line, we changed measure to  $\bar{\mathbb{P}}$ , and instead of simply dropping events under  $\bar{\mathbb{P}}$  we kept a factor of order  $1/\sqrt{k}$  from the first probability. In the third line, we again simply replaced the definition of  $\mu^2$ . By definition of  $b(k)$ , we have  $b(k) \leq \theta \frac{(t-k)}{10} + 1$ , therefore for  $k < t - C \log t$  (for some constant  $C$  large enough),

$$(4.29) \quad (\mathbb{E}[Z_\delta^2])_k \ll e^{-\frac{9}{10}\theta(t-k)} \cdot e^{2\mu b(k)} \ll e^{-\frac{\theta}{2}(t-k)},$$

where we absorbed the  $t^{1/2}$  in the exponential, and we bounded  $\mu$  by 2. The sum over  $k$  is thus of order one. For  $k > t - C \log t$  we observe that  $t^{(\frac{k}{t} + 2\frac{t-k}{t})(1/2+\varepsilon)}/\sqrt{k}$  is of order  $t^\varepsilon$ . Therefore, we get in this range

$$\sum_{k > t - C \log t} (\mathbb{E}[Z_\delta^2])_k \ll t^\varepsilon \sum_{k > t - C \log t} e^{-\theta(t-k)} \cdot e^{2\mu b(k)} \ll t^\varepsilon.$$

The upper bound on the second moment then follows as in the Lemma 4.4. This concludes the proof of Proposition 4.2.  $\square$

## Appendix: Gaussian Ballot Theorem

In the propositions below, it is useful to define for  $0 < \delta \leq 1$  the intervals  $I_x = (x, x + \delta]$  for  $x > 0$ ,  $I_x = (x - \delta, x]$  if  $x < 0$ , and  $I_x = [-\delta, \delta]$  if  $x = 0$ . In the first proposition, we allow the increments of the random walk to have different variances. In the second and third, all increments have variance  $1/2$ .

**Proposition A.1 (Ballot theorem with linear barrier: Lower bound).** *Let  $(S_j, j \geq 1)$  be a random walk with Gaussian increments of mean 0 and variance  $\sigma_j^2$  with  $c^{-1} < \sigma_j < c$  for some fixed  $c \geq 1$ . Define the linear barrier*

$$b(j) = aj + b(0).$$

*Then, for any  $j \geq 1$ ,  $(b(j) - x) \cdot b(0) \leq j$ , we have*

$$(A.1) \quad \mathbb{P}(\forall 1 \leq \ell \leq j, S_\ell \leq b(\ell), S_j \in I_x) \gg \frac{b(0)(b(j) - x)}{j^{3/2}} \delta e^{-\frac{cx^2}{j}}.$$

**Proof.** The probability is bounded below by the probability for a Brownian motion on the continuous-time interval, see for example Lemma 6.2 in [38].  $\square$

**Proposition A.2 (Ballot theorem with linear barrier: Upper bound).** *Let  $(S_j, j \geq 1)$  be a random walk with Gaussian increments of mean 0 and variance  $1/2$ . Define the linear barrier*

$$b(j) = aj + b(0).$$

*Then, for any  $j \geq 1$ ,  $b(0) > 0$  and  $x \leq b(j)$ , we have*

$$(A.2) \quad \mathbb{P}(\forall 1 \leq \ell \leq j, S_\ell \leq b(\ell), S_j \in I_x) \ll \frac{(b(0) + 1)(b(j) - x + 1)}{j^{3/2}}.$$

**Proof.** This is a standard ballot theorem with linear barrier, see Lemma 6.2 in [38].  $\square$



**Proposition A.3 (Ballot theorem with logarithmic barrier).** *Let  $(S_j, j \leq t)$  be a random walk with Gaussian increments of mean 0 and variance 1/2. Define the logarithmic barrier*

$$\psi_j = \log(\min(j, t - j)).$$

*Let  $t/\log t \leq k \leq t$ ,  $y = o(t/\log t)$ ,  $-20k < x \leq \psi_k$ . For  $r = y$ , we have*

$$(A.3) \quad \mathbb{P}(\forall r < j \leq k, S_j \leq y + \psi_j, S_k \in I_x, |S_r| \leq 3r) \ll \frac{(y+1)(y+\psi_k-x+1)}{k^{3/2}} e^{-\frac{x^2}{k}}.$$

**Proof.** See Propositions 4 and 5 in [4]. Proposition 4 there was stated for  $k \geq t/2$  but holds *verbatim* for  $k \geq t/\log t$  by applying Proposition 5. The condition  $r = y = o(t/\log t)$  ensures that  $r$  is small compared to  $k$  so that  $(k-r)^{3/2} \sim k^{3/2}$ . The lower bound on  $x$  is to make sure the distance to the barrier at time  $k$  is  $\ll k$ . Finally, the bound  $|S_r| \leq 3r$  is such that the initial distance at time  $r$  is no more than the order of  $y$ , so that the factor  $(y+1)$  (measuring the distance between the barrier and the starting point at the initial time  $r$ ) on the right side still holds.  $\square$

## Acknowledgements

The authors would like to thank the referee for the numerous insightful comments that led to a substantial improvement of the first version of the paper.

## Funding

The research of L.-P. A. is supported in part by the grant NSF CAREER DMS-1653602 and NSF DMS-2153803.

G. D. gratefully acknowledges support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 754411.

The research of L. H. is supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Project-ID 233630050-TRR 146, Project-ID 443891315 within SPP 2265 and Project-ID 446173099.

## References

- [1] E. Amzallag, L.-P. Arguin, E. Bailey, K. Hui and R. Rao. Evidence of random matrix corrections for the large deviations of Selberg's central limit theorem. *Exp. Math.* **0** (2021) 1–13.
- [2] L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.* **72** (3) (2019) 500–535. [MR3911893](#) <https://doi.org/10.1002/cpa.21791>
- [3] L.-P. Arguin, D. Belius and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.* **27** (1) (2017) 178–215. [MR3619786](#) <https://doi.org/10.1214/16-AAP1201>
- [4] L.-P. Arguin, P. Bourgade and M. Radziwiłł. The Fyodorov–Hiary–Keating Conjecture. I, 2020. Preprint. Available at [arXiv:2007.00988](#).
- [5] L.-P. Arguin and F. Ouimet. Extremes of the two-dimensional Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** (2) (2016) 779–808. [MR3541850](#) <https://doi.org/10.30757/alea.v13-31>
- [6] L.-P. Arguin, F. Ouimet and M. Radziwiłł. Moments of the Riemann zeta function on short intervals of the critical line. *Ann. Probab.* **49** (6) (2021) 3106–3141. [MR4348686](#) <https://doi.org/10.1214/21-aop1524>
- [7] L.-P. Arguin and O. Zindy. Poisson–Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.* **20** (59) (2015) 19. [MR3354619](#) <https://doi.org/10.1214/EJP.v20-3077>
- [8] M. Aymone, W. Heap and J. Zhao. Partial sums of random multiplicative functions and extreme values of a model for the Riemann zeta function. *J. Lond. Math. Soc.* **103** (4) (2021) 1618–1642. [MR4273482](#) <https://doi.org/10.1112/jlms.12421>
- [9] M. A. Belloum. A generalized model interpolating between the random energy model and the branching random walk. Preprint [arXiv:2101.09502](#) (2021).
- [10] J. Berestycki, E. Brunet, A. Cortines and B. Mallein. A simple backward construction of branching Brownian motion with large displacement and applications. Preprint [arXiv:1810.05809](#) (2020). [MR4492972](#) <https://doi.org/10.1214/21-aihp1212>
- [11] M. Biskup. Extrema of the two-dimensional discrete Gaussian free field. In *Random Graphs, Phase Transitions, and the Gaussian Free Field* 163–407. Springer Proc. Math. Stat. **304**. Springer, Cham, 2020. [MR4043225](#) [https://doi.org/10.1007/978-3-030-32011-9\\_3](https://doi.org/10.1007/978-3-030-32011-9_3)
- [12] P. Bourgade. Mesoscopic fluctuations of the zeta zeros. *Probab. Theory Related Fields* **148** (3–4) (2010) 479–500. [MR2678896](#) <https://doi.org/10.1007/s00440-009-0237-3>
- [13] A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.* **19** (2014), no. 18, 28. [MR3164771](#) <https://doi.org/10.1214/EJP.v19-2982>
- [14] A. Bovier and L. Hartung. Variable speed branching Brownian motion I. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** (1) (2015) 261–291. [MR3351476](#)
- [15] A. Bovier and L. Hartung. From 1 to 6: A finer analysis of perturbed branching Brownian motion. *Comm. Pure Appl. Math.* **73** (7) (2020) 1490–1525. [MR4156608](#) <https://doi.org/10.1002/cpa.21893>

- [16] M. Bramson, J. Ding and O. Zeitouni. Convergence in law of the maximum of nonlattice branching random walk. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (4) (2016) 1897–1924. [MR3573300](#) <https://doi.org/10.1214/15-AIHP703>
- [17] M. Bramson, J. Ding and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.* **69** (1) (2016) 62–123. [MR3433630](#) <https://doi.org/10.1002/cpa.21621>
- [18] M. D. Bramson. Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31** (5) (1978) 531–581. [MR494541](#) <https://doi.org/10.1002/cpa.3160310502>
- [19] M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.* **17** (2012), no. 67, 18. [MR2968674](#) <https://doi.org/10.1214/EJP.v17-2253>
- [20] D. W. Farmer, S. M. Gonek and C. P. Hughes. The maximum size of  $L$ -functions. *J. Reine Angew. Math.* **609** (2007) 215–236. [MR2350784](#) <https://doi.org/10.1515/CRELLE.2007.064>
- [21] M. Fels. Extremes of the 2d scale-inhomogeneous discrete Gaussian free field: Sub-leading order and exponential tails. Preprint [arXiv:1910.09915](#) (2020).
- [22] M. Fels and L. Hartung. Extremes of the 2d scale-inhomogeneous discrete Gaussian free field: Convergence of the maximum in the regime of weak correlations. *ALEA Lat. Am. J. Probab. Math. Stat.* **18** (2) (2021) 1891–1930. [MR4332224](#) <https://doi.org/10.30757/alea.v18-71>
- [23] M. Fels and L. Hartung. Extremes of the 2d scale-inhomogeneous discrete Gaussian free field: Extremal process in the weakly correlated regime. *ALEA Lat. Am. J. Probab. Math. Stat.* **18** (2) (2021) 1689–1718. [MR4291038](#) <https://doi.org/10.30757/alea.v18-62>
- [24] 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. <https://doi.org/10.1103/PhysRevLett.108.170601>
- [25] Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: Random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **372** (2007) (2014), 20120503, 32. [MR3151088](#) <https://doi.org/10.1098/rsta.2012.0503>
- [26] M. Gerspach. Pseudomoments of the Riemann Zeta Function. *Ph.D. Thesis* (2020).
- [27] A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. Preprint [arXiv:1304.0677](#) (2013).
- [28] A. J. Harper. On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture. [arXiv:1906.05783](#) (2019).
- [29] N. Kistler. Derrida’s random energy models. From spin glasses to the extremes of correlated random fields. In *Correlated Random Systems: Five Different Methods* 71–120. *Lecture Notes in Math.* **2143**. Springer, Cham, 2015. [MR3380419](#) [https://doi.org/10.1007/978-3-319-17674-1\\_3](https://doi.org/10.1007/978-3-319-17674-1_3)
- [30] N. Kistler and M. A. Schmidt. From Derrida’s random energy model to branching random walks: From 1 to 3. *Electron. Commun. Probab.* **20** (2015), no. 47, 12. [MR3358969](#) <https://doi.org/10.1214/ecp.v20-4189>
- [31] J. E. Littlewood. On the zeros of the Riemann zeta-function. *Proc. Camb. Philos. Soc.* **22** (1924) 295–318. [MR1577156](#) <https://doi.org/10.1112/plms/s2-24.1.175>
- [32] B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stochastic Process. Appl.* **125** (10) (2015) 3958–4019. [MR3373310](#) <https://doi.org/10.1016/j.spa.2015.05.011>
- [33] B. Mallein and P. Miłoś. Maximal displacement of a supercritical branching random walk in a time-inhomogeneous random environment. *Stochastic Process. Appl.* **129** (9) (2019) 3239–3260. [MR3985561](#) <https://doi.org/10.1016/j.spa.2018.09.008>
- [34] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory. I. Classical Theory*. *Cambridge Studies in Advanced Mathematics* **97**, xviii+552. Cambridge University Press, Cambridge, 2007. [MR2378655](#)
- [35] 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. [MR3851835](#) <https://doi.org/10.1007/s00440-017-0812-y>
- [36] F. Ouimet. Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.* **14** (2) (2017) 851–902. [MR3731796](#)
- [37] E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: Statistics on the critical line. *Ann. Probab.* **48** (6) (2020) 2680–2754. [MR4164452](#) <https://doi.org/10.1214/20-AOP1433>
- [38] C. Webb. Exact asymptotics of the freezing transition of a logarithmically correlated random energy model. *J. Stat. Phys.* **145** (6) (2011) 1595–1619. [MR2863721](#) <https://doi.org/10.1007/s10955-011-0359-8>