



**HAL**  
open science

# The free energy in the Derrida–Retaux recursive model

Yueyun Hu, Zhan Shi

► **To cite this version:**

Yueyun Hu, Zhan Shi. The free energy in the Derrida–Retaux recursive model. *Journal of Statistical Physics*, 2018, 172 (718-741), 10.1007/s10955-018-2066-1 . hal-01519983v3

**HAL Id: hal-01519983**

**<https://hal.science/hal-01519983v3>**

Submitted on 16 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The free energy in the Derrida–Retaux recursive model

by

Yueyun Hu<sup>1</sup> and Zhan Shi<sup>2</sup>

*Université Paris XIII & Université Paris VI*

**Summary.** We are interested in a simple max-type recursive model studied by Derrida and Retaux [11] in the context of a physics problem, and find a wide range for the exponent in the free energy in the nearly supercritical regime.

**Keywords.** Max-type recursive model, free energy.

**2010 Mathematics Subject Classification.** 60J80, 82B44.

## 1 Introduction

### 1.1 The Derrida–Retaux model

We are interested in a max-type recursive model investigated in 2014 by Derrida and Retaux [11], as a toy version of the hierarchical pinning model; see Section 1.3. The model can be defined, up to a simple change of variables, as follows: for all  $n \geq 1$ ,

$$(1.1) \quad X_{n+1} \stackrel{\text{law}}{=} (X_n + \tilde{X}_n - 1)^+,$$

---

<sup>1</sup>LAGA, Université Paris XIII, 99 avenue Jean-Baptiste Clément, F-93430 Villetaneuse, France, [yueyun@math.univ-paris13.fr](mailto:yueyun@math.univ-paris13.fr)

<sup>2</sup>LPMA, Université Pierre et Marie Curie, 4 place Jussieu, F-75252 Paris Cedex 05, France, [zhan.shi@upmc.fr](mailto:zhan.shi@upmc.fr)

where  $\tilde{X}_n$  denotes an independent copy of  $X_n$ , and “ $\stackrel{\text{law}}{=}$ ” stands for identity in distribution. We assume that  $X_0$  is a non-negative random variable.

Since  $(X_n + \tilde{X}_n - 1)^+ \leq X_n + \tilde{X}_n$ , we have  $\mathbb{E}(X_{n+1}) \leq 2\mathbb{E}(X_n)$ , which implies the existence of the free energy

$$(1.2) \quad F_\infty := \lim_{n \rightarrow \infty} \downarrow \frac{\mathbb{E}(X_n)}{2^n}.$$

An immediate question is how to separate the two regimes  $F_\infty > 0$  and  $F_\infty = 0$ .

**Example 1.1.** Assume  $\mathbb{P}(X_0 = 2) = p$  and  $\mathbb{P}(X_0 = 0) = 1 - p$ , where  $p \in [0, 1]$  is a parameter. There exists  $p_c \in (0, 1)$  such that  $F_\infty > 0$  if  $p > p_c$ , and that  $F_\infty = 0$  if  $p < p_c$ .

The value of  $p_c$  is known to be  $\frac{1}{5}$  (Collet et al. [8]). □

More generally, we write  $P_X$  for the law of an arbitrary random variable  $X$ , and assume from now on

$$P_{X_0} = (1 - p)\delta_0 + pP_{Y_0},$$

where  $\delta_0$  is the Dirac measure at the origin,  $Y_0$  is a random variable taking values in  $(0, \infty)$ , and  $p \in [0, 1]$  a parameter. In Example 1.1, we have  $Y_0 = 2$ .

We often write  $F_\infty(p)$  instead of  $F_\infty$  in order to make appear the dependence of the free energy in terms of the parameter  $p$ . Clearly  $p \mapsto F_\infty(p)$  is non-decreasing. So there exists a critical parameter  $p_c \in [0, 1]$  such that

$$F_\infty(p) > 0 \text{ if } p > p_c, \quad F_\infty(p) = 0 \text{ if } p < p_c.$$

[The extreme cases:  $p_c = 0$  means  $F_\infty(p) > 0$  for all  $p > 0$ , whereas  $p_c = 1$  means  $F_\infty(p) = 0$  for all  $p < 1$ .]

We can draw the first  $n$  generations of the rooted binary tree leading to the random variable  $X_n$ ; in this sense,  $F_\infty(p)$  can be viewed as a kind of percolation function on the binary tree: when  $F_\infty > 0$ , we say there is percolation, whereas if  $F_\infty = 0$ , we say there is no percolation. From this point of view, two questions are fundamental: (1) What is the critical value

$p_c$ ? (2) What is the behaviour of the free energy  $F_\infty(p)$  when  $p$  is in the neighbourhood of  $p_c$ ?

Concerning the first question, the value of  $p_c$  can be determined if the random variable  $Y_0$  is integer-valued.

**Theorem A (Collet et al. [8]).** Assume  $Y_0$  takes values in  $\{1, 2, \dots\}$ . Then

$$p_c = \frac{1}{\mathbb{E}[(Y_0 - 1)2^{Y_0}] + 1}.$$

Theorem A is proved in [8] assuming  $\mathbb{E}(Y_0 2^{Y_0}) < \infty$ . It is easily seen that it still holds in the case  $\mathbb{E}(Y_0 2^{Y_0}) = \infty$ : Indeed, for  $Z_0 := \min\{Y_0, k\}$  in the place of  $Y_0$ , the corresponding critical value for  $p$  is  $\frac{1}{\mathbb{E}[(Z_0 - 1)2^{Z_0}] + 1}$ , which can be made as close to 0 as possible by choosing  $k$  sufficiently large (by the monotone convergence theorem), so  $p_c = 0$ .

When  $Y_0$  is not integer-valued, Theorem A is not valid any more. The value of  $p_c$  is unknown (see Section 6 for some open problems). However, it is possible to characterise the positivity of  $p_c$ .

**Proposition 1.2.** *We have  $p_c > 0$  if and only if  $\mathbb{E}(Y_0 2^{Y_0}) < \infty$ .*

*Proof.* (1) We first assume that  $Y_0$  takes values in  $\{0, 1, 2, \dots\}$ .

By Theorem A,

$$p_c = \frac{1}{(\mathbb{E}[(Y_0 - 1)2^{Y_0}]^+ + 1)},$$

where  $(\mathbb{E}[(Y_0 - 1)2^{Y_0}]^+)$  is the positive part of  $\mathbb{E}[(Y_0 - 1)2^{Y_0}]$ . This means that if  $\mathbb{E}(Y_0 2^{Y_0}) < \infty$ , then  $p_c = \frac{1}{(\mathbb{E}[(Y_0 - 1)2^{Y_0}]^+ + 1)}$  (which, in particular, is positive), whereas if  $\mathbb{E}(Y_0 2^{Y_0}) = \infty$ , then  $p_c = 0$ .

(2) We now remove the assumption that  $Y_0$  is integer-valued. We write

$$\lfloor Y_0 \rfloor \leq Y_0 \leq \lceil Y_0 \rceil.$$

For both  $\lfloor Y_0 \rfloor$  and  $\lceil Y_0 \rceil$ , we apply the positivity criterion proved in the first step. Since the three conditions  $\mathbb{E}(\lfloor Y_0 \rfloor 2^{\lfloor Y_0 \rfloor}) < \infty$ ,  $\mathbb{E}(\lceil Y_0 \rceil 2^{\lceil Y_0 \rceil}) < \infty$  and  $\mathbb{E}(Y_0 2^{Y_0}) < \infty$  are equivalent, the desired result follows.  $\square$

Proposition 1.2 tells us that the positivity of  $p_c$  does not depend on the exact distribution of  $Y_0$ , but only on its tail behaviour.

We now turn our attention to the second question. For the standard Bernoulli bond percolation problem, the percolation function (i.e., the probability that the origin belongs to the unique infinite cluster) is continuous, but not differentiable, at  $p = p_c$ . For our model, the situation is believed to be very different; in fact, it is predicted ([11]) that the free energy is smooth at  $p = p_c$  and that all the derivatives at  $p_c$  vanish:

**Conjecture 1.3. (Derrida and Retaux [11]).** *Assume  $p_c > 0$ . There exists a constant  $K \in (0, \infty)$  such that*

$$F_\infty(p) = \exp\left(-\frac{K + o(1)}{(p - p_c)^{1/2}}\right), \quad p \downarrow p_c.$$

By Proposition 1.2, the assumption  $p_c > 0$  in Conjecture 1.3 means  $\mathbb{E}(Y_0 2^{Y_0}) < \infty$ .

We have not been able to prove the conjecture. Our aim is to study the influence, on the behaviour of  $F_\infty$  near  $p_c$ , produced by the tail behaviour of  $Y_0$ . It turns out that our main result can be applied to a more general family of recursive models, which we define in the following paragraph.

## 1.2 A generalised max-type recursive model

Let  $\nu$  be a random variable taking values in  $\{1, 2, \dots\}$ , such that  $m := \mathbf{E}(\nu) \in (1, \infty)$ . [We use  $\mathbf{E}$  to denote expectation with respect to the law of  $\nu$ .] For all  $n \geq 1$ , let

$$(1.3) \quad X_{n+1} \stackrel{\text{law}}{=} (X_{n,1} + \dots + X_{n,\nu} - 1)^+,$$

where  $X_{n,1}, X_{n,2}, \dots$  are independent copies of  $X_n$ , and are independent of  $\nu$ . From probabilistic point of view, this is a natural Galton–Watson-type extension of the model in (1.1), which corresponds to the special case  $\nu = 2$  a.s. We do not know whether there would be a physical interpretation of the

randomness of  $\nu$  (including in the related models described in Section 1.3 below), as asked by an anonymous referee.

Let  $\theta > 0$ . Let us consider the following situation: There exist constants  $0 < c_1 \leq c_2 < \infty$  such that for all sufficiently large  $x$ ,

$$(1.4) \quad c_1 e^{-\theta x} \leq \mathbb{P}(Y_0 \geq x) \leq c_2 e^{-\theta x}.$$

When  $\theta > \log m$ , we have  $p_c > 0$  (see Remark 3.1; this is in agreement to Proposition 1.2 if  $\nu$  is deterministic); the behaviour of the system in this case is predicted by Conjecture 1.3. We are interested in the case  $\theta \in (0, \log m]$ .

**Theorem 1.4.** *Assume  $\mathbf{E}(t^\nu) < \infty$  for some  $t > 1$ , and  $m := \mathbf{E}(\nu) > 1$ . Let  $\theta \in (0, \log m)$ . Under the assumption (1.4), we have*

$$F_\infty(p) = p^{\beta+o(1)}, \quad p \downarrow 0,$$

where  $\beta = \beta(\theta) := \frac{\log m}{(\log m) - \theta}$ .

Theorem 1.4, which is not deep, is included in the paper for the sake of completeness. Its analogue in the non-hierarchical pinning setting was known; see [24].

The study of the case  $\theta = \log m$  is the main concern of the paper. It turns out that we are able to say more. Fix  $\alpha \in \mathbb{R}$ . We assume the existence of constants  $0 < c_3 \leq c_4 < \infty$  such that for all sufficiently large  $x$ ,

$$(1.5) \quad c_3 x^\alpha m^{-x} \leq \mathbb{P}(Y_0 \geq x) \leq c_4 x^\alpha m^{-x}.$$

The main result of the paper is as follows.

**Theorem 1.5.** *Let  $\alpha > -2$ . Assume  $\mathbf{E}(t^\nu) < \infty$  for some  $t > 1$ , and  $m := \mathbf{E}(\nu) > 1$ . Under the assumption (1.5), we have*

$$F_\infty(p) = \exp\left(-\frac{1}{p^{\chi+o(1)}}\right), \quad p \downarrow p_c = 0,$$

where  $\chi = \chi(\alpha) := \frac{1}{\alpha+2}$ .

Compared to the original Derrida–Retaux model, additional technical difficulties may appear when  $\nu$  is random. For example, the analogue of the fundamental Theorem A is not known (see Problem 6.3).

The proof of the theorem gives slightly more precision: There exists a constant  $c_5 > 0$  such that for all sufficiently small  $p > 0$ ,

$$F_\infty(p) \leq \exp\left(-\frac{c_5}{p^x}\right).$$

We will regularly use the following elementary inequalities:

$$(1.6) \quad \frac{\mathbb{E}(X_n) - \frac{1}{m-1}}{m^n} \leq F_\infty \leq \frac{\mathbb{E}(X_n)}{m^n}, \quad n \geq 1.$$

The second inequality follows from (1.2). For the first inequality, it suffices to note that by definition,  $\mathbb{E}(X_{n+1}) \geq m\mathbb{E}(X_n) - 1$ , so  $n \mapsto \frac{\mathbb{E}(X_n) - \frac{1}{m-1}}{m^n}$  is non-decreasing, and  $F_\infty = \lim_{n \rightarrow \infty} \uparrow \frac{\mathbb{E}(X_n) - \frac{1}{m-1}}{m^n}$ .

An immediate consequence of (1.6) is the following dichotomy:

- either  $\mathbb{E}(X_n) > \frac{1}{m-1}$  for some  $n \geq 1$ , in which case  $F_\infty > 0$ ;
- or  $\mathbb{E}(X_n) \leq \frac{1}{m-1}$  for all  $n \geq 1$ , in which case  $F_\infty = 0$ .

### 1.3 About the Derrida–Retaux model

The Derrida–Retaux model studied in our paper has appeared in several places in both mathematics and physics literatures.

(a) The recursion in (1.1) belongs to a family of max-type recursive models analysed in the survey paper of Aldous and Bandyopadhyay [1].

(b) The model in (1.1) was investigated by Derrida and Retaux [11] to understand the nature of the pinning/depinning transition on a defect line in presence of strong disorder. The problem of the depinning transition has attracted much attention among mathematicians [2, 3, 5, 9, 20, 13, 14, 15, 16, 17, 18, 24, 31, 32] and physicists [10, 12, 13, 22, 27, 28, 30] over the last thirty years. Much progress has been made in understanding the question of the relevance of a weak disorder [14, 20, 17], i.e., whether a weak disorder is susceptible of modifying the nature of this depinning transition. For strong

disorder or even, for a weak disorder when disorder is relevant, it is known that the transition should always be smooth [18], but the precise nature of the transition is still controversial [30, 11, 27].

It is expected that a similar phase transition should occur in a simplified version of the problem, when the line is constrained to a hierarchical geometry [6, 10, 16, 23]. Even in this hierarchical version, the nature of the transition is poorly understood. This is why Derrida and Retaux [11] came up with a toy model which, they argue, should behave like the hierarchical model. This toy model turns out to be sufficiently complicated that many fundamental questions remain open (we include a final section discussing some of these open problems in Section 6).

(c) The model in (1.1) has also appeared in Collet et al. [8] in their study of spin glass model.

(d) The recursion in (1.1) has led to the so-called parking schema; see Goldschmidt and Przykucki [19].

The rest of the paper and the proofs of the theorems are as follows. In Section 2, we present a (heuristic) outline the proof of Theorem 1.5. Section 3 is devoted to the upper bounds in Theorems 1.4 and 1.5. In Section 4, which is the heart of the paper, we prove the lower bound in Theorem 1.5. The lower bound in Theorem 1.4 is proved in Section 5. Finally, we make some additional discussions and present several open problems in Section 6.

## 2 Proof of Theorem 1.5: an outline

The upper bound in Theorem 1.5, proved in details in Section 3, relies on a simple analysis of the moment generating function. The idea of using the moment generating function goes back to Collet et al. [8] and Derrida and Retaux [11].

The proof of the lower bound in Theorem 1.5, quite involving and based on a multi-scale type of argument, is done in two steps. The first step consists of the following estimate: if the initial distribution  $X_0$  satisfies, for  $t \geq 1$



(say),

$$\mathbb{P}(X_0 \geq t) \text{ “}\geq\text{” } p t^\alpha m^{-t},$$

then

$$(2.1) \quad \mathbb{P}(X_n \geq t) \text{ “}\geq\text{” } \tilde{p} t^\alpha m^{-t},$$

for  $n \geq 1$  and  $t \geq 1$ , where

$$\tilde{p} := p^2 n^{2+\alpha}.$$

[See Lemma 4.1 for a rigorous statement; for a heuristic explanation of (2.1), see below.] This allows us to use inductively the estimate, to arrive at:

$$\mathbb{P}(X_j \geq t) \text{ “}\geq\text{” } p^a j^b t^\alpha m^{-t},$$

for all  $j$  satisfying  $p^a j^b \leq 1$ , with  $a > 0$  and  $b > 0$  such that  $b \approx (2 + \alpha)a$ . By  $\mathbb{E}(X_j) = \int_0^\infty \mathbb{P}(X_j \geq t) dt$ , we get

$$\mathbb{E}(X_j) \text{ “}\geq\text{” } \kappa p^a j^b,$$

where  $\kappa > 0$  is a (small) constant. [This is Lemma 4.2.] As such,  $\mathbb{E}(X_j) \geq \kappa$ , if  $j = p^{-a/b}$ . We are almost home. The rest of the argument consists in replacing  $\kappa$  by a constant greater than  $\frac{1}{m-1}$ . This is done in the second step.

Let  $n \geq 1$ . To see why (2.1) is true, we use a hierarchical representation of the system  $(X_i, 0 \leq i \leq n)$ , due to Collet et al. [7] and Derrida and Reataux [11]. We define a family of random variables  $(X(u), u \in \mathbb{T}^{(n)})$ , indexed by a reversed Galton–Watson tree  $\mathbb{T}^{(n)}$ . Let  $\mathbb{T}_0 = \mathbb{T}_0^{(n)}$  denote the initial generation of  $\mathbb{T}^{(n)}$ . We assume that  $X(u)$ , for  $u \in \mathbb{T}_0$ , are i.i.d. having the distribution of  $X_0$ . For any vertex  $u \in \mathbb{T}^{(n)} \setminus \mathbb{T}_0$ , we set

$$X(u) := (X(u^{(1)}) + \dots + X(u^{(\nu_u)}) - 1)^+,$$

where  $u^{(1)}, \dots, u^{(\nu_u)}$  are the parents of  $u$ . Consider (see (4.6);  $\lambda_1$  being an unimportant constant, and can be taken as  $\frac{1}{3}$ )

$$Z = Z_n \approx \#\{u \in \mathbb{T}_0 : X(u) \geq \lambda_1 n, \mathcal{D}(u)\},$$

where  $\mathcal{D}(u) := \{\exists v \in \mathbb{T}_0 \setminus \{u\}, |u \wedge v| < \lambda_1 n, X(v) \geq t + n - X(u)\}$ , and the symbol “#” denotes cardinality. Here, by  $|u \wedge v| < \lambda_1 n$ , we mean  $u$  and  $v$  have a common descendant before generation  $\lambda_1 n$ . We observe that  $\mathbb{P}(X_n \geq t) \geq \mathbb{P}(Z \geq 1)$ . So by the Cauchy–Schwarz inequality,

$$\mathbb{P}(X_n \geq t) \geq \frac{[\mathbb{E}(Z)]^2}{\mathbb{E}(Z^2)}.$$

The proof of (2.1) is done by proving that  $\mathbb{E}(Z) \approx p^2 n^{2+\alpha} t^\alpha m^{-t}$  (see (4.13)) and that  $\mathbb{E}(Z^2)$  “ $\leq$ ”  $\mathbb{E}(Z)$ .

In the second step, we take  $n \approx (\frac{1}{p})^{1/(2+\alpha)}$ , defined rigorously in (4.24). We use once again the hierarchical representation. Let  $u \in \mathbb{T}_0$  be such that  $X(u) = \max_{v \in \mathbb{T}_0} X(v)$ . Since the values of  $X$  in the initial generation are i.i.d. of the law of  $X_0$ , it is elementary that  $X(u) \geq \ell$  “ $:=$ ”  $n - \log n$  (see (4.23) for the exact value of  $\ell$ ). Let  $k := \frac{\ell}{4}$ . The fact  $k < \ell$  allows us to use the following inequality:

$$X_n \geq X(u) - n + \sum_{j=0}^{k-1} X_j^{(n)} \geq \ell - n + \sum_{j=0}^{k-1} X_j^{(n)},$$

where, for each  $j$ ,  $X_j^{(n)}$  has “approximately” the same law as  $X_j$ . [For a rigorous formulation, see (4.20).] Taking expectation on both sides, and using the fact, proved in the first step, that  $\mathbb{E}(X_j^{(n)})$  is greater than a constant (say  $\kappa_1$ ) if  $j \geq \frac{k}{2}$  (say), we arrive at:

$$\mathbb{E}(X_n) \geq \ell - n + \frac{k}{2} \kappa_1,$$

which is greater than a constant multiple of  $n$ . By the first inequality in (1.6),  $F_\infty \geq \frac{\mathbb{E}(X_n) - \frac{1}{m-1}}{m^n}$ , which implies the desired lower bound in Theorem 1.5.  $\square$

### 3 Upper bounds

Without loss of generality, we assume  $X_0$  is integer valued (otherwise, we consider  $\lceil X_0 \rceil$ ). Consider the generating functions

$$G_n(s) := \mathbb{E}(s^{X_n}), \quad h(s) := \mathbf{E}(s^\nu),$$

where  $\nu$  is the number of independent copies in the convolution relation (1.3):  
 $X_{n+1} \stackrel{\text{law}}{=} (X_{n,1} + \cdots + X_{n,\nu} - 1)^+$ .

The latter can be written as

$$G_{n+1}(s) = \frac{h(G_n(s))}{s} + \frac{s-1}{s} h(G_n(0)).$$

Hence

$$G'_{n+1}(s) = \frac{h'(G_n(s))}{s} G'_n(s) - \frac{h(G_n(s)) - h(G_n(0))}{s^2}.$$

We fix an  $s \in (1, m)$  whose value will be determined later. Write

$$a_n = a_n(s) := G_n(s) - 1.$$

Since  $h(G_n(0)) \leq h(1) = 1$ , we have  $h(G_n(s)) - h(G_n(0)) \geq h(G_n(s)) - h(1) = h(1 + a_n) - h(1) \geq h'(1)a_n = m a_n \geq s a_n$ . Hence

$$G'_{n+1}(s) \leq \frac{h'(G_n(s))}{s} G'_n(s) - \frac{a_n}{s} = \frac{h'(1 + a_n)}{s} G'_n(s) - \frac{a_n}{s}.$$

By the assumption of existence of  $t > 1$  satisfying  $\mathbf{E}(t^\nu) < \infty$ , there exist  $\delta_0 > 0$  and  $c_0 > 0$  such that  $h'(1 + a) \leq h'(1) + c_0 a = m + c_0 a$  for  $a \in (0, \delta_0)$ . Hence, if  $a_n < \delta_0$ , then

$$G'_{n+1}(s) \leq \frac{m + c_0 a_n}{s} G'_n(s) - \frac{a_n}{s}.$$

Let  $N_1 := \inf\{n \geq 0 : a_n \geq \delta_0 \text{ or } G'_n(s) \geq \frac{1}{c_0}\}$  (with  $\inf \emptyset := \infty$ ). As long as  $a_0 < \delta_0$  and  $G'_0(s) < \frac{1}{c_0}$  (which we take for granted from now on), we have

$$G'_{n+1}(s) \leq \frac{m + c_0 a_n}{s} G'_n(s) - \frac{a_n}{s} \leq \frac{m}{s} G'_n(s), \quad 0 \leq n < N_1.$$

Iterating the inequality, we get that

$$(3.1) \quad G'_n(s) \leq \left(\frac{m}{s}\right)^n G'_0(s), \quad 1 \leq n \leq N_1.$$

We now proceed to the proof of the upper bound in Theorem 1.5. By assumption (1.5),  $\mathbb{P}(Y_0 \geq x) \leq c_4 x^\alpha m^{-x}$  for all sufficiently large  $x$ . This

implies, by integration by parts, the existence of a constant  $c_6 > 0$  such that for all  $s \in (1, m)$ ,

$$\mathbb{E}(Y_0 s^{Y_0-1}) \leq \frac{c_6}{(\log m - \log s)^{\alpha+2}}.$$

By definition of  $G_0$ , this yields

$$(3.2) \quad G'_0(s) = \mathbb{E}(X_0 s^{X_0-1}) = p \mathbb{E}(Y_0 s^{Y_0-1}) \leq \frac{c_6 p}{(\log m - \log s)^{\alpha+2}},$$

and for  $n \geq 0$ ,

$$a_n := G_n(s) - 1 \leq \mathbb{E}(s^{X_n} \mathbf{1}_{\{X_n \geq 1\}}) \leq s G'_n(s) \leq m G'_n(s).$$

We choose  $s := m e^{-1/N}$ , where  $N := (c_7 p)^{-1/(2+\alpha)}$ , with  $c_7$  denoting a large constant such that  $e \frac{c_6}{c_7} < \frac{1}{c_0}$  and that  $e m \frac{c_6}{c_7} < \delta_0$ . Then (3.2) ensures that  $G'_0(s) \leq \frac{c_6}{c_7} < \frac{1}{c_0}$ , and  $a_0 \leq m G'_0(s) < \delta_0$ . [In fact,  $a_0$  is *much* smaller than  $m G'_0(s)$ .]

By (3.1), for  $1 \leq n \leq N_1$  (recalling that  $G'_0(s) \leq \frac{c_6}{c_7}$ )

$$G'_n(s) \leq e^{n/N} G'_0(s) \leq e^{n/N} \frac{c_6}{c_7},$$

and

$$a_n \leq m G'_n(s) \leq e^{n/N} \frac{m c_6}{c_7}.$$

Since  $e \frac{c_6}{c_7} < \frac{1}{c_0}$  and  $e m \frac{c_6}{c_7} < \delta_0$  by the choice of  $c_7$ , this yields  $G'_n(s) < \frac{1}{c_0}$  and  $a_n < \delta_0$  for  $1 \leq n \leq \min\{N_1, N\}$ . By definition of  $N_1$ , this implies  $N_1 \geq N$ ; hence  $\min\{N_1, N\} = N$ , which implies  $a_N < \delta_0$ .

By Jensen's inequality,  $\mathbb{E}(X_n) \leq \frac{\log(1+a_n)}{\log s}$ . So  $\mathbb{E}(X_N) \leq \frac{\log(1+\delta_0)}{\log s}$ . In view of the second inequality in (1.6), we get, for all sufficiently small  $p$ ,

$$F_\infty(p) \leq \frac{\mathbb{E}(X_N)}{m^N} \leq \frac{\log(1+\delta_0)}{m^N \log s},$$

proving the upper bound in Theorem 1.5.

The upper bound in Theorem 1.4 is obtained similarly. We choose  $s := e^{\theta-\varepsilon}$  with  $\varepsilon := (\log \frac{1}{p})^{-1}$ . Then

$$G'_0(s) = p \mathbb{E}(Y_0 s^{Y_0-1}) \leq c_8 \frac{p}{\varepsilon^2},$$

for some constant  $c_8 > 0$  and all sufficiently small  $p > 0$ , and  $a_0 \leq m G'_0(s)$ . In particular,  $G'_0(s) < \frac{1}{c_0}$  and  $a_0 < \delta_0$  for all sufficiently small  $p > 0$ . By (3.1), for  $1 \leq n \leq N_1$ ,

$$G'_n(s) \leq G'_0(s) \left( \frac{m}{e^{\theta-\varepsilon}} \right)^n \leq c_8 \frac{p}{\varepsilon^2} \left( \frac{m}{e^{\theta-\varepsilon}} \right)^n,$$

and

$$a_n \leq m G'_n(s) \leq m c_8 \frac{p}{\varepsilon^2} \left( \frac{m}{e^{\theta-\varepsilon}} \right)^n.$$

Let  $N' := \frac{\log(c_9 \varepsilon^2/p)}{\log(\frac{m}{e^{\theta-\varepsilon}})}$ , where  $c_9 > 0$  is a small constant such that  $c_8 \frac{p}{\varepsilon^2} \left( \frac{m}{e^{\theta-\varepsilon}} \right)^{N'} < \frac{1}{c_0}$  and that  $m c_8 \frac{p}{\varepsilon^2} \left( \frac{m}{e^{\theta-\varepsilon}} \right)^{N'} < \delta_0$ . Then  $G'_n(s) < \frac{1}{c_0}$  and  $a_n < \delta_0$  for  $1 \leq n \leq \min\{N_1, N'\}$ . This implies  $a_{N'} < \delta_0$ . Since  $N' = (1 + o(1)) \frac{\log(1/p)}{(\log m) - \theta}$  (for  $p \rightarrow 0$ ), we get that  $F_\infty(p) \leq \frac{\mathbb{E}(X_{N'})}{m^{N'}} \leq \frac{\log(1+a_{N'})}{m^{N'} \log s} \leq p^{\frac{\log m}{(\log m) - \theta} + o(1)}$ ,  $p \rightarrow 0$ , proving the upper bound in Theorem 1.4.  $\square$

**Remark 3.1.** Let  $a_n = a_n(s) := G_n(s) - 1$  as in the proof. Since  $h(G_n(0)) \leq h(1) = 1$ , we have

$$a_{n+1} = G_{n+1}(s) - 1 \leq \frac{h(1 + a_n) - 1}{s}.$$

By assumption on  $\nu$ , there exist  $\delta'_0 > 0$  and  $c'_0 > 0$  such that  $h(1 + a) - 1 \leq ma(1 + c'_0 a)$  for  $a \in (0, \delta'_0)$ . Consequently, if  $0 < a_0 < \delta'_0$  and  $m(1 + c'_0 a_0) < s$ , then inductively for all  $n \geq 0$ ,

$$a_{n+1} \leq \frac{ma_n(1 + c'_0 a_n)}{s} < a_n.$$

In other words, the sequence  $a_n$ ,  $n \geq 1$ , is decreasing.

Assume  $\mathbb{P}(Y_0 > x) \leq c_2 e^{-\theta x}$  for some  $\theta > \log m$  and all sufficiently large  $x$ . Fix  $s \in (m, e^\theta)$ . We have  $\mathbb{E}(s^{Y_0}) < \infty$ . So for sufficiently small  $p > 0$ , we have  $0 < a_0 < \delta'_0$  and  $m(1 + c'_0 a_0) \leq s$ . By the discussions in the last paragraph, the sequence  $a_n$ ,  $n \geq 1$ , is decreasing. This yields  $\sup_{n \geq 0} \mathbb{E}(X_n) < \infty$ ; hence  $F_\infty(p) = 0$ . Consequently,  $p_c > 0$  in this case.

[The discussion here gives a sufficient condition for the positivity of  $p_c$ , not a characterisation.]  $\square$

## 4 Proof of Theorem 1.5: lower bound

Throughout the section, we assume  $\mathbf{E}(\nu^3) < \infty$  and  $m := \mathbf{E}(\nu) > 1$ , which is weaker than the assumption in Theorem 1.5.

We start with a simple hierarchical representation of the system; the idea of this representation already appeared in Collet et al. [7] and in Derrida and Retaux [11].

We define a family of random variables  $(X(u), u \in \mathbb{T})$ , indexed by an infinite tree  $\mathbb{T}$ , in the following way. For any vertex  $u$  in the genealogical tree  $\mathbb{T}$ , we use  $|u|$  to denote the generation of  $u$ ; so  $|u| = 0$  if  $u$  is in the initial generation. We assume that  $X(u)$ , for  $u \in \mathbb{T}$  with  $|u| = 0$  (i.e., in the initial generation of the system), are i.i.d. having the distribution of  $X_0$ . For any vertex  $u \in \mathbb{T}$  with  $|u| \geq 1$ , let  $u^{(1)}, \dots, u^{(\nu_u)}$  denote the  $\nu_u$  parents of  $u$  in generation  $|u| - 1$ , and set

$$X(u) := (X(u^{(1)}) + \dots + X(u^{(\nu_u)}) - 1)^+.$$

We assume that for any  $i \geq 0$ ,  $(\nu_u, |u| = i)$  are i.i.d. having the distribution of  $\nu$ , and are independent of everything else up to generation  $i$ .

Fix an arbitrary vertex  $\mathbf{e}_n$  in the  $n$ -th generation. The set of all vertices, including  $\mathbf{e}_n$  itself, in the first  $n$  generations of  $\mathbb{T}$  having  $\mathbf{e}_n$  as their (unique) descendant at generation  $n$ , is denoted by  $\mathbb{T}^{(n)}$ . Clearly,  $\mathbb{T}^{(n)}$  is (the reverse of the first  $n$  generations of) a Galton–Watson tree, rooted at  $\mathbf{e}_n$ , with reproduction distribution  $\nu$ . Note that  $X(\mathbf{e}_n)$  has the distribution of  $X_n$ .

More generally, for  $v \in \mathbb{T}^{(n)}$  with  $|v| = j \leq n$ , let  $\mathbb{T}(v)$  denote the set of all vertices, including  $v$  itself, in the first  $j$  generations of  $\mathbb{T}$  having  $v$  as their (unique) descendant at generation  $j$ . [So  $\mathbb{T}^{(n)} = \mathbb{T}(\mathbf{e}_n)$ .] Let  $\mathbb{T}_0(v) := \{x \in \mathbb{T}(v) : |x| = 0\}$ . By an abuse of notation, we write  $\mathbb{T}_0 := \mathbb{T}_0(\mathbf{e}_n)$ .

Let  $v \in \mathbb{T}^{(n)}$ . Let  $v_{|v|} = v$ , and for  $|v| < i \leq n$ , let  $v_i$  be the (unique) child of  $v_{i-1}$ ; in particular,  $|v_i| = i$  (for  $|v| \leq i \leq n$ ), and  $v_n = \mathbf{e}_n$ . See Figure 1.

For  $v \in \mathbb{T}^{(n)} \setminus \{\mathbf{e}_n\}$ , let  $\mathbf{bro}(v)$  denote the set of the brothers of  $v$ , i.e., the set of vertices, different from  $v$ , that are in generation  $|v|$  and having the same child as  $v$ . Note that  $\mathbf{bro}(v)$  can be possibly empty.<sup>3</sup>

---

<sup>3</sup>The term “brothers” is with respect to  $\mathbb{T}^{(n)}$ , which is *reversed* Galton–Watson.

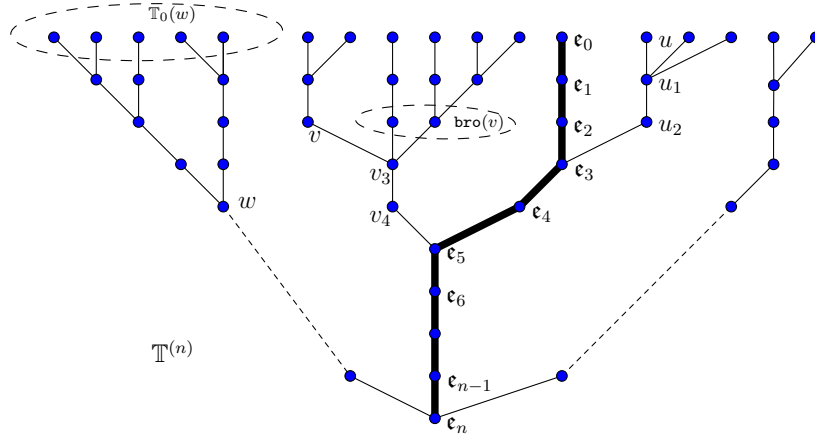


Figure 1: Spine  $(\epsilon_i)_{0 \leq i \leq n}$ ; in the above picture,  $v_2 := v$  and  $v_i = \epsilon_i$  for all  $5 \leq i \leq n$ .

In the rest of the paper, we use  $P_\omega$  the conditional probability given  $\mathbb{T}$ , and its corresponding expectation  $E_\omega$ . The law of  $\mathbb{T}$  is denoted by  $\mathbf{P}$ , the corresponding expectation  $\mathbf{E}$ . We write  $\mathbb{P}(\cdot) := \mathbf{E}[P_\omega(\cdot)]$ , with corresponding expectation  $\mathbb{E}$ .

We now describe the law of the *size-biased Galton–Watson tree*. Let  $\mathbf{Q}$  be the probability measure defined on  $\sigma(\mathbb{T}^{(n)})$ , the sigma-field generated by  $\mathbb{T}^{(n)}$ , by  $\mathbf{Q} := \frac{\#\mathbb{T}_0}{m^n} \bullet \mathbf{P}$ . Under  $\mathbf{Q}$ ,  $\mathbb{T}^{(n)}$  represents (the first  $n$  generations of) a so-called size-biased Galton–Watson tree. There is a simple way to describe the law of the size-biased Galton–Watson tree. Let  $\epsilon_0 = \epsilon_0^{(n)}$  be a random variable taking values in  $\mathbb{T}_0$  (which is not measurable with respect to  $\sigma(\mathbb{T}_0)$ , the sigma-field generated by  $\mathbb{T}_0$ ) whose under  $\mathbf{Q}$ , given  $\sigma(\mathbb{T}_0)$ , is uniformly distributed on  $\mathbb{T}_0$ :  $\mathbf{Q}(\epsilon_0 = u | \mathbb{T}_0) = \frac{1}{\#\mathbb{T}_0}$  for any  $u \in \mathbb{T}_0$ . Let  $\epsilon_i = \epsilon_i^{(n)}$  be the unique descendant at generation  $i$  of  $\epsilon_0$ , for all  $0 \leq i \leq n$ . The collection  $(\epsilon_i, 0 \leq i \leq n)$  is referred to as the spine. The *spinal decomposition theorem* says that under  $\mathbf{Q}$ ,  $\text{bro}(\epsilon_i)$ , for  $0 \leq i \leq n-1$ , are i.i.d., and conditionally on  $\mathcal{G}_n := \sigma(\text{bro}(\epsilon_i), 0 \leq i \leq n-1)$ ,  $(\mathbb{T}_0(v), v \in \text{bro}(\epsilon_i))_{0 \leq i \leq n-1}$  are independent; furthermore, for each  $0 \leq i \leq n-1$ , conditionally on  $\mathcal{G}_n$  and under  $\mathbf{Q}$ ,  $\mathbb{T}_0(v)$ , for  $v \in \text{bro}(\epsilon_i)$  are i.i.d. having the law of  $\mathbb{T}_0(u_i)$  under  $\mathbf{P}$  (for any  $u \in \mathbb{T}_0$  which is  $\sigma(\mathbb{T}_0)$ -measurable). For more details, see Lyons, Pemantle

and Peres [25], or Lyons and Peres ([26], Chap. 12), Shi ([29], Chap. 2) for formalism for tree-valued random variables.

A useful consequence of the spinal decomposition theorem is the many-to-one formula: For any measurable function  $g$ ,

$$(4.1) \quad \mathbf{E} \left( \sum_{u \in \mathbb{T}_0} g(\mathbb{T}(u_1), \dots, \mathbb{T}(u_n)) \right) = m^n \mathbf{E}_{\mathbf{Q}}[g(\mathbb{T}(\mathbf{e}_1), \dots, \mathbb{T}(\mathbf{e}_n))],$$

where  $(\mathbf{e}_i, 0 \leq i \leq n)$  is the spine.

Here is another consequence of the spinal decomposition theorem. Let  $1 \leq i \leq n$ . We have  $\mathbf{E}_{\mathbf{Q}}[\#\mathbb{T}_0(\mathbf{e}_i)] = \mathbf{E}_{\mathbf{Q}}[\#\mathbb{T}_0(\mathbf{e}_{i-1})] + c_{10} m^{i-1}$ , where  $c_{10} := \mathbf{E}_{\mathbf{Q}}[\#\text{bro}(\mathbf{e}_{i-1})] = \mathbf{E}_{\mathbf{Q}}[\#\text{bro}(\mathbf{e}_0)] = \frac{1}{m} \sum_{k=1}^{\infty} k(k-1) \mathbf{P}(\nu = k) < \infty$ . This yields

$$(4.2) \quad \mathbf{E}_{\mathbf{Q}}[\#\mathbb{T}_0(\mathbf{e}_i)] = c_{10} \sum_{j=0}^{i-1} m^j + 1 \leq c_{11} m^i,$$

for some constant  $c_{11} > 0$  and all  $0 \leq i \leq n$ . Similarly, the assumption  $\mathbf{E}(\nu^3) < \infty$  ensuring  $\mathbf{E}_{\mathbf{Q}}[(\#\text{bro}(\mathbf{e}_0))^2] = \frac{1}{m} \sum_{k=1}^{\infty} k(k-1)^2 \mathbf{P}(\nu = k) < \infty$ , we have

$$(4.3) \quad \mathbf{E}_{\mathbf{Q}}\{[\#\mathbb{T}_0(\mathbf{e}_i)]^2\} \leq c_{12} m^{2i},$$

for some constant  $c_{12} > 0$  and all  $0 \leq i \leq n$ .

We now turn to the proof of the lower bound in Theorem 1.5, which is done in two steps. The first step, summarised in Lemma 4.1 below, is a probability estimate that allows for iteration. The second step says that along the spine,  $X_n$  will reach sufficiently high expected values.

## 4.1 First step: Inductive probability estimate

The first step gives a useful inductive probability estimate. In order to make the induction possible, we assume something more general than the assumption (1.5) in Theorem 1.5.



**Lemma 4.1.** *Assume  $\mathbf{E}(\nu^3) < \infty$  and  $m := \mathbf{E}(\nu) > 1$ . Let  $\alpha \in \mathbb{R}$ . Let  $c_{13} > 0$ ,  $c_{14} > 0$  and  $c_{15} > 0$ . There exists  $c > 0$  such that for  $0 < p < 1$  with  $pn^{1+\alpha} \leq c_{13}$ , if the initial distribution of  $X_0$  is such that for some  $1 \leq \gamma \leq c_{14}n$ ,*

$$(4.4) \quad \mathbb{P}(X_0 \geq t) \geq c_{15} p (t + \gamma)^\alpha m^{-t}, \quad \forall t > 0,$$

then

$$\mathbb{P}(X_n \geq t) \geq cp^2 n^{2+\alpha} (t + n)^\alpha m^{-t}, \quad \forall t > 0.$$

*Proof.* Without loss of generality, we assume  $X_0$  is integer valued such that

$$(4.5) \quad \mathbb{P}(X_0 = t) = c_{15} p (t + \gamma)^\alpha m^{-t}, \quad \forall t = 1, 2, \dots$$

[In fact, the distribution in (4.5) is stochastically smaller than or equal to a distribution satisfying (4.4), with a possibly different value of the constant  $c_{15}$ .]

For  $v \in \mathbb{T}^{(n)}$  with  $|v| = j \leq n$ , let

$$M(v) := \max_{r \in \mathbf{bro}(v)} \max_{w \in \mathbb{T}_0(r)} (X(w) - j)^+,$$

where  $r \in \mathbf{bro}(v)$  means, as before, that  $r$  is a brother of  $v$ , and  $X(w)$  is the random variable assigned to the vertex  $w$  on the initial generation. Let  $b > 0$ . Let  $0 < \lambda_1 < \lambda_2 < 1$  be any fixed constants.<sup>4</sup> We consider the integer-valued random variable

$$(4.6) \quad Z := \sum_{u \in \mathbb{T}_0} \mathbf{1}_{A_u},$$

where, for any  $u \in \mathbb{T}_0$ ,

$$A_u := \left\{ X(u) \in [\lambda_1 n, \lambda_2 n], \max_{1 \leq j \leq \lambda_1 n} M(u_j) \geq b + n - X(u) \right\}.$$

Clearly,

$$(4.7) \quad \mathbb{P}(Z \geq 1) \leq \mathbb{P}(X_n \geq b).$$

---

<sup>4</sup>The values of  $\lambda_1$  and  $\lambda_2$  play no significant role in the proof; so we can take, for example,  $\lambda_1 = \frac{1}{3}$  and  $\lambda_2 = \frac{2}{3}$ .

Throughout the proof, we write  $x \lesssim y$  or  $y \gtrsim x$  if  $x \leq cy$  for some constant  $c \in (0, \infty)$  that does not depend on  $(n, p, b)$ , and  $x \asymp y$  if both relations  $x \lesssim y$  and  $y \lesssim x$  hold.

For  $x \geq (1 - \lambda_2)n$  and  $1 \leq j \leq n$ , we have, by (4.5),

$$\mathbb{P}(X_0 \geq x + j) \asymp p(x + j + \gamma)^\alpha m^{-x-j} \asymp px^\alpha m^{-x-j}.$$

[So the parameter  $\gamma$  figuring in the condition (4.4) disappears because  $x + \gamma \asymp x$  if  $x \geq (1 - \lambda_2)n$ .] Note that  $M(u_j)$ , for  $1 \leq j \leq n$ , are independent under  $P_\omega$ . We have, for  $u \in \mathbb{T}_0$ ,  $x \geq (1 - \lambda_2)n$ , and integers  $1 \leq n_1 \leq n_2 \leq n$ ,

$$\begin{aligned} P_\omega \left( \max_{n_1 \leq j \leq n_2} M(u_j) \geq x \right) &= 1 - \prod_{j=n_1}^{n_2} P_\omega(M(u_j) < x) \\ &= 1 - \prod_{j=n_1}^{n_2} \prod_{r \in \text{bro}(u_j)} \prod_{w \in \mathbb{T}_0(r)} [1 - \mathbb{P}(X_0 \geq x + j)] \\ &\asymp \left[ px^\alpha \sum_{j=n_1}^{n_2} m^{-x-j} \Lambda(u_j) \right] \wedge 1, \end{aligned}$$

uniformly in  $1 \leq n_1 \leq n_2 \leq n$ . Here,  $a \wedge b := \min\{a, b\}$  for real numbers, and for all  $v \in \mathbb{T}^{(n)}$ ,

$$\Lambda(v) := \sum_{r \in \text{bro}(v)} \#\mathbb{T}_0(r),$$

with  $\#\mathbb{T}_0(r)$  denoting the cardinality of  $\mathbb{T}_0(r)$ . For future use, we observe that

$$(4.8) \quad P_\omega \left( \max_{n_1 \leq j \leq n_2} M(u_j) \geq x \right) \lesssim px^\alpha \sum_{j=n_1}^{n_2} m^{-x-j} \Lambda(u_j),$$

uniformly in  $1 \leq n_1 \leq n_2 \leq n$ . Taking  $n_1 := 1$  and  $n_2 := \lambda_1 n$ , and by independence of  $X(u)$  and  $\max_{n_1 \leq j \leq n_2} M(u_j)$  under  $P_\omega$ , we arrive at:<sup>5</sup>

$$(4.9) \quad P_\omega(A_u) \asymp \sum_{\ell=\lambda_1 n}^{\lambda_2 n} pn^\alpha m^{-\ell} \left\{ \left[ p(b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-(b+n-\ell)-j} \Lambda(u_j) \right] \wedge 1 \right\}.$$

---

<sup>5</sup>For notational simplification, we treat  $\lambda_1 n$  and  $\lambda_2 n$  as if they were integers.

[We have used  $\ell + \gamma \asymp n$  and  $b + n - \ell \asymp b + n$ .] For future use, we see that by removing “ $\wedge 1$ ” on the right-hand side,

$$(4.10) \quad P_\omega(A_u) \lesssim p^2 n^{1+\alpha} (b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-(b+n)-j} \Lambda(u_j).$$

We now estimate  $\mathbb{E}(Z)$  and  $\mathbb{E}(Z^2)$ .

We first look at the expectation of  $Z$  under  $P_\omega$ : By (4.9),

$$E_\omega(Z) \asymp \sum_{u \in \mathbb{T}_0} \sum_{\ell=\lambda_1 n}^{\lambda_2 n} p n^\alpha m^{-\ell} \left\{ \left[ p(b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-b-n+\ell-j} \Lambda(u_j) \right] \wedge 1 \right\}.$$

We take expectation on both sides with respect to  $\mathbf{P}$ , the law of  $\mathbb{T}$ . By the many-to-one formula (4.1),

$$(4.11) \quad \mathbb{E}(Z) \asymp m^n \sum_{\ell=\lambda_1 n}^{\lambda_2 n} p n^\alpha m^{-\ell} \mathbf{E}_{\mathbf{Q}}(\eta \wedge 1),$$

where

$$\eta = \eta(n, \ell) := p(b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-b-n+\ell} m^{-j} \Lambda(\mathbf{e}_j).$$

By the spinal decomposition theorem, under  $\mathbf{Q}$ ,  $m^{-j} \Lambda(\mathbf{e}_j)$ ,  $1 \leq j \leq n$  are independent, and for each  $j$ ,  $\#\mathbf{bro}(\mathbf{e}_j)$  has the same law as  $\#\mathbf{bro}(\mathbf{e}_1)$ , whereas conditionally on  $\#\mathbf{bro}(\mathbf{e}_j)$ ,  $\Lambda(\mathbf{e}_j)$  is distributed as the sum of  $\#\mathbf{bro}(\mathbf{e}_j)$  independent copies of  $\mathbb{T}_0(u_j)$  under  $\mathbf{P}$  (for any  $u \in \mathbb{T}_0$ ), the latter being the number of individuals in the  $j$ -th generation of a Galton–Watson process with reproduction law  $\nu$  (starting with 1 individual). Accordingly,

$$\mathbf{E}_{\mathbf{Q}}[m^{-j} \Lambda(\mathbf{e}_j) \mid \#\mathbf{bro}(\mathbf{e}_j)] = \#\mathbf{bro}(\mathbf{e}_j),$$

so that

$$\mathbf{E}_{\mathbf{Q}}[m^{-j} \Lambda(\mathbf{e}_j)] = \mathbf{E}_{\mathbf{Q}}[\#\mathbf{bro}(\mathbf{e}_j)] = \mathbf{E}_{\mathbf{Q}}[\#\mathbf{bro}(\mathbf{e}_1)] < \infty.$$

Hence

$$(4.12) \quad \mathbf{E}_{\mathbf{Q}}(\eta) \asymp p(b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-b-n+\ell} \asymp p(b+n)^\alpha n m^{-b-n+\ell}.$$

We now estimate  $\mathbf{E}_{\mathbf{Q}}(\eta \wedge 1)$ . Consider a Galton–Watson process with reproduction law  $\nu$  (starting with 1 individual) under  $\mathbf{P}$ . For each  $j \geq 0$ , let  $m^j W_j$  denote the number of individuals in the  $j$ -th generation. By Athreya and Ney [4] (p. 9, Theorem 2), as long as  $\nu$  has a finite second moment,  $(W_j, j \geq 0)$  is a martingale bounded in  $L^2$ ; in particular,  $W_j$  converges in  $L^2$ , when  $j \rightarrow \infty$ , to a limit denoted by  $W$ . For any  $s > 0$ ,

$$\mathbf{E}_{\mathbf{Q}}(e^{-sm^{-j}\Lambda(\epsilon_j)} \mid \#\mathbf{bro}(\epsilon_j)) = [\mathbf{E}(e^{-sW_j})]^{\#\mathbf{bro}(\epsilon_j)},$$

so that

$$\mathbf{E}_{\mathbf{Q}}(e^{-sm^{-j}\Lambda(\epsilon_j)}) = \mathbf{E}_{\mathbf{Q}}\left\{[\mathbf{E}(e^{-sW_j})]^{\#\mathbf{bro}(\epsilon_j)}\right\} = \mathbf{E}_{\mathbf{Q}}\left\{[\mathbf{E}(e^{-sW_j})]^{\#\mathbf{bro}(\epsilon_0)}\right\}.$$

By conditional Jensen’s inequality,  $e^{-sW_j} = e^{-s\mathbf{E}(W \mid W_j)} \leq \mathbf{E}(e^{-sW} \mid W_j)$ , so  $\mathbf{E}(e^{-sW_j}) \leq \mathbf{E}(e^{-sW})$ . Since  $\mathbf{E}(W) = 1$ , we have  $1 - \mathbf{E}(e^{-sW}) \sim s$ ,  $s \rightarrow 0$ . Hence, there exists a constant  $c_{16} > 0$  such that for all  $s \in (0, 1]$  and all  $j \geq 0$ ,

$$\mathbf{E}_{\mathbf{Q}}(e^{-sm^{-j}\Lambda(\epsilon_j)}) \leq \mathbf{E}_{\mathbf{Q}}\left\{[1 - c_{16}s]^{\#\mathbf{bro}(\epsilon_0)}\right\} \leq 1 - c_{17}c_{16}s,$$

with  $c_{17} := \mathbf{Q}(\#\mathbf{bro}(\epsilon_0) \geq 1) > 0$ . Consequently, with  $c_{18} := c_{17}c_{16}$ ,

$$\begin{aligned} \mathbf{E}_{\mathbf{Q}}(\eta \wedge 1) &\asymp 1 - \mathbf{E}_{\mathbf{Q}}(e^{-\eta}) \\ &= 1 - \prod_{j=1}^{\lambda_1 n} \mathbf{E}_{\mathbf{Q}}\left(e^{-p(b+n)^\alpha m^{-b-n+\ell} m^{-j}\Lambda(\epsilon_j)}\right) \\ &\geq 1 - \prod_{j=1}^{\lambda_1 n} \left(1 - c_{18} p(b+n)^\alpha m^{-b-n+\ell}\right) \\ &\asymp p(b+n)^\alpha n m^{-b-n+\ell}. \end{aligned}$$

Going back to (4.12), this yields  $\mathbf{E}_{\mathbf{Q}}(\eta \wedge 1) \asymp p(b+n)^\alpha n m^{-b-n+\ell}$ . In view of (4.11), we obtain:

$$(4.13) \quad \mathbb{E}(Z) \asymp m^n \sum_{\ell=\lambda_1 n}^{\lambda_2 n} p n^\alpha m^{-\ell} p(b+n)^\alpha n m^{-b-n+\ell} \asymp p^2 n^{2+\alpha} (b+n)^\alpha m^{-b},$$

uniformly in  $b \geq 0$ .

For the second moment of  $Z$ , we write, by an abuse of notation,

$$\mathbb{T}_k := \{x \in \mathbb{T}^{(n)} : |x| = k\}, \quad 0 \leq k \leq n;$$

then

$$E_\omega(Z^2) = E_\omega(Z) + \sum_{k=1}^n \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v),$$

where  $\sum_{(u,v)}$  is over the pairs  $(u, v) \in \mathbb{T}_0 \times \mathbb{T}_0$  with  $u_k = v_k = x$  such that  $u_{k-1} \neq v_{k-1}$ . We take expectation with respect to  $\mathbf{P}$  on both sides, while splitting the sum  $\sum_{k=1}^n$  into  $\sum_{k=\lambda_1 n+1}^n$  and  $\sum_{k=1}^{\lambda_1 n}$ :

$$\begin{aligned} \mathbb{E}(Z^2) &= \mathbb{E}(Z) + \sum_{k=\lambda_1 n+1}^n \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v) \\ (4.14) \quad &+ \sum_{k=1}^{\lambda_1 n} \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v). \end{aligned}$$

We treat the two sums on the right-hand side. See Figure 2.

**First sum:**  $\sum_{k=\lambda_1 n+1}^n \mathbf{E}(\dots)$ . When  $k > \lambda_1 n$ , the events  $A_u$  and  $A_v$  are independent under  $P_\omega$ , so

$$\begin{aligned} P_\omega(A_u \cap A_v) &= P_\omega(A_u) P_\omega(A_v) \\ &\lesssim \left( p^2 n^{1+\alpha} (b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-(b+n)-j} \Lambda(u_j) \right) \times \\ &\quad \left( p^2 n^{1+\alpha} (b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-(b+n)-j} \Lambda(v_j) \right), \end{aligned}$$

the inequality being a consequence of (4.10). We take the expectation with

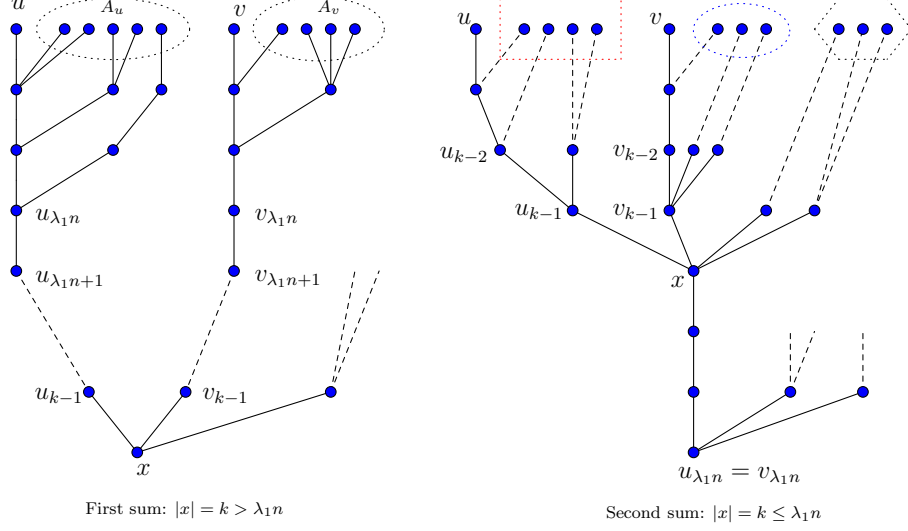


Figure 2: In the first sum,  $A_u$  and  $A_v$  are independent under  $P_\omega$ . In the second sum,  $\max_{1 \leq j \leq k-2} M(u_j)$ ,  $\max_{1 \leq j \leq k-2} M(v_j)$ ,  $M(u_{k-1}, v_{k-1})$  are represented by the rectangle, ellipse and hexagon respectively, and are independent under  $P_\omega$  (for the definition of  $M(u_{k-1}, v_{k-1})$ , see (4.17)).

respect to  $\mathbf{P}$  on both sides, to see that

$$\begin{aligned}
& \sum_{k=\lambda_1 n+1}^n \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v) \\
& \lesssim \sum_{k=\lambda_1 n+1}^n m^{n-k} m^{2k} \left( p^2 n^{1+\alpha} (b+n)^\alpha \sum_{j=1}^{\lambda_1 n} m^{-(b+n)-j} m^j \right)^2 \\
& \asymp \left( p^2 n^{2+\alpha} (b+n)^\alpha m^{-b} \right)^2 \\
(4.15) \quad & \asymp (\mathbf{E}(Z))^2,
\end{aligned}$$

the last line being a consequence of (4.13).

**Second sum:**  $\sum_{k=1}^{\lambda_1 n} \mathbf{E}(\dots)$ . This time, we argue differently, first by conditioning on  $X(u)$  and  $X(v)$ . For  $\ell_u, \ell_v \in [\lambda_1 n, \lambda_2 n]$ , we have, by (4.5), for  $w = u$  or  $v$ ,

$$\mathbb{P}(X(w) = \ell_w) = c_{15} p(\ell_w + \gamma)^\alpha m^{-\ell_w} \asymp p n^\alpha m^{-\ell_w}.$$

So, for  $1 \leq k \leq \lambda_1 n$ ,

$$\mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u, v)} P_\omega(A_u \cap A_v) \asymp \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u, v)} \sum_{\ell_u, \ell_v = \lambda_1 n}^{\lambda_2 n} p^2 n^{2\alpha} m^{-\ell_u - \ell_v} P_\omega(A_u \cap A_v \mid X(u) = \ell_u, X(v) = \ell_v).$$

Write<sup>6</sup>

$$(4.16) \quad B_{x, u, v} := \max_{k \leq j \leq \lambda_1 n} M(x_j) \vee M(u_{k-1}, v_{k-1}),$$

$$(4.17) \quad M(u_{k-1}, v_{k-1}) := \max_{r \in \mathbf{bro}(u_{k-1}, v_{k-1})} \max_{w \in \mathbb{T}_0(r)} (X(w) - (k-1))^+,$$

where  $a \vee b := \max\{a, b\}$ , and  $\mathbf{bro}(u_{k-1}, v_{k-1}) := \mathbf{bro}(u_{k-1}) \setminus \{v_{k-1}\} = \mathbf{bro}(v_{k-1}) \setminus \{u_{k-1}\}$ . [Observe that  $u_j = x_j = v_j$  for  $k \leq j \leq \lambda_1 n$ .] The random variables  $X(u)$ ,  $X(v)$ ,  $\max_{1 \leq j \leq k-2} M(u_j)$ ,  $\max_{1 \leq j \leq k-2} M(v_j)$  and  $B_{x, u, v}$  are independent under  $P_\omega$  (see Figure 2). As such, for  $\ell_u, \ell_v \in [\lambda_1 n, \lambda_2 n]$ ,

$$\begin{aligned} & P_\omega(A_u \cap A_v \mid X(u) = \ell_u, X(v) = \ell_v) \\ &= P_\omega \left( \max_{1 \leq j \leq k-2} M(u_j) \vee B_{x, u, v} \geq b + n - \ell_u, \right. \\ & \quad \left. \max_{1 \leq j \leq k-2} M(v_j) \vee B_{x, u, v} \geq b + n - \ell_v \right) \\ &\leq P_\omega(B_{x, u, v} \geq b + n - \ell_u) + P_\omega(B_{x, u, v} \geq b + n - \ell_v) + \\ & \quad + P_\omega \left( \max_{1 \leq j \leq k-2} M(u_j) \geq b + n - \ell_u \right) P_\omega \left( \max_{1 \leq j \leq k-2} M(v_j) \geq b + n - \ell_v \right). \end{aligned}$$

The first two probability expressions on the right-hand side play the same role by symmetry in  $\ell_u$  and  $\ell_v$ , so let us only look at the first one: By (4.8),

$$\begin{aligned} P_\omega(B_{x, u, v} \geq b + n - \ell_u) &\lesssim p(b + n - \ell_u)^\alpha \sum_{j=k-1}^{\lambda_1 n} m^{-(b+n-\ell_u)-j} \Lambda(u_j) \\ &\lesssim p(b + n)^\alpha \sum_{j=k-1}^{\lambda_1 n} m^{-(b+n-\ell_u)-j} \Lambda(u_j). \end{aligned}$$

---

<sup>6</sup>Notation:  $\max_\emptyset := 0$ .

[Note that  $u_j = x_j$  for  $k \leq j \leq \lambda_1 n$ .] Similarly, for the third probability expression, we have, by (4.8) again, for  $w = u$  or  $v$ ,

$$P_\omega\left(\max_{1 \leq j \leq k-2} M(w_j) \geq b + n - \ell_w\right) \lesssim p(b+n)^\alpha \sum_{j=1}^{k-2} m^{-(b+n-\ell_w)-j} \Lambda(w_j).$$

Assembling these pieces together yields, for  $1 \leq k \leq \lambda_1 n$ ,

$$\begin{aligned} & \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v) \\ \lesssim & \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} \sum_{\ell_u, \ell_v = \lambda_1 n}^{\lambda_2 n} p^2 n^{2\alpha} m^{-\ell_u - \ell_v} p(b+n)^\alpha \sum_{j=k-1}^{\lambda_1 n} m^{-(b+n-\ell_u)-j} \Lambda(u_j) \\ & + \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} \sum_{\ell_u, \ell_v = \lambda_1 n}^{\lambda_2 n} p^2 n^{2\alpha} m^{-\ell_u - \ell_v} p^2 (b+n)^{2\alpha} \\ & \left( \sum_{j=1}^{k-2} m^{-(b+n-\ell_u)-j} \Lambda(u_j) \right) \left( \sum_{j=1}^{k-2} m^{-(b+n-\ell_v)-j} \Lambda(v_j) \right). \end{aligned}$$

On the right-hand side, both expectations can be easily estimated by means of the branching property. The first expectation is

$$\begin{aligned} & \lesssim m^{n-k} m^{2k} \sum_{\ell_u, \ell_v = \lambda_1 n}^{\lambda_2 n} p^2 n^{2\alpha} m^{-(\ell_u + \ell_v)} p(b+n)^\alpha \sum_{j=k-1}^{\lambda_1 n} m^{-(b+n-\ell_u)-j} m^j \\ & \asymp m^{k-\lambda_1 n} p^3 n^{1+2\alpha} (b+n)^\alpha m^{-b} (\lambda_1 n - k), \end{aligned}$$

whereas the second probability expression is

$$\begin{aligned} & \lesssim m^{n-k} m^{2k} \sum_{\ell_u, \ell_v = \lambda_1 n}^{\lambda_2 n} p^2 n^{2\alpha} m^{-\ell_u - \ell_v} p^2 (b+n)^{2\alpha} \\ & \left( \sum_{j=1}^{k-1} m^{-(b+n-\ell_u)-j} m^j \right) \left( \sum_{j=1}^{k-1} m^{-(b+n-\ell_v)-j} m^j \right) \\ & \asymp k^2 m^{k-n-2b} p^4 n^{2+2\alpha} (b+n)^{2\alpha}. \end{aligned}$$



Consequently, for  $1 \leq k \leq \lambda_1 n$ ,

$$\begin{aligned} & \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v) \\ & \lesssim m^{k-\lambda_1 n} p^3 n^{1+2\alpha} (b+n)^\alpha m^{-b} (\lambda_1 n - k) + k^2 m^{k-n-2b} p^4 n^{2+2\alpha} (b+n)^{2\alpha}. \end{aligned}$$

Summing over  $1 \leq k \leq \lambda_1 n$  yields that

$$\begin{aligned} & \sum_{k=1}^{\lambda_1 n} \mathbf{E} \sum_{x \in \mathbb{T}_k} \sum_{(u,v)} P_\omega(A_u \cap A_v) \\ & \lesssim p^3 n^{1+2\alpha} (b+n)^\alpha m^{-b} + m^{-(1-\lambda_1)n-2b} p^4 n^{3+2\alpha} (b+n)^{2\alpha} \\ & \asymp p^3 n^{1+2\alpha} (b+n)^\alpha m^{-b}, \end{aligned}$$

which, by (4.13), is  $\lesssim \mathbb{E}(Z)$  if  $pn^{\alpha-1} \leq 1$ . Combining this with (4.15) and (4.14), we see that  $\mathbb{E}(Z^2) \lesssim \mathbb{E}(Z) + (\mathbb{E}Z)^2$ , as long as  $pn^{\alpha-1} \leq 1$ .

Under the condition  $pn^{1+\alpha} \leq 1$ , we have  $\mathbb{E}(Z) \lesssim 1$  by (4.13), so  $(\mathbb{E}Z)^2 \lesssim \mathbb{E}(Z)$ ; it follows from the Cauchy–Schwarz inequality that  $\mathbb{P}(Z \geq 1) \geq \frac{(\mathbb{E}Z)^2}{\mathbb{E}(Z^2)} \gtrsim \mathbb{E}(Z) \asymp p^2 n^{2+\alpha} (b+n)^\alpha m^{-b}$  (by (4.13)). The lemma follows now from (4.7).  $\square$

## 4.2 Second step: The spinal advantage

Let  $\alpha > -2$  and  $\varepsilon > 0$ .

**Lemma 4.2.** *Assume  $\mathbf{E}(\nu^3) < \infty$  and  $m := \mathbf{E}(\nu) > 1$ . Under the assumption (1.5), for any  $K > \frac{2}{2+\alpha}$ , there exist constants  $s \geq K$  and  $c_{19} > 0$  such that for  $0 < p < 1$  and  $n \leq (\frac{1}{p})^{(s+1)/[(2+\alpha)s+\alpha]}$ ,*

$$(4.18) \quad \mathbb{E}(X_n) \geq c_{19} p^{s+1} n^{(2+\alpha)s+\alpha}.$$

*Proof of Lemma 4.2.* [The condition  $K > \frac{2}{2+\alpha}$  is to make sure that  $(2+\alpha)s + \alpha > 0$  whenever  $s \geq K$ .]

Let  $i \geq 1$  be an integer. Let  $a_i := 2^i$  and  $b_i := (2+\alpha)(2^i - 1)$  (which explains the condition  $\alpha > -2$ : so that  $b_i > 0$ ). Applying Lemma 4.1  $i$

times, we see that for any integer  $i \geq 1$  and any constant  $c > 0$ , there exists a constant  $c(i) > 0$  such that for  $0 < p < 1$  and  $n \geq 1$  with  $p^{a_i} n^{b_i + \alpha} \leq c$ , we have

$$(4.19) \quad \mathbb{P}(X_{in} \geq b) \geq c(i) p^{a_i} n^{b_i} (b+n)^\alpha m^{-b}, \quad \forall b > 0.$$

[For  $i = 1$ , (4.19) follows from Lemma 4.1 with  $\gamma = 1$  and  $t = b$ . Assuming (4.19) holds for  $i$ , it is immediately seen to hold for  $i + 1$ : It suffices to apply Lemma 4.1 to  $\gamma = n$ ,  $t = b$  and to  $p^{a_i} n^{b_i}$  in place of  $p$ .]

For integers  $\ell \in [0, n]$ , we use the above inequality for  $\mathbb{P}(X_{(i-1)n} \geq b)$ , and apply Lemma 4.1 to  $n + \ell$  in place of  $n$ , to see that there exists a constant  $c'(i) > 0$  for  $0 < p < 1$  and  $n \geq 1$  with  $p^{a_i} (n + \ell)^{b_i + \alpha} \leq c$ , and for all integers  $\ell \in [0, n]$ ,

$$\mathbb{P}(X_{in+\ell} \geq b) \geq c'(i) p^{a_i} n^{b_i} (b+n)^\alpha m^{-b}, \quad \forall b > 0.$$

Integrating over  $b$ , this yields the existence of a constant  $c''(i) > 0$ , depending on  $i$ , such that for  $0 < p < 1$  and  $n \geq 1$  with  $p^{a_i} n^{b_i + \alpha} \leq c$ ,

$$\mathbb{E}(X_{in+\ell}) \geq c''(i) p^{a_i} n^{b_i + \alpha}.$$

We choose (and fix)  $i$  sufficiently large, how large depending on  $\alpha$ , such that  $a_i \geq K + 1$ . The lemma follows with  $s := a_i - 1$ .  $\square$

The rest of the proof of the lower bound in Theorem 1.5 consists in improving the lower bound for  $\mathbb{E}(X_n)$  in (4.18), and making it (strictly) greater than  $\frac{1}{m-1}$ , so that by virtue of the first inequality in (1.6), which says that  $F_\infty \geq \frac{\mathbb{E}(X_n) - \frac{1}{m-1}}{m^n}$ , it will give the desired lower bound for the free energy  $F_\infty$  as stated in Theorem 1.5.

Without loss of generality, we assume that the law of  $X_0$ , conditionally on  $X_0 > 0$ , is absolutely continuous.

To improve (4.18), we start with a new lower bound for  $X_n$ . For any vertex  $v \in \mathbb{T}^{(n)}$ , we write  $X(v)$  for the random variable associated with the vertex  $v$ : so if  $|v| = j$ , then  $X(v)$  is distributed as  $X_j$ . Let  $1 \leq k < \ell < n$  be

integers; the values of  $k$  and  $\ell$ , both depending on  $(n, p)$ , will be given later. For  $u \in \mathbb{T}_0$ , let

$$M_k^*(u) := \max_{v \in \mathbb{T}_0(u_k) \setminus \{u\}} X(v), \quad N_k^*(u) := \max_{v \in \mathbb{T}_0 \setminus \mathbb{T}_0(u_k)} X(v).$$

[So  $M_k^*(u) \vee N_k^*(u) \vee X(u)$  coincides with  $\max_{w \in \mathbb{T}_0} X(w)$ ; moreover,  $M_k^*(u)$ ,  $N_k^*(u)$ ,  $X(u)$  are independent random variables under  $P_\omega$ .]

Assume there exists  $u \in \mathbb{T}_0$  such that  $X(u) > N_k^*(u)$  and  $M_k^*(u) \leq \ell < X(u)$ . If such a vertex  $u$  exists (which must be unique, by definition),  $X_n \geq X(u) - n + \sum_{j=0}^{\ell} \sum_{v \in \text{bro}(u_j)} X(v)$ , which is greater than  $\ell - n + \sum_{j=0}^{k-1} \sum_{v \in \text{bro}(u_j)} X(v)$ .<sup>7</sup> [In case  $\ell - n + \sum_{j=0}^{k-1} \sum_{v \in \text{bro}(u_j)} X(v)$  is negative, the statement is, of course, trivial.] We arrive at the following inequality:

$$\begin{aligned} X_n &\geq \sum_{u \in \mathbb{T}_0} \mathbf{1}_{\{X(u) > N_k^*(u) \vee \ell\}} \mathbf{1}_{\{M_k^*(u) \leq \ell\}} \left( \ell - n + \sum_{j=0}^{k-1} \sum_{v \in \text{bro}(u_j)} X(v) \right). \\ (4.20) \quad &=: \sum_{u \in \mathbb{T}_0} \mathbf{1}_{\{X(u) > N_k^*(u) \vee \ell\}} \xi(u), \end{aligned}$$

where, for  $u \in \mathbb{T}_0$ ,

$$\xi(u) := \mathbf{1}_{\{M_k^*(u) \leq \ell\}} \left( \ell - n + \sum_{j=0}^{k-1} \sum_{v \in \text{bro}(u_j)} X(v) \right).$$

Note that  $X(u)$ ,  $N_k^*(u)$  and  $\xi(u)$  are independent under  $P_\omega$ . Hence

$$E_\omega(X_n) \geq \sum_{u \in \mathbb{T}_0} P_\omega\{X(u) > N_k^*(u) \vee \ell\} E_\omega(\xi(u)),$$

Taking expectation with respect to  $\mathbf{P}$ , we obtain, by the many-to-one formula (4.1),

$$\mathbb{E}(X_n) \geq m^n \mathbf{E}_{\mathbf{Q}} \left[ P_\omega\{X(\mathbf{e}_0) > N_k^*(\mathbf{e}_0) \vee \ell\} E_\omega(\xi(\mathbf{e}_0)) \right].$$

By the spinal decomposition theorem, the random variables  $P_\omega\{X(\mathbf{e}_0) > N_k^*(\mathbf{e}_0) \vee \ell\}$  and  $E_\omega(\xi(\mathbf{e}_0))$  are independent under  $\mathbf{Q}$ . See Figure 3. So

$$(4.21) \quad \mathbb{E}(X_n) \geq m^n (\mathbf{Q} \otimes P_\omega)\{X(\mathbf{e}_0) > N_k^*(\mathbf{e}_0) \vee \ell\} (\mathbf{E}_{\mathbf{Q}} \otimes E_\omega)(\xi(\mathbf{e}_0)).$$

<sup>7</sup>Strictly speaking, we should write  $X(\mathbf{e}_n)$  in place of  $X_n$ .

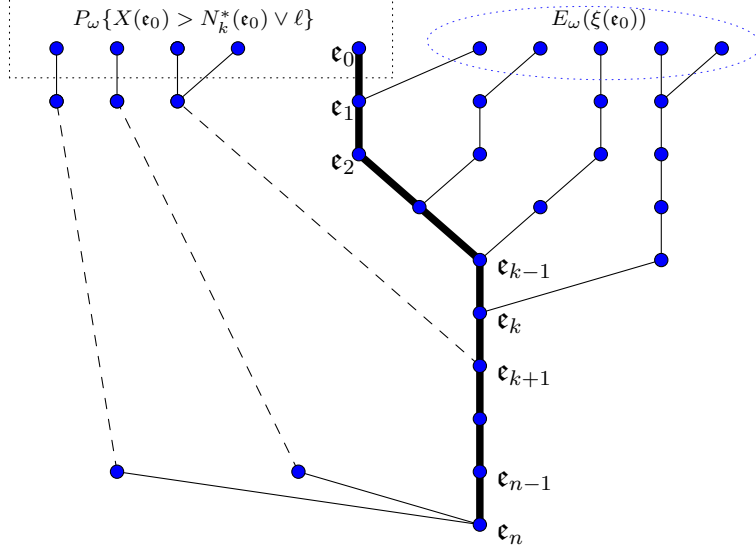


Figure 3: The random variables  $P_\omega\{X(\epsilon_0) > N_k^*(\epsilon_0) \vee \ell\}$  and  $E_\omega(\xi(\epsilon_0))$  are independent under  $\mathbf{Q}$ .

We study  $(\mathbf{E}_\mathbf{Q} \otimes E_\omega)(\xi(\epsilon_0))$  on the right-hand side. Since  $\ell < n$ , we have,

$$\begin{aligned} \xi(\epsilon_0) &\geq \ell - n + \mathbf{1}_{\{M_k^*(\epsilon_0) \leq \ell\}} \sum_{j=0}^{k-1} \sum_{v \in \mathbf{bro}(\epsilon_j)} X(v) \\ &= \ell - n + \sum_{j=0}^{k-1} \sum_{v \in \mathbf{bro}(\epsilon_j)} X(v) - \sum_{j=0}^{k-1} \mathbf{1}_{\{M_k^*(\epsilon_0) > \ell\}} \sum_{v \in \mathbf{bro}(\epsilon_j)} X(v). \end{aligned}$$

We take expectation with respect to  $(\mathbf{Q} \otimes P_\omega)$  on both sides. By the spinal decomposition theorem,

$$\begin{aligned} (\mathbf{E}_\mathbf{Q} \otimes E_\omega)(\xi(\epsilon_0)) &\geq \ell - n + \sum_{j=0}^{k-1} \mathbb{E}(X_j) \mathbf{E}_\mathbf{Q}[\#\mathbf{bro}(\epsilon_0)] \\ (4.22) \quad &- \sum_{j=0}^{k-1} (\mathbf{E}_\mathbf{Q} \otimes E_\omega) \left( \mathbf{1}_{\{M_k^*(\epsilon_0) > \ell\}} \sum_{v \in \mathbf{bro}(\epsilon_j)} X(v) \right). \end{aligned}$$

Let us have a closer look at the last  $(\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega})(\cdots)$  expression on the right-hand side. By the trivial inequality  $X(v) \leq \sum_{r \in \mathbb{T}_0(v)} X(r)$ , we have

$$\sum_{v \in \mathbf{bro}(\mathbf{e}_j)} X(v) \leq \sum_{v \in \mathbf{bro}(\mathbf{e}_j)} \sum_{r \in \mathbb{T}_0(v)} X(r) \leq \sum_{r \in \mathbb{T}_0(\mathbf{e}_{j+1})} X(r).$$

So by the Cauchy–Schwarz inequality,

$$\begin{aligned} & (\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega}) \left( \mathbf{1}_{\{M_k^*(\mathbf{e}_0) > \ell\}} \sum_{v \in \mathbf{bro}(\mathbf{e}_j)} X(v) \right) \\ & \leq [(\mathbf{Q} \otimes P_{\omega})\{M_k^*(\mathbf{e}_0) > \ell\}]^{1/2} \left\{ (\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega}) \left[ \left( \sum_{r \in \mathbb{T}_0(\mathbf{e}_{j+1})} X(r) \right)^2 \right] \right\}^{1/2}. \end{aligned}$$

By definition,  $P_{\omega}\{M_k^*(\mathbf{e}_0) > \ell\} \leq (\#\mathbb{T}_0(\mathbf{e}_k)) \mathbb{P}(X_0 > \ell)$ , which, by the assumption (1.5), is  $\lesssim (\#\mathbb{T}_0(\mathbf{e}_k)) p \ell^{\alpha} m^{-\ell}$ . Hence

$$(\mathbf{Q} \otimes P_{\omega})\{M_k^*(\mathbf{e}_0) > \ell\} \leq \mathbf{E}_{\mathbf{Q}}(\#\mathbb{T}_0(\mathbf{e}_k)) \mathbb{P}(X_0 > \ell) \lesssim \mathbf{E}_{\mathbf{Q}}(\#\mathbb{T}_0(\mathbf{e}_k)) p \ell^{\alpha} m^{-\ell}.$$

By (4.2), this yields

$$(\mathbf{Q} \otimes P_{\omega})\{M_k^*(\mathbf{e}_0) > \ell\} \lesssim m^k p \ell^{\alpha} m^{-\ell}.$$

On the other hand,

$$(\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega}) \left[ \left( \sum_{r \in \mathbb{T}_0(\mathbf{e}_{j+1})} X(r) \right)^2 \right] = \mathbf{E}_{\mathbf{Q}} \left[ \left( \sum_{r \in \mathbb{T}_0(\mathbf{e}_{j+1})} \mathbb{E}(X_0) \right)^2 \right] + \mathbf{E}_{\mathbf{Q}} \left[ \sum_{r \in \mathbb{T}_0(\mathbf{e}_{j+1})} \sigma^2 \right],$$

where  $\sigma^2 := \text{Var}(X_0) \leq c_{20} p$ , and  $\mathbb{E}(X_0) = c_{21} p$ . As such,

$$\begin{aligned} & (\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega}) \left[ \left( \sum_{r \in \mathbb{T}_0(\mathbf{e}_{j+1})} X(r) \right)^2 \right] \\ & \leq (c_{21} p)^2 \mathbf{E}_{\mathbf{Q}}[(\#\mathbb{T}_0(\mathbf{e}_{j+1}))^2] + c_{20} p \mathbf{E}_{\mathbf{Q}}(\#\mathbb{T}_0(\mathbf{e}_{j+1})) \lesssim p^2 m^{2j} + p m^j, \end{aligned}$$

the last inequality following from (4.2) and (4.3). Consequently,

$$(\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega}) \left( \mathbf{1}_{\{M_k^*(\mathbf{e}_0) > \ell\}} \sum_{v \in \mathbf{bro}(\mathbf{e}_j)} X(v) \right) \lesssim \left[ m^k p \ell^{\alpha} m^{-\ell} (p^2 m^{2j} + p m^j) \right]^{1/2}.$$

As such, as long as we take

$$(4.23) \quad k := \lfloor \frac{\ell}{4} \rfloor, \quad \ell := n - c_{22} \log n,$$

we have, for some  $c_{23} > 0$ ,

$$\sum_{j=0}^{k-1} (\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega}) \left( \mathbf{1}_{\{M_k^*(\mathbf{e}_0) > \ell\}} \sum_{v \in \mathbf{bro}(\mathbf{e}_j)} X(v) \right) \leq c_{23}.$$

Going back to (4.22), we obtain, with  $c_{24} := \mathbf{E}_{\mathbf{Q}}[\#\mathbf{bro}(\mathbf{e}_0)] < \infty$ ,

$$(\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega})(\xi(\mathbf{e}_0)) \geq \ell - n + c_{24} \sum_{j=0}^{k-1} \mathbb{E}(X_j) - c_{23}.$$

Let  $\varepsilon > 0$ . Let  $s \geq K$  be the constants in Lemma 4.2. We choose  $K$  so large that  $\frac{s+1}{(2+\alpha)s+\alpha} \leq \frac{1}{2+\alpha} + \varepsilon$ . If we take

$$(4.24) \quad n = n(p) := \left( \frac{1}{p} \right)^{\frac{s+1}{(2+\alpha)s+\alpha}} \leq \left( \frac{1}{p} \right)^{\frac{1}{2+\alpha} + \varepsilon},$$

(the last inequality holding for all sufficiently small  $p$ ), then by Lemma 4.2,  $\sum_{j=0}^{\lfloor \frac{\ell}{4} \rfloor - 1} \mathbb{E}(X_j) \geq c_{25} n$  for some constant  $c_{25} > 0$ . With our choice of  $\ell$  in (4.23), this implies  $(\mathbf{E}_{\mathbf{Q}} \otimes E_{\omega})(\xi(\mathbf{e}_0)) \geq c_{26} n$ , with  $c_{26} > 0$ . Going back to (4.21), we obtain, for all sufficiently small  $p > 0$ ,

$$(4.25) \quad \mathbb{E}(X_n) \geq m^n (\mathbf{Q} \otimes P_{\omega}) \{X(\mathbf{e}_0) > N_k^*(\mathbf{e}_0) \vee \ell\} c_{25} n.$$

By the many-to-one formula (4.1) again, this yields

$$\begin{aligned} \mathbb{E}(X_n) &\geq \mathbf{E} \left( \sum_{u \in \mathbb{T}_0} P_{\omega} \{X(u) > N_k^*(u) \vee \ell\} \right) c_{25} n \\ &\geq \mathbf{E} \left( \sum_{u \in \mathbb{T}_0} \mathbf{1}_{\{c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n\}} P_{\omega} \{X(u) > N_k^*(u) \vee \ell\} \right) c_{25} n, \end{aligned}$$

for any constants  $c_{28} > c_{27} > 0$ . For any  $u \in \mathbb{T}_0$ , since the conditional law of  $X_0$  given  $X_0 > 0$  is assumed to be absolutely continuous, we have

$$P_{\omega} \{X(u) > N_k^*(u) \vee \ell\} \geq P_{\omega} \{X(u) = \max_{r \in \mathbb{T}_0} X(r)\} - P_{\omega} \{\max_{r \in \mathbb{T}_0} X(r) \leq \ell\}.$$

We have  $\sum_{u \in \mathbb{T}_0} P_\omega \{X(u) = \max_{r \in \mathbb{T}_0} X(r)\} = 1$ , whereas  $P_\omega \{\max_{r \in \mathbb{T}_0} X(r) \leq \ell\} = [1 - \mathbb{P}(X_0 > \ell)]^{\#\mathbb{T}_0} \leq [1 - c_3 \ell^\alpha m^{-\ell}]^{\#\mathbb{T}_0}$  by (1.5); hence

$$\begin{aligned} & \mathbf{E} \left( \sum_{u \in \mathbb{T}_0} \mathbf{1}_{\{c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n\}} P_\omega \{X(u) > N_k^*(u) \vee \ell\} \right) \\ & \geq \mathbf{P}(c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n) - \mathbf{E} \left( \sum_{u \in \mathbb{T}_0} \mathbf{1}_{\{c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n\}} [1 - c_3 \ell^\alpha m^{-\ell}]^{\#\mathbb{T}_0} \right) \\ & = \mathbf{P}(c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n) - \mathbf{E} \left( (\#\mathbb{T}_0) \mathbf{1}_{\{c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n\}} [1 - c_3 \ell^\alpha m^{-\ell}]^{\#\mathbb{T}_0} \right). \end{aligned}$$

On the event  $\{c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n\}$ , we have  $(\#\mathbb{T}_0) [1 - c_3 \ell^\alpha m^{-\ell}]^{\#\mathbb{T}_0} \leq c_{28} m^n \exp(-c_3 \ell^\alpha m^{-\ell} c_{27} m^n)$ , which tends to 0 with the choice of  $\ell := n - c_{22} \log n$ , as long as  $c_{22} > \frac{1-\alpha}{\log m}$ . On the other hand, the constants  $c_{28} > c_{27} > 0$  can be chosen such that  $\mathbf{P}(c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n) \rightarrow c_{29} > 0$ ,  $n \rightarrow \infty$ . Hence

$$\liminf_{n \rightarrow \infty} \mathbf{E} \left( \sum_{u \in \mathbb{T}_0} \mathbf{1}_{\{c_{27} m^n \leq \#\mathbb{T}_0 \leq c_{28} m^n\}} P_\omega \{X(u) > N_k^*(u) \vee \ell\} \right) \geq c_{29} > 0.$$

In view of (4.25), we obtain: for some constant  $c_{30} > 0$  and all sufficiently small  $p$ , with  $n = n(p)$  given in (4.24),

$$\mathbb{E}(X_n) \geq c_{30} n.$$

By the first inequality in (1.6), we get  $F_\infty \geq \frac{\mathbb{E}(X_n) - \frac{1}{m-1}}{m^n} \geq \frac{c_{30} n - \frac{1}{m-1}}{m^n}$ . The definition of  $n$  in (4.24) yields that for an arbitrary  $\varepsilon > 0$  and all sufficiently small  $p$ ,  $F_\infty \geq \exp(-(\frac{1}{p})^{\frac{1}{2+\alpha} + \varepsilon})$ , proving the lower bound in Theorem 1.5.

[We mention that the lower bound is proved under the assumption  $\mathbf{E}(\nu^3) < \infty$ , instead of  $\mathbf{E}(t^\nu) < \infty$  for some  $t > 1$ .]  $\square$

## 5 Proof of Theorem 1.4: lower bound

We use the obvious stochastic inequality that  $X_n$  is stochastically greater than or equal to  $\max_{u \in \mathbb{T}_0} X(u) - n$ . Hence for all  $b \geq 0$ ,

$$\begin{aligned} P_\omega(X_n > b) & \geq P_\omega \left( \max_{u \in \mathbb{T}_0} X(u) - n > b \right) \\ & = 1 - [1 - \mathbb{P}(X_0 \leq n + b)]^{\#\mathbb{T}_0}. \end{aligned}$$

By assumption, for all sufficiently large  $n$  (say  $n \geq n_1$ ),  $\mathbb{P}(X_0 \leq n + b) \geq p c_1 e^{-\theta(n+b)}$ . Thus, for some constant  $c_{31} > 0$ , all  $n \geq n_1$  and all  $b \geq 0$ ,

$$\begin{aligned} P_\omega(X_n > b) &\geq 1 - \left(1 - p c_1 e^{-\theta(n+b)}\right)^{\#\mathbb{T}_0} \\ &\geq c_{31} \min\{p e^{-\theta(n+b)} \#\mathbb{T}_0, 1\}. \end{aligned}$$

We take  $n = N(p) := \lceil \frac{c_{32} + \log(1/p)}{(\log m) - \theta} \rceil$ , where  $c_{32} > 0$  is a sufficiently large constant. Note that  $p e^{-\theta(N(p)+b)} m^{N(p)} \geq 1$  if (and only if)  $b \leq B(p)$  where  $B(p) := \frac{1}{\theta} [((\log m) - \theta)N(p) - \log(\frac{1}{p}) + \theta]$ . Therefore, for all sufficiently small  $p$  (such that  $N(p) \geq n_1$ ) and all  $0 < b \leq B(p)$ , we have<sup>8</sup>

$$\mathbb{P}(X_{N(p)} > b) \geq c_{31} \mathbf{P}(\#\mathbb{T}_0 \geq m^{N(p)}) = c_{31} \mathbf{P}(\#\mathbb{T}_0 \geq m^n) \geq c_{33},$$

where  $c_{33} > 0$  does not depend on  $p$ . This implies that  $\mathbf{E}(X_n) = \int_0^\infty \mathbf{P}(X_n > b) db \geq c_{33} B(p)$ , which, by definition of  $B(p)$  and  $N(p)$ , is greater than  $\frac{c_{33}(c_{32} + \theta)}{\theta}$ . The latter is greater than 2 if we choose  $c_{32} := \frac{2\theta}{c_{33}}$ . By the first inequality in (1.6), we get, for all sufficiently small  $p$ ,

$$F_\infty(p) \geq \frac{2 - \frac{1}{m-1}}{m^{N(p)}} = \left(2 - \frac{1}{m-1}\right) \exp\left(-\lceil \frac{c_{32} + \log(1/p)}{(\log m) - \theta} \rceil \log m\right),$$

proving the lower bound in Theorem 1.4.

[The lower bound only requires  $\mathbf{E}(\nu \log^+ \nu) < \infty$ , instead of  $\mathbf{E}(t^\nu) < \infty$  for some  $t > 1$ .] □

## 6 Comments and questions

We present some remarks and open problems.

**(a) Change of measures?** Theorem A and Proposition 1.2 in the introduction reveal the importance of  $\mathbf{E}(X_0 m^{X_0})$ . The latter strongly indicates that there could be a change-of-measures story hidden in the model.

---

<sup>8</sup>To ensure that  $\mathbf{P}(\#\mathbb{T}_0 \geq m^n)$  is greater than a positive constant, uniformly in  $n$ , it suffices to have  $\mathbf{E}(\nu \log^+ \nu) := \sum_{k=1}^\infty k(\log k) \mathbf{P}(\nu = k) < \infty$ ; see [21] or [25].



**Problem 6.1.** *Is it possible to prove Theorem A by means of a change-of-measures argument?*

**(b) About the value of  $p_c$  for non integer valued distributions.**

Theorem A in the introduction gives the value of  $p_c$  when  $\nu = 2$  a.s. and  $Y_0$  is integer valued. [They are valid whenever  $\nu$  is deterministic (i.e.,  $\nu = m$  a.s.), with  $\mathbf{E}[(Y_0 - 1)2^{Y_0}]$  and  $\mathbf{E}(Y_0 2^{Y_0})$  replaced by  $\mathbf{E}[(m - 1)Y_0 - 1)m^{Y_0}]$  and  $\mathbf{E}(Y_0 m^{Y_0})$ , respectively.] The following problem looks important to us.

**Problem 6.2.** *Assume  $\nu$  is deterministic. What can be said about  $p_c$  without the assumption that  $Y_0$  is integer valued?*

Problem 6.2, which is borrowed from Derrida and Retaux [11], seems challenging. For example, even assuming that  $\nu = 2$  a.s., and that  $Y_0$  takes values in  $\{\frac{1}{2}, 1, \frac{3}{2}, 2, \dots\}$ , we do not know what the value of  $p_c$  should be in general.

**(c) More about the value of  $p_c$ .** When  $\nu$  is not deterministic, even if assuming  $Y_0$  is integer valued as in Theorem A (see the introduction), it is not clear what  $p_c$  should be. It is possible to have some bounds, but it seems to be hard to have an analytical expression.

**Problem 6.3.** *Assume  $\nu$  is not deterministic and  $Y_0$  takes values in  $\{1, 2, \dots\}$ . What can be said about  $p_c$ ?*

## Acknowledgements

We are grateful to Bernard Derrida who introduced us to the problem, and with whom we have had regular discussions for two years. We wish to thank Nina Gantert for many discussions, Quentin Berger for enlightenment on renormalisation models, and Chunhua Ma, Bastien Mallein and Quan Shi for pointing out [19] to us. Two anonymous referees have carefully read the manuscript; their insightful comments have led to improvements in the paper. The project was partly supported by ANR MALIN (ANR-16-CE93-0003); Y.H. also acknowledges support from ANR SWiWS (ANR-17-CE40-0032).

## References

- [1] Aldous, D.J. and Bandyopadhyay, A. (2005). A survey of max-type recursive distributional equations. *Ann. Appl. Probab.* **15**, 1047–1110.
- [2] Alexander, K.S. and Sidoravicius, V. (2006). Pinning of polymers and interfaces by random potentials. *Ann. Appl. Probab.* **16**, 636–669.
- [3] Alexander, K.S. and Zygouras, N. (2009). Quenched and annealed critical points in polymer pinning models. *Commun. Math. Phys.* **291**, 659–689.
- [4] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*. Springer, New York.
- [5] Berger, Q. and Lacoïn, H. (2015). Pinning on a defect line: characterization of marginal disorder relevance and sharp asymptotics for the critical point shift. [arXiv:1503.07315](https://arxiv.org/abs/1503.07315).
- [6] Berger, Q. and Toninelli, F.L. (2013). Hierarchical pinning model in correlated random environment. *Ann. Inst. H. Poincaré Probab. Statist.* **49**, 781–816.
- [7] Collet, P., Eckmann, J.P., Glaser, V. and Martin, A. (1984). A spin glass with random couplings. *J. Statist. Phys.* **36**, 89–106.
- [8] Collet, P., Eckmann, J.P., Glaser, V. and Martin, A. (1984). Study of the iterations of a mapping associated to a spin-glass model. *Commun. Math. Phys.* **94**, 353–370.
- [9] Derrida, B., Giacomin, G., Lacoïn, H. and Toninelli, F.L. (2009). Fractional moment bounds and disorder relevance for pinning models. *Commun. Math. Phys.* **287**, 867–887.
- [10] Derrida, B., Hakim, V. and Vannimenus, J. (1992). Effect of disorder on two-dimensional wetting. *J. Statist. Phys.* **66**, 1189–1213.
- [11] Derrida, B. and Retaux, M. (2014). The depinning transition in presence of disorder: a toy model. *J. Statist. Phys.* **156**, 268–290.
- [12] Forgacs, G., Luck, J.M., Nieuwenhuizen, Th.M. and Orland, H. (1986). Wetting of a disordered substrate: Exact critical behavior in two dimensions. *Phys. Rev. Lett.* **57**, 2184–2187.
- [13] Garel, T. and Monthus, C. (2005). Numerical study of the disordered Poland-Scheraga model of DNA denaturation. *J. Statist. Mech. Theory Exper.* 2005(06), P06004.

- [14] Giacomin, G. (2007). *Random Polymer Models*. Imperial College Press, World Scientific.
- [15] Giacomin, G. (2011). Disorder and critical phenomena through basic probability models. École d'Été de Probabilités de Saint-Flour XL. Lecture Notes in Mathematics **2025**, Springer, Berlin.
- [16] Giacomin, G., Lacoïn, H. and Toninelli, F.L. (2010). Hierarchical pinning models, quadratic maps and quenched disorder. *Probab. Theor. Related Fields* **147**, 185–216.
- [17] Giacomin, G., Lacoïn, H. and Toninelli, F.L. (2010). Marginal relevance of disorder for pinning models. *Commun. Pure Appl. Math.* **63**, 233–265.
- [18] Giacomin, G. and Toninelli, F.L. (2006). Smoothing effect of quenched disorder on polymer depinning transitions. *Commun. Math. Phys.* **266**, 1–16.
- [19] Goldschmidt, C. and Przykucki, M. (2016). Parking on a random tree. [arXiv:1610.08786](https://arxiv.org/abs/1610.08786).
- [20] den Hollander, F. (2009). *Random Polymers*. École d'Été de Probabilités de Saint-Flour XXXVII. Lecture Notes in Mathematics **1974**, Springer, Berlin.
- [21] Kesten, H. and Stigum, B.P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37**, 1211–1223.
- [22] Kunz, H. and Livi, R. (2012). DNA denaturation and wetting in the presence of disorder. *Eur. Phys. Lett.* **99**, 30001.
- [23] Lacoïn, H. (2010). Hierarchical pinning model with site disorder: disorder is marginally relevant. *Probab. Theor. Related Fields* **148**, 159–175.
- [24] Lerouillois, V. (2015). Polymer pinning models and condition on the existence of a phase transition. (preprint)
- [25] Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of  $L \ln L$  criteria for mean behavior of branching processes. *Ann. Probab.* **23**, 1125–1138.
- [26] Lyons, R. and Peres, Y. (2016). *Probability on Trees and Networks*. Cambridge University Press.
- [27] Monthus, C. (2016). Strong disorder renewal approach to DNA denaturation and wetting: typical and large deviation properties of the free energy. [arXiv:1611.00501](https://arxiv.org/abs/1611.00501)

- [28] Monthus, C. and Garel, T. (2008). Critical points of quadratic renormalizations of random variables and phase transitions of disordered polymer models on diamond lattices. *Phys. Rev. E* **77(2)**, 021132.
- [29] Shi, Z. (2015). *Branching Random Walks*, École d'été Saint-Flour XLII (2012), *Lecture Notes in Mathematics* **2151**. Springer, Berlin.
- [30] Tang, L.H. and Chaté, H. (2001). Rare-event induced binding transition of heteropolymers. *Phys. Rev. Lett.* **86**, 830–833.
- [31] Toninelli, F.L. (2008). A replica-coupling approach to disordered pinning models. *Commun. Math. Phys.* **280**, 389–401.
- [32] Toninelli, F.L. (2008). Disordered pinning models and copolymers: beyond annealed bounds. *Ann. Appl. Probab.* **18**, 1569–1587.