



Ising model on a Galton-Watson tree with a sparse random external field

Irene Ayuso Ventura, Quentin Berger

► To cite this version:

Irene Ayuso Ventura, Quentin Berger. Ising model on a Galton-Watson tree with a sparse random external field. 2023. hal-04253083

HAL Id: hal-04253083

<https://hal.science/hal-04253083>

Preprint submitted on 21 Oct 2023

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.

Ising model on a Galton–Watson tree with a sparse random external field

Irene Ayuso Ventura^{*} and Quentin Berger^{†,‡}

^{*}Laboratoire d'Analyse et de Mathématiques Appliquées, Université Paris-Est Créteil.

[†]Laboratoire de Probabilité Statistique et Modélisation, Sorbonne Université.

[‡]Département de Mathématiques et Applications, École Normale Supérieure, Université PSL.

Abstract

We consider the Ising model on a supercritical Galton–Watson tree \mathbf{T}_n of depth n with a sparse random external field, given by a collection of i.i.d. Bernoulli random variables with vanishing parameter p_n . This may be viewed as a toy model for the Ising model on a configuration model with a few interfering external vertices carrying a plus spin: the question is to know how many (or how few) interfering vertices are enough to influence the whole graph. Our main result consists in providing a necessary and sufficient condition on the parameters $(p_n)_{n \geq 0}$ for the root of \mathbf{T}_n to remain magnetized in the large n limit. Our model is closely related to the Ising model on a (random) *pruned* sub-tree \mathbf{T}_n^* with plus boundary condition; one key result is that this pruned tree turns out to be an inhomogeneous, n -dependent, Branching Process. We then use standard tools such as tree recursions and non-linear capacities to study the Ising model on this sequence of Galton–Watson trees; one difficulty is that the offspring distributions of \mathbf{T}_n^* , in addition to vary along the generations $0 \leq k \leq n-1$, also depend on n .

Keywords: Ising model, random graphs, random external field, phase transition.

MSC2020 AMS classification: Primary: 82B20, 60K35, 82B44, 82B26.

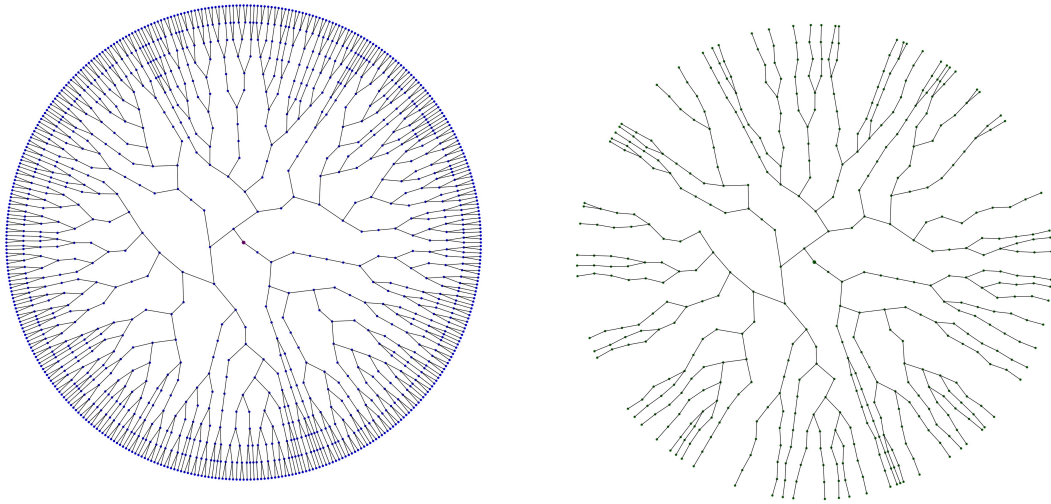


Figure 1: A tree of depth $n = 15$ and its *pruned* version by a Bernoulli process on its leaves.

Contents

1	Introduction of the model and main results	2
2	A few preliminaries	13
3	Upper bound and main arguments for the lower bound	16
4	Pruning the dead branches of a Galton–Watson tree	20
5	Transition in the shape of the pruned tree T_n^*	25
6	Ising model on T_n^* with plus boundary external field	35
A	Some technical proofs and comments	44

1 Introduction of the model and main results

The Ising model is a celebrated model, studied in depths for over 100 years. It was first introduced by Wilhelm Lenz and Ernst Ising as a model for magnetism, and was originally defined on regular lattices; we refer for instance to the books [4, 15] for a general introduction. Later, the Ising model was also studied on different types of graphs, starting with regular structures such as the Bethe lattice or Cayley tree, see [28] for an extensive summary on the subject. The seminal article [19], in which Lyons identifies the critical temperature of the model on an arbitrary infinite tree, opened the way to the study of the Ising model and other statistical mechanics models on tree-like graphs, and, more generally, on random graphs (see for instance [22]).

Motivated by the interest of the model to describe complex networks (see [13] for a review), the literature on the Ising model on random graphs has grown considerably in recent years. Let us mention a few relevant results on the subject. First, for the Ising model on quenched random graphs, the thermodynamic limit of has been studied in [10, 9, 11] as well as its critical behavior in [12, 17]. Thereafter, the annealed Ising model has also gotten some attention, see for instance [5, 6] references therein. More recently, the local weak limit of the Ising model on locally tree-like random graphs was considered in [1, 23].

However, most of the literature on the Ising model on random graphs considers free or plus boundary conditions or a *homogeneous* external field, but there does not seem to be many results when the boundary conditions or the external field are random (and depend on the size of the graph). In the present paper, we consider the Ising model on the simplest random graph possible, a Galton–Watson tree of depth n , but with a sparse random external field (which may be restricted to the boundary, which is close to being a boundary condition) whose distribution depends on n . We see our results as a first step towards the study of the Ising model on a random graph with a few external interfering vertices.

1.1 General setting of the paper

For a finite graph $G = (V, E)$, we consider the following Gibbs (ferromagnetic) Ising measure on spins $\sigma \in \{-1, +1\}^{|V|}$, with inverse temperature β and external field $h = (h_v)_{v \in V}$:

$$P_{G,\beta}^h(\sigma) := \frac{1}{Z_{G,\beta}^h} \exp \left(\beta \left(\sum_{\substack{u,v \in V \\ u \sim v}} \sigma_u \sigma_v + \sum_{v \in V} h_v \sigma_v \right) \right), \quad (1.1)$$

where we denoted $u \sim v$ if $\{u, v\} \in E$. To simplify the statements and without loss of generality, we have assumed here that the coupling parameter is $J = 1$.

In many cases, a natural boundary of V , denoted $\partial V \subset V$, can be identified¹. Then, we can consider the Ising model on the graph G with boundary condition $\xi \in \{+1, -1\}^{\partial V}$ by considering the Gibbs measure

$$P_{G,\beta}^\xi(\sigma) := \frac{1}{Z_{G,\beta}^\xi} \exp \left(\beta \left(\sum_{\substack{u,v \in V \setminus \partial V \\ u \sim v}} \sigma_u \sigma_v + \sum_{\substack{u \in V \setminus \partial V, v \in \partial V \\ u \sim v}} \xi_v \sigma_v \right) \right). \quad (1.2)$$

For the Ising model with external field (1.1), in the case where $h_v = 0$ for all $v \in V \setminus \partial V$, we will say that the Ising model has *boundary external field*.

Remark 1.1 (Exterior boundary). It might also be natural to consider an exterior boundary of G , denoted $\partial^{\text{ex}} G = (\partial^{\text{ex}} V, \partial^{\text{ex}} E)$, where $\partial^{\text{ex}} V$ is a set of external vertices (disjoint from V) and $\partial^{\text{ex}} E$ is a set of boundary edges $\{x, y\}$ with $x \in V$, $y \in \partial^{\text{ex}} V$. We can then consider the Ising model on G with *exterior* boundary condition $\xi \in \{-1, +1\}^{\partial^{\text{ex}} V}$ by considering the Gibbs measure (1.2) on the graph $\bar{G} = (\bar{V}, \bar{E})$ with $\bar{V} = V \cup \partial^{\text{ex}} V$, $\bar{E} = E \cup \partial^{\text{ex}} E$ and with boundary condition ξ on $\partial \bar{V} = \partial^{\text{ex}} V$.

Notice that, in the definition (1.1), if the external field has value $h_v \in \mathbb{Z}$, we may interpret the external field as some exterior boundary condition, where the set $\{v, h_v \neq 0\}$ can be interpreted as the boundary. Indeed, it corresponds to adding $|h_v|$ extra edges to v , all leading to vertices with assigned value $\text{sign}(h_v)$.

Setting of the paper. In the following, we focus on the Ising model (1.1) with external field with $h_v \in \{0, 1\}$; in fact, we will consider a random external field with either $(h_v)_{v \in V}$ or $(h_v)_{v \in \partial V}$ given by i.i.d. Bernoulli random variables of parameter $p \in (0, 1)$. With an abuse of terminology, this corresponds to adding a plus (exterior) boundary condition to the vertices $v \in V$ with $h_v = +1$; it indeed corresponds to adding exactly one extra edge to v , with a plus on the other side of the edge. We refer to Figure 2 for an illustration.

¹For instance, if G is a subgraph of a graph $\hat{G} = (\hat{V}, \hat{E})$, the natural boundary is $\partial V = \{v \in V, \exists u \in \hat{V} \setminus V \text{ s.t. } u \sim v\}$. If G is a finite tree, one usually takes ∂V as the set of leaves; if $V = \{-n, \dots, n\}^d$, one usually takes $\partial V = \{(v_1, \dots, v_d) \in V, \exists i | v_i| = n\}$.

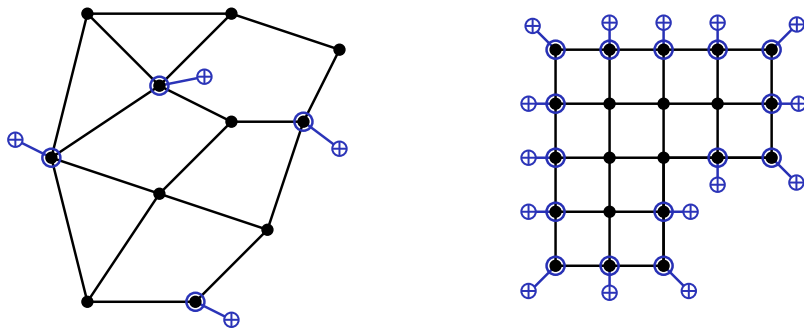


Figure 2: Representation of graphs with external fields $h_v \in \{0, 1\}$. On each graph, we have circled the vertices where $h_v = +1$. The graph on the left has no natural boundary: we can think of the vertices with $h_v = +1$ as the boundary, each vertex being connected to a '+' through an extra edge. On the right, the graph is a subset of \mathbb{Z}^2 and $h_v = +1$ for all $v \in \partial V$: this is our *boundary external field* and in the present case it does not exactly correspond to the Ising model with (exterior) plus boundary condition, because the corner vertices are connected to a '+' through only one extra edge (to obtain plus boundary condition one should take $h_v = +2$ for the corner vertices).

One key physical quantity that we are going to study is the magnetization of a vertex $v \in V$:

$$\mathbf{m}_{G,\beta}^h(v) := \mathbb{E}_{G,\beta}^h[\sigma_v].$$

A closely related quantity is the probability that a given spin is in the plus state, namely $\mathbb{P}_{G,\beta}^h(\sigma_v = +1) = \frac{1}{2}(1 + \mathbf{m}_{G,\beta}^h(v))$ and the following log-likelihood ratio

$$r_{G,\beta}^h(v) = \log \left(\frac{\mathbb{P}_{G,\beta}^h(\sigma_v = +1)}{\mathbb{P}_{G,\beta}^h(\sigma_v = -1)} \right) = \log \left(\frac{1 + \mathbf{m}_{G,\beta}^h(v)}{1 - \mathbf{m}_{G,\beta}^h(v)} \right).$$

Considering a sequence of growing graphs $(G_n)_{n \geq 1}$ with associated non-negative external fields (that may depend on n), we say that there is spontaneous magnetization at inverse temperature β if, choosing for each n a vertex v_n uniformly at random in the graph, the magnetization $\mathbf{m}_{G_n,\beta}^h(v_n)$ (or equivalently the log-likelihood ratio $r_{G_n,\beta}^h(v_n)$) remains bounded away from 0 as $n \rightarrow \infty$ (either almost surely or in probability).

1.2 Ising model on the Configuration Model with interfering vertices

Let us now introduce one of our main motivation for considering a random sparse external field on a Galton–Watson tree: the Ising model on a random graph given by the configuration model, with a small proportion of additional interfering individuals.

The Configuration Model is a random graph in which edges are placed randomly between vertices whose degrees are fixed beforehand. We refer to [29, Ch. 7] for a complete introduction but let us briefly present the construction. Let N be the number of vertices of the graph, and let $\mathbf{d} = (d_i)_{i \in [N]}$ be a sequence of degrees, verifying $d_i \geq 1$; we also use the notation $[N] = \{1, \dots, N\}$. Then, the configuration model, noted $\text{CM}_N(\mathbf{d})$ is an undirected random (multi)graph such that each vertex $i \in [N]$ has degree d_i (self-loops and multiple edges between pairs of vertices are allowed). It is

constructed inductively as follows. As a preliminary to the construction, attach d_i half-edges to each vertex $i \in \llbracket N \rrbracket$, so that there is a collection \mathcal{C}_0 of $\ell_N := \sum_{i=1}^N d_i$ available half-edges. Then, construct the first edge of the graph by choosing two half-edges uniformly at random from \mathcal{C}_0 and by pairing them; afterwards, remove these two half-edges from the set \mathcal{C}_0 . After this first step, the new set \mathcal{C}_1 of available half-edges contains $\ell_N - 2$ elements. This procedure is iterated $\frac{1}{2}\ell_N$ times, until there are no more half-edges available; notice that ℓ_N must be even.

Let us stress that the Configuration Model has no natural boundary, but one may think of having a few additional external vertices that are “interfering” with the graph. To model this, add M_N vertices to the initial N vertices of the model (we think of having $M_N \ll N$), all with degree 1, and call these extra vertices *interfering*. One can then proceed to construct the graph as described above, *i.e.* a configuration model with both original and interfering vertices². Notice that, even if interfering vertices have degree one, an original vertex might have more than one interfering vertex attached to it. The M_N interfering vertices may also be interpreted as some (external) boundary of the graph: in the context of the Ising model, one may consider the model where interfering vertices all have a plus spin, and try to determine a condition whether there is spontaneous magnetization on a sequence of configuration models, depending on M_N, N and the degree sequence \mathbf{d} .

In the case where $M_N \leq N$, another natural (and closely related) way of adding interfering external vertices is to consider the graph $\text{CM}_N(\mathbf{d})$ and, to each vertex $i \in \llbracket N \rrbracket$, attach an extra (interfering) vertex of degree one, with probability $p_N := M_N/N$. One obvious difference from the previous construction is that each vertex has at most one interfering vertex attached; also, in the first construction, interfering vertices change the distribution of the original configuration model. Indeed, an interfering vertex is taking over from an original one in the first construction, instead of just being added afterwards, as in the second construction. In the context of the Ising model, this version corresponds to having a (random) external field given by the spin of interfering vertices, say i.i.d. Bernoulli. Here again, one may ask whether there is spontaneous magnetization on a sequence of configuration models, depending on p_N, N and the degree sequence \mathbf{d} .

To summarize, the general question is to determine how many (or how few) interfering vertices are enough to have some influence on a random individual. Since the configuration model rooted at some randomly chosen vertex o locally behaves like a branching process see [31, Sec. 4.2] (and also Section 1.5.4 below), we consider in the present paper, as a toy model, the Ising model on a Galton–Watson tree with randomly attached interfering vertices, *i.e.* a sparse external Bernoulli field.

1.3 Ising model on a Galton–Watson tree: main results

Let μ be a distribution on \mathbb{N} and consider \mathbf{T}_n a random tree of depth $n \in \mathbb{N}$ generated by a Branching Process with offspring distribution μ , stopped at generation n ; we will

²The fact that interfering vertices have degree 1 ensures that these vertices cannot interfere with each other; but one could naturally consider a degree sequence $(\tilde{d}_i)_{i \in \llbracket M_N \rrbracket}$ for these vertices.

write $\mathbf{T} = \mathbf{T}_n$ if there is no confusion possible. We will denote by ρ its root and by $\partial\mathbf{T}$ the set of its leaves. Also, we denote by \mathbf{P} the law of \mathbf{T} , and we make the following assumption, which ensures in particular that the tree is super-critical.

Assumption 1. *The offspring distribution μ satisfy $\mu(0) = 0$ and $\mu(1) < 1$. In particular, $\nu := \sum_{d=1}^{\infty} d\mu(d) > 1$.*

The condition $\mu(0) = 0$ ensures that a.s. there is no extinction, so the tree \mathbf{T}_n a.s. reaches depth n ; additionally, it has no leaves except at generation n . We could weaken this assumption and work conditionally on having no extinction, using for instance [21, §5.7], but we work with Assumption 1 for technical simplicity.

We consider the Ising model (1.1) on a tree \mathbf{T} with different possibilities for the external field or boundary condition:

- (a) With ‘+’ boundary condition, meaning that $\sigma_v = +1$ (corresponding to an external field $h_v = +\infty$) if $v \in \partial\mathbf{T}$; the results are well-known since the work of Lyons [19], see Theorem 1.3 below. We denote the Ising Gibbs measure $P_{n,\beta}^+$ in this case.
- (b) With (sparse) Bernoulli external field, in which $(h_v)_{v \in \mathbf{T}}$ are i.i.d. Bernoulli random variables with parameter p_n , independent of \mathbf{T} , whose law we denote by \mathbb{P} . We denote by $P_{n,\beta}^{(p_n)}$ the Ising Gibbs measure in this case.
- (c) With (sparse) Bernoulli boundary external field, in which $(h_v)_{v \in \mathbf{T}}$ are i.i.d. Bernoulli variables with parameter $\hat{p}_n := p_n \mathbf{1}_{\{v \in \partial\mathbf{T}\}}$; again, we denote their law by \mathbb{P} , by a slight abuse of notation. We denote by $P_{n,\beta}^{(\hat{p}_n)}$ the Ising Gibbs measure in this case.

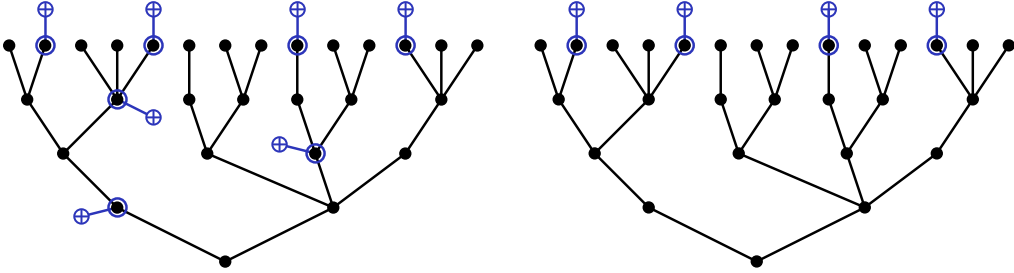


Figure 3: Two Ising models on a Galton–Watson tree with Bernoulli external field $h_v \in \{0,1\}$: the vertices with $h_v = +1$ are circle and a ‘+’ has been added to them. On the left, the external field $h_v \in \{0,1\}$ lives inside the whole tree (model $P_{n,\beta}^{(p_n)}$); on the right, the external field $h_v \in \{0,1\}$ lives only on the leaves (model $P_{n,\beta}^{(\hat{p}_n)}$).

Remark 1.2. The second model (b) mimics the Ising model on a configuration model with sparse interfering vertices³; the third model (c) is interesting in itself and will serve as a point of comparison. We stress that the parameter p_n depends on the depth n of the tree, but is constant among vertices in the tree (no matter their generation).

³Notice that this is not exactly how the local limit of the Ising model on the configuration model would look like. For instance, the configuration model converges locally to a uni-modular branching process, that is, the root has a different offspring distribution from the rest of the vertices.

We then consider the *root magnetization* with external field (or boundary condition) $\mathbf{a} \in \{+, (p_n), (\hat{p}_n)\}$:

$$\mathbf{m}_{n,\beta}^{\mathbf{a}} := \mathbf{m}_{n,\beta}^{\mathbf{a}}(\rho) := E_{n,\beta}^{\mathbf{a}}[\sigma_\rho], \quad (1.3)$$

which is a random variable that depends on the realization of the tree \mathbf{T} and of the field $(h_v)_{v \in \mathbf{T}}$ (if the latter is random). We say that the *root is asymptotically (positively) magnetized* for the model $\mathbf{a} \in \{+, (p_n), (\hat{p}_n)\}$ at inverse temperature β if

$$\lim_{\varepsilon \downarrow 0} \liminf_{n \rightarrow \infty} \mathbf{P} \otimes \mathbb{P}(\mathbf{m}_{n,\beta}^{\mathbf{a}} > \varepsilon) = 1, \quad (1.4)$$

where $\mathbf{P} \otimes \mathbb{P}$ denotes the joint law of \mathbf{T} and $(h_v)_{v \in \mathbf{T}}$. Conversely, the root is *asymptotically not magnetized* if $\mathbf{m}_{n,\beta}^{\mathbf{a}}$ goes to 0 in $\mathbf{P} \otimes \mathbb{P}$ -probability. Finally, note that these statements are equivalent if we replace the root magnetization by the log-likelihood ratio $r_{n,\beta}^{\mathbf{a}} := r_{n,\beta}^{\mathbf{a}}(\rho)$ of the root.

Obviously, one can compare the three models **(a)**, **(b)**, **(c)** above, *i.e.* external field or boundary condition $\mathbf{a} \in \{+, (p_n), (\hat{p}_n)\}$: indeed, we clearly have that $\mathbf{m}_{n,\beta}^+ \geq \mathbf{m}_{n,\beta}^{(\hat{p}_n)}$ and $\mathbf{m}_{n,\beta}^{(p_n)} \geq \mathbf{m}_{n,\beta}^{(\hat{p}_n)}$. We now state our main results.

(a) With a plus boundary condition on the leaves. First of all, we recall the seminal result from Lyons [19] about the phase transition of the Ising model on a tree; we state here only in our simpler context of a Galton–Watson tree.

Theorem 1.3 ([19]). *Consider the Ising model on a Galton–Watson tree \mathbf{T} with plus boundary condition. Then we have root asymptotic magnetization (1.4) (with $\mathbf{a} = +$) at inverse temperature β if and only if $\nu \tanh(\beta) > 1$, where we recall that ν is the mean offspring distribution.*

The general result holds for a generic infinite tree \mathbf{T} : there is root magnetization if and only if $\text{br}(\mathbf{T}) \tanh(\beta) > 1$, where $\text{br}(\mathbf{T})$ is the branching number of the tree \mathbf{T} ; we have $\text{br}(\mathbf{T}) = \nu$ for branching processes. In other words, Theorem 1.3 identifies the critical temperature for the Ising model on a tree: $\beta_c = \tanh^{-1}(\frac{1}{\nu})$.

(b) With a sparse Bernoulli external field inside the tree. Let $(p_n)_{n \geq 0}$ be a sequence of parameters in $[0, 1]$. For each n , we let \mathbf{T}_n be a GW tree up to generation n and we let $(h_v)_{v \in \mathbf{T}_n}$ be i.i.d. Bernoulli random variables of parameter p_n , independent of \mathbf{T}_n . Then, we have the following result, which is the main goal of this article.

Theorem 1.4. *Suppose that Assumption 1 holds and that μ has a finite second moment. Consider the Ising model on \mathbf{T} with sparse Bernoulli external field, with $\lim_{n \rightarrow \infty} p_n = 0$. Then we have asymptotic root magnetization (1.4) (with $\mathbf{a} = (p_n)$) at inverse temperature β if and only if*

$$\nu \tanh(\beta) > 1 \quad \text{and} \quad \liminf_{n \rightarrow \infty} (\nu \tanh(\beta))^n p_n > 0.$$

In the case where $\liminf_{n \rightarrow \infty} p_n > 0$, the root is asymptotically magnetized.

In other words, this theorem gives the exact speed at which $(p_n)_{n \geq 0}$ should decrease in order not to have root magnetization. The first condition in the theorem comes from Lyons' Theorem 1.3; the second condition shows that the sparsity of the Bernoulli field may somehow shift the critical point. For instance, if $p_n = \alpha^n$ for some $\alpha < 1$, then one has root magnetization if and only if $\alpha \nu \tanh(\beta) \geq 1$, so the new critical inverse temperature is $\beta_c = \tanh^{-1}(\frac{1}{\alpha \nu})$; note that in that case, the root is also magnetized at the critical temperature, contrary to what happens in Theorem 1.3.

(c) With a sparse Bernoulli boundary external field on the leaves. We have a similar result when we put the sparse Bernoulli field only on the leaves. As above, for each n , we let \mathbf{T}_n be a GW tree up to generation n and we let $(h_v)_{v \in \mathbf{T}_n}$ be i.i.d. Bernoulli random variables of parameter $\hat{p}_n = p_n \mathbb{1}_{\{v \in \mathbf{T}_n\}}$, independent of \mathbf{T}_n .

Theorem 1.5. *Suppose that Assumption 1 holds and that μ has a finite second moment. Consider the Ising model on \mathbf{T} with sparse Bernoulli boundary external field, with $\lim_{n \rightarrow \infty} p_n = 0$. Then we have asymptotic root magnetization (1.4) (with $a = (\hat{p}_n)$) at inverse temperature β if and only if*

$$\nu \tanh(\beta) > 1 \quad \text{and} \quad \liminf_{n \rightarrow \infty} (\nu \tanh(\beta))^n p_n > 0.$$

In the case where $\liminf_{n \rightarrow \infty} p_n > 0$, the root is asymptotically magnetized if and only if $\nu \tanh(\beta) > 1$.

Remark 1.6. In Theorems 1.4-1.5, the second moment assumption on the offspring distribution is actually not needed to show that if $\nu \tanh(\beta) \leq 1$ or $\lim_{n \rightarrow \infty} (\nu \tanh(\beta))^n p_n = 0$ then the root is asymptotically not magnetized.

Remark 1.7. We could obviously consider a more general sparse random external field. For instance if $(h_v)_{v \in \mathbf{T}}$ are i.i.d. with distribution $(1 - p_n)\delta_0 + p_n\mu_Y$ for some positive random variable Y , one can easily compare this model with a Bernoulli external field and obtain identical results (provided that $\mathbb{E}[Y] < +\infty$). We have chosen to focus on Bernoulli external fields for the simplicity of exposition.

1.4 Outline of the proof and organisation of the paper

Let us outline our strategy of proof, which relies on standard tools for the Ising model on trees, namely Lyons' iteration [19] for the log-likelihood ratio (we recall it in Section 2.2, see (2.5)), and Pemantle–Peres [27] relation between the log-likelihood ratio and the non-linear 3-capacity of the tree, see Section 6.1 for an overview.

After a few preliminaries in Section 2, we prove in Section 3 the upper bound on the log-likelihood ratio and give the starting point of the proof of the lower bound. More precisely, in Section 3.1 we use Lyons' iteration to derive an upper bound on the log-likelihood ratio of the root for the Ising model with Bernoulli external field inside the whole tree (which dominates the case where the external field is only on the leaves). The starting point of the lower bound is presented in Section 3.2: we show that the log-likelihood ratio of the root for the Ising model with Bernoulli boundary external field is

equivalent to that of the Ising model with plus boundary external field on a modified tree, that we call *pruned tree* since it corresponds to removing all branches that do not lead to some $h_v = +1$. In other words, the Ising model with random (Bernoulli) boundary external field on \mathbf{T}_n corresponds to an Ising model with plus boundary condition on a random subtree \mathbf{T}_n^* of \mathbf{T}_n . Then, the rest of the paper consists in studying the Ising model with plus boundary condition on the pruned tree, that we denote \mathbf{T}_n^* .

First of all, in Section 4, we show that, under $\mathbf{P} \otimes \mathbb{P}$ the pruned tree is actually an *inhomogeneous* Branching Process, whose offspring distributions $(\mu_k^*)_{0 \leq k \leq n-1}$ are explicit (see (4.2)) and depend on the generation k but also on the depth n of the tree — in other words, we have a triangular array of offspring distributions.

Then, in Section 5 we show that the pruned tree somehow exhibits a sharp phase transition. More precisely, there exists some $k^* := \log(p_n \nu^n) / \log \nu$ (which depends on n and go to $+\infty$ if $\liminf_{n \rightarrow \infty} p_n (\nu \tanh \beta)^n > 0$), such that:

- if $k^* - k$ is large, then μ_k^* is very close to the original offspring distribution μ ;
- if $k - k^*$ is large, then μ_k^* is very close to being a Dirac mass at 1.

This statement is made precise (and quantitative) in Proposition 5.7. Additionally, Section 5 contains several technical results quantifying this phase transition for various quantities of the pruned tree (for instance the mean and variances of μ_k^*), that turn out to be important for the last part of the proof.

Finally, Section 6 concludes the proof of the lower bound on the log-likelihood ratio in Theorem 1.5 (which implies the lower bound in Theorem 1.4). Relying on the work of Pemantle–Peres [27], the log-likelihood ratio of the root for the Ising model on \mathbf{T}_n^* with plus boundary condition is of the same order as the (non-linear) L^3 - or 3-capacity of \mathbf{T}_n^* , see Theorem 6.4. Hence, Section 6 focuses on estimating $\text{Cap}_3(\mathbf{T}_n^*)$; the estimates are collected in Proposition 6.5. Using that the L^3 -capacity is bounded from below by the (linear) 2-capacity, which coincide with the usual notion of conductance for resistor networks, Thomson’s principle enables us to obtain (after a few technical estimates) a lower bound for $\text{Cap}_3(\mathbf{T}_n^*)$, which concludes the proof. For the sake of completeness, we also provide in Proposition 6.5 an upper bound on $\text{Cap}_3(\mathbf{T}_n^*)$; in particular, we aim at giving a general scheme of proof to estimate the L^3 -capacity of any inhomogeneous Galton–Watson tree. As a side result of independent interest, we find for instance that for the Ising model on a Galton–Watson tree (with offspring distribution of finite variance) with plus boundary condition, then at the critical temperature $\beta_c = \tanh^{-1}(\frac{1}{\nu})$ the magnetization of the root lies between n^{-1} and $n^{-1/2}$, see Remark 6.7, which seems to be a new result (we believe that the correct order is $n^{-1/2}$, at least if the offspring distribution admits enough moments, but we leave this to another work since it is not the main focus of the paper).

1.5 Some further comments

Let us now conclude this section with several comments on our results, suggesting for instances possible directions for further investigations.

1.5.1 Comparison with a homogeneous but vanishing external field

A first natural question is to know whether the same results as in Theorems 1.4-1.5 would hold if one replaced the random Bernoulli external field $(h_v)_{v \in \mathbf{T}}$ by its mean $(\mathbb{E}[h_v])_{v \in \mathbf{T}}$. This corresponds to considering the Ising model on a tree with a *homogeneous* but vanishing external field (either inside the whole tree or only on the boundary).

In particular, we want to look at the two following (sequences of) Ising models on the tree \mathbf{T} of depth n : (i) homogeneous external field $h_v \equiv p_n$ for all $v \in \mathbf{T}$. (ii) homogeneous boundary external field $h_v \equiv p_n$ for all $v \in \partial \mathbf{T}$.

Using Lyons iteration as in Section 3.1 would yield that there is no root magnetization in the first model (hence in the second model) whenever $p_n \rightarrow 0$ and $p_n(\nu \tanh(\beta))^n \rightarrow 0$. We actually believe that if $p_n \rightarrow 0$ and $\liminf_{n \rightarrow \infty} p_n(\nu \tanh(\beta))^n > 0$, then the root is asymptotically magnetized in the second model (hence in the first one). This does not appear straightforward, but one should be able to adapt our proof in Section 6 to derive such a result. Indeed in view of Lyons' iteration (2.5), the model corresponds to some Ising model with plus boundary condition on some modified *elongated* tree $\tilde{\mathbf{T}}_n$, where $\tilde{\mathbf{T}}_n$ is obtained from \mathbf{T}_n by adding “straight branches” of length ℓ to each leaf of \mathbf{T}_n (with ℓ roughly chosen so that $\tanh(\beta)^\ell = p_n$), *i.e.* $\tilde{\mathbf{T}}_n$ is an inhomogeneous Branching Process of depth $n + \ell$ with offspring distribution $\mu_k = \mu$ for generations $k \leq n - 1$ and $\mu_k = \delta_1$ for $n \leq k \leq n + \ell - 1$. Then, thanks to Pemantle–Peres [27] comparison theorem, the problem is reduced to estimating the 3-capacity of $\tilde{\mathbf{T}}_n$. For this, one may use a similar approach as in Section 6; in particular, we believe that the same estimates obtained for the 3-capacity of our pruned tree \mathbf{T}_n^* (cf. Proposition 6.5) hold for the elongated tree $\tilde{\mathbf{T}}_n$. we do not develop further on this issue since it is not the main purpose of our paper.

1.5.2 About the moment condition on the offspring distribution

Obviously, there are a few limitations on the generality of the offspring distribution μ that we consider on our Galton–Watson tree. The main restriction that we have is that μ admits a finite second moment. This assumption is useful when we estimate the effective resistance of the pruned tree \mathbf{T}_n^* via a Thomson's principle, see Section 6.2.2; in particular, by applying Markov's inequality, we are reduced to estimating second moments of the size of the tree.

Adapting our proof to the case of an infinite variance (and possibly of an infinite mean) is an interesting question to consider. It is reasonable to expect that our main results remain valid in the case of an infinite variance, but this seems technically demanding; for instance one would need to control the tail of the random variables appearing in (6.11).

In the same spirit, another interesting question would be to obtain sharper bounds on the 3-capacity of the pruned tree \mathbf{T}_n^* (or even of \mathbf{T}_n), hence of the log-likelihood ratio of the root, for instance using a Thomson's principle for the non-linear capacity. We believe that this improvement would for instance show that for the Ising model with plus boundary condition, at the critical temperature $\beta_c = \tanh^{-1}(1/\nu)$, the log-likelihood ratio (hence the magnetization) is of order $1/\sqrt{n}$, at least in the case where

μ admits a finite *third* moment (since the L^3 Thomson principle should make third moments appear). What happens when the offspring distribution μ fails to have a third moment is an extremely interesting question.

1.5.3 Starting with an inhomogeneous, n -dependent, Galton–Watson tree

Keeping in mind our application to the Configuration Model (see Section 1.5.4 below), let us mention that our method appear to be robust to the case where the initial Galton–Watson tree \mathbf{T}_n of depth n is homogeneous but with an offspring distribution $\mu^{(n)}$ that depends on n . In particular, we believe that our results hold simply by replacing ν with $\nu^{(n)}$, the mean of the offspring distribution $\mu^{(n)}$, provided that the means $(\nu^{(n)})_{n \geq 0}$, resp. the variances $(\sigma_{\mu^{(n)}}^2)_{n \geq 0}$, remain bounded away from 1, resp. 0, and $+\infty$.

The case where one starts with a Galton–Watson tree \mathbf{T}_n of depth n which is inhomogeneous with offspring distribution $(\mu_k^{(n)})_{0 \leq k \leq n-1}$ should be analogous. One would need to replace the quantity ν^n with $\prod_{k=0}^{n-1} \nu_k^{(n)}$, where $\nu_k^{(n)}$ is the mean of the offspring distribution $\mu_k^{(n)}$; again, one should for instance assume that the means $(\nu_k^{(n)})_{0 \leq k \leq n-1, n \geq 1}$, resp. the variances $(\sigma_{\mu_k^{(n)}}^2)_{0 \leq k \leq n-1, n \geq 1}$, remain bounded away from 1, resp. 0, and $+\infty$.

1.5.4 Back to the configuration model

Coming back to the configuration model, we may try to use our toy model to make some predictions. Consider the configuration model recalled in Section 1.2, with N vertices and degree sequence $\mathbf{d} = (d_i)_{i \in [N]}$. Denote $N_k := |\{i \in [N], d_i = k\}|$ the number of vertices of degree k and let D_N be a random variable whose distribution is given by $P(D_N = k) = N_k/N$, which corresponds to the degree of a vertex chosen uniformly at random. A standard and natural assumption on the model, see [31, § 1.3.3], is that, as $N \rightarrow \infty$, D_N converges in distribution to some random variable D and that we also have convergence of the first two moments of D_N to those of D (assuming that $E[D^2] < +\infty$).

Then, it is known from [31, Thm. 4.1] that, choosing a vertex o uniformly at random and rooting $\text{CM}(\mathbf{d})$ at o , the rooted $\text{CM}(\mathbf{d})$ locally converges (*i.e.* can be locally coupled with) to a branching process with offspring distribution μ (except the root which has offspring distribution D), where μ is the distribution of $D^* - 1$, with D^* is the size-biased of D ; more precisely $\mu(k) := \frac{k+1}{E[D]} P(D = k+1)$.

Therefore, Assumption 1 corresponds to having $P(D = 1) = 0$, $P(D = 2) < 1$ and also $\nu := \sum_{k=1}^{\infty} k\mu(k) = \frac{1}{E[D]} E[D(D-1)] > 1$ (note that $\nu < +\infty$ if $E[D^2] < +\infty$). The assumption that μ admits a finite second moment translates into the requirement that $E[D^3] < +\infty$. Then, we can try to apply our results simply by analogy, *i.e.* identifying the configuration model rooted at a random vertex v with a Galton–Watson tree \mathbf{T}_n with offspring distribution μ (except at its root) and depth $n = \frac{1}{\log \nu} \log N$; the choice of n is such that the number of vertices in \mathbf{T}_n is roughly $\nu^n = N$. Our results then translate into the fact that the vertex o is asymptotically magnetized if and only if $\liminf_{n \rightarrow \infty} p_n \nu^n \tanh(\beta)^n > 0$, or since $p_n = M_N/N$, if and only if $\liminf_{N \rightarrow \infty} M_N N^\alpha > 0$, where $\alpha := \frac{1}{\log \nu} \log \tanh(\beta)$.

However, the approximation of the configuration model rooted at a random vertex o with a Galton–Watson tree only works up to depth $\tilde{n} = c \log N$ for some constant $c > 0$, provided that the maximal degree $d_{\max} = \max_{v \in \llbracket N \rrbracket} d_v$ verifies $d_{\max} = O(N^a)$ for some $a < 1$. This is a reformulation of [30, Lem. 3.3 and Rem. 3.4] in our context: a coupling can be made between the configuration model $\text{CM}(\mathbf{d})$ rooted at a random vertex o and a (N -dependent) Branching Process up to $m_N = o(\sqrt{N/d_{\max}})$ vertices; if $d_{\max} = N^{o(1)}$, this corresponds to a depth $\tilde{n} = \frac{1}{2 \log \nu} \log N$ for the Galton–Watson tree. The fact that loops start to appear in the graph at some point breaks Lyons iteration’s argument and new ideas are needed. However, one should be able to obtain at least some bounds on the magnetization of o ; in particular, the fact that our result is robust to the case of a n -dependent Galton–Watson tree could prove useful when working with the coupling mentioned above.

A natural (weak form of the) conjecture is the following.

Conjecture 1. *For the Ising model on the configuration model $\text{CM}(\mathbf{d})$ with M_N interfering external ‘+’ vertices, there exists some $\tilde{\alpha} > 0$, depending only on the inverse temperature β and the mean ν , such that a randomly chosen vertex o in $\text{CM}(\mathbf{d})$ is:*

- *asymptotically magnetized if $M_N \geq N^{-\tilde{\alpha}+\varepsilon}$ for some $\varepsilon > 0$;*
- *asymptotically not magnetized if $M_N \leq N^{-\tilde{\alpha}-\varepsilon}$ for some $\varepsilon > 0$.*

In other words, the threshold for having asymptotic root magnetization should be at a polynomial number of “interfering” vertices $M_N = N^{-\tilde{\alpha}+o(1)}$; it is natural to guess that $\tilde{\alpha} \leq \alpha$ with $\alpha := \frac{1}{\log \nu} \log \tanh(\beta)$, but it is not clear whether $\tilde{\alpha} = \alpha$ or not, which is an interesting question.

1.5.5 About free boundary conditions and extremal Gibbs states

In the present paper, we focus on a non-negative (boundary) external field. Indeed, it is natural to consider such a condition to break the $+/-$ symmetry of the model.

Another setting, that has been extensively studied both in the physics and mathematics literature, is to consider the (nearest-neighbor, ferromagnetic) Ising model on an infinite tree with zero external field and free boundary condition. One question is then to determine whether the free measure \mathbf{P}_β^f is equal to $\frac{1}{2}(\mathbf{P}_\beta^+ + \mathbf{P}_\beta^-)$, with \mathbf{P}_β^+ , \mathbf{P}_β^- the measures with ‘+’ and ‘−’ boundary conditions respectively. More generally, the question is to understand the set of extremal Gibbs measures (also called pure states), and in particular whether this set is reduced to \mathbf{P}_β^+ , \mathbf{P}_β^- .

On a k -regular tree \mathbb{T}_k (or more generally on hyperbolic graphs, see e.g. [33, 34]), it has been shown that at sufficiently low temperature, there are uncountably many extremal Gibbs states⁴, see [18] and [2, 16] for rigorous results (we also refer to [8] for the case of other finite-spin models). In particular, there are multiple phase transitions;

⁴In [25], the authors consider a branching plane lattice $\mathbb{Z} \times \mathbb{T}_k$ and outline the difference between the multiplicity of extremal Gibbs states on a tree in contrast with \mathbb{Z}^d , where all translation invariant Gibbs states are mixtures of \mathbf{P}_β^+ , \mathbf{P}_β^- , see [3].

in this context, let us mention [24] which shows that the free Ising measure is a factor of IID beyond the uniqueness regime (see also [20] for a wider introduction to the problem).

It is then reasonable to ask whether the results on k -regular trees remain valid on random trees, or on tree-like graphs. As an example of such study, in [23], the authors consider the free Ising measure on a growing sequence of graphs $(G_n)_{n \geq 1}$ that locally converge to a k -regular tree: their main result is that the Ising measure locally weakly converges to the mixture $\frac{1}{2}(P_\beta^+ + P_\beta^-)$. Our present work raises the natural question to determine whether (and to which point) this result continues to hold if one adds a *signed* sparse (boundary) random external field, *i.e.* if $(h_v)_{v \in \mathbf{T}}$ (or $(h_v)_{v \in \partial \mathbf{T}}$) are i.i.d. with law $\mathbb{P}(h_v = \pm 1) = \frac{1}{2}p_n$, $\mathbb{P}(h_v = 0) = 1 - p_n$, with $\lim_{n \rightarrow \infty} p_n = 0$. This is in fact related to the question of the effect of a random boundary condition on the Ising model (in particular on the coexistence of pure states), in the spirit of [32], see also [14] for a recent overview. We believe that such questions are natural and promising directions of research, in continuity of the present paper.

2 A few preliminaries

2.1 Some notation for trees

Let $\mathbf{t} = (V(\mathbf{t}), E(\mathbf{t}), \rho)$ be a tree⁵ with root ρ ; with an abuse of notation, we also write $v \in \mathbf{t}$ as a shorthand for $v \in V(\mathbf{t})$. The distance between two vertices in the tree is the number of edges of the unique path connecting them.

Given two vertices $v, w \in \mathbf{t}$, we say that w is a *descendant* of v , and we denote $v \leq w$, if the vertex v is on the shortest path from the root ρ to the vertex w . For a vertex $v \in \mathbf{t}$, we let $|v|$ denote the distance from v to the root ρ ; note that if we have $v \leq w$, we have $|v| \leq |w|$.

We consider a tree of depth n , that is such that $\max_{v \in \mathbf{t}} |v| = n$.

- We denote $t_k = \{v \in \mathbf{t} : |v| = k\}$ the k -th generation of the tree \mathbf{t} , for $k \leq n$.
- Given two vertices $v, w \in \mathbf{t}$ we say w is a (direct) successor of v , and we note $v \rightarrow w$, if $v < w$ and $|v| = |w| - 1$.
- For a vertex $v \in \mathbf{t}$, we denote $S(v) = \{w \in \mathbf{t} : v \rightarrow w\}$ the set of (direct) successors of v , and $d(v) = |S(v)|$ its cardinal, *i.e.* the number of descendants of v .
- We say that a vertex $v \in \mathbf{t}$ is a leaf if it has no successor, namely if $d(v) = 0$, and we denote $\partial \mathbf{t} = \{v \in \mathbf{t}, d(v) = 0\}$ the set of leaves of the tree.
- For a vertex $v \in \mathbf{t}$, we let $\mathbf{t}(v)$ be the subtree of \mathbf{t} consisting of v (as a root) and all vertices w such that $v \leq w$; if $v \in t_k$ is in generation k , we may also use the notation $\mathbf{t}_k(v)$ to make the dependence on k explicit.

In the following, thanks to Assumption 1, we only consider trees that have no leaves except at generation n ($\partial \mathbf{t} = t_n$), *i.e.* such that all vertices $|v| < n$ have at least one successor, $d(v) \geq 1$.

⁵Recall that a tree is a connected, acyclic and undirected graph, or equivalently, a graph where every two vertices are connected by exactly one path.

2.2 Log-likelihood ratio and Lyons iteration

One fundamental element of the proof of Theorem 1.3 is the fact that the log-likelihood ratio of the root magnetization can be expressed recursively on the tree. In this section, we introduce a few notation and state this recursive formula, that we call Lyons iteration; for the sake of completeness and because we adapt the iteration in the following, we recall how to obtain it.

Lyons iteration with plus boundary condition. Let us start by giving the definition of the log-likelihood ratio $r_{\mathbf{t},\beta}^+(\rho)$ of the root on a tree \mathbf{t} of depth n , with (classical) plus boundary conditions at temperature β :

$$r_{n,\beta}^+(\rho) := \log \left(\frac{P_{n,\beta}^+(\sigma_\rho = +1)}{P_{n,\beta}^+(\sigma_\rho = -1)} \right)$$

We also introduce the log-likelihood ratio $r_{n,\beta}^+(u)$ of a vertex $u \in \mathbf{t}$. We define the partition function of the Ising model on the sub-tree $\mathbf{t}(u)$ with plus boundary conditions at temperature β , conditioned on the vertex u (the root of $\mathbf{t}(u)$) having spin $a \in \{-1, 1\}$:

$$Z_{\beta}^{+,a}(u) = \sum_{\substack{\sigma \in \{-1,1\}^{|\mathbf{t}(u)|} \\ \sigma_u = a}} \exp \left(\beta \left(\sum_{\substack{v,w \in \mathbf{t}(u) \\ v \sim w}} \sigma_v \sigma_w + \sum_{v \in \partial \mathbf{t}(u)} \sigma_v \right) \right). \quad (2.1)$$

Then, we let

$$r_{n,\beta}^+(u) = \log \left(\frac{Z_{\beta}^{+,+1}(u)}{Z_{\beta}^{+,-1}(u)} \right), \quad (2.2)$$

which corresponds to the log-likelihood ratio of the root of the Ising model in the sub-tree $\mathbf{t}(u)$, of depth $n - |u|$, with plus boundary condition (we keep the subscript n to remember that $u \in \mathbf{t}$ with \mathbf{t} of depth n). Let us notice that, for any vertex $u \in \mathbf{t}$ different from the root, $r_{n,\beta}^+(u)$ does not correspond to $\log(P_{n,\beta}^+(\sigma_v = +1)/P_{n,\beta}^+(\sigma_v = -1))$.

After a straightforward computation, for $u \in \mathbf{t} \setminus \partial \mathbf{t}$, we can write $(Z_{n,\beta}^{+,a}(u))_{a \in \{-1, +1\}}$ in terms of the partition functions $(Z_{n,\beta}^{+,a}(v))$ with v successors of u : we have

$$Z_{\beta}^{+,a}(u) = \prod_{v, u \rightarrow v} \left(e^{+\beta a} Z_{\beta}^{+,+1}(v) + e^{-\beta a} Z_{\beta}^{+,-1}(v) \right). \quad (2.3)$$

We can therefore express the log-likelihood ratio $r_{n,\beta}^+(u)$ in terms of partition functions:

$$r_{n,\beta}^+(u) = \log \left(\frac{Z_{\beta}^{+,+1}(u)}{Z_{\beta}^{+,-1}(u)} \right) = \sum_{u \rightarrow v} \log \left(\frac{e^{+\beta} Z_{n,\beta}^{+,+1}(v) + e^{-\beta} Z_{n,\beta}^{+,-1}(v)}{e^{-\beta} Z_{n,\beta}^{+,+1}(v) + e^{+\beta} Z_{n,\beta}^{+,-1}(v)} \right). \quad (2.4)$$

Defining $g_{\beta}(x) := \log \left(\frac{e^{2\beta} e^x + 1}{e^{2\beta} + e^x} \right)$, we therefore end up with the following crucial recursion:

$$r_{n,\beta}^+(u) = \begin{cases} \sum_{u \rightarrow v} g_{\beta}(r_{n,\beta}^+(v)) & \text{if } u \notin \partial \mathbf{t}, \\ +\infty & \text{if } u \in \partial \mathbf{t}. \end{cases} \quad (2.5)$$

Lyons iteration with an external field. We can also define, analogously to (2.2), the log-likelihood ratio of a vertex $u \in \mathbf{t}$ for the Ising model with external field $h = (h_v)_{v \in \mathbf{t}}$.

For $u \in \mathbf{t}$, let $Z_\beta^{h,a}(u)$ denote the partition function of the Ising model on $\mathbf{t}(u)$ at temperature β with external field $(h_v)_{v \in \mathbf{t}(u)}$, conditioned on the root having spin $a \in \{-1, +1\}$. In the same way as in (2.3), we have for $u \in \mathbf{t} \setminus \partial \mathbf{t}$

$$Z_\beta^{h,a}(u) = e^{\beta h_u a} \prod_{v, u \rightarrow v} \left(e^{+\beta a} Z_\beta^{h,+1}(v) + e^{-\beta a} Z_\beta^{h,-1}(v) \right).$$

Then, as in (2.4), the log-likelihood ratio $r_{n,\beta}^h(u)$ can be expressed in terms of the conditional partition functions, and we end up with the following recursion, analogous to (2.5):

$$r_{n,\beta}^h(u) = \begin{cases} 2\beta h_u + \sum_{u \rightarrow v} g_\beta(r_{n,\beta}^h(v)) & \text{if } u \notin \partial \mathbf{t}, \\ 2\beta h_u & \text{if } u \in \partial \mathbf{t}. \end{cases} \quad (2.6)$$

Here, we have also used that if $u \in \partial \mathbf{t}$ is a leaf, the log-likelihood ratio $r_{n,\beta}^h(u)$ is easily seen to be $2\beta h_u$.

Remark 2.1. Note that the Ising model on \mathbf{t} with a plus boundary external field $h_+ = (h_v)_{v \in \partial \mathbf{t}}$, i.e. $h_v = +1$ for all $v \in \partial \mathbf{t}$, corresponds to the Ising model with plus boundary condition on some extended tree, obtained by adding an extra generation to \mathbf{t} with exactly one descendant to all $v \in \partial \mathbf{t}$, recall Figure 2. In this setting, the log-likelihood ratios $r_{n,\beta}^{h_+}(u)$ verify the same recursion as in (2.5) but with a different initial condition on the leaves:

$$r_{n,\beta}^{h_+}(v) = 2\beta \quad \text{if } v \in \partial \mathbf{t}. \quad (2.7)$$

2.3 The case of a non-vanishing sequence $(p_n)_{n \geq 0}$

Before we move to the proofs, let us evacuate the case of a non vanishing sequence $(p_n)_{n \geq 0}$ in the different models of Section 1.3.

Lemma 2.2. *Let \mathbf{T} be a branching process satisfying Assumption 1 and consider the Ising model with sparse Bernoulli external field inside the tree, i.e. model (b) of Section 1.3, with parameters $(p_n)_{n \geq 0}$ in $[0, 1]$. If $\liminf_{n \rightarrow \infty} p_n > 0$, then the root is asymptotically magnetized.*

Proof. To prove this, it is enough to consider the event where $h_\rho = +1$. On the event that $h_\rho = 1$ and since we have $r_n^{(p_n)}(v) \geq 0$ for all $v \leftarrow \rho$, we get that $r_n(\rho) \geq 2\beta$. Hence, for $\varepsilon < 2\beta$ we have $\mathbf{P} \otimes \mathbb{P}(r_n(\rho) > \varepsilon) \geq \mathbb{P}(h_\rho = 1) = p_n$. For such ε , we get that

$$\liminf_{n \rightarrow \infty} \mathbf{P} \otimes \mathbb{P}(r_n(\rho) > \varepsilon) > 0,$$

which concludes the proof. □

In the case of the Ising model with non-vanishing Bernoulli boundary external field, *i.e.* model (c) of Section 1.3 with $\liminf_{n \rightarrow \infty} p_n > 0$, one obtains a similar result as with plus boundary conditions, that is Theorem 1.3: the root is asymptotically magnetized if and only if $\nu \tanh(\beta) > 1$. This is actually a corollary of Theorem 6.4 (from [27]) and our bounds on the 3-capacity of the pruned tree, see Proposition 6.5 which is also valid for a non-vanishing sequence $(p_n)_{n \geq 0}$.

3 Upper bound and main arguments for the lower bound

In the section, we consider the Ising model with sparse external field inside the tree, *i.e.* model (b) of Section 1.3. We let $(p_n)_{n \geq 0}$ be a sequence in $[0, 1]$ and we give an upper bound on the log-likelihood ratio and a lower bound in terms of an Ising model on a *pruned* tree with plus boundary external field.

3.1 Upper bound: Lyons argument

We show here the following proposition, using Lyons' iteration (2.6), which gives a sufficient condition for having no root magnetization.

Proposition 3.1. *Suppose that Assumption 1 holds. Consider the Ising model on \mathbf{T} with sparse Bernoulli external field, with $\lim_{n \rightarrow \infty} p_n = 0$.*

Then if $\lim_{n \rightarrow \infty} (\nu \tanh(\beta))^n p_n = 0$, there is no asymptotic root magnetization. More precisely, there is a constant $C = C_{\beta, \nu}$ such that

$$\mathbf{E} \otimes \mathbf{E} [r_{n, \beta}^{p_n}(\rho)] \leq C \begin{cases} p_n & \text{if } \tanh(\beta) \nu < 1, \\ \sqrt{p_n} & \text{if } \tanh(\beta) \nu = 1, \\ p_n (\nu \tanh(\beta))^n & \text{if } \tanh(\beta) \nu > 1. \end{cases}$$

Proof. This is a simple application of the recursion formula (2.6). Let us first give a simple bound, valid for a generic tree \mathbf{t} and external field $(h_v)_{v \in \mathbf{t}}$.

Using the easy bound $g_\beta(x) \leq \tanh(\beta)x$ for all $x \geq 1$ (note that $\tanh(\beta)$ is the derivative of g_β at 0), we obtain from (2.6) that for any $u \in \mathbf{t} \setminus \partial \mathbf{t}$,

$$r_{n, \beta}^h(u) \leq 2\beta h_u + \tanh(\beta) \sum_{u \rightarrow v} r_{n, \beta}^h(v).$$

Applying this inequality recursively we obtain the following upper bound for the log-likelihood ratio of the root ρ :

$$r_{n, \beta}^h(\rho) \leq \sum_{k=0}^n \tanh(\beta)^k \sum_{u \in t_k} 2\beta h_u.$$

In our specific setting where $(h_v)_{v \in \mathbf{T}}$ is an i.i.d. Bernoulli field with parameter p_n , we need to show that the upper bound goes to 0 in $\mathbb{P} \otimes \mathbf{P}$ probability. We simply take

its expectation with respect to $\mathbf{P} \otimes \mathbb{P}$: we obtain

$$\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(\rho)] \leq 2\beta p_n \sum_{k=0}^n (\tanh(\beta)\nu)^k. \quad (3.1)$$

- In the case $\tanh(\beta)\nu < 1$, the last sum is bounded by a constant and we have that $\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(\rho)] \leq cp_n$, which goes to 0.
- In the case $\tanh(\beta)\nu > 1$, we have

$$\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(\rho)] \leq \frac{2\beta}{\tanh(\beta) - 1} p_n ((\tanh(\beta)\nu)^n - 1) \leq c_{\beta,\nu} p_n (\tanh(\beta)\nu)^n,$$

which goes to 0 under the assumptions of the proposition.

- The case $\tanh(\beta)\nu = 1$ is a bit more subtle: taking the expectation in (3.1) gives the upper bound $\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(\rho)] \leq 2\beta np_n$, which goes to 0 only if $p_n = o(\frac{1}{n})$. In the general case, we need some extra work. Going back to Lyons' iteration (2.6) and using that $x \mapsto g_\beta(x)$ is concave, we get by Jensen's inequality (recalling that ν is the mean offspring distribution) that

$$\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(u)] \leq 2\beta p_n + \nu g_\beta(\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(v)]) \quad \text{for } u \notin \partial \mathbf{T},$$

where u, v are generic vertices at successive generations; note that the expectation $\mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(u)]$ depends only on (the distribution of) the subtree $\mathbf{t}(u)$ and $(h_v)_{v \in \mathbf{t}(u)}$, so in fact only on $|u|$. Hence, setting $y_k := \mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(u)]$ with $|u| = k$, we end up with the following iteration

$$y_k \leq \begin{cases} f_{n,\beta}(y_{k+1}) & \text{if } 0 \leq k < n, \\ 2\beta p_n & \text{if } k = n. \end{cases} \quad (3.2)$$

with $f_{n,\beta}(x) = 2\beta p_n + \nu g_\beta(x)$, which is a concave increasing function. Let us denote x_n the unique solution of $f_{n,\beta}(x_n) = x_n$. Then, it is clear that either the initial condition verifies $y_n \leq x_n$ and then $y_k \leq x_n$ for all $k \leq n$, or it verifies $y_n > x_n$ and then $y_k \leq y_n = 2\beta p_n$ for all $k \leq n$. We have therefore proven that

$$y_0 = \mathbf{E} \otimes \mathbb{E}[r_{n,\beta}^{(p_n)}(\rho)] \leq \max\{2\beta p_n, x_n\},$$

and it remains to show that the fixed point x_n goes to 0 as $n \rightarrow \infty$. But this should be clear, since $x \mapsto \nu g_\beta(x)$ is a strictly concave function with slope $\nu \tanh(\beta) = 1$ at the origin.

In fact, let us prove that $x_n = O(\sqrt{p_n})$. First, note that $\nu g_\beta(x) \leq x - c$ for any $x \geq 1$ with the constant $c = 1 - \nu g_\beta(1) > 0$ (by concavity), so that $f_{n,\beta}(x) \leq 2\beta p_n + x - c$ for all $x \geq 1$. If n is large enough so that $2\beta p_n < c$, this inequality cannot be verified at $x = x_n$, so we must have $x_n < 1$. Then, writing that there is some constant $c_\beta > 0$ such that $\nu g_\beta(x) \leq x - c_\beta x^2$ for all $x \in (0, 1)$, we get that $x_n = f_{n,\beta}(x_n) \leq 2\beta p_n + x_n - c_\beta x_n^2$ for all n large enough. We conclude easily that $x_n \leq c'_\beta \sqrt{p_n}$ for n large, which is what was claimed. \square

3.2 Lower bound: comparison with a pruned tree

First of all, we use the Ising model with Bernoulli boundary external field $(h_v)_{v \in \partial \mathbf{T}}$, *i.e.* model (c) in Section 1.3, to get a lower bound for the magnetization of the Ising model with Bernoulli external field in the whole tree, since the magnetization is clearly higher in the second case.

Now, let us show in this section that, conditionally on the realization \mathbf{t} of the Galton–Watson tree \mathbf{T} and on the realization h of a Bernoulli external field $(h_v)_{v \in \partial \mathbf{T}}$, the magnetization is equal to the magnetization on a *pruned* (sub)-tree \mathbf{t}^* with plus boundary external field $h_+ = (h_v)_{v \in \partial \mathbf{t}^*}$, *i.e.* with $h_v = +1$ for all $v \in \partial \mathbf{t}^*$. Informally, the pruned tree \mathbf{t}^* is obtained by removing all branches in \mathbf{t} that do not lead to a leaf with $h_v = +1$; we call this procedure the *pruning of dead branches*, see Figure 4 for an illustration.

To be more formal, given a tree \mathbf{t} of depth n and $(h_v)_{v \in \partial \mathbf{t}}$ an external field with value in $\{0, 1\}$, let us define indicator variables $(Y_u)_{u \in \mathbf{t}}$ as follows:

$$Y_u = \begin{cases} +1 & \text{if there exists } v \geq u, v \in \partial \mathbf{t} \text{ such that } h_v = +1 \\ 0 & \text{otherwise.} \end{cases} \quad (3.3)$$

We interpret having $Y_u = +1$ as the fact that the vertex u belongs to a living branch, and so has “survived to the pruning”. Alternatively, we can construct $(Y_u)_{u \in \mathbf{t}}$ iteratively starting from the leaves:

- for $v \in \partial \mathbf{t} = t_n$, set $X_v := h_v$;
- iteratively, for $u \in \mathbf{t} \setminus \partial \mathbf{t}$, we set $Y_u = +1$ if and only if there is some $v \leftarrow u$ with $Y_v = +1$.

We then define the *pruned tree* as

$$\mathbf{t}^* = \text{Pruned}_h(\mathbf{t}) := \{v \in \mathbf{t}, Y_v = +1\}. \quad (3.4)$$

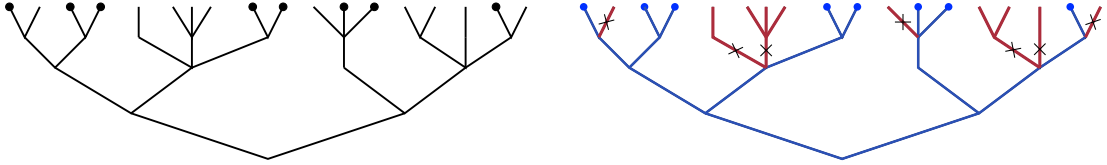


Figure 4: Illustration of the pruning procedure. On the left, we have represented the tree \mathbf{t} and the leaves with $h_v = +1$ with dots. The picture on the right represents how the pruning procedure proceeds: in blue are all the “living branches” leading to a leaf with $h_v = +1$; in red are all the “dead branches” leading to $h_v = 0$, that have to be pruned. The pruned tree \mathbf{t}^* consists in keeping only the living (blue) branches.

Let us consider two following log-likelihood ratios, at inverse temperature β :

- For $u \in \mathbf{t}$, $r_{\mathbf{t}, \beta}^h(u)$ is the log-likelihood ratio for the Ising model on \mathbf{t} with external field $h \in \{0, 1\}^{\partial \mathbf{t}}$ on the boundary;
- For $u \in \mathbf{t}^*$, $r_{\mathbf{t}^*, \beta}^{h_+}(u)$ is the log-likelihood ratio for the Ising model on \mathbf{t}^* with external field $h_+ \equiv 1$ on the boundary $\partial \mathbf{t}^*$ (note that we have $h_+ = h|_{\partial \mathbf{t}^*}$).

We now show the following lemma.

Lemma 3.2. *We have $r_{\mathbf{t},\beta}^h(u) = 0$ for any $u \in \mathbf{t} \setminus \mathbf{t}^*$ and $r_{\mathbf{t},\beta}^h(u) = r_{\mathbf{t}^*,\beta}^{h+}(u)$ for any $u \in \mathbf{t}^*$. In particular, the log-likelihood ratio of the root verifies $r_{\mathbf{t},\beta}^h(\rho) = r_{\mathbf{t}^*}^{h+}(\rho)$.*

Proof. We simply need to write the following Lyons recursions (see also Remark 1.1). On the tree \mathbf{t} , we have

$$r_{\mathbf{t},\beta}^h(u) = \begin{cases} \sum_{u \rightarrow v} g_\beta(r_{\mathbf{t},\beta}^h(v)), & \text{if } u \notin \partial \mathbf{t}, \\ 2\beta h_v, & \text{if } v \in \partial \mathbf{t}. \end{cases}$$

Since $g_\beta(0) = 0$, this readily proves that $r_{\mathbf{t},\beta}^h(u) = 0$ if all $v \in \partial \mathbf{t}$ descendants of u verify $h_v = 0$, i.e. if $Y_u = 0$ and u has been pruned ($u \in \mathbf{t} \setminus \mathbf{t}^*$). On the tree \mathbf{t}^* , we have

$$r_{\mathbf{t}^*,\beta}^{h+}(u) = \begin{cases} \sum_{u \rightarrow v} g_\beta(r_{\mathbf{t}^*,\beta}^{h+}(v)), & \text{if } v \notin \partial \mathbf{t}^*, \\ 2\beta, & \text{if } v \in \partial \mathbf{t}^*. \end{cases}$$

With a slight abuse of notation, we can extend the definition to the tree \mathbf{t} by setting $r_{\mathbf{t}^*,\beta}^{h+}(u) = 0$ if $u \in \mathbf{t} \setminus \mathbf{t}^*$. Now, this extended definition of $r_{\mathbf{t}^*,\beta}^{h+}$ yields exactly the same recursion as $r_{\mathbf{t},\beta}^h(u)$ so $r_{\mathbf{t}^*,\beta}^{h+}(u) = r_{\mathbf{t},\beta}^h(u)$ for all $u \in \mathbf{t}$. \square

With Lemma 3.2, the rest of the paper then consist in studying the Ising model on the pruned version \mathbf{T}_n^* of a Galton–Watson tree \mathbf{T}_n , with plus boundary external field. More precisely, we show the following.

Theorem 3.3. *Let \mathbf{T}_n be a Galton–Watson tree of depth n whose offspring distribution μ satisfies Assumption 1 and has a finite second moment, and let $(p_n)_{n \geq 0}$ be a vanishing sequence. Let \mathbf{T}_n^* be the pruned version of \mathbf{T}_n with $(h_v)_{v \in \partial \mathbf{T}}$ given by i.i.d. Bernoulli random variables with parameter p_n , i.e. $\mathbf{T}_n^* = \text{Pruned}_h(\mathbf{T}_n)$. Then, for the Ising model on \mathbf{T}_n^* with plus boundary external field on $\partial \mathbf{T}_n^*$, the root is asymptotically magnetized if and only if*

$$\nu \tanh(\beta) > 1 \quad \text{and} \quad \liminf_{n \rightarrow \infty} (\nu \tanh(\beta))^n p_n > 0.$$

We proceed in several steps, which are split into the next three sections:

- First, we prove that, under $\mathbf{P} \otimes \mathbb{P}$, the pruned version of a Galton–Watson tree is an inhomogeneous branching process, whose distribution is explicit.
- Second, we study the *shape* of \mathbf{T}_n^* . More precisely, we observe that the pruned tree exhibits a sort of a phase transition: we identify some $k^* = k^*(\mu, p_n)$ such that the pruned tree looks like the original Galton–Watson tree up to $k^* - O(1)$ and then looks like very thin branches from $k^* + O(1)$ to n .
- Lastly, we use our previous observation to estimate the (non-linear) 3-capacity of \mathbf{T}_n^* , which is known to be related to the root magnetization, see [27].

4 Pruning the dead branches of a Galton–Watson tree

In this section, we consider \mathbf{T}_n a random tree of depth $n \in \mathbb{N}$, generated by a branching process with offspring distribution μ such that $\mu(0) = 0$. We let $p_n \in [0, 1]$ be a given parameter and we let \mathbf{T}_n^* be the pruned version of \mathbf{T} with $(h_v)_{v \in \partial \mathbf{T}}$ given by i.i.d. Bernoulli random variables with parameter p_n , as defined above in Section 3.2. Recall that we denote by \mathbf{P} the law of \mathbf{T} and by \mathbb{P} the law of the i.i.d. Bernoulli random variables $(h_v)_{v \in \partial \mathbf{T}}$.

The main result of this section is that \mathbf{T}_n^* , under $\mathbf{P} \otimes \mathbb{P}$, is a Branching Process with *inhomogeneous n -dependent* generation offspring distributions, that we denote $(\mu_k^*)_{0 \leq k \leq n-1}$. The fact that \mathbf{T}_n^* is a branching process is a priori not obvious, since the pruning of a dead branch in the tree depends on random variables at the leaves.

4.1 Notation, statement of the result

Before we state the main result of this section, let us introduce some notation. For $k \in \{0, \dots, n\}$, we denote T_k , respectively T_k^* , the k -th generation of \mathbf{T} , respectively \mathbf{T}_n^* .

We set, for $k \in \{0, \dots, n\}$,

$$\gamma_k := \mathbf{P} \otimes \mathbb{P}(Y_u = 0) \quad \text{for a generic } u \in T_k, \quad \text{and} \quad \delta_k := 1 - \gamma_k, \quad (4.1)$$

where $Y_u = \mathbb{1}_{\{u \in \mathbf{T}_n^*\}}$ is a Bernoulli random variable that indicates whether the vertex u has survived the pruning, see (3.3).

Remark 4.1. Here, γ, δ are indexed by their generations starting from the root (*i.e.* from the bottom of the tree to the top), but it is also helpful to index quantities by their distances to the leaves (*i.e.* from the top of the tree to the bottom). In the rest of the paper, we will denote with a bar quantities that are indexed by the distance to the leaves: for instance, for $k \in \{0, \dots, n\}$, we set

$$\bar{\gamma}_k := \gamma_{n-k}, \quad \bar{\delta}_k = \delta_{n-k} = 1 - \bar{\gamma}_k.$$

Let G be the generating function for the offspring distribution, that is

$$G(s) := \mathbf{E}[s^X] = \sum_{d=1}^{\infty} \mu(d) s^d \quad s \in [0, 1],$$

where X is a random variable of distribution μ . Then, the parameters γ_k can easily be expressed iteratively in terms of the generating function G .

Lemma 4.2. *The sequence $(\bar{\gamma}_k)_{0 \leq k \leq n}$ is characterized by the following iteration:*

$$\bar{\gamma}_0 = 1 - p_n \quad \text{and for } k \in \{1, \dots, n\}, \quad \bar{\gamma}_k = G(\bar{\gamma}_{k-1}).$$

Equivalently, we have

$$\bar{\gamma}_k = \mathbf{E}[(1 - p_n)^{|T_k|}].$$

Proof. First of all, we have $\gamma_n = \mathbb{P} \otimes \mathbf{P}(Y_v = 0)$ for v a leaf, so by definition (3.3) of Y_v , we have $\gamma_n = \mathbb{P}(h_v = 0) = 1 - p_n$, i.e. $\bar{\gamma}_0 = 1 - p_n$.

Now, notice that for a generic $u \in T_k$ with $k < n$, by definition of Y_u we have that $Y_u = 0$ if and only if $Y_{u'} = 0$ for all $u' \leftarrow u$. In other words, for a fixed realization of \mathbf{T} we have

$$\mathbb{P}(Y_u = 0) = \prod_{u', u \rightarrow u'} \mathbb{P}(Y_{u'} = 0),$$

since the variables (h_v) are independent on the different sub-trees $\mathbf{T}(u')$, $u' \leftarrow u$. Taking the expectation with respect to the tree and using the branching property we therefore get that for $k \leq n - 1$,

$$\gamma_k = \mathbb{E}[(\gamma_{k+1})^X] = G(\gamma_{k+1}).$$

Using that $\bar{\gamma}_k = \gamma_{n-k}$, this gives the desired iteration.

For the other formula for $\bar{\gamma}_k$, we notice that $\bar{\gamma}_k = G^{\circ k}(1 - p_n)$, with $G^{\circ k} = G \circ \dots \circ G$ (k times). We obtain the desired formula since $G^{\circ k}$ is the generating function of $|T_k|$. \square

We can now state our main result on the random pruned tree \mathbf{T}_n^* . For any $k \in \{0, \dots, n\}$, define μ_k^* the distribution

$$\mu_k^*(d) = \frac{1}{\delta_k} \sum_{\ell \geq 0} \mu(d + \ell) \binom{d + \ell}{\ell} (\gamma_{k+1})^\ell (\delta_{k+1})^d, \quad d \in \mathbb{N}, \quad (4.2)$$

and note that μ_k^* also depends on n , through the parameters γ and δ (see Lemma 4.2).

Proposition 4.3. *The tree $\mathbf{T}_n^* = \text{Pruned}_h(\mathbf{T}_n)$ obtained by the pruning of the Galton–Watson tree \mathbf{T}_n by i.i.d. Bernoulli random variables $(h_v)_{v \in \partial \mathbf{T}}$ is, under $\mathbf{P} \otimes \mathbb{P}$, an inhomogeneous branching process. Its offspring distributions are $(\tilde{\mu}_0^*, \mu_1^*, \dots, \mu_{n-1}^*)$, where μ_k^* is defined in (4.2) above and $\tilde{\mu}_0^*(d) = \delta_0 \mu_0^*(d)$ for $d \geq 1$, $\tilde{\mu}_0^*(0) = \gamma_0$.*

4.2 Preliminary observations

Before we turn to the proof of Proposition 4.3, let us comment on the offspring distribution. First of all, $\tilde{\mu}_0^*(0)$ is the probability that the whole tree is pruned, i.e. γ_0 , and μ_0^* is the law $\tilde{\mu}_0^*$ conditioned on being non-zero. Hence, conditionally on the whole tree not being pruned, \mathbf{T}_n^* is an *inhomogeneous* branching process with offspring distribution $(\mu_k^*)_{0 \leq k \leq n-1}$.

We also have some nice interpretation of the distribution μ_k^* . Conditionally on the tree not being pruned, we know that each vertex $u \in \mathbf{T}^*$ has at least one descendant in $\partial \mathbf{T}$ that has not been pruned; in other words, μ_k^* is supported on \mathbb{N} . It turns out that for $u \in T_k^*$, the number of children of vertex u in \mathbf{T}^* can be constructed as follows: take X a random variable with distribution μ , so u has X descendants in \mathbf{T} ; prune these descendants with probability γ_{k+1} independently, but conditionally on having at least one surviving descendant.

Definition 4.4. Let $n \in \mathbb{N}$ and $p \in (0, 1]$. A random variable \hat{B} is said to follow a *zero-truncated binomial* of parameters n and p , and we write $\hat{B} \sim \widehat{\text{Bin}}(n, p)$ if

$$\mathbb{P}(\hat{B} = k) = \frac{1}{1 - (1 - p)^n} \binom{n}{k} p^k (1 - p)^{n-k} \quad k \in \{1, \dots, n\}.$$

Put otherwise, we have $\mathbb{P}(\hat{B} = k) = \mathbb{P}(B = k \mid B > 0)$ with $B \sim \text{Bin}(n, p)$.

For a random variable X , we also write $W \sim \text{Bin}(X, p)$ if for any $z \in \mathbb{N}$,

$$\mathbb{P}(W = z) = \sum_{d=1}^{\infty} \mathbb{P}(X = d) \mathbb{P}(B = z), \quad \text{with } B \sim \text{Bin}(d, p),$$

i.e. $W \sim \text{Bin}(d, p)$ conditionally on $X = d$. A similar notation holds for $W \sim \widehat{\text{Bin}}(X, p)$.

Lemma 4.5. For all $k \in \{0, \dots, n-1\}$, let X_k^* be a random variable with law μ_k^* . Then we have that $X_k^* \sim \widehat{\text{Bin}}(X, \delta_{k+1})$, where X is a random variable of distribution μ . Put otherwise, if $B_k \sim \text{Bin}(X, \delta_{k+1})$, then $\mathbb{P}(X_k^* = z) = \mathbb{P}(B_k = z \mid B_k > 0)$ for $z \in \mathbb{N}$.

Proof. For $k \in \{0, \dots, n-1\}$, let $\tilde{X}_k \sim \widehat{\text{Bin}}(X, \delta_{k+1})$. Letting $B_k \sim \text{Bin}(X, \delta_{k+1})$, we have for $z \in \mathbb{N}$,

$$\mathbb{P}(\tilde{X}_k = z) = \mathbb{P}(B_k = z \mid B_k > 0) = \frac{\mathbb{P}(B_k = z)}{1 - \mathbb{P}(B_k = 0)},$$

with

$$\mathbb{P}(B_k = 0) = \sum_{d=1}^{\infty} \mu(d) \mathbb{P}(\text{Bin}(d, \delta_{k+1}) = 0) = \sum_{d=1}^{\infty} \mu(d) (1 - \delta_{k+1})^d = \sum_{d=1}^{\infty} \mu(d) (\gamma_{k+1})^d.$$

Then $\mathbb{P}(B_k = 0) = \mathbf{E}[(\gamma_{k+1})^X] = \gamma_k$, thanks to Lemma 4.2. Therefore, we end up with

$$\mathbb{P}(\tilde{X}_k = z) = \frac{1}{1 - \gamma_k} \mathbb{P}(B_k = z) = \frac{1}{\delta_k} \sum_{d=1}^{\infty} \mu(d) \binom{d}{z} (\gamma_{k+1})^{d-z} (\delta_{k+1})^z = \mu_k^*(z),$$

which concludes the proof. \square

4.3 Proof of Proposition 4.3

Our goal is to write $\mathbf{P} \otimes \mathbb{P}(\mathbf{T}_n^* = \mathbf{t}^*)$ in terms of the offspring distributions $(\mu_k^*)_{0 \leq k \leq n-1}$. The approach for doing this is to consider all the possible trees \mathbf{t} (sampled from \mathbf{T}) such that after pruning we may obtain \mathbf{t}^* .

First of all, notice that we clearly have $\mathbb{P} \otimes \mathbf{P}(\mathbf{T}_n^* = \emptyset) = \gamma_0$. Indeed, conditionally on the tree \mathbf{T}_n , the probability that the whole tree is pruned is $\mathbb{P}(h_v = 0 \forall v \in T_n) = (1 - p_n)^{|T_n|}$, so $\mathbb{P} \otimes \mathbf{P}(\mathbf{T}_n^* = \emptyset) = \mathbf{E}[(1 - p_n)^{|T_n|}] = \gamma_0$, see Lemma 4.2.

We now work with $\mathbf{t}^* \neq \emptyset$ of depth n and we write

$$\mathbb{P} \otimes \mathbf{P}(\mathbf{T}_n^* = \mathbf{t}^*) = \sum_{\mathbf{t} \in \text{BP}_n(\mu)} \mathbb{P} \otimes \mathbf{P}(\mathbf{T} = \mathbf{t}, \text{Pruned}_h(\mathbf{t}) = \mathbf{t}^*).$$

Here, $\text{BP}_n(\mu)$ is the set of all trees \mathbf{t} of depth n generated by a Branching Process of offspring distribution μ .

Preliminary calculation: offspring distribution. For $u \in \mathbf{T}_n$, we denote X_u^* the number of descendants of u inside the pruned tree \mathbf{T}_n^* . Notice that the number of successor of the vertex u on the tree \mathbf{T}_n (*i.e.* before pruning), that we denote X_u , has to verify $X_u \geq X_u^*$. We have

$$\{X_u^* = d\} = \bigcup_{\ell \geq 0} \bigcup_{V \subset S(u), |V|=\ell} \{X_u = d + \ell, Y_v = 0 \text{ for } v \in V, Y_v = 1 \text{ for } v \in S(u) \setminus V\},$$

meaning that the vertex u has $d + \ell$ successors (on \mathbf{T}) for some $\ell \geq 0$, and exactly ℓ of them get pruned.

Using this representation, we can compute the offspring distribution of a vertex u at generation k . Using that vertices $v \leftarrow u$ in generation $k + 1$ have a probability γ_{k+1} of being pruned and that the events $\{Y_v = 0\}_{v \leftarrow u}$ are independent (they depend on Bernoulli variables (h_v) in different sub-trees), we get for $d \geq 0$

$$\mathbb{P} \otimes \mathbf{P}(X_u^* = d) = \sum_{\ell \geq 0} \mu(d + \ell) \binom{d + \ell}{\ell} (\gamma_{k+1})^\ell (\delta_{k+1})^d. \quad (4.3)$$

Notice that (4.3) generalizes the iteration $\gamma_k := \mathbb{P} \otimes \mathbf{P}(Y_u = 0) = G(\gamma_{k+1})$ of Lemma 4.2.

Now, notice that conditionally on $\mathbf{T}_n^* \neq \emptyset$, a vertex $u \in \mathbf{T}_n^*$ has at least one descendant in \mathbf{T}_n^* . Note that we have $\mathbf{P} \otimes \mathbb{P}(X_u^* > 0) = \mathbf{P} \otimes \mathbb{P}(Y_u = 1) = \delta_{|u|}$, since u has at least one descendant in the pruned tree if and only if it survives the pruning, *i.e.* $Y_u = 1$. We therefore get the offspring distribution of a vertex $u \in T_k^*$ in \mathbf{T}_n^* (conditioned on having $\mathbf{T}_n^* \neq \emptyset$):

$$\mathbf{P} \otimes \mathbb{P}(X_u^* = d \mid X_u^* > 0) = \frac{1}{\delta_k} \sum_{\ell \geq 0} \mu(d + \ell) \binom{d + \ell}{\ell} (\gamma_{k+1})^\ell (\delta_{k+1})^d = \mu_k^*(d).$$

It remains to show that the number of descendants in the different generations are independent.

Main calculation: probability of having a given pruned tree. Let \mathbf{t}^* be a non-empty tree of depth n which is a possible candidate for being a pruned version of a Galton–Watson tree \mathbf{T} . We let d_u^* denote the number of descendants of $u \in \mathbf{t}^*$; recall that each vertex $u \in \mathbf{t}^*$ (in generation $k < n$) has to be such that it has at least one descendant, that is $d_u^* \geq 1$.

In order to have $\mathbf{T}_n^* = \mathbf{t}^*$, one must have \mathbf{t}^* as a *skeleton* for \mathbf{T} , see Figure 5. Then, similarly as above, we can write the event $\{\mathbf{T}_n^* = \mathbf{t}^*\}$ as

$$\bigcap_{k=0}^{n-1} \bigcap_{u \in t_k^*} \bigcup_{\ell_u \geq 0} \bigcup_{V \subset S(u), |V|=\ell_u} \{X_u = d_u^* + \ell_u, Y_v = 0 \text{ for } v \in V, Y_v = 1 \text{ for } v \in S(u) \setminus V\},$$

meaning that each vertex $u \in \mathbf{t}^*$ must have $d_u^* + \ell_u$ successors in \mathbf{T} for some $\ell_u \geq 0$, with exactly ℓ_u of them getting pruned; the descendants of u inside \mathbf{t}^* are the v with $Y_v = 1$.

Now, note that the events $\{Y_v = 0\}$ in the above are all independent because they depend on different sub-trees of \mathbf{T} (that lead to leaves with only $h \equiv 0$); recall that $\mathbf{P} \otimes \mathbb{P}(Y_v = 0) = \gamma_{k+1}$ for $v \in T_k$. On the other hand, the events $\{Y_v = 1\}$ are not independent, since all vertices $v \in \mathbf{T}$ such that $Y_v = 1$ are ancestors of leaves with $h. = +1$. In particular, having $Y_w = 1$ for a leaf $w \in \partial \mathbf{T}_n$ implies that $Y_v = 1$ for all ancestors v of w , see Figure 5. However, all the events $\{Y_v = 1\}$ can be regrouped into a simpler event $\{h_w = 1, \forall w \in t_n^*\}$, which is independent of all events $\{Y_v = 0\}$; note that $\mathbf{P} \otimes \mathbb{P}(h_w = 1, \forall w \in t_n^*) = (p_n)^{|t_n^*|}$.

All together, referring to Figure 5 for an illustration of the computation, and using also that the $(X_v)_{v \in \mathbf{T}}$ are independent with distribution μ , we obtain that

$$\mathbb{P} \otimes \mathbf{P}(\mathbf{T}_n^* = \mathbf{t}^*) = \prod_{k=0}^{n-1} \prod_{u \in t_k^*} \sum_{\ell_u \geq 0} \mu(d_u^* + \ell_u) \binom{d_u^* + \ell_u}{\ell_u} (\gamma_{k+1})^{\ell_u} \times (p_n)^{|t_n^*|}. \quad (4.4)$$

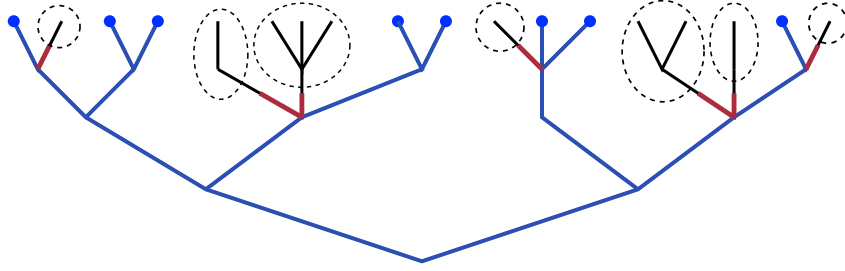


Figure 5: Illustration of the computation in (4.4): the tree \mathbf{t}^* is represented in blue and the tree \mathbf{T} whose pruned version is \mathbf{t}^* is represented in black. For every $h_w = 1$ in the leaves (represented by a blue dot), all the ancestors v of w automatically have $Y_v = 1$: this contributes to (4.4) by a factor $(p_n)^{|t_n^*|}$. A branch that is pruned just above generation k contributes to the probability by a factor γ_{k+1} , and all these pruning events (represented by a red segment) are independent since they depend on distinct subtrees of \mathbf{T} (the corresponding subtrees are circled with dashed lines).

We can now reformulate (4.4). Notice that $p_n = \delta_n$, and that $|t_n^*| = \sum_{u \in t_{n-1}^*} d_u^*$: we therefore get that $(p_n)^{|t_n^*|} \prod_{u \in t_{n-1}^*} \sum_{\ell_u \geq 0} \mu(d_u^* + \ell_u) \binom{d_u^* + \ell_u}{\ell_u} (\gamma_{k+1})^{\ell_u}$ is equal to

$$\prod_{u \in t_{n-1}^*} \sum_{\ell_u \geq 0} \mu(d_u^* + \ell_u) \binom{d_u^* + \ell_u}{\ell_u} (\gamma_{k+1})^{\ell_u} (\delta_n)^{d_u^*} = (\delta_{n-1})^{|t_{n-1}^*|} \prod_{u \in t_{n-1}^*} \mu_{n-1}^*(d_u^*),$$

where we have used the definition (4.2) of μ_k^* to get that $\sum_{\ell \geq 0} \mu(d + \ell) \binom{d + \ell}{\ell} (\gamma_{k+1})^\ell (\delta_n)^d = \delta_{n-1} \mu_{n-1}^*(d)$. Similarly, we get that

$$(\delta_{n-1})^{|t_{n-1}^*|} \prod_{u \in t_{n-2}^*} \sum_{\ell_u \geq 0} \mu(d_u^* + \ell_u) \binom{d_u^* + \ell_u}{\ell_u} (\gamma_{k+1})^{\ell_u} = (\delta_{n-2})^{|t_{n-2}^*|} \prod_{u \in t_{n-2}^*} \mu_{n-2}^*(d_u^*).$$

Therefore, iterating, we finally obtain from (4.4) that for any non-empty \mathbf{t}^* ,

$$\mathbb{P} \otimes \mathbf{P}(\mathbf{T}_n^* = \mathbf{t}^*) = \delta_0 \prod_{k=0}^{n-1} \prod_{u \in t_k^*} \mu_k^*(d_u^*).$$

This concludes the proof since this gives that for any \mathbf{t}^*

$$\mathbb{P} \otimes \mathbf{P}(\mathbf{T}_n^* = \mathbf{t}^*) = \tilde{\mu}_0^*(d_\rho^*) \prod_{k=1}^{n-1} \prod_{u \in t_k^*} \mu_k^*(d_u^*) =: \mathbf{P}^*(\mathbf{T}_n^* = \mathbf{t}^*),$$

where \mathbf{P}^* is the law of some inhomogeneous branching process of depth n , with offspring distribution $(\tilde{\mu}_0^*, \mu_1^*, \dots, \mu_n^*)$. \square

5 Transition in the shape of the pruned tree \mathbf{T}_n^*

5.1 Some notation and preliminaries

Recall that X denotes a random variable with law μ and that $G(s) := \mathbb{E}[s^X]$ denotes its generating function. In all this section, we assume that $\mu(0) = 0$, $\mu(1) < 1$ and μ admits a finite second moment. For $k \geq 1$, we denote $m_k := \mathbb{E}[X^k] \in (1, \infty]$ and we assume that $\nu := m_1 < +\infty$ and $m_2 < +\infty$; we denote $\sigma^2 = \text{Var}(X) = m_2 - m_1^2$. We also let

$$d_0 := \min \{d \geq 1, \mu(d) > 0\} \quad (d_0 \geq 1).$$

As the parameters $(\gamma_k)_{0 \leq k \leq n}$ are defined recursively in terms of the function G , see Lemma 4.2, let us state the following useful lemma on the function G , whose proof is elementary (we include it in Appendix A.1 for completeness).

Lemma 5.1. *We have the following bounds on G : for all $s \in [0, 1]$,*

$$1 + m_1(s - 1) \leq G(s) \leq 1 + m_1(s - 1) + C_\mu(s - 1)^2, \quad (5.1)$$

$$\mu(d_0)s^{d_0} \leq G(s) \leq \mu(d_0)s^{d_0} + s^{d_0+1}, \quad (5.2)$$

where $C_\mu := \frac{1}{2}(m_2 - m_1)$.

As a direct application of (5.1), we obtain the following bounds. For all $t \in [0, 1]$,

$$\nu t(1 - \tilde{C}_\mu t) \leq F(t) := 1 - G(1 - t) \leq \nu t, \quad (5.3)$$

where $\tilde{C}_\mu = C_\mu/\nu$.

We also state the following lemma on recursively defined sequences $u_{j+1} = G(u_j)$. Again, its proof is elementary but included in Appendix A.1 for completeness.

Lemma 5.2. *Let $\alpha \in (0, 1)$ and $u_0 \leq 1 - \alpha$. Define recursively the sequence $(u_j)_{j \geq 0}$ by the relation $u_{j+1} = G(u_j)$. Then:*

(i) If $d_0 = 1$, i.e. $\mu(1) \in (0, 1)$, there is a constant $C_\alpha = C_\alpha(\mu)$ such that for all $j \geq 1$

$$u_0 \mu(1)^j \leq u_j \leq C_\alpha u_0 \mu(1)^j. \quad (5.4)$$

(ii) If $d_0 \geq 2$, there is a constant $c_\alpha > 0$ such that for all $j \geq 1$,

$$(u_0)^{d_0^{j-1}} \mu(d_0)^{\frac{d_0^j}{d_0-1}} \leq u_j \leq e^{-c_\alpha d_0^j}. \quad (5.5)$$

(Note that the decay is doubly exponential in the second case.)

5.2 Some estimates on the parameters $(\gamma_k)_{0 \leq k \leq n}$

We now study how the parameters $(\gamma_k)_{0 \leq k \leq n}$ vary. Recall that γ_k is defined in (4.1) as the probability that a vertex at generation k is pruned, and also that we have set $\bar{\gamma}_k = \gamma_{n-k}$ and $\delta_k = 1 - \gamma_k$, $\bar{\delta}_k = \delta_{n-k}$.

Lemma 4.2 shows that $\bar{\gamma}_k$ is defined by the iteration $\bar{\gamma}_{k+1} = G(\bar{\gamma}_k)$. Hence, we easily get that the parameters $(\bar{\delta}_k)_{0 \leq k \leq n}$ can be recursively determined in terms of the function $F(t) := 1 - G(1 - t)$ in (5.3). Indeed, for $0 \leq k \leq n$, we have

$$\bar{\delta}_{k+1} = 1 - \bar{\gamma}_{k+1} = 1 - G(\bar{\gamma}_k) = 1 - G(1 - \bar{\delta}_k) = F(\bar{\delta}_k).$$

A first observation from these iterative definitions is that by convexity of G , and since $G(1) = 1$, we get that $\bar{\gamma}_{k+1} = G(\bar{\gamma}_k) \leq \bar{\gamma}_k$. We therefore have the following:

Lemma 5.3. *The sequence $(\gamma_k)_{0 \leq k \leq n}$ is non-decreasing; equivalently, $(\delta_k)_{0 \leq k \leq n}$ is non-increasing.*

The main result of this section is that the parameters $(\gamma_k)_{0 \leq k \leq n}$, or equivalently $(\delta_k)_{0 \leq k \leq n}$, exhibit a (sharp) phase transition. Let us define:

$$k^* = k^*(p_n) := \log_\nu(p_n \nu^n), \quad \text{and} \quad \bar{k}^* = n - k^* = \log_\nu\left(\frac{1}{p_n}\right), \quad (5.6)$$

where $\log_\nu(x) = \frac{1}{\log \nu} \log x$. In the following, we omit the integer part to simplify notation and often treat k^*, \bar{k}^* as integers. Let us note that $k^* \rightarrow \infty$ as soon as $p_n \nu^n \rightarrow \infty$ and that $\bar{k}^* \rightarrow \infty$ as soon as $p_n \rightarrow 0$.

Proposition 5.4. *There are constants $c_1 \in (0, 1)$ and $0 < c_2 \leq c_3$ (that depend only on the distribution μ) such that:*

$$\text{for all } k \geq k^* \quad c_1 \left(\frac{1}{\nu}\right)^{k-k^*} \leq \delta_k \leq \left(\frac{1}{\nu}\right)^{k-k^*},$$

and, depending on whether $d_0 = 1$ (i.e. $\mu(1) \in (0, 1)$) or $d_0 \geq 2$,

$$\text{for all } k \leq k^* \quad \begin{cases} \text{if } d_0 = 1, & c_2 \mu(1)^{k^*-k} \leq \gamma_k \leq c_3 \mu(1)^{k^*-k}, \\ \text{if } d_0 \geq 2, & e^{-c_3(d_0)^{k^*-k}} \leq \gamma_k \leq e^{-c_2(d_0)^{k^*-k}}. \end{cases}$$

The lower bounds in the above are not particularly relevant to our purpose (they however have some technical use in the proofs), so let us write more compactly (whether $d_0 = 1$ or $d_0 \geq 2$) upper bounds that we will use repeatedly in the proof. There is a constant $c_4 > 0$ such that

$$\gamma_k \leq e^{-c_4(k^*-k)} \quad \text{for } k \leq k^* \quad \text{and} \quad \delta_k \leq \left(\frac{1}{\nu}\right)^{k-k^*} \quad \text{for } k \geq k^*. \quad (5.7)$$

The interpretation of Proposition 5.4 (or of (5.7)) is as follows. On one hand, when k is much smaller than k^* in the sense that $k^* - k \gg 1$, then γ_k is very small (or δ_k is close to 1), meaning that vertices have an exponentially small chance of being pruned. On the other hand, when k is much larger than k^* in the sense that $k - k^* \gg 1$, then δ_k is very small (or γ_k is close to 1), meaning that vertices are very likely being pruned. The transition is *sharp*, in the sense that it occurs in a window of size $O(1)$.

One may therefore have the following picture in mind for the pruned tree \mathbf{T}_n^* : for $k \leq k^*$, \mathbf{T}_n^* resembles the original Galton–Watson tree, while for $k \geq k^*$, \mathbf{T}_n^* has thin long branches (*i.e.* with only one descendant per vertex) recalling that the vertices of the pruned tree are conditioned on having at least one descendant. This picture is made more formal in Proposition 5.7 below, see also Figure 1 on page 1 for an illustration.

Remark 5.5. In the case where $p_n = \alpha^n$ with $\alpha \in (0, 1)$, then we have that $\bar{k}^* = n \log_\nu(\frac{1}{\alpha})$ with $\eta_\alpha := \log_\nu(\frac{1}{\alpha}) \in (0, 1)$. This shows that the transition occurs sharply around generation $k^* = (1 - \eta_\alpha)n$.

Proof of Proposition 5.4. For our purposes, it will be more convenient to work with the iteration from Lemma 4.2 which goes from the leaves to the root. We therefore prefer to work with $\bar{\gamma}_k, \bar{\delta}_k$ instead of γ_k, δ_k , and we want to show the following:

$$\begin{aligned} \frac{c_1}{(\nu)^{\bar{k}^*-k}} &\leq \bar{\delta}_k \leq \frac{1}{(\nu)^{\bar{k}^*-k}} && \text{for all } k \leq \bar{k}^*, \\ c_2\mu(1)^{k-\bar{k}^*} &\leq \bar{\gamma}_k \leq c_3\mu(1)^{k-\bar{k}^*} && \text{for all } k \geq \bar{k}^*. \end{aligned}$$

The second line being valid if $d_0 = 1$ (*i.e.* $\mu(1) \in (0, 1)$), and replaced with

$$e^{-c_3(d_0)^{k-\bar{k}^*}} \leq \bar{\gamma}_k \leq e^{-c_2(d_0)^{k-\bar{k}^*}}$$

in the case where $d_0 \geq 2$. The main idea of the proof is to use the iteration $\bar{\delta}_{k+1} = F(\bar{\delta}_k)$ together with the estimate (5.3), until $\bar{\delta}_k$ is not small enough; we show that this occurs around $k = \bar{k}^*$. Then, we apply Lemma 5.2 to the iteration $\bar{\gamma}_{k+1} = G(\bar{\gamma}_k)$ starting from $k = \bar{k}^*$ for which $\bar{\gamma}_{\bar{k}^*}$ is not too small.

Step 1. We start with some intermediate result which makes use of (5.3).

Lemma 5.6. *Let $\bar{k}_1^* = \min\{0 \leq k \leq n : \sum_{i=0}^k \tilde{C}_\mu \bar{\delta}_i > 1/2\}$, with \tilde{C}_μ as in (5.3). Then, for all $k \leq \bar{k}_1^*$ we have*

$$\frac{1}{2}(\nu)^k p_n \leq \bar{\delta}_k \leq (\nu)^k p_n. \quad (5.8)$$

The upper bound is valid for any $0 \leq k \leq n$.

Proof. Let $0 \leq k \leq n$. As $\bar{\delta}_k = F(\bar{\delta}_{k-1})$, the right-hand side of the inequality (5.3) gives us $\bar{\delta}_k = F(\bar{\delta}_{k-1}) \leq \nu \bar{\delta}_{k-1}$. Applying this inequality iteratively to all $\bar{\delta}_i$ with $i \leq k-1$, we obtain

$$\bar{\delta}_k \leq (\nu)^k \bar{\delta}_0 = (\nu)^k p_n,$$

where we have used that $\bar{\delta}_0 = p_n$. This proves the right-hand side of the inequality.

Using the left-hand side of (5.3), we have $\bar{\delta}_k \geq \nu \bar{\delta}_{k-1} (1 - \tilde{C}_\mu \bar{\delta}_{k-1})$. Applying this inequality iteratively to all $\bar{\delta}_i$ such that $i \leq k-1$, we have

$$\bar{\delta}_k \geq (\nu)^k \bar{\delta}_0 \prod_{i=0}^{k-1} (1 - \tilde{C}_\mu \bar{\delta}_i). \quad (5.9)$$

By definition of \bar{k}_1^* , and the fact that \tilde{C}_μ and $(\bar{\delta}_j)_{0 \leq j \leq n}$ are non-negative, we have that $\tilde{C}_\mu \bar{\delta}_k \leq 1/2$ for all $k \leq \bar{k}_1^*$. Thus, applying the Weierstrass inequality $\prod_{i=0}^{k-1} (1 - x_i) \geq 1 - \sum_{i=0}^{k-1} x_i$ for $x_i \in [0, 1]$, we get that for $0 \leq k \leq \bar{k}_1^*$,

$$\prod_{i=0}^{k-1} (1 - \tilde{C}_\mu \bar{\delta}_i) \geq 1 - \sum_{i=0}^{k-1} \tilde{C}_\mu \bar{\delta}_i \geq \frac{1}{2},$$

where the last inequality follows again thanks to the definition of \bar{k}_1^* . This concludes the left-hand side of the inequality, using again that $\bar{\delta}_0 = p_n$. \square

Step 2. We now show the bounds on $\bar{\delta}_k$. First of all, since the bound $\bar{\delta}_k \leq (\nu)^k p_n$ in Lemma 5.6 is valid for any $0 \leq k \leq n$, this gives the desired upper bound, recalling that \bar{k}^* is such that $(\nu)^{\bar{k}^*} p_n = 1$, see (5.6).

For the lower bound on $\bar{\delta}_k$, thanks to Lemma 5.6 we have a lower bound for all $k \leq \bar{k}_1^*$. If $\bar{k}^* \leq \bar{k}_1^*$ this concludes the proof, otherwise we need to control $\bar{\delta}_k$ for $\bar{k}_1^* < k \leq \bar{k}^*$. In fact, we show that \bar{k}_1^* is comparable to \bar{k}^* , in the sense that there exist two constants L_1, L_2 , that only depend on the law μ , such that

$$L_1 + \bar{k}^* \leq \bar{k}_1^* \leq \bar{k}^* + L_2. \quad (5.10)$$

To get the upper bound in (5.10), we use that by the upper bound in Lemma 5.6, we have

$$\sum_{i=0}^{\bar{k}_1^*} \bar{\delta}_i \leq p_n \sum_{i=0}^{\bar{k}_1^*} (\nu)^i = p_n \frac{(\nu)^{\bar{k}_1^*+1} - 1}{\nu - 1} \leq \frac{\nu}{\nu - 1} (\nu)^{\bar{k}_1^* - \bar{k}^*},$$

where we have again used that $p_n = (\nu)^{\bar{k}^*}$. Thus, by definition of \bar{k}_1^* , we have that $\frac{1}{2\tilde{C}_\mu} \leq \frac{\nu}{\nu-1} (\nu)^{\bar{k}_1^* - \bar{k}^*}$, which yields $\bar{k}_1^* \geq \bar{k}^* + L_1$, for $L_1 := \log_\nu \left(\frac{\nu-1}{2\nu\tilde{C}_\mu} \right)$.

For the lower bound in (5.10), using the lower bound in Lemma 5.6, we have similarly as above

$$\sum_{i=0}^{\bar{k}_1^*-1} \bar{\delta}_i \geq \frac{1}{2} p_n \sum_{i=0}^{\bar{k}_1^*-1} (\nu)^i \geq \frac{1}{2} p_n (\nu)^{\bar{k}_1^*-1} = \frac{1}{2\nu} (\nu)^{\bar{k}_1^* - \bar{k}^*}.$$

Thus, by definition of \bar{k}_1^* we have $\frac{1}{2\nu}(\nu)^{\bar{k}_1^* - \bar{k}^*} \leq \frac{1}{2C_\mu}$, which yields the bound $\bar{k}_1^* \leq \bar{k}^* + L_2$, for $L_2 := \log_\nu \left(\frac{\nu}{C_\mu} \right)$.

Therefore, by the lower bound in Lemma 5.6 and recalling the definition of \bar{k}^* , we obtain that $\bar{\delta}_{\bar{k}_1^*} \geq \alpha_1$, for $\alpha_1 := \frac{1}{2}(\nu)^{L_1} \leq \frac{1}{2}$ (recall we are treating the case $\bar{k}_1^* < \bar{k}^*$ so $L_1 \leq 0$). Using that $(\bar{\delta}_k)_{k \geq 0}$ is non-decreasing, we get that $\bar{\delta}_k \geq \alpha_1$ for all $\bar{k}_1^* \leq k \leq \bar{k}^*$. Adjusting the constant (and since $(\nu)^k p_n$ is of order one for $\bar{k}_1^* \leq k \leq \bar{k}^*$), we therefore get that for all $k \leq \bar{k}^*$,

$$\bar{\delta}_k \geq c_1(\nu)^k p_n, \quad (5.11)$$

which gives the desired bound, using the definition of \bar{k}^* .

Step 3. We now conclude the proof by showing the bounds on $\bar{\gamma}_k$ for $k \geq \bar{k}^*$. We have proven above that $\bar{\delta}_{\bar{k}^*} \geq \alpha_1$ for some $\alpha_1 < 1$ that depends only on the law μ . Indeed, this is in the previous paragraph if $\bar{k}_1^* < \bar{k}^*$, and follows simply from Lemma 5.6 if $\bar{k}_1^* \geq \bar{k}^*$ since then $\bar{\delta}_{\bar{k}^*} \geq \frac{1}{2}(\nu)^{\bar{k}^*} p_n = \frac{1}{2}$.

We therefore get that $\bar{\gamma}_{\bar{k}^*} \leq 1 - \alpha_1$.

We can then apply Lemma 5.2 to $u_j := \bar{\gamma}_{\bar{k}^*+j}$, for $j = k - \bar{k}^* \geq 0$: in the case $d_0 = 1$ ($\mu(1) \in (0, 1)$), we obtain

$$\bar{\gamma}_{\bar{k}^*} \mu(1)^{k - \bar{k}^*} \leq \bar{\gamma}_k \leq C_{\alpha_1} \mu(1)^{k - \bar{k}^*},$$

and a similar application of Lemma 5.2 gives the correct upper bound in the case $d_0 \geq 2$.

This concludes the upper bound on $\bar{\gamma}_k$ (the important part), but for the lower bound, we need to show that $\bar{\gamma}_{\bar{k}^*} \geq \alpha_2$ for some universal constant $\alpha_2 > 0$ that depends only on the law μ . To see this, note that the upper bound in Lemma 5.6 gives that $\bar{\delta}_{\bar{k}^*-1} \leq \nu^{\bar{k}^*-1} p_n = \nu^{-1}$, so $\bar{\gamma}_{\bar{k}^*} = G(\bar{\gamma}_{\bar{k}^*-1}) \leq G(1 - \nu^{-1}) =: \alpha_2$. One can then apply Lemma 5.2 to obtain the correct lower bounds. \square

5.3 Transition in the shape of the pruned tree

We now provide a statement that clarifies the intuition that the offspring distribution μ_k^* of the pruned tree \mathbf{T}_n^* is close to μ for generations $0 \leq k \leq k^*$ and close to a Dirac mass at 1 for generations $k^* \leq k \leq n - 1$. We refer to Figure 6 for an illustration of this transition in the shape of the pruned tree.

Recall that if μ and $\tilde{\mu}$ are two probability measures on a measurable space (Ω, \mathcal{A}) , the total variation distance between μ and $\tilde{\mu}$ is defined by

$$d_{\text{TV}}(\mu, \tilde{\mu}) := \sup_{A \in \mathcal{A}} |\mu(A) - \tilde{\mu}(A)|.$$

An important property of the total variation is that it can be rewritten as $d_{\text{TV}}(\mu, \nu) = \inf_{X \sim \mu, Y \sim \tilde{\mu}} \mathbb{P}(X \neq Y)$, where the infimum is taken over all couplings of $\mu, \tilde{\mu}$, *i.e.* all (joint) distributions for pairs of random variables (X, Y) whose marginal distributions are $X \sim \mu$ and $Y \sim \tilde{\mu}$.

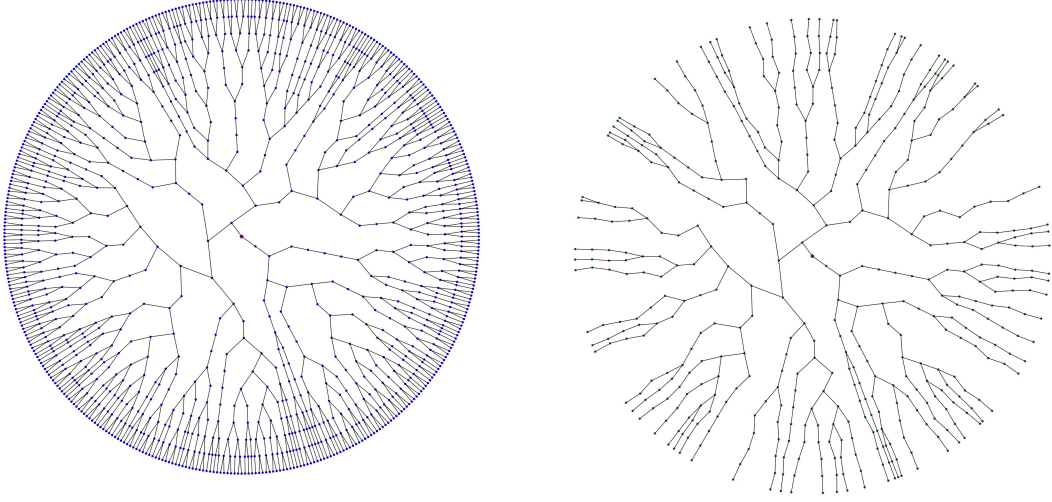


Figure 6: Illustration of the transition in the shape of the pruned tree. On the left, we have represented a tree \mathbf{t} with mean offspring $\nu = 1.5$ and $n = 15$ generations. On the right, the picture represents the tree \mathbf{t}^* obtained by pruning \mathbf{t} with Bernoulli random variables on the leaves of parameters $p_n = 0.2$. We can observe a change in the pruned tree \mathbf{t}^* at generation $k^* \simeq 11$; for $k \leq k^*$, the tree \mathbf{t}^* looks like the original tree \mathbf{t} , whereas for $k > k^*$ mainly thin branches remain.

Proposition 5.7. *There are some constants $c_4, c'_4 > 0$ such that*

$$\begin{aligned} d_{\text{TV}}(\mu_k^*, \mu) &\leq c'_4 e^{-c_4(k^* - k)} && \text{for } k \leq k^*, \\ d_{\text{TV}}(\mu_k^*, \mu_{\{1\}}) &\leq 2 \left(\frac{1}{\nu}\right)^{k - k^*} && \text{for } k \geq k^*, \end{aligned}$$

where we denoted by $\mu_{\{1\}}$ the Dirac mass at 1.

Proof. Let us start with the case $k \leq k^*$. Recall the interpretation of μ_k^* given by Lemma 4.5: it provides a natural coupling (X_k^*, X) with $X_k^* \sim \mu_k^*$ and $X \sim \mu$, where the conditional law of X_k^* given X is a zero-truncated Binomial $\widehat{\text{Bin}}(X, \delta_k)$. Then, by the interpretation of the total variation distance in terms of coupling, we obtain

$$d_{\text{TV}}(\mu_k^*, \mu) \leq \mathbb{P}(X_k^* \neq X) = \sum_{d \geq 1} \mu(d) \mathbb{P}(B_{k,d} \neq d \mid B_{k,d} > 0),$$

with $B_{k,d} \sim \text{Bin}(d, \delta_{k+1})$. As $\gamma_{k+1} \in [0, 1]$ and $d \geq 1$, we have

$$\begin{aligned} \mathbb{P}(B_{k,d} = d) &= (\delta_{k+1})^d = (1 - \gamma_{k+1})^d \geq 1 - d\gamma_{k+1} \\ \mathbb{P}(B_{k,d} = 0) &= (1 - \delta_{k+1})^d = (\gamma_{k+1})^d \leq \gamma_{k+1} \leq \gamma_{k^*+1} \quad \text{for } k \leq k^*, \end{aligned}$$

recalling also that $(\gamma_k)_{k \geq 0}$ is non-decreasing for the last inequality (see Lemma 5.3). Since $\gamma_{k^*+1} = \bar{\gamma}_{k^*-1} \leq 1 - \alpha_1$ for some $\alpha_1 > 0$ (see Step 3 of the proof of Proposition 5.4), we therefore have

$$\mathbb{P}(B_{k,d} \neq d \mid B_{k,d} > 0) = \frac{1 - (1 - \gamma_{k+1})^d}{1 - (\gamma_{k+1})^d} \leq \frac{d}{\alpha_1} \gamma_{k+1},$$

so $d_{\text{TV}}(\mu_k^*, \mu) \leq \frac{1}{\alpha_1} \nu \gamma_{k+1}$. This yields the desired upper bound thanks to Proposition 5.4; see also the general bound (5.7).

For the case $k \geq k^*$, note that for any coupling (X_k^*, Y) such that the marginal distributions are $X_k^* \sim \mu_k^*$ and $Y = 1$ we have $d_{\text{TV}}(\mu_k^*, \mu_{\{1\}}) \leq \mathbb{P}(X_k^* \neq 1)$. Using again Lemma 4.5 that describes the law of X_k^* , we have

$$\mathbb{P}(X_k^* \neq 1) = \sum_{d \geq 2} \mu(d) \mathbb{P}(B_{k,d} \geq 2 \mid B_{k,d} > 0),$$

with $B_{k,d} \sim \text{Bin}(d, \delta_{k+1})$. By sub-additivity and Bonferroni's inequality, we have

$$\mathbb{P}(B_{k,d} \geq 2) \leq \frac{d(d-1)}{2} (\delta_{k+1})^2, \quad \mathbb{P}(B_{k,d} \geq 1) \geq d\delta_k - \frac{d(d-1)}{2} (\delta_{k+1})^2.$$

Thus, we have the bound

$$\mathbb{P}(B_{k,d} \geq 2 \mid B_{k,d} > 0) \leq \frac{(d-1)(\delta_{k+1})^2}{\delta_k(2 - (d-1)\delta_{k+1})}.$$

This gives that $\mathbb{P}(B_{k,d} \geq 2 \mid B_{k,d} > 0) \leq (d-1)\delta_{k+1}$ for $d \leq (\delta_{k+1})^{-1}$, using also that $(\delta_k)_{k \geq 0}$ is non-increasing. Bounding $\mathbb{P}(B_{k,d} \geq 2 \mid B_{k,d} \geq 1)$ by 1 when $d > (\delta_{k+1})^{-1}$, we get

$$\begin{aligned} \sum_{d \geq 2} \mu(d) \mathbb{P}(B_{k,d} \geq 2 \mid B_{k,d} > 0) &\leq \sum_{d=2}^{(\delta_{k+1})^{-1}} \mu(d) d \delta_{k+1} + \sum_{d > (\delta_{k+1})^{-1}} \mu(d) \\ &\leq \nu \delta_{k+1} + \mathbb{P}(X > (\delta_{k+1})^{-1}) \leq 2\nu \delta_{k+1}, \end{aligned}$$

where we have used Markov's inequality in the last term. This gives the desired upper bound, using Proposition 5.4. \square

5.4 Mean of the offspring distribution and size of a generation

In this section we estimate the mean and the variance of the offspring distribution of the pruned tree, and we also observe the phase transition in these quantities. This is not a direct consequence of Proposition 5.7 since the total variation distance does not allow one to control moments. We then estimate the mean size of the generation k of the pruned tree \mathbf{T}_n^* , which will reveal useful in Section 6. Some of the technical estimates of this section are postponed to Appendix A.1.

5.4.1 Mean of μ_k^*

For $0 \leq k < n$, we let X_k^* be a random variable with distribution μ_k^* and $\nu_k^* := \mathbb{E}[X_k^*]$ and $(\sigma_k^*)^2 = \text{Var}(X_k^*)$. We now give several estimates on ν_k^* , and in particular we observe the phase transition around k^* . We are also able to estimate the variances $(\sigma_k^*)^2$ (which bring other useful information on μ_k^*), but it is of no particular use for the sequel so we postpone estimates on $(\sigma_k^*)^2$ to Appendix A.2.

Remark 5.8. Recall that in \mathbf{T}_n^* the offspring distribution of the root is $\tilde{\mu}_0^*$, see Proposition 4.3. However, when the root is conditioned on having at least one descendant (*i.e.* $\mathbf{T}_n^* \neq \emptyset$), which is an event of probability δ_0 , the offspring distribution of the root is μ_0^* .

Let us stress that one can easily obtain a formula for ν_k^* (we prove it in Appendix A.2, see Lemma A.1):

$$\nu_k^* := \mathbb{E}[X_k^*] = \frac{\delta_{k+1}}{\delta_k} \nu. \quad (5.12)$$

We can also show that $(\nu_k^*)_{0 \leq k < n}$ is non-increasing, see Lemma A.2 (also proven in Appendix A.1). We now show that the phase transition observed in Proposition 5.4 translates into the means $(\nu_k^*)_{0 \leq k < n}$. Recall the definition (5.6) of $k^* := \log_\nu(\nu^n p_n)$.

Lemma 5.9. *There are constants c_4, c_5 (that depend only on the law μ) such that*

$$0 \leq \nu - \nu_k^* \leq c_5 e^{-c_4(k^* - k)} \quad \text{for } k \leq k^*, \quad (5.13)$$

$$0 \leq \nu_k^* - 1 \leq 1 + c_5 \left(\frac{1}{\nu}\right)^{k - k^*} \quad \text{for } k \geq k^*, \quad (5.14)$$

where c_4 is the constant appearing in (5.7).

This lemma shows that ν_k^* is very close to ν if $k^* - k \gg 1$ whereas ν_k^* is very close to 1 if $k - k^* \gg 1$. Lemma A.3 complements this information by controlling the variance of μ_k^* , which is close to the variance of μ if $k^* - k \gg 1$ whereas it is close to 0 if $k - k^* \gg 1$. This confirms that the pruned tree roughly grows as a Galton–Watson tree with mean $\nu > 1$ for $k \leq k^*$ (*i.e.* with an exponential growth rate ν) and then the tree does not grow much more since the mean is roughly 1 for $k \geq k^*$. We make this statement precise by controlling the mean size of a generation, see Lemma 5.10 below.

Proof of Lemma 5.9. Let us start with the case $k \leq k^*$. First of all, using that the parameters $(\delta_k)_{0 \leq k < n}$ are non-increasing (see Lemma 5.3), we get from the expression (5.12) of ν_k^* the easy bound $\nu_k^* \leq \nu$. On the other hand, using that $\delta_k \leq 1$, we get from (5.12) that $\nu_k^* \geq \nu \delta_{k+1}$. Since $\delta_{k+1} = 1 - \gamma_{k+1}$, applying the bound in Proposition 5.4, we obtain (5.13).

For the case $k \geq k^*$, $\delta_k = \bar{\delta}_{n-k}$ verifies $\bar{\delta}_{k+1} = F(\bar{\delta}_k)$ with $F(t) = 1 - G(1 - t)$, we have thanks to (5.3)

$$\nu \delta_{k+1} (1 - \tilde{C}_\mu \delta_{k+1}) \leq \delta_k \leq \nu \delta_{k+1}.$$

Therefore, using the expression (5.12) for ν_k^* , we get that

$$1 \leq \nu_k^* = \frac{\delta_{k+1}}{\delta_k} \nu \leq \frac{1}{1 - \tilde{C}_\mu \delta_{k+1}} \leq 1 + 2C_\mu \delta_{k+1}, \quad (5.15)$$

for all k such that $\tilde{C}_\mu \delta_{k+1} \leq \frac{1}{2}$, using that $(1 - x)^{-1} \leq 1 + 2x$ for $x \in [0, \frac{1}{2}]$. Hence, using the bound in Proposition 5.4, this yields (5.14) for $k \geq k_2^* := \min\{k, \delta_{k+1} \leq 1/(2\tilde{C}_\mu)\}$. Similarly as in the proof of Step 2 of Proposition 5.4, one easily gets that $k_2^* \leq k^* + L'_1$ for some constant L'_1 that depends only on the law μ ; adapting the constant, this yields (5.14). \square

5.4.2 Size of the generations

Recall that T_k^* designates the set of vertices in the k -th generation of the pruned tree \mathbf{T}_n^* and $|T_k^*|$ its size. For ease of notations, let us note $Z_k = |T_k^*|$, for $0 \leq k \leq n$. By Proposition 4.3, we have that when the pruned tree is conditioned to be non-empty, $(Z_k)_{0 \leq k \leq n}$ is an inhomogeneous branching process with offspring distribution μ_k^* . We can construct $(Z_k)_{0 \leq k \leq n}$ as follows: set $Z_0 = 1$ and for $0 \leq k < n$

$$Z_{k+1} = \sum_{a=1}^{Z_k} X_{k,a},$$

where $(X_{k,a})_{a \geq 1}$ is a family of i.i.d. random variables with distribution μ_k^* .

More generally and for future use, we may want to study for any $0 \leq i \leq j \leq n$ the size $Z_{i,j}$ of the population generated by an individual of generation i up to generation j . The distribution of $Z_{i,j}$ can be constructed as follows: $Z_{i,i} = 1$ and iteratively, for $j \in \{i, \dots, n-1\}$, $Z_{i,j+1} = \sum_{a=1}^{Z_{i,j}} X_{j,a}$. Let us now define, for $0 \leq i < j \leq n$,

$$M_{i,j}^* := \mathbb{E}[Z_{i,j}] = \prod_{k=i}^{j-1} \nu_k^*, \quad (5.16)$$

where $M_{i,i} = 1$ by convention. We can now show how the pruned tree growth stabilizes at generation k^* ; recall the definition (5.6) of $k^* := \log_\nu(\nu^n p_n)$.

Lemma 5.10. *There are constants c_4 (from (5.7)), c_6, c_7 , that depend only on the law μ , such that*

$$(1 - e^{-c_4(k^* - k)})(\nu)^k \leq M_{0,k}^* \leq (\nu)^k \quad \text{for } k \leq k^*, \quad (5.17)$$

$$c_6(\nu)^{k^*} \leq M_{0,k}^* \leq c_7(\nu)^{k^*} \quad \text{for } k \geq k^*. \quad (5.18)$$

This can be summarized in a slightly weaker form as follows: for all $0 \leq k \leq n$,

$$c'_6(\nu)^{k \wedge k^*} \leq M_{0,k}^* \leq c_7(\nu)^{k \wedge k^*}.$$

This lemma properly shows that $M_{0,k}$ grows as $(\nu)^k$ up to $k = k^*$ and then almost completely stops growing.

Proof. Notice that the expression of ν_i^* in (5.12) yields the following expression of $M_{0,k}^*$: for all $0 \leq k \leq n$,

$$M_{0,k}^* = \prod_{i=0}^{k-1} \frac{\delta_{i+1}}{\delta_i} \nu = \frac{\delta_k}{\delta_0} (\nu)^k.$$

For $k \leq k^*$, as the parameters $(\delta_k)_{0 \leq k < n}$ are non-increasing (see Lemma 5.3) and $\delta_0 \leq 1$, we have

$$\delta_k(\nu)^k \leq M_{0,k}^* = \frac{\delta_k}{\delta_0} (\nu)^k \leq (\nu)^k.$$

Since $\delta_k = 1 - \gamma_k$, applying the bound in Proposition 5.4 (or the general bound (5.7)), we get the desired bound.

For $k \geq k^*$, the upper bound of δ_k from Proposition 5.4 yields

$$M_{0,k}^* \leq (\nu)^k \frac{1}{(\nu)^{k-k^*}} \frac{1}{\delta_0}.$$

Since δ_0 is close to 1 (see e.g. Proposition 5.4), we have the right-hand side of (5.18). For the left-hand side, using $\delta_0 \leq 1$ and Proposition 5.4, we have

$$M_{0,k}^* \geq (\nu)^k \delta_k \geq c_1 (\nu)^{k^*},$$

which concludes the proof. \square

To conclude this section, we also give some bounds on $M_{n-k,n}^*$, i.e. the growth of generations at the top of the tree. This completes the overall picture of the pruned tree. Recall that $\bar{k}^* := n - k^* = \log_\nu(\frac{1}{p_n})$.

Lemma 5.11. *There are constants c_8, c_9 , that depend only on the law μ , so that*

$$\begin{aligned} 1 &\leq M_{n-k,n}^* \leq 1 + c_8 \left(\frac{1}{\nu}\right)^{\bar{k}^*-k}, & \text{for } k \leq \bar{k}^*, \\ (\nu)^{k-\bar{k}^*} &\leq M_{n-k,n}^* \leq c_9 (\nu)^{k-\bar{k}^*}, & \text{for } k \geq \bar{k}^*, \end{aligned} \quad (5.19)$$

which can also be summarized as $\max\{1, (\nu)^{k-\bar{k}^*}\} \leq M_{n-k,n}^* \leq 1 + c(\nu)^{k-\bar{k}^*}$.

Proof. The proof is similar to the above. First of all, we clearly have $M_{n-k,n}^* \geq 1$. For the upper bounds, we can use that

$$M_{n-k,n}^* = \frac{\delta_n}{\delta_{n-k}} (\nu)^k = \frac{\bar{\delta}_0}{\bar{\delta}_k} (\nu)^k \leq \prod_{i=0}^{k-1} (1 - \tilde{C}_\mu \bar{\delta}_i)^{-1},$$

where we have used (5.9) for the last inequality. As $(1-x)^{-1} \leq 1+x \leq e^x$, we get that

$$M_{n-k,n}^* \leq \exp\left(\tilde{C}_\mu \sum_{i=0}^{k-1} \bar{\delta}_i\right) \leq \exp\left(\frac{\tilde{C}_\mu}{\nu-1} p_n (\nu)^k\right)$$

where we have used that $\bar{\delta}_i \leq p_n (\nu)^i$ for all $i \geq 0$, see Lemma 5.6. Since by definition of \bar{k}^* we have $p_n = (\nu)^{-\bar{k}^*}$, we obtain $M_{n-k,n}^* \leq \exp(c(\nu)^{k-\bar{k}^*}) \leq 1 + c_9 (\nu)^{k-\bar{k}^*}$ for $k \leq \bar{k}^*$.

For $k \geq \bar{k}^*$, we use that $1 \geq \bar{\delta}_k \geq \bar{\delta}_{\bar{k}^*} \geq c_1$ (see (5.11)) and the fact that $\bar{\delta}_0 = p_n = (\nu)^{-\bar{k}^*}$, to obtain

$$(\nu)^{k-\bar{k}^*} \leq M_{n-k,n}^* = \frac{\bar{\delta}_0}{\bar{\delta}_k} (\nu)^k \leq \frac{1}{c_1} (\nu)^{k-\bar{k}^*}.$$

This concludes the proof. \square

6 Ising model on \mathbf{T}_n^* with plus boundary external field

6.1 Ising model on trees and non-linear capacity

For the Ising model on a tree \mathbf{t} with plus boundary condition, Pemantle and Peres [27] observed that the magnetization of the root (more precisely the log-likelihood ratio) is comparable to the L^3 -capacity, or 3-capacity, of the tree, equipped with a specific set of resistances. We state this as Theorem 6.4 below, but let us introduce the necessary notation first.

Non-linear L^p -capacity. Let us start by defining the L^p - or p -capacity for a given (finite or infinite) tree $\mathbf{t} = (V, E)$ rooted at a vertex ρ ; we use analogous notation as in [27]. We let $\partial\mathbf{t}$ be the set of maximal paths oriented away from the root. If the tree is finite, we can identify $\partial\mathbf{t}$ with the set of leaves of \mathbf{t} . For infinite trees, we assume that there is no leaf and that all paths in $\partial\mathbf{t}$ are infinite.

The tree \mathbf{t} is equipped with a set of resistances on its edges: to each edge $e = uv$ with $u \rightarrow v$, assign a resistance $R_u := R(e) \in \mathbb{R}_+$ and a conductance $C_u := C(e) = R(e)^{-1}$. We say that a function $\theta : V \mapsto \mathbb{R}_+$ is a flow on the tree \mathbf{t} if for every $u \in V$ it verifies $\theta(u) = \sum_{u \rightarrow v} \theta(v)$, i.e. if the inflow is equal to the outflow at every vertex of the tree. Additionally, we define the strength $|\theta|$ of a flow θ as $|\theta| := \sum_{\rho \rightarrow v} \theta(v)$, that is, the total outflow from the root ρ . A flow θ with $|\theta| = 1$ is called a *unit* flow.

Definition 6.1 (L^p -capacity). Let $p > 1$ and set $s = p - 1$. Then we define the p -capacity of the tree \mathbf{t} with resistances $(R_e)_{e \in V}$ as

$$\text{Cap}_p(\mathbf{t}) := \sup_{\theta: \mathcal{V}_p(\theta)=1} |\theta|, \quad \text{with} \quad \mathcal{V}_p(\theta) := \sup_{y \in \partial\mathbf{t}} \sum_{e \in y} (\theta(e)R(e))^s.$$

Remark 6.2. We stress that when $p = 2$, by Thomson's principle, the 2-capacity $\text{Cap}_2(T)$ reduces to the usual electrical effective conductance between the root ρ and the leaves of the tree; we refer to [21, Ch. 2, 3 & 9] for an extensive introduction on electrical networks on graphs.

Then, Pemantle and Peres [27] establish a recursive expression for the p -capacity on a tree. Recall that for a vertex $u \in V$, $\mathbf{t}(u)$ denotes the sub-tree of \mathbf{t} consisting of u (as a root) and all descendants of u .

Lemma 6.3 (Lemma 3.1 in [27]). *Let \mathbf{t} be a locally finite tree with root ρ . Fix $p > 1$ and $s = p - 1$. For any vertex $u \in V$, define*

$$\phi_p(u) := R_u \text{Cap}_p(\mathbf{t}(u)),$$

where $R_\rho = 1$ by convention; in particular, $\phi(u) = R_u$ if u is a leaf. Then, for any vertex $u \in V$, we have

$$\phi_p(u) = \sum_{u \rightarrow v} \frac{R_u}{R_v} \frac{\phi_p(v)}{(1 + \phi_p(v)^s)^{1/s}}. \quad (6.1)$$

Relation to Lyons' iteration. Lemma 6.3 is reminiscent of the iteration (2.5) for the log-likelihood ratio, namely $r_\beta(u) = \sum_{u \rightarrow v} g_\beta(r_\beta(v))$. Additionally, [27, Lem. 4.2] observes that the function $g_\beta(x) = \log\left(\frac{e^{2\beta}e^x + 1}{e^{2\beta} + e^x}\right)$ verifies

$$\frac{\tanh(\beta)x}{(1 + c_2x^2)^{1/2}} \leq g_\beta(x) \leq \frac{\tanh(\beta)x}{(1 + c_1x^2)^{1/2}},$$

for some $c_1, c_2 > 0$. Therefore, a direct consequence of Theorem 3.2 in [27] and of the recursion (2.5) is the following.

Theorem 6.4 (Theorem 3.2 in [27]). *Let \mathbf{t} be a tree with a set of leaves $\partial\mathbf{t}$. There are constants κ_1, κ_2 such that, for the Ising model on \mathbf{t} with plus boundary condition, the log-likelihood ratio of the root verifies*

$$\kappa_1 \text{Cap}_3(\mathbf{t}) \leq r_{\mathbf{t},\beta}^+(\rho) \leq \kappa_2 \text{Cap}_3(\mathbf{t}).$$

Here $\text{Cap}_3(\mathbf{t})$ is the 3-capacity of the tree \mathbf{t} equipped with resistances $R_u := \tanh(\beta)^{-|u|}$.

This result is used in [27] to obtain a criterion for the magnetization of the Ising Model with plus boundary conditions on an infinite tree \mathbf{t} : Theorem 2.2 in [27] shows that the root ρ is magnetized if and only if $\text{Cap}_3(\mathbf{t}) > 0$.

The goal of the rest of the section is therefore to estimate the 3-capacity of the pruned tree \mathbf{T}_n^* . We will prove the following result, which thanks to Theorem 6.4 will conclude the proof of Theorem 3.3; note that Theorem 6.4 applies with a plus boundary condition, but can be adapted to treat the case of a plus boundary external field, see in particular Remark 1.1.

Proposition 6.5. *Let \mathbf{T}_n be a Galton–Watson tree with offspring distribution μ satisfying Assumption 1, and let $\mathbf{T}_n^* = \text{Pruned}_n(\mathbf{T}_n)$ be the pruned version of \mathbf{T}_n with $(h_v)_{v \in \partial\mathbf{T}}$ given by i.i.d. Bernoulli random variables with parameter $p_n \in (0, 1]$, as defined in Section 3.2.*

- *Upper bound. There is a constant $c > 0$ such that*

$$\mathbf{E} \otimes \mathbb{E}[\text{Cap}_3(\mathbf{T}_n^*)] \leq c \alpha_n \tag{6.2}$$

with

$$\alpha_n := \begin{cases} p_n(\tanh(\beta)\nu)^n & \text{if } \tanh(\beta)\nu \neq 1, \\ \min\{n^{-1/2}, p_n\} & \text{if } \tanh(\beta)\nu = 1. \end{cases}$$

In particular, if $\tanh(\beta)\nu \leq 1$ or if $\lim_{n \rightarrow \infty} p_n(\tanh(\beta)\nu)^n = 0$, then $\text{Cap}_3(\mathbf{T}_n^*)$ converges to 0 in L^1 , hence in probability.

- *Lower bound. If the offspring distribution μ admits a finite second moment and if $\liminf_{n \rightarrow \infty} p_n \nu^n = +\infty$, we have that*

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \otimes \mathbb{P}(\text{Cap}_3(\mathbf{T}_n^*) \leq \varepsilon \tilde{\alpha}_n) = 0, \tag{6.3}$$

with

$$\tilde{\alpha}_n := \begin{cases} \min\{1, p_n(\tanh(\beta)\nu)^n\} & \text{if } \tanh(\beta)\nu \neq 1, \\ \min\{n^{-1}, p_n\} & \text{if } \tanh(\beta)\nu = 1. \end{cases}$$

In particular, if $\tanh(\beta)\nu > 1$ and $\liminf_{n \rightarrow \infty} p_n(\tanh(\beta)\nu)^n > 0$, then $\text{Cap}_3(\mathbf{T}_n^*)$ remains asymptotically bounded away from 0; more precisely, $\{\text{Cap}_3(\mathbf{T}_n^*)^{-1}\}_{n \geq 0}$ is tight.

Remark 6.6. In the case where $\tanh(\beta)\nu \neq 1$, then we identify the correct order for $\text{Cap}_3(\mathbf{T}_n^*)$, hence for the root magnetization: $\min\{1, p_n(\tanh(\beta)\nu)^n\}$. In the critical case $\tanh(\beta)\nu = 1$, we have identified the correct order for the root magnetization only in the case where $p_n \leq n^{-1}$, but the upper and lower bound differ otherwise.

Remark 6.7. For the critical Ising model on the Galton–Watson tree, *i.e.* when one has $\tanh(\beta)\nu = 1$ and $p_n \equiv 1$, then Proposition 6.5 and Theorem 6.4 give (also applying Markov’s inequality) that for any $\varepsilon > 0$

$$\lim_{\varepsilon \downarrow 0} \liminf_{n \rightarrow \infty} \mathbf{P}\left(\frac{\varepsilon}{n} \leq r_n^+(\rho) \leq \frac{\varepsilon^{-1}}{\sqrt{n}}\right) = 1$$

Roughly speaking, it says that the magnetization of the root $r_n^+(\rho)$ is bounded from above by $O(1/\sqrt{n})$ and from below by $O(1/n)$. To our knowledge, such bounds do not appear in the literature; our lower bound seems not to be optimal so we believe that the root magnetization is of order $1/\sqrt{n}$, at least when the offspring distribution admits a finite second moment.

Proposition 6.5 amounts to estimating the 3-capacity of an inhomogeneous branching process. We find this question interesting in its own so we try to study this problem in the most general terms. We will use the structure on the pruned tree only at the end of the proof, to conclude the argument; the estimates in Section 5 turn out crucial for this last step.

6.2 Capacity of an inhomogeneous branching process

Let us consider a tree T of depth n generated by a branching process with inhomogeneous, n -dependent, offspring distributions $(\mu_k^{(n)})_{0 \leq k < n}$. Note that the offspring distributions may also depend on n , as it is the case for the pruned tree \mathbf{T}_n^* , see Proposition 4.3 and the definition (4.2) of μ_k^* .

For notational convenience, we will simply write $\mu_k := \mu_k^{(n)}$, and we assume that $\mu_k(0) = 0$ for all k . We denote ν_k the mean of μ_k (by convention $\nu_n = 1$) and for $0 \leq i < j \leq n+1$,

$$M_{i,j} := \prod_{k=i}^{j-1} \nu_k, \tag{6.4}$$

with $M_{i,i} = 1$ by convention. Hence, $M_{i,j}$ is the mean size of a population generated by a branching process with offspring distribution $(\mu_k)_{i \leq k < j}$. We then have the following result.

Proposition 6.8. *Let T be a tree of depth n generated by an inhomogeneous branching process with offspring distributions $(\mu_k^{(n)})_{0 \leq k < n}$, and associated resistances $(R_u)_{u \in T}$ given by $R_u = R^{-|u|}$ for some fixed $R > 0$. Recalling the definition (6.4) of $M_{i,j}$, we have the following upper bound*

$$\mathbf{E}[\text{Cap}_3(T)] \leq \left(\sum_{k=1}^n \left(\frac{1}{R^k M_{0,k}} \right)^2 \right)^{-1/2}. \quad (6.5)$$

For the lower bound, define for $0 \leq k \leq n-1$ the quantity $v_{k,n} := 1 + \sum_{i=k}^{n-1} \frac{\sigma_i^2}{M_{k,i}}$, where σ_i^2 is the variance of the law μ_i . Then we have that for any $\varepsilon > 0$,

$$\mathbf{P} \left(\text{Cap}_3(T) \leq \varepsilon \left(\sum_{k=0}^n \frac{v_{k,n}}{R^k M_{0,k}} \right)^{-1} \right) \leq \mathbf{P}(|T_n| \leq \varepsilon^{1/3} M_{0,n}) + \varepsilon^{1/3}. \quad (6.6)$$

6.2.1 Upper bound in Proposition 6.8: proof of (6.5)

For $u \in T$, recall the definition $\phi_p(u) := R_u \text{Cap}_p(T(u))$ in Lemma 6.3. Now, consider the random variable $\Phi_k := \phi_p(u)$ for some vertex $u \in T_k$ chosen uniformly at random from generation k . Then, as the number of offspring of vertices in the same generation are independent, by Lemma 6.3 we get for $0 \leq k < n$, with $s := p-1$,

$$\Phi_k = R \sum_{i=1}^{X_k} \frac{\Phi_{k+1}^{(i)}}{(1 + (\Phi_{k+1}^{(i)})^s)^{1/s}},$$

where X_k is a random variable with distribution μ_k and $(\Phi_{k+1}^{(i)})_{i \geq 1}$ are i.i.d. copies of Φ_{k+1} . Note that we also have used that $R_v = R^{-|v|}$ here.

Taking the expectation we obtain by the branching property

$$\mathbf{E}[\Phi_k] = R \nu_k \mathbf{E} \left[\frac{\Phi_{k+1}}{(1 + \Phi_{k+1}^s)^{1/s}} \right].$$

As the function $x \mapsto x/(1+x^s)^{1/s}$ is concave, we can apply Jensen's inequality to the previous expression: we obtain

$$\mathbf{E}[\Phi_k] \leq R \nu_k \frac{\mathbf{E}[\Phi_{k+1}]}{(1 + \mathbf{E}[\Phi_{k+1}]^s)^{1/s}} = \frac{R \nu_k}{(1 + \mathbf{E}[\Phi_{k+1}]^{-s})^{1/s}}. \quad (6.7)$$

Now, defining $z_k := \mathbf{E}[\Phi_k]^{-s}$, we can rewrite (6.7) as $z_k \geq (R \nu_k)^{-s} (1 + z_{k+1})$. By applying this inequality recursively, we obtain that for all $k \leq n$,

$$z_k \geq \sum_{i=k}^n \prod_{j=k}^i (R \nu_j)^{-s} + z_n \prod_{j=k}^{n-1} (R \nu_j)^{-s} \geq \sum_{i=k}^n (R^{i+1-k} M_{k,i+1})^{-s},$$

where we have used the definition (6.4) $M_{k,i+1} := \prod_{j=k}^i \nu_j$ for the last inequality. Since $z_0 = \mathbf{E}[\Phi_0]^{-s}$ and $\Phi_0 = \text{Cap}_p(T)$, we get that

$$\mathbf{E} [\text{Cap}_p(T)] \leq \left(\sum_{i=0}^{n-1} \left(\frac{1}{R^{i+1} M_{0,i+1}} \right)^s \right)^{-1/s}.$$

This gives a general bound on the p -capacity $\text{Cap}_p(T)$: with $p = 3$, $s = 2$, this is the bound (6.5), up to an index change. \square

6.2.2 Lower bound in Proposition 6.8: proof of (6.6)

Step 1. We start by showing that the 2-capacity of a tree is a lower bound for its 3-capacity.

Lemma 6.9. *Let T be a tree of depth n with no leaf excepts the vertices at generation n and equipped with resistances $(R_v)_{v \in T}$. Then, for any $1 < p \leq p'$ we have $\text{Cap}_p(T) \leq \text{Cap}_{p'}(T)$.*

Proof. Our starting point is Lemma 6.3. Recall that, for a vertex u on the tree T we defined $\phi_p(u) = R_u \text{Cap}_p(T(u))$ and that, for any $p > 1$, we have $\text{Cap}_p(T(u)) = 1$ when u is a leaf. Then, Lemma 6.3 shows that for any vertex $u \in T$ that is not a leaf we have

$$\phi_p(T(u)) = \sum_{u \rightarrow v} \frac{R_u}{R_v} \frac{\phi_p(T(v))}{(1 + \phi_p(T(v))^s)^{1/s}}, \quad \phi_{p'}(T(u)) = \sum_{u \rightarrow v} \frac{R_u}{R_v} \frac{\phi_{p'}(T(v))}{(1 + \phi_{p'}(T(v))^{s'})^{1/s'}},$$

with $s = p - 1$, $s' = p' - 1$. To conclude the proof of the lemma, it suffices to show that for any $0 < s \leq s'$ we have $(1 + x^s)^{1/s} \geq (1 + x^{s'})^{1/s'}$ for any $x \geq 0$ and apply this inequality iteratively. But this simply follows from the fact that for $\alpha \in [0, 1]$ and any $a, b \geq 0$ we have $(a + b)^\alpha \leq a^\alpha + b^\alpha$, so that $(1 + x^{s'})^{s/s'} \leq 1 + x^s$. \square

Let us define $A_n := \sum_{k=0}^n \frac{v_{k,n}}{R^k M_{0,k}}$. Using Lemma 6.9, we can bound

$$\mathbf{P} (\text{Cap}_3(T) \leq \varepsilon (A_n)^{-1}) \leq \mathbf{P} (\text{Cap}_2(T) \leq \varepsilon (A_n)^{-1}).$$

We can therefore focus on the 2-capacity $\text{Cap}_2(T)$ of the tree; recall that this is the effective conductance between the root ρ and the leaves ∂T (see e.g. [21, Ch. 2] for an overview of the theory of electrical networks on graphs). We then write $\text{Cap}_2(T) = \mathcal{C}(\rho \leftrightarrow \partial T) = \mathcal{R}(\rho \leftrightarrow \partial T)^{-1}$, where $\mathcal{R}(\rho \leftrightarrow \partial T)$ is the effective resistance between the root ρ and the leaves ∂T . We are now reduced to showing the following inequality

$$\mathbf{P} (\mathcal{R}(\rho \leftrightarrow \partial T) \geq \varepsilon^{-1} A_n) \leq \mathbf{P} (W_{0,n} \leq \varepsilon^{1/3}) + \varepsilon^{1/3}, \quad (6.8)$$

where $W_{0,n} = |T_n|/M_{0,n}$, see (6.10) below.

Step 2. Our second step is to find an upper bound on the resistance $\mathcal{R}(\rho \leftrightarrow \partial T)$. By Thomson's principle (see [21, §2.4]), we have

$$\mathcal{R}(\rho \leftrightarrow T_n) = \inf_{\theta: |\theta|=1} \sum_{u \in T} R_u \theta(u)^2, \quad (6.9)$$

where θ is a flow on the tree T .

An upper bound is therefore obtained simply by choosing a specific flow θ on T : similarly to what is done in [26, Lem. 2.2] (see also [7, Lem. 3.3]), we use the uniform flow $\hat{\theta}$ on T . For a vertex $u \in T_k$ in the k -th generation, we let

$$\hat{\theta}(u) := \frac{|T_{k,n}(u)|}{|T_n|},$$

where we note $T_{k,n}(u)$ the set of descendants in generation n of the vertex u , *i.e.* of individuals in generation $n - k$ of the sub-tree $T(u)$. We can easily see that the uniform flow $\hat{\theta}$ is a unitary flow: indeed, we have

$$\hat{\theta}(\rho) = \frac{1}{|T_n|} \sum_{\rho \rightarrow v} |T_{1,n}(v)| = 1.$$

Therefore, by the Thomson principle (6.9), we have

$$\mathcal{R}(\rho \leftrightarrow T) \leq \sum_{v \in V} R_v \hat{\theta}(v)^2 = \frac{1}{|T_n|^2} \sum_{k=0}^n \sum_{u \in T_k} R^{-k} |T_{k,n}(u)|^2.$$

Before we work on this upper bound, let us rewrite it using some notation. For $0 \leq k \leq n$ and $u \in T_k$, let us define for $\ell \in \{k, \dots, n\}$

$$W_{k,\ell}(u) = \frac{1}{M_{k,\ell}} |T_{k,\ell}(u)|, \quad (6.10)$$

and notice that $(W_{k,\ell}(u))_{k \leq \ell \leq n}$ is a martingale (with mean 1). We also denote $W_{0,k} := |T_k|/M_{0,k}$ the martingale $W_{0,k}(\rho)$. Then, we have

$$\begin{aligned} \mathcal{R}(\rho \leftrightarrow T) &\leq \frac{1}{(W_{0,n})^2} \sum_{k=0}^n R^{-k} \left(\frac{M_{0,n}}{M_{k,n}} \right)^2 \sum_{v \in T_k} W_{k,n}(v)^2 \\ &\leq \frac{1}{(W_{0,n})^2} \sum_{k=0}^n \frac{R^{-k}}{M_{0,k}} \frac{1}{M_{0,k}} \sum_{v \in T_k} W_{k,n}(v)^2, \end{aligned} \quad (6.11)$$

where we have used that $M_{0,n} = M_{0,k} M_{k,n}$.

Step 3. We are now ready to conclude the proof. Using the above inequality and decomposing according to whether $W_{0,n}$ is small or not, we have

$$\begin{aligned} \mathbf{P}\left(\mathcal{R}(\rho \leftrightarrow T) \geq \varepsilon^{-1} A_n\right) &\leq \mathbf{P}\left(\sum_{k=0}^n \frac{R^{-k}}{M_{0,k}} \frac{1}{M_{0,k}} \sum_{v \in T_k} W_{k,n}(v)^2 \geq \varepsilon^{-1} (W_{0,n})^2 A_n\right) \\ &\leq \mathbf{P}(W_{0,n} \leq \varepsilon^{1/3}) + \mathbf{P}\left(\sum_{k=0}^n \frac{R^{-k}}{M_{0,k}} \frac{1}{M_{0,k}} \sum_{v \in T_k} W_{k,n}(v)^2 \geq \varepsilon^{-1/3} A_n\right) \\ &\leq \mathbf{P}(W_{0,n} \leq \varepsilon^{1/3}) + \varepsilon^{1/3} A_n^{-1} \sum_{k=0}^n \frac{R^{-k}}{M_{0,k}} \mathbf{E}[(W_{k,n})^2], \end{aligned}$$

where we have used Markov's inequality for the last part, together with the fact that $(W_{k,n}(v))_{v \in T_k}$ are i.i.d.; here $W_{k,n}$ denotes a random variable with the same distribution. Now, $(|T_{k,\ell}|)_{0 \leq \ell \leq n}$ is an inhomogeneous branching process with offspring distribution $(\mu_\ell)_{k \leq \ell < n}$: it is standard to get that

$$\mathbf{E}[|T_{k,\ell+1}|^2] = (\nu_\ell)^2 \mathbf{E}[|T_{k,\ell}|^2] + \sigma_\ell^2 \mathbf{E}[|T_{k,\ell}|] = (\nu_\ell)^2 \mathbf{E}[|T_{k,\ell}|^2] + \sigma_\ell^2 M_{k,\ell},$$

where σ_ℓ^2 is the variance of the law μ_ℓ . Iterating, we get that

$$\mathbf{E}[|T_{k,n}|^2] = (M_{k,n})^2 + \sum_{i=k}^{n-1} (M_{i,n})^2 \sigma_i^2 M_{k,i}, \quad \text{so} \quad \mathbf{E}[(W_{k,n})^2] = 1 + \sum_{i=k}^{n-1} \frac{\sigma_i^2}{M_{k,i}} =: v_{k,n}.$$

This concludes the proof of (6.6), recalling the definition of A_n . \square

6.3 Conclusion of the proof of Proposition 6.5

We now apply Proposition 6.8 to obtain an upper and a lower bound on $\text{Cap}_3(\mathbf{T}_n^*)$, which is the content of Proposition 6.5.

6.3.1 Upper bound on $\text{Cap}_3(\mathbf{T}_n^*)$: proof of (6.2)

By Proposition 6.8, we have that

$$\mathbf{E} \otimes \mathbb{E}[\text{Cap}_3(\mathbf{T}_n^*)] \leq \left(\sum_{k=1}^n \left(\frac{1}{R^k M_{0,k}^*} \right)^2 \right)^{-1/2},$$

with $R := \tanh(\beta) < 1$ and $M_{i,j}^*$ is defined in (5.16). We now use Lemma 5.10 to get that

$$c_7^{-1} K_n \leq \sum_{k=1}^n \left(\frac{1}{R^k M_{0,k}^*} \right)^2 \leq c_6^{-1} K_n, \quad \text{with } K_n := \sum_{k=1}^n R^{-2k} (\nu)^{-2(k \wedge k^*)}.$$

Now, we can easily study K_n depending on the value of $R\nu$; note that we only need a lower bound on K_n since the above shows that $\mathbf{E} \otimes \mathbb{E}[\text{Cap}_3(\mathbf{T}_n^*)] \leq c K_n^{-1/2}$.

- If $R\nu \neq 1$, then keeping only the term $k = n$ in the sum, we have

$$K_n \geq (\nu)^{-2k^*} R^{-2n} = \left(\frac{1}{p_n(R\nu)^n} \right)^2,$$

where we have used that $n - k^* = \log_\nu(\frac{1}{p_n})$. We end up with

$$\mathbf{E} \otimes \mathbb{E}[\text{Cap}_3(\mathbf{T}_n^*)] \leq c p_n(R\nu)^n,$$

which gives (6.2) in the case $R\nu \neq 1$. Note that it goes to zero if $\lim_{n \rightarrow \infty} p_n(R\nu)^n = 0$.

- If $R\nu = 1$, then we have

$$K_n \geq \sum_{k=1}^{k^*} 1 + \sum_{k=k^*+1}^n R^{-2(k-k^*)} \geq k^* + R^{-2(n-k^*)} \geq \max\{k^*, p_n^{-2}\},$$

where we have used that $R^{n-k^*} = 1/\nu^{n-k^*}$ with $\nu^{n-k^*} = \nu^{\bar{k}^*} = \frac{1}{p_n}$, by the definition (5.6) of \bar{k}^* . Now, notice that $k^* = n - \log_\nu(\frac{1}{p_n})$ so either $p_n \leq n^{-1/2}$, or $p_n \geq n^{-1/2}$ and then $k^* \geq n/2$. We therefore end up with $K_n \geq c \max\{n, p_n^{-2}\}$, so that

$$\mathbf{E} \otimes \mathbb{E}[\text{Cap}_3(\mathbf{T}_n^*)] \leq c' \min\{n^{-1/2}, p_n\}.$$

This concludes the proof of (6.2) in the case $R\nu = 1$. \square

6.3.2 Lower bound on $\text{Cap}_3(\mathbf{T}_n^*)$: proof of (6.3)

Applying Proposition 6.8, we need to control the two quantities in (6.6). We treat them in the two following lemmas.

Lemma 6.10. *Let σ_k^* be the variance of the law μ_k^* and $M_{k,i} = \prod_{j=k}^{i-1} \nu_j^*$ as defined in (5.16). Then, there exists a constant $C > 0$ such that for all $0 \leq k \leq n$*

$$v_{k,n}^* := 1 + \sum_{i=k}^{n-1} \frac{(\sigma_i^*)^2}{M_{k,i}^*} \leq C.$$

Lemma 6.11. *Let $W_{0,n}^* = \frac{1}{M_{0,n}^*} |T_n^*|$. Then, if $\liminf_{n \rightarrow \infty} p_n \nu^n = +\infty$, we have*

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(W_{0,n} \leq \varepsilon) = 0.$$

Using these two lemmas, we get from Proposition 6.8 that

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \otimes \mathbb{P} \left(\text{Cap}_3(\mathbf{T}_n^*) \leq \varepsilon \left(\sum_{k=1}^n \frac{1}{R^k M_{0,k}^*} \right)^{-1} \right) = 0,$$

Using again Lemma 5.10 to bound $M_{0,k}^*$, we get that

$$c\tilde{K}_n \leq \sum_{k=1}^n \frac{1}{R^k M_{0,k}^*} \leq c'\tilde{K}_n \quad \text{with} \quad \tilde{K}_n := \sum_{k=1}^{k^*} (R\nu)^{-k} + (\nu)^{-k^*} \sum_{k=k^*+1}^n R^{-k},$$

so that $\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \otimes \mathbb{P}(\text{Cap}_3(\mathbf{T}_n^*) \leq \varepsilon(\tilde{K}_n)^{-1}) = 0$ and we now need to control \tilde{K}_n depending on the value of $R\nu$; note that we only need an upper bound.

- If $R\nu \neq 1$, since $R = \tanh(\beta) < 1$, we have that

$$\tilde{K}_n \leq \frac{1}{R\nu - 1} (1 - (R\nu)^{-k^*}) + c(\nu)^{-k^*} R^{-n} \leq c' + \frac{c''}{(R\nu)^{k^*}} + \frac{c}{p_n(R\nu)^n},$$

where we have used that $(\nu)^{n-k^*} = 1/p_n$ by definition (5.6) of k^* . In the case $R\nu < 1$ (where we already know that $\text{Cap}_3(\mathbf{T}_n^*)$ goes to 0 in probability) we get that $\tilde{K}_n \leq \frac{c}{p_n(R\nu)^n}$; note that this upper bound diverges. In the case $R\nu > 1$ the upper bound is a constant times $\max\{1, \frac{1}{p_n(R\nu)^n}\}$. This concludes the proof of (6.3) in the case $R\nu \neq 1$.

Note that in the case where $R\nu > 1$ and $\liminf_{n \rightarrow \infty} p_n(R\nu)^n > 0$, then $\limsup_{n \rightarrow \infty} (\tilde{K}_n)^{-1} < +\infty$, so it proves in particular that $\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \otimes \mathbb{P}(\text{Cap}_3(\mathbf{T}_n^*) \leq \varepsilon) = 0$.

- If $R\nu = 1$, then similarly to the above, we have

$$\tilde{K}_n \leq k^* + c(\nu)^{-k^*} R^{-n} \leq c' \max\{k^*, (\nu)^{n-k^*}\} = c' \max\{k^*, p_n^{-1}\}$$

recalling that $(\nu)^{n-k^*} = 1/p_n$. This concludes the proof of (6.3) in the case $R\nu = 1$. \square

6.3.3 Last technical lemmas: proof of Lemmas 6.10 and 6.11

Proof of Lemma 6.10. We start with the case where $k \geq k^*$. Then, splitting the sum in the definition of $v_{k,n}^*$ (see 6.8) according to whether $i \leq k^*$ or $i \geq k^*$, we obtain

$$v_{k,n}^* = 1 + \sum_{i=k}^{k^*} \frac{(\sigma_i^*)^2}{M_{k,i}^*} + \sum_{i=k^*+1}^{n-1} \frac{(\sigma_i^*)^2}{M_{k,i}^*} \leq 1 + c \sum_{i=k}^{k^*} \nu^{-(i-k)} + \sum_{i=k^*+1}^{n-1} (\sigma_i^*)^2,$$

where we have used that $(\sigma_i^*)^2 \leq m_2$, see Remark A.4, that $M_{k,i}^* = \frac{M_{0,i}^*}{M_{0,k}^*} \geq c(\nu)^{i-k}$ for $k \leq i \leq k^*$ thanks to Lemma 5.10, and finally that $M_{k,i}^* \geq 1$ in the case $i \geq k^*$. The first sum is finite since $\nu > 1$, and the second sum is also finite, using that $(\sigma_i^*)^2 \leq c\nu^{-(i-k^*)}$ for $i > k^*$, see Lemma A.3.

In the case where $k < k^*$, then we have similarly

$$v_{k,n}^* = 1 + \sum_{i=k}^{n-1} \frac{(\sigma_i^*)^2}{M_{k,i}^*} \leq 1 + \sum_{i=k}^{n-1} (\sigma_i^*)^2,$$

that last sum being also bounded by a universal constant. \square

Proof of Lemma 6.11. First of all, let us observe that in the general case of a branching process with inhomogeneous, n -dependent, offspring distributions $(\mu_k^{(n)})_{1 \leq k \leq n}$, one cannot use a martingale convergence as one would for a homogeneous branching process. For the pruned tree, we now give some *ad-hoc* proof, which uses the structure of the tree. For completeness, we give in Appendix A.3 and alternative line of proof that one should follow in the general case.

In the case of the pruned tree \mathbf{T}_n^* , we use the fact that the distribution of $|T_n^*|$ is explicit: conditionally on $|T_n|$, it is a binomial $\text{Bin}(|T_n|, p_n)$. In particular,

$$M_{0,n}^* := \mathbf{E}[|T_n^*| \mid |T_n^*| > 1] = \delta_0^{-1} \mathbf{E}[|T_n^*|] = \delta_0^{-1} p_n \nu^n,$$

with δ_0 close to 1. We therefore get that

$$\begin{aligned} \mathbf{P}(|T_n^*| \leq \varepsilon M_{0,n}) &\leq \mathbf{P}(|T_n| \leq 2\varepsilon \nu^n) + \mathbf{P}(|T_n| \geq 2\varepsilon \nu^n, \text{Bin}(|T_n|, p_n) \leq \varepsilon \delta_0^{-1} p_n \nu^n) \\ &\leq \mathbf{P}(|T_n| \leq 2\varepsilon \nu^n) + \mathbf{P}(\text{Bin}(2\varepsilon \nu^n, p_n) \leq \varepsilon \delta_0^{-1} p_n \nu^n). \end{aligned}$$

The second probability clearly goes to 0, as long as $\liminf_{n \rightarrow \infty} p_n \nu^n = +\infty$ (using also that δ_0 goes to 1). For the first probability, we use that \mathbf{T} is a homogeneous branching process: we have that $(\frac{1}{\nu^n} |T_n|)_{n \geq 0}$ is a martingale that converges almost surely to a *positive* random variable \mathcal{W} (recall we assumed that $\mu(0) = 0$). Hence, we get that

$$\limsup_{n \rightarrow \infty} \mathbf{P}(|T_n^*| \leq \varepsilon M_{0,n}) \leq \mathbf{P}(\mathcal{W} \leq 2\varepsilon),$$

which goes to 0 as $\varepsilon \downarrow 0$ and concludes the proof. \square

A Some technical proofs and comments

A.1 About recursions: proofs of Lemmas 5.1 and 5.2

Proof of Lemma 5.1. Recall that G is the generating function of a random variable $X \in \mathbb{N}$ with mean m_1 and second moment m_2 .

For (5.1), by the Taylor-Lagrange theorem, for all $s \in [0, 1]$ there exists $x_1 \in [s, 1]$ such that:

$$G(s) = G(1) + G'(1)(s-1) + \frac{G''(x_1)}{2}(s-1)^2. \quad (\text{A.1})$$

By convexity of the function G , we have that $0 \leq G''(x_1) \leq G''(1) = m_2 - m_1$. Since $G(1) = 1$, $G'(1) = m_1$, we therefore obtain from (A.1) that for all $s \in [0, 1]$

$$1 + m_1(s-1) \leq G(s) \leq 1 + m_1(s-1) + \frac{1}{2}(m_2 - m_1)(s-1)^2,$$

which is the desired bound.

The bounds of (5.2) are easily deduced from the definition $G(s) = \sum_{i \geq 1} \mu(i) s^i$, using that $s^i \leq s^{d_0+1}$ for $i \geq d_0+1$ and $\sum_{i \geq d_0+1} \mu(i) \leq 1$ for the upper bound. \square

Proof of Lemma 5.2. Thanks to the bounds (5.2), we have for any $j \geq 1$

$$\mu(d_0) (u_{j-1})^{d_0} \leq u_j = G(u_{j-1}) \leq \mu(d_0) (u_{j-1})^{d_0} + (u_{j-1})^{d_0+1}.$$

Thus, by iteration, we obtain for $j \geq 1$,

$$\prod_{i=0}^{j-1} \mu(d_0)^{d_0^i} \times (u_0)^{d_0^j} \leq u_j \leq \prod_{i=0}^{j-1} \mu(d_0)^{d_0^i} \times (u_0)^{d_0^{j-1}} \times \prod_{i=0}^{j-1} \left(1 + \frac{u_i}{\mu(d_0)}\right)^{d_0^{j-1-i}}. \quad (\text{A.2})$$

(i) In the case $d_0 = 1$, the lower bound is directly given by (A.2). For the upper bound, we only need to control $\prod_{i=0}^{j-1} (1 + \frac{1}{\mu(1)} u_i)$. We use that, G being convex, $G(s) \leq p_\alpha s$ for all $s \in [0, 1 - \alpha]$, with $p_\alpha = \frac{G(1-\alpha)}{1-\alpha} < 1$: we easily deduce by iteration that $u_i \leq (p_\alpha)^i u_0$ for all $i \geq 0$. This leads to the following:

$$\prod_{i=0}^{j-1} \left(1 + \frac{u_i}{\mu(1)}\right) \leq \exp\left(\frac{1}{\mu(1)} \sum_{i=0}^{j-1} u_i\right) \leq \exp\left(\frac{u_0}{\mu(1)} \frac{1}{1-p_\alpha}\right) \leq C_\alpha,$$

with $C_\alpha := e^{\frac{1-\alpha}{\mu(1)(1-p_\alpha)}}$, which proves (5.4).

(ii) In the case $d_0 \geq 2$, the lower bound is also immediate from (A.2), since $\sum_{i=0}^{j-1} d_0^i = \frac{d_0^j - 1}{d_0 - 1}$. For the upper bound, let us start by noting that if $\mu(d_0) = 1$, then we have the exact formula $u_j = (u_0)^{d_0^{j-1}}$, with $u_0 \leq 1 - \alpha$. If $\mu(d_0) < 1$, then there is some $v_0 \in (0, 1)$ such that $G(s) \leq s^{d_0}$ for all $s \leq v_0$. Since there is some j_α such that $u_j \leq v_0$ for all $j \geq j_\alpha$, we get as above that for $j \geq j_\alpha$

$$u_j \leq (u_{j_\alpha})^{d_0^{j-j_\alpha}} \leq \exp\left(- (d_0^{-j_\alpha} \log v_0) d_0^j\right).$$

Adjusting the constant to deal with the terms $j < j_\alpha$, this gives the desired bound. \square

A.2 Complementary estimates on the offspring distributions μ_k^*

For $0 \leq k < n$, let X_k^* be a random variable with distribution μ_k^* (the offspring distribution of the pruned tree) and let ν_k^* be its mean and $(\sigma_k^*)^2$ its variance.

Recall also Lemma 4.5 which says that $X_k^* \sim \widehat{\text{Bin}}(X, \delta_{k+1})$, where $\widehat{\text{Bin}}$ denotes a zero-truncated binomial, *i.e.* a binomial conditioned on being strictly positive, see Definition 4.4, and X is a random variable with law μ . For later reference, if $\hat{B} \sim \widehat{\text{Bin}}(n, p)$, its mean and variance are given by

$$\mathbb{E}[\hat{B}] = \frac{np}{1 - (1-p)^n}, \quad \text{Var}(\hat{B}) = \frac{np}{(1 - (1-p)^n)^2} (1-p - (1-p+np)(1-p)^n). \quad (\text{A.3})$$

Recall that ν, m_2 denote the first and second moment of the offspring distribution μ .

Lemma A.1. *Let X_k^* be a random variable with distribution μ_k^* . then*

$$\begin{aligned} \nu_k^* &:= \mathbb{E}[X_k^*] = \frac{\delta_{k+1}}{\delta_k} \nu, \\ (\sigma_k^*)^2 &:= \text{Var}(X_k^*) = \nu_k^* \left(1 + \delta_{k+1} \left(\frac{m_2}{\nu} - 1\right) - \nu_k^*\right). \end{aligned}$$

Proof of Lemma A.1. Letting $B_k \sim \text{Bin}(X, \delta_{k+1})$, we have thanks to Lemma 4.5 that

$$\nu_k^* = \sum_{z \geq 1} z \mathbb{P}(X_k^* = z) = \sum_{z \geq 1} z \frac{\mathbb{P}(B_k = z)}{\mathbb{P}(B_k > 0)} = \frac{1}{\delta_k} \sum_{z \geq 1} z \mathbb{P}(B_k = z),$$

where we have used that $P(B_k = 0) = \gamma_k$ so $P(B_k > 0) = \delta_k$, see the proof of Lemma 4.5. Working with the definition of B_k , we easily see that

$$\sum_{z \geq 1} z P(B_k = z) = \sum_{d \geq 1} \mu(d) E[\text{Bin}(d, \delta_{k+1})] = \delta_{k+1} \sum_{d \geq 0} \mu(d) d = \delta_{k+1} \nu.$$

For the expression of the variance, let us compute the second moment of X_k^* : for $0 \leq k < n$, similarly as above we have

$$\begin{aligned} E[(X_k^*)^2] &= \frac{1}{\delta_k} \sum_{z \geq 1} z^2 P(B_k = z) = \frac{1}{\delta_k} \sum_{d \geq 0} \mu(d) E[\text{Bin}(d, \delta_{k+1})^2] \\ &= \frac{1}{\delta_k} (\delta_{k+1} \gamma_{k+1} \nu + \delta_{k+1}^2 m_2). \end{aligned}$$

Thus, by computing $E[(X_k^*)^2] - E[X_k^*]^2$, we get the following expression for the variance:

$$(\sigma_k^*)^2 = \frac{\delta_{k+1}}{\delta_k} \left(\gamma_{k+1} \nu + \delta_{k+1} m_2 - \frac{\delta_{k+1}}{\delta_k} (\nu)^2 \right),$$

which gives the correct expression, using the formula for ν_k^* . \square

We now give some property of the means $(\nu_k^*)_{0 \leq k < n}$.

Lemma A.2. *The sequence of means $(\nu_k^*)_{0 \leq k < n}$ is non-increasing.*

Proof. We start by writing ν_k^* only in terms of the random variable X with distribution μ and of the parameter γ_{k+1} . Note that, since conditionally on X the random variable B_k is a zero-truncated binomial $\widehat{\text{Bin}}(X, 1 - \gamma_{k+1})$, we have

$$E[X_k^* | X] = \frac{X(1 - \gamma_{k+1})}{1 - (\gamma_{k+1})^X}.$$

Thus, $\nu_k^* = E[f(\gamma_{k+1}, X)]$ for $f(x, y) := \frac{y(1-x)}{1-x^y}$. As $X \geq 1$ and $\gamma_{k+1} \in [0, 1]$ is non-decreasing, it is enough to prove that $x \mapsto f(x, y)$ is non-increasing on $[0, 1]$, for any $y \geq 1$.

Fix $y \geq 1$ and let $f_y(x) := f(x, y)$, $g_y(x) := x^y - 1 + yx^{y-1}(1-x)$ for $x \in [0, 1]$. Thus, we can write

$$f'_y(x) = \frac{y}{(1-x^y)^2} g_y(x).$$

We can easily see that, for $y \geq 1$ and $x \in [0, 1]$, we have $g_y(x) \leq 0$, as $g_y(0) = -1$, $g_y(1) = 0$ and $x \mapsto g_y(x)$ is non-decreasing. Therefore, $f'_y(x) \leq 0$ for $x \in [0, 1]$, which concludes the proof. \square

Let us now study the variances $((\sigma_k^*)^2)_{0 \leq k < n}$ and prove a statement analogous to Lemma 5.9; in particular we also observe the phase transition from Proposition 5.4 in terms of the variances. Recall the definition (5.6) of $k^* := \log_\nu(\nu^n p_n)$.

Lemma A.3. *There are constants c, c' (that depend only on the law μ) such that:*

$$(\sigma_k^*)^2 \leq c \left(\frac{1}{\nu} \right)^{k-k^*} \quad \text{for } k \geq k^*, \quad (\text{A.4})$$

$$|(\sigma_k^*)^2 - \sigma^2| \leq c' e^{-c_4(k^*-k)} \quad \text{for } k \leq k^*. \quad (\text{A.5})$$

Remark A.4. Note that we have no monotonicity for the sequence $(\sigma_k^*)_{0 \leq k \leq n-1}$, nor the upper bound $(\sigma_k^*)^2 \leq \sigma^2$ in general. Indeed, choosing $\mu(d_0) = 1$ for some $d_0 \geq 2$, then $\sigma^2 = 0$ but $(\sigma_k^*)^2 > 0$, since then μ_k^* is a zero-truncated binomial $\widehat{\text{Bin}}(d, \bar{\delta}_{k+1})$. On the other hand, we have a general (easy) bound: $(\sigma_k^*)^2 \leq \mathbb{E}[\widehat{\text{Bin}}(X, \delta_{k+1})^2] \leq \mathbb{E}[X^2] =: m_2$.

Proof. Recall from Lemma A.1 the formula: $(\sigma_k^*)^2 = \nu_k^* (1 + \delta_{k+1}(\frac{m_2}{\nu} - 1) - \nu_k^*)$.

For $k \geq k^*$, using Lemma 5.9, we get that $\nu_k^* \leq m_1$ and $1 - \nu_k^* \leq 0$, so

$$(\sigma_k^*)^2 \leq \delta_{k+1}(m_2 - \nu),$$

and we directly get the upper bound (A.4) thanks to Proposition 5.4.

For $k \leq k^*$, using that $\delta_{k+1} = 1 - \gamma_{k+1}$ and setting $\varepsilon_k := \nu - \nu_k^* \geq 0$, we have (bounding also $\nu_k^* \leq \nu$),

$$(\sigma_k^*)^2 \leq \nu \left(1 - \nu + \varepsilon_k + \frac{m_2}{\nu} - 1 - \gamma_{k+1} \left(\frac{m_2}{\nu} - 1 \right) \right) \leq \sigma^2 + \varepsilon_k \nu,$$

where we have used that $\frac{m_2}{\nu} - 1 \geq 0$. Using Lemma 5.9, we get the desired upper bound. For the lower bound, we write similarly

$$(\sigma_k^*)^2 \geq (\nu - \varepsilon_k) \left(1 - \nu + \frac{m_2}{\nu} - 1 - \gamma_{k+1} \left(\frac{m_2}{\nu} - 1 \right) \right) \geq \sigma^2 - \frac{\sigma^2}{\nu} \varepsilon_k - \gamma_{k+1}(m_2 - \nu).$$

Using Lemma 5.9 and Proposition 5.4, this concludes the proof. \square

A.3 About general inhomogeneous, n -dependent, branching processes: alternative proof of Lemma 6.11

In this section, we give an alternative proof of Lemma 6.11, following a different scheme of proof that remains valid for general inhomogeneous, n -dependent, branching processes T , with offspring distribution $(\mu_k^{(n)})_{0 \leq k < n}$; note that what makes the result non-trivial is precisely the n -dependence since one cannot use almost sure martingale convergence results. We denote ν_k, σ_k^2 the mean and variance of $\mu_k = \mu_k^{(n)}$.

We define $W_{0,n} := \frac{1}{M_{0,n}} |T_n|$ as in Section 6.2 and we want to control $\mathbf{P}(W_{0,n} \leq \varepsilon)$. Equivalently to Lemma 6.11, we want to prove the following:

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbf{E}[e^{-\lambda W_{0,n}}] = 0. \quad (\text{A.6})$$

We let $G_k = G_k^{(n)}$ be the generating function of μ_k (recall we assumed $\mu_k(0) = 0$),

$$G_k(s) := \sum_{d=1}^{\infty} \mu_k(d) s^d \quad s \in [0, 1]$$

and we note that $G_{T_n} = G_0 \circ G_1 \circ \dots \circ G_{n-1}$ is the generating function of T_n . Hence, we can rewrite $\mathbf{E}[e^{-\lambda W_{0,n}}] = G_{T_n}(e^{-\lambda/M_{0,n}})$. Since, for any fixed $\lambda > 0$ and n large enough, we have $e^{-\lambda/M_{0,n}} \leq 1 - \frac{\lambda}{2M_{0,n}}$, we get that (A.6), hence Lemma 6.11 is equivalent to having

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} G_{T_n} \left(1 - \frac{\lambda}{M_{0,n}} \right) = 0. \quad (\text{A.7})$$

Let us describe the general strategy to prove (A.7). For $s \in (0, 1)$, define the sequence $(\bar{\Gamma}_k(s))_{0 \leq k \leq n}$ as

$$\begin{aligned} \bar{\Gamma}_0(s) &:= \bar{\Gamma}_0 = s \\ \bar{\Gamma}_k(s) &:= \bar{\Gamma}_k = G_{n-k} \circ \dots \circ G_{n-1}(s) = G_{n-k}(\bar{\Gamma}_{k-1}). \end{aligned}$$

Note that we keep the notation $\bar{\Gamma}_k$ (instead of Γ_k) in analogy with Section 5, since we index our quantities in terms of the distance to the leaves.

The general strategy of proof can then be decomposed into the following steps:

- 1) Find $\bar{\kappa} = \bar{\kappa}(\lambda, n)$ such that for $s := 1 - \frac{\lambda}{M_{0,n}}$ we have $\bar{\Gamma}_{\bar{\kappa}}(s) \leq 1 - \eta$, for some universal constant $\eta > 0$.

Informally, $\bar{\kappa}$ is the number of steps needed for $\bar{\Gamma}_k(s)$ to start getting away from 1.

- 2) Show that $n - \bar{\kappa} \geq \kappa_0(\lambda)$ uniformly in n , with $\lim_{n \rightarrow \infty} \kappa_0(\lambda) = +\infty$.

Informally, this tells that after $\bar{\kappa}$ iterations there are still many iterations remaining.

- 3) Show that for any $\eta > 0$ and any $\varepsilon > 0$, there exists some universal $k_{\eta, \varepsilon} > 0$ (i.e. that does not depend on n) such that $G_0 \circ \dots \circ G_k(1 - \eta) \leq \varepsilon$ for $k \geq k_{\eta, \varepsilon}$.

Informally, this tells that if one has reached $\bar{\Gamma}_k \leq 1 - \eta$ and that a large number of iterations remains, then $\bar{\Gamma}_n := G_{T_n}(1 - \frac{\lambda}{M_{0,n}})$ is small.

These three steps indeed allow us to prove (A.7): applying Step 1, we have by monotonicity of the generating functions,

$$G_{T_n} \left(1 - \frac{\lambda}{M_{0,n}} \right) \leq G_0 \circ \dots \circ G_{n-\bar{\kappa}}(1 - \eta).$$

Then, applying Step 2 in the second inequality (together with the fact that $G_i(s) \leq s$) for any $s \in [0, 1]$, we get that

$$G_{T_n} \left(1 - \frac{\lambda}{M_{0,n}} \right) \leq G_0 \circ \dots \circ G_{\kappa_0(\lambda)}(1 - \eta)$$

Step 3 above, together with the fact that $\kappa_0(\lambda) \rightarrow \infty$ as $\lambda \rightarrow \infty$ concludes the proof.

Remark A.5. More generally, one could simply define some $\bar{\kappa} = \bar{\kappa}(n, \lambda)$, show some upper bound $\bar{\Gamma}_{\bar{\kappa}}(s) \leq 1 - \bar{\eta}$ with $\bar{\eta} := \bar{\eta}(n, \lambda)$, and then prove that $G_0 \circ \dots \circ G_{n-\bar{\kappa}}(1 - \eta) \rightarrow 0$. One should think of the three steps described above as the general scheme one should follow to prove (A.7). We refer to (A.15) for a general sufficient condition to get (A.7).

We now perform these three steps in a general framework: for the first two steps, we give an explicit candidate for $\bar{\kappa}$, see (A.10), and a bound for $\bar{\Gamma}_{\bar{\kappa}}$ in terms of $\bar{\kappa}$, see (A.11).

We apply these to the inhomogeneous branching process \mathbf{T}_n^* , where all quantities are explicit.

Step 1. For $s \in (0, 1)$, we define

$$\bar{\kappa} := \bar{\kappa}(n, s) := \max \left\{ 0 \leq k \leq n, \sum_{i=n-k+1}^n \frac{\tilde{C}_{i-1}}{M_{0,i}} < \frac{1}{2M_{0,n}(1-s)} \right\},$$

where $\tilde{C}_j := \frac{1}{2}(\frac{\sigma_j^2}{\nu_j} + \nu_j - 1)$. We then have the following result, which helps us deal with the first step of the proof (its proof is analogous to that of Proposition 5.4).

Lemma A.6. *For $s \in (0, 1)$, there is a constant c_9 that depend only on μ , such that*

$$c_9 M_{n-k,n}(1-s) \leq 1 - \bar{\Gamma}_k(s) \leq M_{n-k,n}(1-s) \quad \text{for all } k \leq \bar{\kappa}. \quad (\text{A.8})$$

The upper bound holds for any $0 \leq k \leq n$.

Proof. Let us define $\bar{\Delta}_k := 1 - \bar{\Gamma}_k$, which satisfies the recursion: $\bar{\Delta}_0 = 1 - s$, $\bar{\Delta}_k = F_{n-k}(\bar{\Delta}_{k-1})$, with $F_j(s) := 1 - G_j(1-s)$. Note that as in Lemma 5.1 or (5.3), we have some general bounds for the generating functions $(G_k)_{0 \leq k < n}$ or $(F_k)_{0 \leq k < n}$, in terms of their first and second moment: for all $0 \leq k < n$ and for $t \in [0, 1]$, we have

$$\nu_j t(1 - \tilde{C}_j t) \leq F_j(t) \leq \nu_j t, \quad (\text{A.9})$$

where $\tilde{C}_j := \frac{1}{2\nu_j} G_j''(1) = \frac{1}{2}(\frac{\sigma_j^2}{\nu_j} + \nu_j - 1)$. Now, the upper bound in (A.8) is a direct recursive application of the right-hand side bound in (A.9).

For the lower bound, we use the following result, which is analogous to Lemma 5.6; we omit its proof as it is very similar. Let $s \in (0, 1]$ and set $\bar{\kappa}_1(s) := \max\{1 \leq j \leq n : \sum_{i=0}^{j-1} \tilde{C}_{n-i-1} \bar{\Delta}_i(s) < 1/2\}$. Then, for all $k \leq \bar{\kappa}_1$, we have

$$\frac{1}{2} M_{n-k,n}(1-s) \leq \bar{\Delta}_k(s) \leq M_{n-k,n}(1-s).$$

Now, notice that by the upper bound on $\bar{\Delta}_i(s)$, we have

$$\sum_{i=0}^{j-1} \tilde{C}_{n-i-1} \bar{\Delta}_i(s) \leq \sum_{i=0}^{j-1} \tilde{C}_{n-i-1} M_{n-i,n}(1-s) = (1-s) M_{0,n} \sum_{i=0}^{j-1} \frac{\tilde{C}_{n-i-1}}{M_{0,n-i}}.$$

Hence, $\bar{\kappa}_1 \geq \max\{1 \leq j \leq n : \sum_{i=0}^{j-1} \frac{\tilde{C}_{n-i-1}}{M_{0,n-i}} \geq (2M_{0,n}(1-s))^{-1}\} =: \bar{\kappa}$, which concludes the proof. \square

Step 2. In the case where $s = 1 - \frac{\lambda}{M_{0,n}}$, i.e. $M_{0,n}(1-s) = \lambda$, then $\bar{\kappa}$ reduces to

$$\bar{\kappa} := \bar{\kappa}(n, \lambda) := \max \left\{ 0 \leq k \leq n, \sum_{i=n-k}^n \frac{\tilde{C}_{i-1}}{M_{0,i}} < \frac{1}{2\lambda} \right\}, \quad (\text{A.10})$$

and (A.8) in Proposition A.6 gives that for $k = \bar{\kappa}$,

$$\bar{\Gamma}_{\bar{\kappa}}(s) \leq 1 - M_{n-\bar{\kappa},n}(1-s) = 1 - \frac{\lambda}{M_{0,n-\bar{\kappa}}}. \quad (\text{A.11})$$

Steps 1-2. Application to the pruned tree \mathbf{T}_n^ .* For the inhomogeneous, n -dependent, branching process \mathbf{T}_n^* , notice that we are able to control both $\tilde{C}_i^* = \frac{1}{2}(\frac{(\sigma_i^*)^2}{\nu_i^*} + \nu_i^* - 1)$ and $M_{0,i}^*$ (we add \cdot^* to the notation to make clear that we are dealing with this specific case).

Using Lemma 5.10, we get that $M_{0,i}^*$ is of order $(\nu)^{i \wedge k^*}$. Also, using Lemma 5.9 to bound $\nu_i^* - 1$ and Lemma A.3 to bound $(\sigma_i^*)^2$, we get that

$$\tilde{C}_i^* \leq \tilde{C} \text{ for } i \leq k^*, \quad \tilde{C}_i^* \leq c(\nu)^{-(i-k^*)} \text{ for } i \geq k^*.$$

Then, we obtain, for $n - k \geq k^*$

$$\sum_{i=n-k+1}^{n-1} \frac{\tilde{C}_{i-1}^*}{M_{0,i}^*} \leq C' \sum_{i=n-k+1}^{k^*} (\nu)^{-i} + c(\nu)^{-k^*} \sum_{i=k^*+1}^n (\nu)^{-(i-k^*)} \leq c'(\nu)^{-(n-k)} + c''(\nu)^{-k^*}.$$

In the case where $p_n \nu^n \rightarrow \infty$ then we have that $k^* \rightarrow \infty$, so we get that for n large enough

$$\bar{\kappa}^* \geq \max \left\{ 0 \leq k \leq n, (\nu)^{-(n-k)} < \frac{1}{4c'\lambda} \right\},$$

so we end up with $\bar{\kappa}^* \geq n - \log_\nu(c\lambda)$. It is easily seen that an analogous upper bound also holds.

This enables us to conclude both Steps 1-2. Indeed, we have that $n - \bar{\kappa}^* \geq \kappa_0(\lambda) := \log_\nu(c\lambda)$, which goes to $+\infty$ as $\lambda \rightarrow \infty$, which is what is required in Step 2. Also, using that $n - \bar{\kappa}^* \leq \kappa'_0(\lambda) := \log_\nu(c'\lambda)$, we get for $s = 1 - \frac{\lambda}{M_{0,n}^*}$,

$$\bar{\Gamma}_{\bar{\kappa}^*}(s) \leq 1 - \frac{\lambda}{M_{0,n-\bar{\kappa}^*}^*} \leq 1 - \frac{\lambda}{M_{0,\kappa'_0}^*} \leq 1 - c\lambda(\nu)^{-\kappa'_0},$$

where we have used Lemma 5.10 to bound M_{0,κ'_0}^* (note that $\kappa'_0 \leq k^*$, at least for n large enough since $k^* \rightarrow \infty$). By definition of κ'_0 , we get that $\bar{\Gamma}_{\bar{\kappa}^*}(s) \leq 1 - \eta$ for some universal constant $\eta > 0$: this ensures that Step 1 is verified with $\bar{\kappa}^*$ defined above.

Step 3. For any $\eta > 0$ and any $\varepsilon > 0$, let us define the following quantity

$$K(\eta, \varepsilon) := \min \{ k \geq 1, G_0 \circ G_1 \circ \dots \circ G_k(1 - \eta) \leq \varepsilon \}. \quad (\text{A.12})$$

If Steps 1-2 have been verified, Step 3 consists in showing that $K(\eta, \varepsilon)$ is bounded by a universal constant, *i.e.* a quantity that does not depend on n ; recall that the offspring distributions $(\mu_k)_{0 \leq k < n}$ may depend on n .

Note that the order of composition in (A.12) is important and there is no easy iteration for the quantity $G_0 \circ \dots \circ G_{k-1}(1 - \eta)$. However, following a similar argument

as in Lemma 5.2 allows us to deduce a good upper bound. Indeed, by convexity, we have that for any $u \in (0, 1 - \eta)$,

$$G_i(u) \leq S_i u, \quad \text{where } S_i = S_i(\eta) = \frac{1}{1 - \eta} G_i(1 - \eta).$$

Iterating, we therefore get that $G_0 \circ G_1 \circ \dots \circ G_{k-1}(1 - \eta) \leq (1 - \eta) \prod_{i=0}^{k-1} S_i$, so we obtain an upper bound for $K(\eta, \varepsilon)$:

$$K(\eta, \varepsilon) \leq \tilde{K}(\eta, \varepsilon) := \min \left\{ k \geq 1, \prod_{i=0}^{k-1} S_i \leq \frac{\varepsilon}{1 - \eta} \right\}. \quad (\text{A.13})$$

One then simply needs to show that for any $\varepsilon > 0$ there is a (universal) constant $k_{\varepsilon, \eta} > 0$ such that

$$\prod_{i=0}^{k-1} \frac{G_k(1 - \eta)}{1 - \eta} \leq \varepsilon \quad \text{for all } k \geq k_{\varepsilon, \eta}.$$

This follows for instance if we prove that there exists a constant $S_\eta < 1$ and a sequence k_n with $\lim_{n \rightarrow \infty} k_n = +\infty$ such that for n large enough

$$S_k = \frac{G_k(1 - \eta)}{1 - \eta} \leq S_\eta, \quad \text{uniformly for } k \leq k_n. \quad (\text{A.14})$$

Indeed, in that case, for $k_{\varepsilon, \eta} = \frac{\log \varepsilon}{\log S_\eta}$, we get that $k_{\varepsilon, \eta} \leq k_n$ for n large enough, so

$$\prod_{i=0}^{k_{\varepsilon, \eta}-1} \frac{G_k(1 - \eta)}{1 - \eta} \leq (S_\eta)^{k_{\varepsilon, \eta}} = \varepsilon,$$

for n large enough. This concludes Step 3 since $k_{\varepsilon, \eta}$ is a universal constant.

Step 3. Application to the pruned tree \mathbf{T}_n^ .* We now to prove the sufficient condition (A.14) for the pruned tree \mathbf{T}_n^* . A simple calculation shows that for any $u \in [0, 1]$,

$$|G_k^*(u) - G(u)| \leq d_{\text{TV}}(\mu_k^*, \mu) \leq c'_4 e^{-c_4(k^* - k)},$$

where the second inequality holds for $k \leq k^*$, using Proposition 5.7. Now, if we set $k_n := \frac{1}{2}k^*$, which goes to $+\infty$ as $n \rightarrow \infty$ as soon as $p_n \nu^n \rightarrow +\infty$, we get that

$$\frac{G_k^*(1 - \eta)}{1 - \eta} \leq \frac{G(1 - \eta)}{1 - \eta} + \frac{c'_4}{1 - \eta} e^{-c_4 k_n} \quad \text{uniformly for } k \leq k_n.$$

This shows (A.14), for instance with the constant $S_\eta := 1 - \frac{1}{2}(1 - \frac{G(1 - \eta)}{1 - \eta})$, and concludes the proof of Step 3. \square

Remark A.7. Note that we have derived a slightly more general sufficient condition for (A.7). Indeed, defining $\bar{\kappa}$ as in (A.10) and $\kappa := n - \bar{\kappa}$, the bound (A.11) gives that $\bar{\Gamma}_{\bar{\kappa}} \leq 1 - \eta$ with $\eta = \eta_{n, \lambda} = \frac{\lambda}{M_{0, \kappa}}$. One then simply needs to verify that for

any $\varepsilon > 0$, we have $G_0 \circ \dots \circ G_\kappa(1 - \eta_{n,\lambda}) \leq \varepsilon$ at least for λ large, uniformly in n . One therefore needs to show that $K(\frac{\lambda}{M_{0,n-\bar{\kappa}}}, \varepsilon) \leq \kappa$ or even $\tilde{K}(\frac{\lambda}{M_{0,n-\bar{\kappa}}}, \varepsilon) \leq \kappa$ (recall the definitions (A.12)-(A.13)), at least for λ large enough. Put otherwise, a sufficient condition for having (A.7), hence Lemma 6.11, is the following:

$$\lim_{\lambda \rightarrow +\infty} \limsup_{n \rightarrow \infty} \prod_{i=0}^{\kappa-1} \frac{G_i(1 - \frac{\lambda}{M_{0,\kappa}})}{1 - \frac{M_{0,\kappa}}{\lambda}} = 0, \quad (\text{A.15})$$

where $\kappa := n - \bar{\kappa}$, with $\bar{\kappa} = \bar{\kappa}(n, \lambda)$ defined in (A.10).

Acknowledgements. The authors warmly thank Remco van der Hofstad and Arnaud Le Ny for many enlightening discussions. I.A.V. acknowledges the support of CNRS IRP Bezout - Eurandom. Q.B. acknowledges the support of grant ANR-22-CE40-0012.

References

- [1] A. Basak and A. Dembo. Ferromagnetic Ising measures on large locally tree-like graphs. *The Annals of Probability*, 45(2):780 – 823, 2017.
- [2] P. Blekher and N. Ganikhodgaev. On pure phases of the ising model on the bethe lattices. *Theory of Probability & Its Applications*, 35(2):216–227, 1991.
- [3] T. Bodineau. Translation invariant gibbs states for the ising model. *Probability theory and related fields*, 135:153–168, 2006.
- [4] A. Bovier. *Statistical Mechanics of Disordered Systems: A Mathematical Perspective*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2006.
- [5] V. H. Can. Critical behavior of the annealed Ising model on random regular graphs. *Journal of Statistical Physics*, 169:480–503, 2017.
- [6] V. H. Can, C. Giardinà, C. Giberti, and R. van der Hofstad. Annealed Ising model on configuration models. *Annales de l’Institut Henri Poincaré (B) Probabilités et Statistiques*, 58(1):134–163, 2022.
- [7] D. Chen, Y. Hu, and S. Lin. Resistance growth of branching random networks. *Electronic Journal of Probability*, 23:1–17, Jan. 2018.
- [8] L. Coquille, C. Külske, and A. Le Ny. Extremal inhomogeneous gibbs states for sos-models and finite-spin models on trees. *Journal of Statistical Physics*, 190(4):71, 2023.
- [9] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Brazilian Journal of Probability and Statistics*, 24(2):137 – 211, 2010.
- [10] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *The Annals of Applied Probability*, 20(2):565–592, Apr. 2010.
- [11] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising models on power-law random graphs. *Journal of Statistical Physics*, 141(4):638–660, Nov. 2010.
- [12] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising critical exponents on random trees and graphs. *Communications in Mathematical Physics*, 328:355–395, 2014.
- [13] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275–1335, Oct. 2008.

- [14] E. O. Endo, A. C. van Enter, and A. Le Ny. The roles of random boundary conditions in spin systems. *In and Out of Equilibrium 3: Celebrating Vladas Sidoravicius*, pages 371–381, 2021.
- [15] S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, 1 edition, Nov. 2017.
- [16] D. Gandolfo, J. Ruiz, and S. Shlosman. A manifold of pure gibbs states of the ising model on a cayley tree. *Journal of Statistical Physics*, 148:999–1005, 2012.
- [17] C. Giardinà, C. Giberti, R. van der Hofstad, and M. Prioriello. Quenched central limit theorems for the Ising model on random graphs. *Journal of Statistical Physics*, 160(6):1623–1657, jul 2015.
- [18] Y. Higuchi. Remarks on the limiting Gibbs states on a $(d + 1)$ -tree. *Publications of the Research Institute for Mathematical Sciences*, 13(2):335–348, 1977.
- [19] R. Lyons. The Ising model and percolation on trees and tree-like graphs. *Communications in Mathematical Physics*, 125(2):337–353, Jan. 1989.
- [20] R. Lyons. Factors of IID on trees. *Combinatorics, Probability and Computing*, 26(2):285–300, 2017.
- [21] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2017.
- [22] A. Montanari. Statistical mechanics and algorithms on sparse and random graphs. *Lectures on Probability Theory and Statistics*. Saint-Flour, 2013.
- [23] A. Montanari, E. Mossel, and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probability Theory and Related Fields*, 152:31–51, 2012.
- [24] D. Nam, A. Sly, and L. Zhang. Ising model on trees and factors of iid. *Communications in Mathematical Physics*, pages 1–38, 2022.
- [25] C. M. Newman and C. C. Wu. Markov fields on branching planes. *Probability Theory and Related Fields*, 85(4):539–552, 1990.
- [26] R. Pemantle and Y. Peres. Galton-Watson trees with the same mean have the same polar sets. *The Annals of Probability*, pages 1102–1124, 1995.
- [27] R. Pemantle and Y. Peres. The critical Ising model on trees, concave recursions and non-linear capacity. *The Annals of Probability*, 38(1):184–206, Jan. 2010.
- [28] U. A. Rozikov. *Gibbs Measures On Cayley Trees*. World Scientific, July 2013.
- [29] R. van der Hofstad. *Random Graphs and Complex Networks, Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.
- [30] R. van der Hofstad. The giant in random graphs is almost local. *preprint arXiv:2103.11733*, 2021.
- [31] R. van der Hofstad. *Random Graphs and Complex Networks, Vol. 2*. accessible on R. van der Hofstad website, 2023.
- [32] A. C. van Enter, K. Netočný, and H. G. Schaap. Incoherent boundary conditions and metastates. *Lecture Notes-Monograph Series*, pages 144–153, 2006.
- [33] C. C. Wu. Ising models on hyperbolic graphs. *Journal of Statistical Physics*, 85:251–259, 1996.
- [34] C. C. Wu. Ising models on hyperbolic graphs II. *Journal of Statistical Physics*, 100:893–904, 2000.

I. AYUSO VENTURA, LABORATOIRE D'ANALYSE ET DE MATHÉMATIQUES APPLIQUÉES, UNIVERSITÉ PARIS-EST CRÉTEIL, 61 AVENUE DU GÉNÉRAL DE GAULLE, BÂTIMENT P, 94010 CRÉTEIL, FRANCE.

E-mail address: `irene.ayuso-ventura@u-pec.fr`

Q. BERGER, SORBONNE UNIVERSITÉ, LABORATOIRE DE PROBABILITÉS, STATISTIQUE ET MODÉLISATION, 75005 PARIS, FRANCE and DMA, ÉCOLE NORMALE SUPÉRIEURE, UNIVERSITÉ PSL, 75005 PARIS, FRANCE.

E-mail address: `quentin.berger@sorbonne-universite.fr`