The longest increasing subsequence of Brownian separable permutons

Arka Adhikari
University of Maryland

Jacopo Borga
MIT

Thomas Budzinski
ENS de Lyon

William Da Silva
University of Vienna

Delphin Sénizergues
Université Paris Nanterre

Abstract

We establish a scaling limit result for the length $\mathrm{LIS}(\sigma_n)$ of the longest increasing subsequence of a permutation σ_n of size n sampled from the Brownian separable permuton $\boldsymbol{\mu}_p$ of parameter $p \in (0,1)$, which is the universal limit of pattern-avoiding permutations. Specifically, we prove that

$$\frac{\mathrm{LIS}(\sigma_n)}{n^\alpha} \xrightarrow[n\to\infty]{\text{a.s.}} X,$$

where $\alpha = \alpha(p)$ is the unique solution in the interval $(1/2, 1)$ to the equation

$$\frac{1}{4^{\frac{1}{2\alpha}}\sqrt{\pi}} \frac{\Gamma\left(\frac{1}{2} - \frac{1}{2\alpha}\right)}{\Gamma\left(1 - \frac{1}{2\alpha}\right)} = \frac{p}{p-1},$$

and $X = X(p)$ is a non-deterministic and a.s. positive and finite random variable, which is a measurable function of the Brownian separable permuton. Notably, the exponent $\alpha(p)$ is an increasing continuous function of p with $\alpha(0^+) = 1/2$, $\alpha(1^-) = 1$ and $\alpha(1/2) \approx 0.815226$, which corresponds to the permuton limit of uniform separable permutations. We prove analogous results for the size of the largest clique of a graph sampled from the Brownian cographon of parameter $p \in (0,1)$.

Contents

1 Introduction

The Brownian separable permuton is the universal random limiting permuton describing the scaling limit of many families of random pattern-avoiding permutations. The Brownian cographon is the universal random limiting graphon describing the scaling limit of various dense random graphs sampled from graph classes. In this paper, we solve the following open problems:

- What is the length $\mathrm{LIS}(\sigma_n)$ of the longest increasing subsequence of a random permutation σ_n of size n sampled from the Brownian separable permuton?

- What is the size $\mathrm{LCL}(G_n)$ of the largest clique of a random graph G_n on n vertices sampled from the Brownian cographon?

Specifically, we prove a scaling limit result for both quantities, that is,[1]

$$\frac{\mathrm{LIS}(\sigma_n)}{n^\alpha} \xrightarrow[n\to\infty]{\text{a.s.}} X \qquad \text{and} \qquad \frac{\mathrm{LCL}(G_n)}{n^\alpha} \xrightarrow[n\to\infty]{\text{a.s.}} X,$$

where α is explicitly determined through a simple equation and the limiting random variable X is non-deterministic, almost surely positive and finite, and is a deterministic measurable function of the Brownian separable permuton and the Brownian cographon, respectively. In particular, our main results – Theorems 1.1 and 1.3 – solve [BDSG24, Conjecture 1.5] and go significantly further, as they establish precise scaling limits.

In the remainder of this section, we introduce both the Brownian separable permuton and the Brownian cographon, present precise statements of our results, and discuss the motivations behind these questions, including relevant references.

1.1 Scaling limit for the length of the longest increasing subsequence

A **permuton** is a Borel probability measure μ on the unit square $[0,1]^2$ with two uniform marginals, that is, $\mu([a,b] \times [0,1]) = \mu([0,1] \times [a,b]) = b-a$ for all $0 \le a < b \le 1$. Permutons are natural objects to describe the scaling limits of random permutations and have been studied quite intensively in the past decade; see [Bor21, Section 2.1] for an introduction to the theory of permutons and, in particular, to the notion of permuton limit. The Brownian separable permutons $\boldsymbol{\mu}_p$ are a one-parameter family of *random* permutons, indexed by $p \in (0,1)$, which are known to describe the permuton limit of many families of pattern-avoiding permutations [BBF+18,Maa20,BBF+22b]. Moreover, the Brownian separable permutons represent the critical regime for the skew Brownian permutons introduced in [Bor23], a larger universality class for generalized pattern-avoiding permutations which enjoy several connections with classical objects studied in random geometry, such as Liouville quantum gravity and Schramm–Loewner evolutions [Bor23,BHSY23,BGS25].

Given a (possibly random) permuton μ, there is a natural way to sample permutations of size n from it: conditional on μ, sample a sequence $(Z_i)_{i\ge 1}$ of independent points in the unit square $[0,1]^2$ with distribution μ. The first n points $(Z_i)_{i\le n}$ induce a random permutation $\mathrm{Perm}[\mu, n]$: for

---

[1]The natural couplings underlying the almost sure convergence will be described in detail in Sections 1.1 and 1.2.

any $i, j \in \{1, \ldots, n\}$, let $\text{Perm}[\mu, n](i) = j$ if the point with i-th lowest x-coordinate has j-th lowest y-coordinate. Note that $\text{Perm}[\mu, n]$ are naturally coupled for different values of n.

We are interested in studying the longest increasing subsequence $\text{LIS}(\sigma_n)$ of the permutations

$$\sigma_n = \sigma_n(p) := \text{Perm}[\boldsymbol{\mu}_p, n] \tag{1.1}$$

sampled from the Brownian separable permuton $\boldsymbol{\mu}_p$ of parameter $p \in (0, 1)$. The precise definition of the Brownian separable permuton is given in Section 3.2. We stress that the definition is given for the sake of completeness, but for the purposes of this paper, it will be enough to know the distribution of σ_n, which we now describe.

Let $T_n = T_n(p)$ be a uniform planar rooted binary tree with n leaves, decorated (given the tree) with i.i.d. Bernoulli(p) $\oplus/\ominus$ signs on its nodes, *i.e.* $\mathbb{P}(\oplus) = 1 - \mathbb{P}(\ominus) = p$. We will say that T_n is a *p*-**signed uniform binary tree**. See, for instance, the left-hand side of Figure 1. Given T_n, we construct π_n via the following procedure:

- Number the leaves of T_n from 1 to n from left to right.

- For each node v of T_n decorated by $\ominus$, swap the order of the two children of v (together with their subtrees), but keep the subtrees oriented in the same way.

- Finally, define $\pi_n = \pi_n(1) \ldots \pi_n(n)$ as the sequence of numbers for the leaves of the reordered tree from left to right.

Then, as explained for instance in [BBF$^+$20, Section 5.1], the permutation π_n has the same law as σ_n in (1.1). An example of this construction is given in Figure 1. Moreover, as we will carefully explain in Remark 3.2, if the trees T_n are coupled for different values of n via the Rémy algorithm introduced in Section 2.1, then the corresponding permutations σ_n are coupled exactly as in (1.1).

$$\pi_n = 321564$$
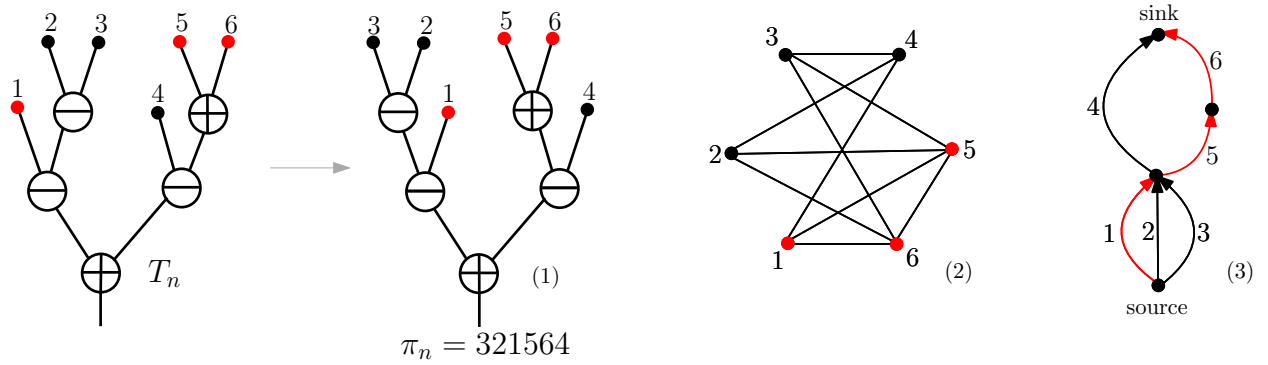
Figure 1: **Left:** A planar rooted binary tree with 6 leaves, decorated with $\oplus/\ominus$ signs on its nodes. **Right:** (1) The tree is obtained from the tree on the left by swapping the children of each $\ominus$ decorated node as explained below (1.1). The corresponding permutation π_n is obtained by reading the labels of the leaves from left to right. (2) The graph is obtained from the tree on the left by following the strategy detailed below (1.4). (3) The rooted planar map is obtained from the tree on the left by operating parallel compositions at $\ominus$ nodes and series compositions at $\oplus$ nodes. Here the leaves of the tree on the left correspond to the edges of the map. The map has a natural orientation from the *source* to the *sink*, inherited from the tree. Note that the tree leaves highlighted in red on the tree on the left induce a largest positive subtree (see Section 2.1 for terminology). These three leaves induce a longest increasing subsequence of the corresponding permutation π_n, a largest clique of the corresponding graph, and a longest directed path of the corresponding planar map.

Our interest in studying the longest increasing subsequence $\mathrm{LIS}(\sigma_n)$ is motivated by various factors, including the general study of longest increasing subsequences in non-uniform permutation models and the connections between the longest increasing subsequence of permutations sampled from skew Brownian permutons and certain models of directed metrics on random planar maps. We refer the reader to [BDSG24, Sections 1.1.2 and 1.3.1] for further details. Our main result on the Brownian separable permutons is the following. We denote by Γ the standard Gamma function.

Theorem 1.1 (Scaling limit for the length of the longest increasing subsequence of the Brownian separable permutons). *For all $p \in (0, 1)$, let $\sigma_n = \mathrm{Perm}[\boldsymbol{\mu}_p, n]$ be a random permutation of size n sampled from the Brownian separable permuton $\boldsymbol{\mu}_p$ of parameter p, as introduced in* (1.1). *Recall that the permutations σ_n are naturally coupled for different values of n. We have the almost sure convergence*

$$\frac{\mathrm{LIS}(\sigma_n)}{n^\alpha} \xrightarrow[n \to \infty]{a.s.} X, \tag{1.2}$$

where $\alpha = \alpha(p)$ is the unique solution in the interval $(1/2, 1)$ to the equation

$$\frac{1}{4^{\frac{1}{2\alpha}} \sqrt{\pi}} \frac{\Gamma(\frac{1}{2} - \frac{1}{2\alpha})}{\Gamma(1 - \frac{1}{2\alpha})} = \frac{p}{p-1}, \tag{1.3}$$

and the limiting random variable $X = X(p)$ is non-deterministic and almost surely positive and finite. Moreover, X is a deterministic measurable function of the Brownian separable permuton $\boldsymbol{\mu}_p$.

Setting $\zeta_{\mathbb{Z}}(s) := \frac{1}{4^s \sqrt{\pi}} \frac{\Gamma(1/2-s)}{\Gamma(1-s)}$, we can write the equation in (1.3) as $\zeta_{\mathbb{Z}}(\frac{1}{2\alpha}) = \frac{p}{p-1}$. The function $\zeta_{\mathbb{Z}}(s)$ is the spectral zeta function of $\mathbb{Z}$ introduced in [FK17]. The beautiful formula in [KP22, Theorem 2] allows us to rewrite our equation in (1.3) in the following elegant product form:

$$\prod_{i=1}^{\infty} \frac{(i - \frac{1}{2\alpha})^2}{i(i - \frac{1}{\alpha})} = \frac{p}{p-1}.$$

We refer the reader to [Kar20,KP22] for more background on spectral zeta functions. The graph of $\alpha(p)$ is shown in Figure 2. A heuristic explanation for why the exponent α satisfies the equation (1.3) is given in Section 2.3. However, at present, we have no explanation for the appearance of the number-theoretic function $\zeta_{\mathbb{Z}}$. Further details on the fact that X is a deterministic measurable function of the Brownian separable permuton are given in Section 9.5.
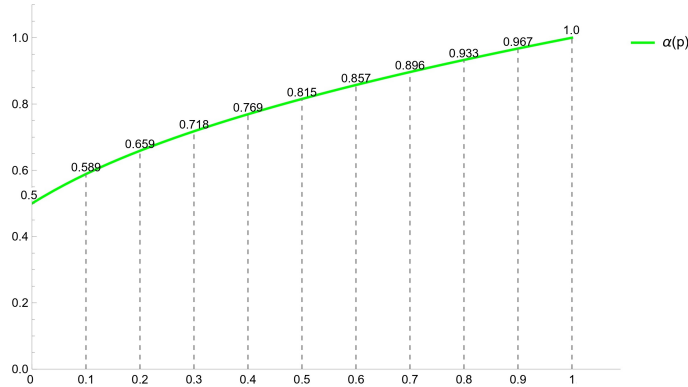
Figure 2: The graph of the function $\alpha(p)$ from Theorem 1.1 for $p \in (0, 1)$. The values of $\alpha(i/10)$ for $i = 0, 1, \ldots, 10$ are shown on top of the function, where $\alpha(0) = \alpha(0^+)$ and $\alpha(1) = \alpha(1^-)$.
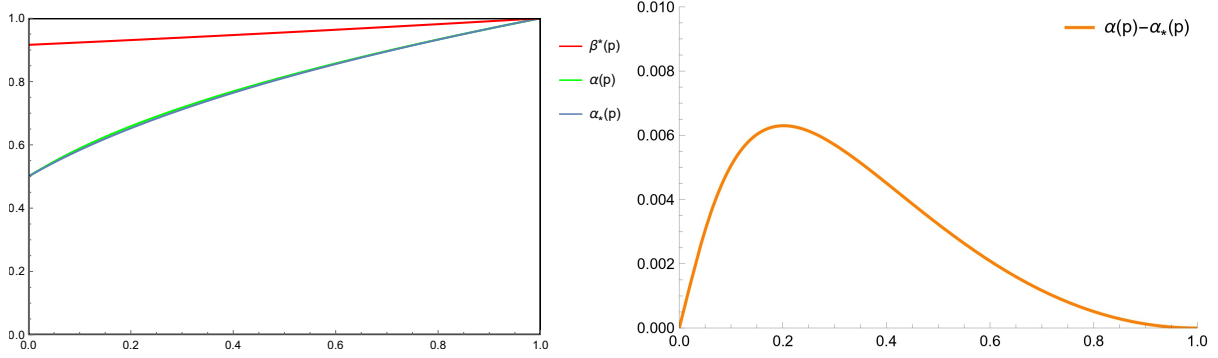
Figure 3: **Left:** The graph of the functions $\alpha_*(p)$, $\alpha(p)$ and $\beta^*(p)$ from Theorem 1.1 and Theorem 1.2. **Right:** The graphs of $\alpha(p) - \alpha_*(p)$. The maximum is reached at $p \approx 0.202$ where it takes the value $\alpha(p) - \alpha_*(p) \approx 0.0063$. Moreover, $\alpha(p) - \alpha_*(p) > 0$ for all $p \in (0,1)$.

We believe that the law of $X(p)$ is absolutely continuous with respect to the Lebesgue measure and that for all $p \neq p'$, the laws of $X(p)$ and $X(p')$ are distinct – although we do not prove this here. This belief is also supported by the numerical simulations in Appendix A.

We recall that the study of the longest increasing subsequence of Brownian separable permutons was initiated in [BBD$^+$22] and the best previously known result on the behavior of LIS(σ_n) was the following.

Theorem 1.2 ([BDSG24, Theorem 1.1]). *There exist two explicit functions $\alpha_* : (0,1) \to (1/2,1)$ and $\beta^* : (0,1) \to (1/2,1)$ such that for all $p \in (0,1)$:*

- $1/2 < \alpha_*(p) \leq \beta^*(p) < 1$.

- *For all $\varepsilon > 0$, it holds with probability tending to one as $n \to \infty$ that*

$$n^{\alpha_*(p)-\varepsilon} \leq \mathrm{LIS}(\sigma_n) \leq n^{\beta^*(p)+\varepsilon}.$$

We emphasize that, prior to the present work, even the existence of the exponent $\alpha(p)$ was an open problem. The graphs of $\alpha_*(p)$, $\alpha(p)$ and $\beta^*(p)$ are shown on the left-hand side of Figure 3 for comparison. We refer to [BDSG24, Remark 1.2] for the exact equations characterizing $\alpha_*(p)$ and $\beta^*(p)$. It is interesting to note that the previously known lower bound $\alpha_*(p)$ for the exponent $\alpha(p)$ is very close to the exact value for all $p \in (0,1)$, even though $\alpha_*(p) < \alpha(p)$ for all $p \in (0,1)$ as shown on the right-hand side of Figure 3. We also note that the values of $\alpha(p)$ given in Figure 2 are consistent with the numerical simulations in [BDSG24, Appendix A].

Uniform separable permutations vs. permutations sampled from the permuton. For the reader familiar with the notion of separable permutation (a beautiful introduction, including connections to algebraic geometry and Kontsevich's polynomial interchange problem, can be found in Ghys's book [Ghy16]), we stress that $\sigma_n(p)$ is a.s. a separable permutation [BBF$^+$20, Definition 5.1] for all $p \in (0,1)$, but is *not* a uniform separable permutation, even when $p = 1/2$. Indeed, there are multiple rooted binary trees with n leaves decorated with $\oplus/\ominus$ signs on their nodes that are mapped to the same separable permutation by the construction described below (1.1) and in Figure 1. Nevertheless, when $p = 1/2$, the random permutations σ_n have *the same* permuton limit as uniform separable permutations [BBF$^+$18], and the two models are conjectured to display exactly the same LIS behavior; see Conjecture 1.4 below and [BDSG24, Conjecture 1.6].

6

Moreover, for all $p \in (0,1)$, the permutation $\sigma_n(p)$ defined as in (1.1) has the following consistency property: for $k < n$, the permutation induced by k uniform indices in σ_n has the same law as σ_k; see [BBF$^+$20, Proposition 2.9]. In addition, the model $\sigma_n(1/2)$ can be viewed as the unique model in the universality class of uniform separable permutations that satisfies this consistency property, which makes it perhaps more natural than the uniform model.

Largest positive subtrees: local convergence and scaling limit. Due to the connection presented in Figure 1 between trees and permutations, Theorem 1.1 has a natural translation in terms of the p-signed uniform binary tree T_n. With a slight abuse of notation, let $\mathrm{LIS}(T_n)$ denote the maximal number of leaves of a subtree[2] of T_n whose pairwise highest common ancestors all carry $\oplus$ signs. We will refer to this as the size of the "largest positive subtree" of T_n. Then, by construction of the permutation π_n from T_n, we have $\mathrm{LIS}(T_n) = \mathrm{LIS}(\pi_n) \stackrel{\mathrm{d}}{=} \mathrm{LIS}(\sigma_n)$, so that Theorem 1.1 may be interpreted as a scaling limit result for $\mathrm{LIS}(T_n)$. This is precisely the point of view we adopt to prove Theorem 1.1, as explained in the proof strategy outlined in Section 2. As a key intermediate step, we will prove in Theorem 6.2 the local convergence of a largest positive subtree in a variant of the model T_n. As a by-product of our arguments, we will also obtain in Theorem 2.6 a description of the limiting *shape* of a largest positive subtree in the same variant of T_n. See Section 2.4 for a more detailed discussion.

1.2 Scaling limit for the size of the largest clique and independent set

A **graphon** is an equivalence class of symmetric measurable functions $W : [0,1]^2 \to \{0,1\}$ (*i.e.* $W(x,y) = W(y,x)$ for all $x, y \in [0,1]$) under the equivalence relation $\sim$, where $W \sim U$ if there exists a measurable, invertible, Lebesgue measure preserving function $\phi : [0,1] \to [0,1]$ such that $W(\phi(x), \phi(y)) = U(x,y)$ for almost every $x, y \in [0,1]$. A graphon can be thought of as a continuous analog of the adjacency matrix of a graph, viewed up to relabeling its continuous vertex set. Graphons have been used for many different purposes, with the original goal of having a notion of scaling limit for dense graphs [Lov12].

The **Brownian cographon** $\boldsymbol{W}_p$ of parameter $p \in (0,1)$ is a *random* universal graphon describing, for instance, the graphon limit of uniform random cographs when the number of vertices tends to infinity [BBF$^+$22a,Stu21] – in this case $p = 1/2$. In [Len23a,Len23b], the author proves that $\boldsymbol{W}_p$ is also the graphon limit of many other natural families of graphs, showing the universality of $\boldsymbol{W}_p$.

Let $(U_i)_{i \geq 1}$ be i.i.d. uniform random variables on $[0,1]$. Given a graphon W, the random graph $\mathrm{Graph}[W, n]$ induced by W of size n is defined as follows: consider the first n random variables $(U_i)_{i \leq n}$ and consider n vertices $\{v_1, v_2, \ldots, v_n\}$. We connect the vertices v_i and v_j with an edge if and only if $W(U_i, U_j) = 1$. This definition can be naturally extended to the case of random graphons (see [BBF$^+$22a, Section 3.2] for further details) and gives a coupling for the graphs $\mathrm{Graph}[W, n]$ for different values of n.

An **independent set** of a graph G is a subset of its vertices such that any two distinct vertices in the subset are not adjacent, and a **clique** of a graph G is a subset of its vertices such that every two distinct vertices in the subset are adjacent.

We are interested in the size of the largest independent set $\mathrm{LIN}(G_n)$ and the size of the largest clique[3] $\mathrm{LCL}(G_n)$ of the graphs

$$G_n = G_n(p) := \mathrm{Graph}[\boldsymbol{W}_p, n] \tag{1.4}$$

---

[2]See Section 2.1 for the precise notion of subtrees.

[3]These two quantities are usually denoted by $\alpha(\cdot)$ and $\omega(\cdot)$ in the literature, but we preferred to adopt a different notation since it is more consistent with the one used for permutations.

sampled from the **Brownian cographon** $\boldsymbol{W}_p$ of parameter $p \in [0, 1]$. Our main motivation comes from connections with the *Erdős–Hajnal conjecture*, as explained in [BDSG24, Section 1.2.2].

As in the case of permutons, the exact law of $\boldsymbol{W}_p$ is not important for our results and is provided in Section 3.2 for completeness, but the law of G_n can be described in terms of a uniform rooted binary tree T_n with n leaves, decorated with i.i.d. Bernoulli(p) $\oplus/\ominus$ signs on its nodes. Specifically, numbering the leaves of T_n from 1 to n from left to right, the law of G_n is described as follows: there is an edge between the vertices v_i and v_j if and only if the highest common ancestor of the i-th and j-th leaves in T_n is decorated by a $\oplus$. Note that by flipping the $\oplus/\ominus$ signs, $\mathrm{LIN}(G_n(p))$ has the same law as $\mathrm{LCL}(G_n(1-p))$ for all $p \in (0, 1)$.

Theorem 1.3 (Scaling limit for the size of the largest clique and independent set of the Brownian cographon)**.** *For all $p \in (0, 1)$, let $G_n = \mathrm{Graph}[\boldsymbol{W}_p, n]$ be a random graph of size n sampled from the Brownian cographon $\boldsymbol{W}_p$ of parameter p, as explained in (1.4). Recall that the graphs G_n are naturally coupled for different values of n. We have the almost sure convergence*

$$\frac{\mathrm{LCL}(G_n)}{n^{\alpha(p)}} \xrightarrow[n\to\infty]{a.s.} X(p) \qquad and \qquad \frac{\mathrm{LIN}(G_n)}{n^{\alpha(1-p)}} \xrightarrow[n\to\infty]{a.s.} X(1-p),$$

where $\alpha(p) \in (1/2, 1)$ and the law of $X(p)$ are as in Theorem 1.1. Moreover, $X(p)$ and $X(1-p)$ are deterministic measurable functions of the Brownian cographon $\boldsymbol{W}_p$.

The study of the largest clique and independent sets of the Brownian cographons was initiated in [BBD$^+$22] and the best previously known result on the behavior of $\mathrm{LCL}(G_n)$ and $\mathrm{LIN}(G_n)$ was [BDSG24, Theorem 1.3], which is the counterpart of Theorem 1.2 above. We note right away that just like for permutations, if G_n is built from T_n as in Figure 1, then $\mathrm{LCL}(G_n) = \mathrm{LIS}(T_n)$ by construction. In particular, the convergence statements of Theorems 1.1 and 1.3 are equivalent.

1.3 Some conjectures and additional comments

In this final section, we outline a few conjectures and related comments that extend and contextualize our main results.

Uniform separable permutations and cographs. We believe that our results can be extended to the length of the longest increasing subsequence in uniform separable permutations and the size of the largest clique and independent set in uniform cographs (and, more generally, to most natural models known to converge to the Brownian separable permuton or the Brownian cographon).

Conjecture 1.4. *Let $\alpha(1/2)$ and $X(1/2)$ be as in Theorem 1.1. Let $\overline{\sigma}_n$ and $\overline{G}_n$ be a uniform separable permutation and a uniform separable cograph of size n, respectively. Then there is a deterministic constant $c > 0$ such that we have the convergences in distribution*

$$\frac{\mathrm{LIS}(\overline{\sigma}_n)}{n^{\alpha(1/2)}} \xrightarrow[n\to\infty]{d} c \cdot X(1/2) \qquad and \qquad \frac{\mathrm{LIN}(\overline{G}_n)}{n^{\alpha(1/2)}} \xrightarrow[n\to\infty]{d} c \cdot X(1/2).$$

The numerical simulations in Appendix A suggest that $c \approx 0.901$. Note that Conjecture 1.4 is more precise than [BDSG24, Conjecture 1.6].

We also expect that some of the ideas in our paper could be adapted to study the length of the longest increasing subsequence in random recursive separable permutations [FRL23]. We emphasize that these permutations have a random permuton limit that differs from the Brownian separable one.

8

The limiting random variable $X(p)$. In this work, we do not determine the exact law of the limiting random variable $X(p)$ appearing in Theorem 1.1. It would be interesting to identify the distribution of $X(p)$ or at least establish further properties. For instance, three natural questions are whether the laws of $X(p)$ and $X(p')$ are distinct for all $p \neq p'$, what the expectation of $X(p)$ is, or what the tail behavior of $X(p)$ is. Some numerical simulations of these random variables are presented in Appendix A.

Hausdorff dimension of the largest increasing subset of the Browian separable permuton. The (closed) support of the Brownian separable permuton $\boldsymbol{\mu}_p$ is the (random) set $\mathrm{supp}(\boldsymbol{\mu}_p)$ defined as the intersection of all closed sets $K \subset [0,1]^2$ with $\boldsymbol{\mu}_p(K) = 1$. We say that a set $A \subset [0,1]^2$ is increasing if

$$A \subset ([0,t] \times [0,s]) \cup ([t,1] \times [s,1]) \qquad \text{for all } (t,s) \in A.$$

An example of an increasing set is the graph of a non-decreasing function $[0,1] \to [0,1]$, or a subset of such a graph. Let $\mathcal{I}_p$ denote the collection of increasing subsets of $\mathrm{supp}(\boldsymbol{\mu}_p)$. Then, we conjecture that for all $p \in (0,1)$, almost surely,

$$\sup_{A \in \mathcal{I}_p} \{ \dim A \} = \alpha(p),$$

where $\alpha(p)$ is defined as in Theorem 1.1, and $\dim(\cdot)$ denotes the Hausdorff dimension. We also believe that the limiting random variable $X(p)$ of Theorem 1.1 is connected to the maximal $\alpha(p)$-dimensional Hausdorff measure (with a potential gauge correction) of such a largest increasing subset.

A directed version of the Liouville quantum gravity metric. Our results also describe the length of the longest directed path from the source to the sink in a natural model[4] of rooted series–parallel maps [BBMF11, Proposition 6]; see also Figure 1. Rooted series-parallel maps are graphs with two distinguished vertices called source and sink, formed recursively by two simple composition operations (*parallel composition* and *series composition*). They have been used to model series and parallel electric circuits, and are a natural model of last-passage percolation in random geometry. The scaling limit of different models of series-parallel maps (in the Gromov–Hausdorff sense) has been studied in [Stu20,ABS$^+$25]. See the discussion in [BDSG24, Section 1.3.1] for more connections between longest increasing subsequences in the skew Brownian permuton and models of last-passage percolation in random geometry, including Baxter permutations and bipolar orientations [BM22].

In the continuum, skew Brownian permutons were constructed in [Bor23] using Liouville quantum gravity and a pair of Schramm–Loewner evolutions. This connection was recently further strengthened in [BG24]. In this correspondence, it is conjectured that Brownian separable permutons can be constructed using $\mathrm{SLE}_4/\mathrm{CLE}_4$ and critical Liouville quantum gravity ($\gamma = 2$). In connection to the aforementioned discrete picture, we expect the exponent $\alpha(p)^{-1}$ to coincide with the Hausdorff dimension of a directed version of the critical Liouville quantum gravity metric space; a space whose construction still represents a major challenge but which, at least in the critical regime, appears significantly more approachable in light of the results presented in this paper.

---

[4]As in the case of separable permutations, this natural model of series–parallel maps is the unique one within the universality class of uniform series–parallel maps that satisfies a certain consistency property; recall the discussion on uniform separable permutations on page 6. Nevertheless, we expect the asymptotic behavior of the longest directed path to be the same in both the uniform and non-uniform models.

2 Proof strategy

We now explain how we prove Theorem 1.1 and Theorem 1.3. Readers only interested in a rough idea of the proof strategy can simply browse through Section 2.1 and take a look at the diagram in Figure 6. Reading through this entire section will provide a good overview of all main ideas that go into the proof of our main results.

2.1 Main results for the size of the largest positive subtree of a p-signed uniform binary tree

From the description of the law of σ_n and G_n in terms of a uniform rooted binary tree T_n with i.i.d. Bernoulli(p) $\oplus/\ominus$ signs on its internal vertices, the convergences of Theorem 1.1 and Theorem 1.3 immediately follow if we prove an analogous scaling limit result for the size of the "largest positive subtree" of T_n, when the trees $(T_n)_{n\geq 1}$ are coupled via the Rémy algorithm. We are now going to precisely define the "largest positive subtree" and introduce the Rémy algorithm. Then, we will state our main results on the "largest positive subtree" of T_n.

Notation for trees. We always consider planar rooted (planted) trees, drawn with the root at the bottom. By planted, we mean that the root has an additional edge extending below it, which is not connected to any other vertex; see Figure 4. All trees in this paper are planted (sometimes without explicit mention). The **degree** $\deg(v)$ of a vertex v is the number of edges incident to v (in particular, provided the tree is not a single leaf, the root of a rooted (planted) binary tree has degree 3). The vertices of a tree are of two types: the **leaves**, which are the vertices of degree one, and the **nodes**, which are all the other vertices (*i.e.* the internal vertices). The **size** $|t|$ of a tree t is equal to the number of leaves in t. Moreover, all the **subtrees** of any tree t considered in this paper are obtained by first selecting a subset $\mathcal{L}$ of the leaves of t, then removing from t all the vertices that do not have a descendant in $\mathcal{L}$, and finally contracting all the vertices of degree 2 that may appear. We denote such a subtree by $t|_{\mathcal{L}}$. Moreover, we denote by $u(t|_{\mathcal{L}})$ the **uncontracted version** of $t|_{\mathcal{L}}$, *i.e.* the tree obtained by only removing from t all the vertices that do not have a descendant in $\mathcal{L}$.
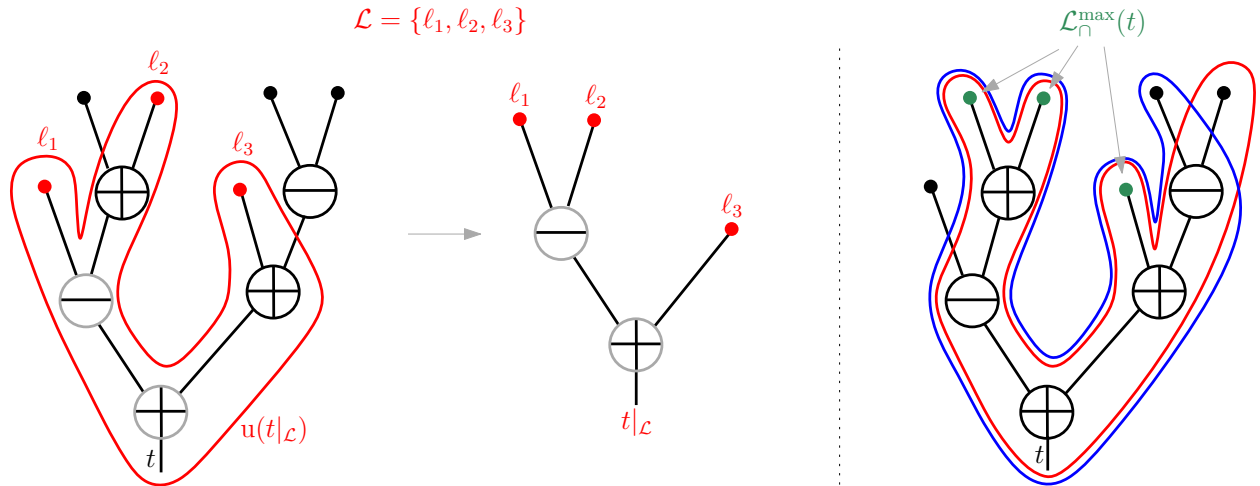
Figure 4: **Left:** A sign-decorated binary tree t with a subtree $t|_{\mathcal{L}}$ induced by the three leaves ℓ_1, ℓ_2, ℓ_3. The uncontracted version $u(t|_{\mathcal{L}})$ of $t|_{\mathcal{L}}$ is highlighted by the red curve on t. **Right:** The same sign-decorated binary tree t with all (*i.e.* two) maximal positive (uncontracted) subtrees of t highlighted in red and blue. We colored in green the leaves in $\mathcal{L}_{\cap}^{\max}(t)$.

Both notions of subtrees are seen as planar rooted (planted) trees, where the root is given by the vertex with the lowest height in the original tree t.

A **sign-decorated tree** t is a planar rooted (planted) tree such that all nodes are decorated with a $\oplus$ or $\ominus$ sign.[5] Given a sign-decorated tree t and one of its nodes v, we denote by $s(v)$ the sign decorating v. Nodes carrying $\oplus$ (resp. $\ominus$) signs are sometimes called positive (resp. negative). A subtree of a sign-decorated tree t is still a sign-decorated tree where all the signs are inherited from t. A sign-decorated tree t is **positive** if all the nodes are decorated with $\oplus$ signs. For any finite sign-decorated tree t, we write $\mathrm{LIS}(t)$ for the **maximal size** (*i.e.* the maximal number of leaves) of a positive subtree of t. A **maximal positive subtree** of t is a subtree of t of size $\mathrm{LIS}(t)$. The notation $\mathrm{LIS}(t)$ is motivated by the following fact: $\mathrm{LIS}(t)$ is equal to the size of the longest increasing subsequence of the permutation obtained from t using the procedure described below (1.1) and in Figure 1. Similarly, $\mathrm{LIS}(t)$ is equal to the size of the largest clique of the graph obtained from t using the procedure described below (1.4) and the length of the longest directed path from the source to the sink of the corresponding series-parallel planar map. Finally, we denote by $\mathcal{L}_{\cap}^{\max}(t)$ the set of leaves in the intersection of *all* maximal positive subtrees of t.

The Rémy algorithm. We describe a coupling between the p-signed uniform binary trees $T_n = T_n(p)$ for all $n \geq 1$, which will be of fundamental importance in this paper. This coupling is a natural adaptation of the Rémy algorithm [Rém85] to sign-decorated trees. To build this coupling, we define T_n by induction on n as follows (*cf.* Figure 5):

- The tree T_1 is just a single leaf with an additional edge extending below it (recall that our trees are planted).

- Conditionally on $T_1, \ldots, T_n$, let E_n be chosen uniformly at random among the $2n - 1$ edges of T_n, and let S_n be an independent Bernoulli(p) $\oplus/\ominus$ sign. We build T_{n+1} by adding in the middle of E_n a new vertex with sign S_n, and grafting a new leaf on the left of this vertex with probability $1/2$ or on the right of this vertex with probability $1/2$. We see T_{n+1} as a rooted (planted) tree: when E_n is the special edge extending below the root of T_n, the root of T_{n+1} is the new vertex; otherwise it is the same as that of T_n.

We also note that the transitions of the Rémy algorithm are even simpler when it is run in reverse: in the above coupling, for any $1 < m \leq n$, the law of T_{m-1} conditionally on $(T_m, T_{m+1}, \ldots, T_n)$ is obtained by removing a uniform random leaf L_m of T_m and contracting the corresponding node.

We are now ready to state our three main results on $\mathrm{LIS}(T_n)$.

Theorem 2.1 (Existence of the exponent)**.** *For all $p \in (0, 1)$, there exists $\alpha = \alpha(p) \in (1/2, 1)$ such that we have the convergence in probability*

$$\frac{\log \mathrm{LIS}(T_n)}{\log n} \xrightarrow[n \to \infty]{\mathbb{P}} \alpha. \tag{2.1}$$

Moreover,

$$\alpha = \limsup_{n \to \infty} \frac{\log \mathbb{E}[\mathrm{LIS}(T_n)]}{\log n}. \tag{2.2}$$

The fact that $\alpha(p) \in (1/2, 1)$, provided the exponent exists, is a simple consequence of Theorem 1.2. The latter actually yields the better bounds $\alpha(p) \in (\alpha_*(p), \beta^*(p))$, but the bounds

---

$\alpha(p) \in (1/2, 1)$ are sufficient for our arguments. Another important note is that the description of α in (2.2) and Markov's inequality provide a quantitative upper bound on the probability that $\text{LIS}(T_n)$ exceeds n^α by a large amount. This fact will be helpful later to obtain certain important estimates (Proposition 5.1) used to prove the next main result (more explanations are given in Section 2.4).

Theorem 2.2 (Determination of the exponent)**.** *For all $p \in (0, 1)$, let $\alpha = \alpha(p) \in (1/2, 1)$ be as in Theorem 2.1. Then α is the unique solution in the interval $(1/2, 1)$ to the equation in (1.3), that is,*

$$\frac{1}{4^{\frac{1}{2\alpha}}\sqrt{\pi}} \frac{\Gamma\left(\frac{1}{2} - \frac{1}{2\alpha}\right)}{\Gamma\left(1 - \frac{1}{2\alpha}\right)} = \frac{p}{p-1}.$$

Our third main result gives the scaling limit of $\text{LIS}(T_n)$. Recall that the trees $(T_n)_{n\geq 1}$ are coupled via the Rémy algorithm.

Theorem 2.3 (Scaling limit of $\text{LIS}(T_n)$)**.** *For all $p \in (0, 1)$, we have the almost sure convergence*

$$\frac{\text{LIS}(T_n)}{n^\alpha} \xrightarrow[n\to\infty]{a.s.} X,$$

where the value $\alpha = \alpha(p)$ is as in Theorems 2.1 and 2.2 and the limiting random variable $X = X(p)$ is non-deterministic and almost surely positive and finite.

Note that Theorems 2.2 and 2.3 immediately imply Theorems 1.1 and 1.3 due to the description of the law of σ_n and G_n. The only fact that does not immediately follow from the above results is that the random variable X in Theorem 1.1 (resp. Theorem 1.3) is a deterministic measurable function of the Brownian separable permuton (resp. the Brownian cographon). This fact will be justified in Section 9.5.

The rest of this section is organized as follows. We first explain our strategy to prove the first main result, that is, the existence of the exponent α in Theorem 2.1. This is done via a supermultiplicativity argument explained in Section 2.2. Then, in a second phase, we determine
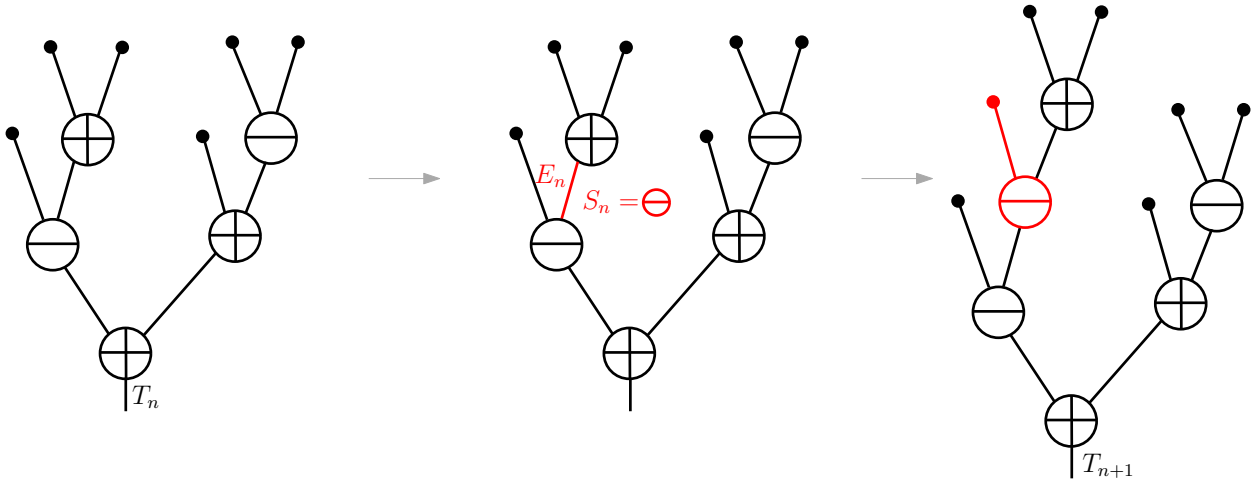
Figure 5: Our version of the Rémy algorithm for p-signed uniform binary trees. In this realization, we start with a p-signed uniform binary tree T_n with n leaves, then we select a uniform edge E_n, sample a $\ominus$ sign (this happens with probability $1 - p$) and finally, we graft a new leaf to the *left* (this happens with probability $1/2$) of the new vertex on E_n decorated by $\ominus$.

the exact value of $\alpha(p)$ by showing that $\alpha(p)$ is the unique solution to the equation in (1.3) in the interval $(1/2, 1)$. The latter step is the most involved part of the paper and will be explained in Sections 2.3 and 2.4. At the end of Section 2.4, we will also explain how we prove our last main result, that is, the scaling limit in Theorem 2.3.

We invite the reader to compare the explanations in the next sections with the diagram in Figure 6 which should be helpful to visualize the structure of the full proof.

Important note. From now on, we consider $p \in (0, 1)$ as a fixed parameter and omit its dependency in our notation. We note that our constants may still depend on p.

2.2 Existence of the exponent α via supermultiplicativity

We introduce the compact notation

$$X_n := \mathrm{LIS}(T_n). \tag{2.3}$$

Our first objective will be to show the existence of the exponent α, proving Theorem 2.1. This is done in Section 4 using a supermultiplicativity argument. The basic idea here would be to decompose the trees T_n in some subregions to obtain a supermultiplicativity equation for the random variables X_n of the form

$$X_{m \cdot n} \succeq \sum_{i=1}^{X_m} X_n^{(i)}, \tag{2.4}$$

where the $X_n^{(i)}$ are i.i.d. copies of X_n and are independent of X_m. Here $\succeq$ denotes stochastic domination (see Section 3.1 for further details if necessary). This domination could be interpreted as a "super-branching property" for the sequence $(X_{2^k})_{k \geq 0}$. Comparison with branching processes could then be applied to get that $(X_{2^k})^{\frac{1}{k}}$ converges in probability to some constant, and hence by monotonicity we would have

$$\frac{\log X_n}{\log n} \xrightarrow[n \to +\infty]{\mathbb{P}} \alpha.$$

The issue is that we will not be able to write an equation as simple as (2.4) because most sensible ways to decompose uniform binary trees into several regions result in regions with *random* and *different* sizes (Lemma 3.8). This will have the consequence that our supermultiplicativity equation (see (4.3)) will be more complicated than (2.4). However, we can exactly describe the joint law of such random regions (Lemma 3.8). This will enable us to compare our supermultiplicativity equation with a more standard one (see (4.7)) and to prove Theorem 2.1.

2.3 The sequences q and Q and a heuristic derivation of the exponent α

The next step is to determine the exact value of α by showing that α is the unique solution to the equation (1.3) in the interval $(1/2, 1)$. For the remainder of Section 2, we omit most technical details and make frequent use of the symbol $\approx$ to indicate that two quantities behave similarly, in a deliberately imprecise sense.

2.3.1 The sequences q and Q

Here, our first idea is to randomize the size of T_n by introducing a critical binary Bienaymé–Galton–Watson tree T with i.i.d. Bernoulli(p) signs on its nodes (sampled independently of the tree), so

13

Theorem 2.1 — Existence of the exponent α — Section 4

Proposition 5.1 — **Rough** regularity estimates for the sequences q and Q — Section 5

Proposition 5.1 — $q(k) = k^{-\gamma+o(1)}$ with $\gamma = \frac{1}{2\alpha} + 1$

Lemma 2.4 — Recursive relation for the sequence q

Equation (2.9) — **Optimal** regularity estimates for the sequences q and Q — (Not proved)

Theorem 7.1 — **Good** regularity estimates for the sequences q and Q — Section 7.2

Theorems 6.2 and 7.2 — Law of large numbers for $\#\mathcal{L}_\cap^{\max}(T^k)$ from the local limit — Sections 6 and 7.1

Theorem 9.1 — Quantitative law of large numbers for $\#\mathcal{L}_\cap^{\max}(T^k)$ — Section 9

Lemmas 8.1, 8.2 and 8.3 — From equalities to inequalities — Section 8

Theorem 2.2 — Exact value for the exponent α — Section 8

Theorem 2.3 — Scaling limit $\frac{\mathrm{LIS}(T_n)}{n^\alpha} \xrightarrow[n\to\infty]{\mathrm{d}} X$ — Section 9

T is a p-signed critical binary BGW-tree
T^k is T conditioned on $\mathrm{LIS}(T) = k$
T_n is a uniform p-signed tree with n leaves

$q(k) = \mathbb{P}(\mathrm{LIS}(T) = k)$
$Q(k) = \mathbb{P}(\mathrm{LIS}(T) \geq k) = \sum_{i \geq k} q(i)$

Figure 6: A diagram for the strategy of the proof of the three main results of this paper, *i.e.* Theorems 2.1, 2.2 and 2.3. The existence of the exponent α (Theorem 2.1) is proved in Section 4 via a supermultiplicativity argument. **The gray path:** It is relatively simple to deduce from the existence of α that $q(k) = \mathbb{P}(\mathrm{LIS}(T) = k) = k^{-\frac{1}{2\alpha}-1+o(1)}$ (Proposition 5.1). To determine the exact value of α (Theorem 2.2), we first find an appropriate recursive relation (Lemma 2.4) for the sequence $q(k)$. Assuming certain optimal regularity estimates (2.9), one can determine the value of α from the recursive relation for $q(k)$; but unfortunately such optimal regularity estimates are out of reach with our methods. **The green path:** To circumvent this issue, in Section 5, we first deduce from the existence of the exponent α certain rough regularity estimates (Proposition 5.1). Then, from these estimates, we obtain in Section 6 a local convergence result for a maximal positive subtree seen from a uniformly chosen leaf in the related tree model T^k (Theorem 6.2) and in Section 7.1 a corresponding law of large numbers for $\#\mathcal{L}_\cap^{\max}(T^k)$ (Theorem 7.2). Finally, in Section 7.2, we obtain from this law of large numbers certain good regularity estimates (Theorem 7.1). The latter estimates, combined with a new technique which converts the recursive relation for $q(k)$ into certain auxiliary inequalities (Lemmas 8.1, 8.2 and 8.3), allow us to determine α in Section 8. **The orange path:** Starting from the good regularity estimates and running again part of the proof of the local convergence theorem, we obtain in Section 9 a quantitative version of the law of large numbers for $\#\mathcal{L}_\cap^{\max}(T^k)$ (Theorem 9.1) and finally deduce our scaling limit result for $X_n = \mathrm{LIS}(T_n)$ (Theorem 2.3).

that T conditioned to have size n is distributed as T_n. We introduce the two fundamental sequences $q = (q(k))_{k \geq 1}$ and $Q = (Q(k))_{k \geq 1}$, defined for all $k \geq 1$ by

$$q(k) := \mathbb{P}(\mathrm{LIS}(T) = k) \qquad \text{and} \qquad Q(k) := \mathbb{P}(\mathrm{LIS}(T) \geq k) = \sum_{i \geq k} q(i). \tag{2.5}$$

We first explain intuitively why the sequences q and Q are related to the exponent α that we aim to compute. Write $Q(k) = \sum_{n \geq 1} \mathbb{P}(|T| = n) \cdot \mathbb{P}(\mathrm{LIS}(T_n) \geq k)$. Using Theorem 2.1, we get that $\mathbb{P}(\mathrm{LIS}(T_n) \geq k)$ is very small if $k \gg n^\alpha$ and very close to 1 if $k \ll n^\alpha$. This suggests that

$$Q(k) \approx \mathbb{P}\left(|T| \geq k^{\frac{1}{\alpha}}\right) \approx k^{-\frac{1}{2\alpha}},$$

using the standard fact that $\mathbb{P}(|T| \geq n) \approx n^{-\frac{1}{2}}$ (Lemma 3.3). A rigorous version of this argument (Proposition 5.1) yields

$$q(k) = k^{-\frac{1}{2\alpha} - 1 + o(1)} \qquad \text{and} \qquad Q(k) = k^{-\frac{1}{2\alpha} + o(1)} \qquad \text{as } k \to \infty. \tag{2.6}$$

To simplify the notation, it will be convenient to set

$$\gamma := \frac{1}{2\alpha} + 1 \in \left(\frac{3}{2}, 2\right), \tag{2.7}$$

so that $q(k) = k^{-\gamma + o(1)}$ and $Q(k) = k^{-\gamma + 1 + o(1)}$ by (2.6). To determine γ (or equivalently α), we write a recursion for the sequence q by decomposing the critical binary Bienaymé–Galton–Watson tree T at the root. We obtain the following equation.

Lemma 2.4 (Recursive relation for the sequence q). *Let $q(k)$ be as in* (2.5). *Then for $k \geq 2$,*

$$q(k) = (1-p)q(k)\sum_{i=1}^{k-1} q(i) + p \sum_{i=1}^{(k-1)/2} q(i)q(k-i) + \frac{1-p}{2}q(k)^2 + \mathbb{1}_{\{k\,\mathrm{even}\}}\frac{p}{2}q(k/2)^2. \tag{2.8}$$

Proof. Decompose the critical binary Bienaymé–Galton–Watson tree T at the root. The four summands correspond to the following four cases: the root (which is necessarily a node because $k \geq 2$) is decorated by a $\ominus/\oplus$ sign and the left and right subtrees have maximal positive subtrees of different/equal sizes. $\square$

Note that one can explicitly compute that $q(1) = \mathbb{P}(\mathrm{LIS}(T) = 1) = (1 + \sqrt{p})^{-1}$ and then use (2.8) to numerically compute the first few values of the sequence q. Unfortunately, although the recursive relation in (2.8) entirely determines the sequence q, it is very sensitive to initial conditions, see for instance Figure 7. As a result, it seems quite hard to deduce anything accurate on the sequence q directly from (2.8), without any additional input.

2.3.2 A heuristic calculation for the exponent α assuming optimal regularity estimates

We now explain how one could determine the value of γ assuming the following **optimal regularity estimates** about the sequences q and Q (which we will never be able to establish): For some constant $A > 0$, assume

$$Q(k) \sim \frac{A}{\gamma - 1} \cdot k^{-\gamma + 1}, \qquad q(k) \sim Ak^{-\gamma}, \qquad q(k-1) - q(k) \sim \gamma Ak^{-\gamma - 1} \qquad \text{as } k \to \infty. \tag{2.9}$$

15

Rearranging the terms of (2.8), we get

$$q(k)\left(1 + \frac{1-p}{p}Q(k)\right) = \left(\sum_{i=1}^{(k-1)/2} q(i)q(k-i) + \frac{1-p}{2p}q(k)^2 + \mathbb{1}_{\{k \text{ even}\}}\frac{1}{2}q(k/2)^2\right),$$

so that

$$\frac{1-p}{p}Q(k)q(k) = \left(\sum_{i=1}^{(k-1)/2} q(i)q(k-i)\right) - q(k) + \left(\frac{1-p}{2p}q(k)^2 + \mathbb{1}_{\{k \text{ even}\}}\frac{1}{2}q(k/2)^2\right). \quad (2.10)$$

Then we could determine the value of γ by inserting the asymptotics (2.9) into the equation (2.10) and equating their leading term as $k \to \infty$. First, the left-hand side would be

$$\frac{1-p}{p}Q(k)q(k) = \frac{1-p}{p}\frac{A^2}{\gamma-1}k^{-2\gamma+1} + o(k^{-2\gamma+1}).$$

Note already that some terms on the right-hand side will not contribute to the leading term of order $k^{-2\gamma+1}$, because

$$\frac{1-p}{2p}q(k)^2 + \mathbb{1}_{\{k \text{ even}\}}\frac{1}{2}q(k/2)^2 = O(k^{-2\gamma}) = o(k^{-2\gamma+1}).$$

By Abel summation, the sum on the right-hand side of (2.10) can be rewritten as

$$\sum_{i=1}^{(k-1)/2} q(i)q(k-i) = \sum_{i=1}^{(k-1)/2} Q(i)q(k-i) - \sum_{i=2}^{(k+1)/2} Q(i)q(k-i+1)$$

$$= q(k-1) - Q\left(\left\lfloor\frac{k+1}{2}\right\rfloor\right)q\left(\left\lceil\frac{k+1}{2}\right\rceil\right) + \sum_{i=2}^{(k-1)/2} Q(i)(q(k-i) - q(k-i+1)).$$
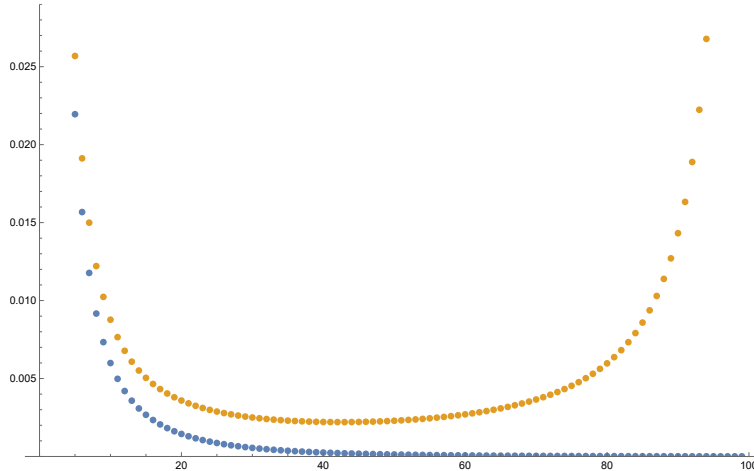
Figure 7: We fix $p = 1/2$. In blue, the first 100 values of the sequence $q(k)$ computed using the recursive relation in (2.8) with initial condition $q(1) = 2 - \sqrt{2}$. In orange, the first 100 values of the sequence $\widetilde{q}(k)$ computed using the same recursive relation in (2.8) with initial condition $\widetilde{q}(1) = 2 - \sqrt{2} + 0.005$. The first values of each sequence are not shown because they are too large.

We use this to express the right-hand side of (2.10) and allow ourselves to replace any of the terms $q(k)$, $Q(k)$ by their asymptotics given in (2.9) and neglect any term that would be $o(k^{-2\gamma+1})$. We get (the entire argument could be turned into an actual proof with some additional details that we spare to the reader)

$$
\left(\sum_{i=1}^{(k-1)/2} q(i)q(k-i) \right) - q(k)
$$

$$
\approx q(k-1) - q(k) - \frac{A}{\gamma-1}\left(\frac{k}{2}\right)^{-\gamma+1} \cdot A\left(\frac{k}{2}\right)^{-\gamma} + \sum_{i=2}^{(k-1)/2} \frac{A}{\gamma-1} i^{-\gamma+1} \cdot \gamma A(k-i)^{-\gamma-1}
$$

$$
\approx -\frac{A^2 \cdot 2^{2\gamma-1}}{\gamma-1} \cdot k^{-2\gamma+1} + \frac{\gamma A^2}{\gamma-1} \cdot k^{-2\gamma+1} \cdot \sum_{i=2}^{(k-1)/2} \frac{1}{k} \cdot \left(\frac{i}{k}\right)^{-\gamma+1}\left(1-\frac{i}{k}\right)^{-\gamma-1}
$$

$$
\approx \frac{A^2}{\gamma-1} \cdot k^{-2\gamma+1} \cdot \left(-2^{2\gamma-1} + \gamma \int_0^{\frac{1}{2}} x^{-\gamma+1}(1-x)^{-\gamma-1} \mathrm{d}x \right),
$$

where the integral comes as the limit of a Riemann sum. Equating the leading term on the left-hand side and right-hand side of (2.10) would ensure that

$$
\frac{1-p}{p} = -2^{2\gamma-1} + \gamma \int_0^{\frac{1}{2}} x^{-\gamma+1}(1-x)^{-\gamma-1} \mathrm{d}x, \tag{2.11}
$$

which after the substitution $\alpha = \frac{1}{2(\gamma-1)}$ from (2.7) and some manipulations[6] would lead to the equation in (1.3).

Remark 2.5. It is possible to run a different version of the above argument and identify the exponent γ (and α) from a set of assumptions that are weaker than (2.9), namely that there exists a slowly varying function φ such that

$$
q(k) = k^{-\gamma}\varphi(k) \qquad \text{and that} \qquad \frac{q(k+1)}{q(k)} = 1 + O\left(\frac{1}{k}\right).
$$

We will be able to prove that the first asymptotic holds thanks to Theorem 7.1 below, but the second one will remain out of reach using our methods. Hence, we will need a different strategy explained in the next subsection.

2.4 Our proof strategy: good but not optimal regularity estimates

We now describe the strategy used in this paper. Most of the paper will be devoted to proving some **good regularity estimates** for the sequences q and Q. Specifically, we show that there exist slowly varying functions φ and Φ (see Section 3.1 for definitions) such that

$$
q(k) = k^{-\frac{1}{2\alpha}-1}\varphi(k) \qquad \text{and} \qquad Q(k) = k^{-\frac{1}{2\alpha}}\Phi(k), \tag{2.12}
$$

where α is the exponent of Theorem 2.1. Although these estimates are not as precise as (2.9), they will still be sufficient to conclude with some additional effort.

The rest of this section is dedicated to outlining the key ideas behind the structure of this paper.

---

[6]We will formally do these computations in the second part of the proof of Lemma 8.3 on page 82.

The dynamical picture. The control (2.12) on the sequences q and Q will come from understanding a dynamical version of the problem. We will work on a probability space where the whole sequence $(T_n)_{n\geq 1}$ is coupled under Rémy's algorithm described in Section 2.1: under this construction, conditionally on $(T_i)_{i\geq n}$, the tree T_{n-1} can be obtained by removing a uniform leaf L_n from T_n and contracting the corresponding node. This ensures that the corresponding sequence $(X_n)_{n\geq 1}$ defined in (2.3) is non-decreasing and changes at each step by either 0 or 1. We denote by τ_k the first time that X_n attains any number $k \geq 1$. Note that studying the asymptotic behavior of the process $(\tau_k)_{k\geq 1}$ is equivalent to studying that of $(X_n)_{n\geq 1}$ since

$$X_n = k \qquad \text{for all } n \in [\![\tau_k\,, \tau_{k+1} - 1]\!]. \tag{2.13}$$

It just happens that the sequences q and Q are more naturally controlled in terms of the stopping times $(\tau_k)_{k\geq 1}$, as we can see below.

How the sequence $(\tau_k)_{k\geq 1}$ is related to q and Q. Let the p-signed critical binary Bienaymé–Galton–Watson tree T be defined on the same probability space as, and independent of, the sequence $(T_n)_{n\geq 1}$. Since T conditioned to have size n is distributed as T_n, we have that $\mathrm{LIS}(T) \overset{\mathrm{d}}{=} X_{|T|}$. Combining this with the asymptotics $\mathbb{P}(|T| \geq n) \approx c n^{-\frac{1}{2}}$ that holds for some constant $c > 0$ (Lemma 3.3), we get that

$$Q(k) = \mathbb{P}(\mathrm{LIS}(T) \geq k) = \mathbb{P}\big(X_{|T|} \geq k\big) \overset{(2.13)}{=} \mathbb{P}(|T| \geq \tau_k) \approx c \cdot \mathbb{E}\Big[\tau_k^{-1/2}\Big]. \tag{2.14}$$

Similarly, we have $q(k) = Q(k) - Q(k+1) \approx c \cdot \mathbb{E}\Big[\tau_k^{-1/2} - \tau_{k+1}^{-1/2}\Big]$.

Regularity of q, Q and $(\tau_k)_{k\geq 1}$. By Karamata's characterization theorem [BGT89, Theorem 1.4.1] for slowly varying functions, in order to prove the estimates (2.12), it is enough to show that, for any $x \geq 1$, we have $\frac{Q(xk)}{Q(k)} \to x^{-\frac{1}{2\alpha}}$ as $k \to \infty$ and the analogous estimate for q. Our objective is hence to control the ratios $\frac{q(xk)}{q(k)}$ and $\frac{Q(xk)}{Q(k)}$. From the previous paragraph, this could be obtained by showing that in some sense, when k is large, we have

$$\tau_{\lceil xk \rceil} \approx x^{\frac{1}{\alpha}} \tau_k. \tag{2.15}$$

We start by controlling what happens between the two consecutive increase times τ_k and τ_{k+1}.

Controlling the time $(\tau_{k+1} - \tau_k)$ between successive increases of X. First note that for a given $n \geq 2$, we have $X_{n-1} = X_n - 1$ if, and only if, the leaf L_n removed from T_n to obtain T_{n-1} belongs to the set $\mathcal{L}_\cap^{\max}(T_n)$ of leaves that are in *all* positive subtrees of T_n of maximal size. Hence, we can write

$$X_{n-1} = X_n - \mathbb{1}_{L_n \in \mathcal{L}_\cap^{\max}(T_n)}, \tag{2.16}$$

where conditionally on $(T_i)_{i\geq n}$, the probability of the event $\{L_n \in \mathcal{L}_\cap^{\max}(T_n)\}$ is $\frac{1}{n}\#\mathcal{L}_\cap^{\max}(T_n)$.

Now let us apply this argument in a cascading manner: we start at time $n = \tau_{k+1} - 1$, so that $X_n = k$, and remove one leaf after another from the tree until its LIS drops below k, that is, until

we reach $\tau_k - 1$. Heuristically, we get for all $j \geq 0$,

$$
\begin{aligned}
\mathbb{P}\big(\tau_{k+1} - \tau_k > j \mid T_{\tau_{k+1}}\big) &= \mathbb{P}\big(L_{\tau_{k+1}-1} \notin \mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-1}), \ldots, L_{\tau_{k+1}-j} \notin \mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-j}) \mid T_{\tau_{k+1}}\big) \\
&= \mathbb{E}\Bigg[\bigg(1 - \frac{\#\mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-1})}{\tau_{k+1}-1}\bigg) \cdot \ldots \cdot \bigg(1 - \frac{\#\mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-j})}{\tau_{k+1}-j}\bigg) \;\Bigg|\; T_{\tau_{k+1}} \Bigg] \\
&\approx \bigg(1 - \frac{\#\mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-1})}{\tau_{k+1}}\bigg)^j,
\end{aligned}
$$

which suggests the approximation in distribution

$$
\mathrm{Law}(\tau_{k+1} - \tau_k | T_{\tau_{k+1}}) \overset{\mathrm{d}}{\approx} \mathrm{Geom}\bigg(\frac{\#\mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-1})}{\tau_{k+1}} \bigg). \tag{2.17}
$$

A key insight. The key insight (justified in the next paragraphs) underlying the strategy that we follow for the rest of the paper is that, in a suitable probabilistic sense, we have

$$
\#\mathcal{L}_\cap^{\max}(T_n) \approx \lambda X_n \qquad \text{and hence} \qquad \#\mathcal{L}_\cap^{\max}(T_{\tau_{k+1}-1}) \approx \lambda k, \tag{2.18}
$$

for some *a priori* unknown $\lambda \in (0,1)$ depending only on p. Hence, from (2.17) we can compare the law of $\tau_{k+1} - \tau_k$ conditional on τ_{k+1} with a geometric random variable of parameter $\frac{\lambda k}{\tau_{k+1}}$, which has expectation $\frac{\tau_{k+1}}{\lambda k}$. From this, we can deduce that $\mathbb{E}[\tau_j \mid \tau_{j+1}] \approx \tau_{j+1}\big(1 - \frac{1}{\lambda j}\big)$. By the Markov property (and assuming some concentration of τ_k conditionally on $\tau_{\lceil xk \rceil}$), for $x \geq 1$, we get

$$
\tau_k \approx \mathbb{E}\Big[\tau_k \;\Big|\; \tau_{\lceil xk \rceil}\Big] \approx \tau_{\lceil xk \rceil} \cdot \prod_{j=k}^{\lceil xk \rceil - 1} \bigg(1 - \frac{1}{\lambda j}\bigg) \approx \tau_{\lceil xk \rceil} \cdot x^{-\frac{1}{\lambda}}. \tag{2.19}
$$

Since from (2.6) we expect that $\frac{Q(xk)}{Q(k)} \approx x^{-\frac{1}{2\alpha}}$ and from (2.14) that $Q(k) \approx c \cdot \mathbb{E}\big[\tau_k^{-1/2}\big]$, the last display implies in particular that the constant λ is actually equal to α (see Remark 7.9). Moreover, recalling (2.15), this is exactly what we need to derive the regularity estimates for q and Q. The justification for the key insight in (2.18) will be developed in the next paragraphs.

A first rough estimate. It is possible to use the above line of reasoning using the inequality $\#\mathcal{L}_\cap^{\max}(T_n) \leq X_n$, valid deterministically, instead of the more precise (2.18). This (together with a symmetric argument with time going forwards) leads to the following estimates that hold for some constant $C > 0$, for all $k \geq 1$ and all $x \geq 1$:

$$
Cx^{-C} \leq \frac{q(xk)}{q(k)} \leq C \qquad \text{and} \qquad (1 + o_{k\to\infty}(1))x^{-C} \leq \frac{Q(xk)}{Q(k)} \leq 1. \tag{2.20}
$$

These, together with the asymptotics (2.6), are referred to as the **rough regularity estimates** and constitute the content of Proposition 5.1, proved in Section 5. These estimates are not precise enough to obtain (2.12), nor to extract the exponent γ from (2.8). However, they will still prove useful in the next part of the argument, which is itself used to prove (2.18).

A local convergence argument. In order to get (2.18), we rely on a local convergence argument developed in Section 6. There, we work with a slightly different tree model, denoted by T^k, which has the distribution of the p-signed critical binary Bienaymé–Galton–Watson tree T conditioned on the event $\{\mathrm{LIS}(T) = k\}$.[7] More precisely, for any finite sign-decorated binary tree t, we write t^{lmax} for the *leftmost* maximal positive subtree of t. This is the maximal positive subtree of t informally defined by going left at any negative node v where the left and right subtrees above v have equal LIS. We will study the local convergence of T^k around a leaf L^k chosen uniformly at random among the k leaves of $(T^k)^{\mathrm{lmax}}$; see Theorem 6.2 and the discussion below it.

The main tool that we use is some *exploration process* that records (among other things) the LIS of the subtree above node v, as v travels along the spine of the tree from the root to the marked leaf L^k. It turns out that this process is a Markov chain and this is precisely our reason for studying the conditioned model T^k with fixed LIS instead of T_n.

The proof of local convergence relies on the construction of a coupling between the exploration processes associated with T^k and $T^{k'}$, for different large values of k and k', ensuring that the two processes coalesce very fast as $k, k' \to \infty$. Since the transitions of the Markov chain are given in terms of the sequence $q(k)$, the rough regularity estimates (2.20) play an important role in this proof.

Motivation for the local convergence. The point of the former result is to prove that the probability that the uniform leaf L^k belongs to $\mathcal{L}_\cap^{\mathrm{max}}(T^k)$ tends to some deterministic limit $\lambda \in (0,1)$, which can be interpreted as the probability of some event happening in the local limit. This in turn ensures that, as $k \to \infty$,

$$\frac{\#\mathcal{L}_\cap^{\mathrm{max}}(T^k)}{\mathrm{LIS}(T^k)} = \frac{\#\mathcal{L}_\cap^{\mathrm{max}}(T^k)}{k} \xrightarrow{\mathbb{P}} \lambda. \tag{2.21}$$

This is the content of Theorem 7.2 proved in Section 7.1. With some additional work, it is possible to transfer this result to our original model T_n, which has the same distribution as T conditionally on $\{|T| = n\}$. This ensures that for any large $n \geq 1$, we have

$$\#\mathcal{L}_\cap^{\mathrm{max}}(T_n) \approx \lambda \, \mathrm{LIS}(T_n) = \lambda X_n, \tag{2.22}$$

obtaining the key insight from (2.18). While the existence of the local limit is motivated by the concentration result in (2.21), it also provides valuable insights into the landscape around a uniformly chosen leaf in a maximal positive subtree.

Obtaining the good regularity estimates and the exponent α. To summarize, we used the rough regularity estimate (2.20), obtained earlier, to derive (2.22). This, in turn, allows us to rigorously justify the arguments leading to (2.19), and thereby obtain the **good regularity estimates** in (2.12) for the sequences q and Q. This last step actually turns out to be quite challenging and will take up the whole of Section 7.2.

Once (2.12) is established, we will need a new technique in Section 8 to extract the value of the exponent α. Here the main idea is to compare the recursive relation for q from Lemma 2.4 with two related *inequalities*, each satisfied by a more regular test sequence. These test sequences, combined with (2.12), will then yield sharp enough upper and lower bounds for the sequence q to extract the exact exponent α.

---

[7]Beware that T^k and T_n stand for two very different conditionings; to avoid any potential risk of confusion we shall always stick to k for the conditioning on the LIS and n for that on the number of leaves.

As a by-product: the scaling limit result. Using as an input the newly obtained good regularity estimates (2.12), we can revisit the local convergence argument and obtain in Theorem 9.1 a more quantitative version of (2.21). We find that there exists $\varepsilon > 0$ such that for k large we have

$$\mathbb{E}\left[\left|\#\mathcal{L}_{\cap}^{\max}(T^k) - \alpha k\right|\right] \leq k^{-\varepsilon}. \tag{2.23}$$

Going back to (2.16) and assuming that (2.23) can be leveraged to get some quantitative control over $|\#\mathcal{L}_{\cap}^{\max}(T_n) - \alpha X_n|$, we get that roughly

$$\mathbb{E}[X_{n-1} \mid (T_i)_{i \geq n}] = X_n - \frac{\#\mathcal{L}_{\cap}^{\max}(T_n)}{n} \approx X_n - \frac{\alpha X_n}{n} \approx \left(1 - \frac{\alpha}{n}\right) X_n.$$

If the last display held with an equal sign, it would be easy by some (backward) martingale argument to show that $n^{-\alpha} X_n$ converges almost surely to some random variable. The proof of Theorem 2.3 in Section 9 follows this idea, by writing $n^{-\alpha} X_n$ as some (backward) martingale part plus some error term, and showing that the estimates that we have are sufficient to control the error term. Some additional work will then be needed to ensure that all the properties of the limit X claimed in Theorem 2.3 hold.

Another by-product: scaling limit of largest positive subtrees of T^k. Let us finally mention that another by-product of our proof is a scaling limit result for the shape of the leftmost maximal positive subtree of T^k, which complements the local convergence mentioned before (Theorem 6.2). If t is a finite sign-decorated binary tree, we recall that t^{lmax} stands for its leftmost maximal positive subtree. We can equip it with the metric coming from the restriction to t^{lmax} of the graph distance on t (equivalently, this is the graph distance on the uncontracted version of t^{lmax}). If $a > 0$ and S is a metric space, we also denote by $a \cdot S$ the metric space obtained by multiplying the distances in S by a factor a.

Theorem 2.6 (Scaling limit of the leftmost positive subtree of T^k). *Let α be the unique solution in the interval $(1/2, 1)$ to the equation in (1.3) and recall that $\gamma - 1 \overset{(2.7)}{=} \frac{1}{2\alpha} \in (1/2, 1)$. Recall also the slowly varying function φ of (2.12). Then the following convergence in distribution holds for the Gromov–Hausdorff topology:*

$$\frac{\varphi(k)}{k^{\gamma-1}} \cdot (T^k)^{\text{lmax}} \xrightarrow[k\to\infty]{d} \mathcal{T}_\gamma,$$

where $\mathcal{T}_\gamma$ is the self-similar fragmentation tree with self-similarity index $1 - \gamma$ and with binary dislocation measure ν_γ given by

$$\int f(s_1, s_2)\nu_\gamma(\mathrm{d}\mathbf{s}) = \int_0^{1/2} f(1-x, x)\, x^{-\gamma}(1-x)^{-\gamma}\, \mathrm{d}x, \tag{2.24}$$

for any measurable function $f : [0, 1]^2 \to \mathbb{R}_{>0}$.

Since this theorem is not needed anywhere in the proof of the main results, we omit precise definitions of continuous fragmentation trees and refer to [HM12] instead. We will give a short proof of Theorem 2.6 in Section 9.6 as an easy consequence of the good regularity estimates on q, together with a general convergence criterion established by Haas and Miermont in [HM12]. Since Theorem 2.6 is the result that comes naturally as a by-product of our arguments, we do not try to show the analogous convergence for $(T_n)^{\text{lmax}}$, which appears to be more challenging.

3 Notation and preliminary results

3.1 Basic notation and definitions

We denote by $\mathbb{N} = \{0, 1, 2, \dots\}$ the set of non-negative integers, and for all real numbers $a < b$, we write the integer interval $[a, b] \cap \mathbb{Z}$ as $[\![a, b]\!]$. The set of positive real numbers is denoted by $\mathbb{R}_{>0}$, while the set of non-negative real numbers is denoted by $\mathbb{R}_{\geq 0}$. Moreover, we often write $a \wedge b$ for $\min(a, b)$ and $a \vee b$ for $\max(a, b)$.

For sums, we use the convention $\sum_{k=x}^{y} := \sum_{k=\lceil x \rceil}^{\lfloor y \rfloor}$, for all $x < y \in \mathbb{R}_{\geq 0}$. The same convention is naturally extended to the sequences $q = (q(k))$ and $Q = (Q(k))$ defined in (2.5) for which we write $Q(x)$ instead of $Q(\lceil x \rceil)$ when $x \geq 1$.

For two real-valued functions f and g with domain D, we write $f(x) = o(g(x))$ if $\frac{f(x)}{g(x)} \xrightarrow[x \to \infty]{} 0$. We write $f(x) = O(g(x))$ if there exists a constant $C > 0$ such that $f(x) \leq C\, g(x)$ for all $x \in D$. Moreover, given a parameter $K > 0$, we write $f(x) = o_K(g(x))$ if $\frac{f(x)}{g(x)} \xrightarrow[x \to \infty]{} 0$, where the rate of convergence depends on K (but not on x). Similarly, given a parameter $K > 0$, we write $f(x) = O_K(g(x))$ if there exists a constant $C_K > 0$ depending on K (but not on x) such that $f(x) \leq C_K\, g(x)$ for all $x \in D$.

We recall that a **slowly varying function** is a measurable function $\varphi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ such that for all $x > 0$, we have $\frac{\varphi(xz)}{\varphi(z)} \to 1$ as $z \to \infty$. A measurable function $\varphi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is called **regularly varying** of index β if for all $x > 0$, we have $\frac{\varphi(xz)}{\varphi(z)} \to x^{\beta}$ as $z \to \infty$. A comprehensive reference for the theory of regular variation is [BGT89]. In particular, we will need the following standard results on slowly varying functions. The first item is the Uniform Convergence Theorem [BGT89, Theorem 1.2.1] and the second is a particular case of the Potter bounds [BGT89, Theorem 1.5.6, item (ii)].

Lemma 3.1. *Let $\varphi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be a slowly varying function.*

1. *For any constant $K > 0$, the ratio $\frac{\varphi(a)}{\varphi(b)}$ converges to 1 as $a, b \to \infty$ uniformly in a, b as long as $K^{-1} \leq \frac{a}{b} \leq K$.*

2. *For any $\delta > 0$, there is an absolute constant $K = K(\delta)$ such that for all integers $1 \leq a < b$, we have $\frac{\varphi(a)}{\varphi(b)} \leq K \left(\frac{a}{b}\right)^{-\delta}$.*

For a sequence of real-valued random variables, we write $Y_n \xrightarrow[n \to \infty]{\mathbb{P}} Y$ to denote convergence in probability to the random variable Y, we write $Y_n \xrightarrow[n \to \infty]{\text{a.s.}} Y$ to denote almost sure convergence, and $Y_n \xrightarrow[n \to \infty]{\text{d}} Y$ to denote convergence in distribution. When two random variables Y and Z have the same law, we write $Y \stackrel{\text{d}}{=} Z$.

We also recall that a real-valued random variable X **stochastically dominates** another real-valued random variable Y if there exists a coupling of X and Y such that $X \geq Y$ a.s., and we write $X \succeq Y$.

A summary of the main notation used throughout this paper can be found in the index on page 97.

3.2 The Brownian separable permuton and the Brownian cographon

For completeness, we describe the constructions of the Brownian separable permutons and the Brownian cographons. We refer the reader to [Maa20, BBF+20, BBF+22a] for more details.

Under $\mathbb{P}$, we call **signed Brownian excursion** a triplet $(\mathfrak{e}, \mathfrak{s}, p)$ consisting of a (normalized) Brownian excursion $\mathfrak{e}$, together with an independent sequence $\mathfrak{s}$ of i.i.d. Bernoulli(p) $\oplus/\ominus$ signs, *i.e.* $\mathbb{P}(\oplus) = 1 - \mathbb{P}(\ominus) = p$. One should think of the sequence $\mathfrak{s}$ as being indexed by the local minima of $\mathfrak{e}$.

We define the following random relation $\lhd_{\mathfrak{e},\mathfrak{s},p}$ on $[0,1]$: conditional on $\mathfrak{e}$, if $x, y \in [0,1]$ with $x < y$, and if $\min_{[x,y]} \mathfrak{e}$ is reached at a unique point which is a strict local minimum $m_{x,y} \in (x, y)$, then

$$\begin{cases} x \lhd_{\mathfrak{e},\mathfrak{s},p} y & \text{if} \quad \mathfrak{s}(m_{x,y}) = \oplus, \\ y \lhd_{\mathfrak{e},\mathfrak{s},p} x & \text{if} \quad \mathfrak{s}(m_{x,y}) = \ominus. \end{cases} \tag{3.1}$$

Standard properties of the Brownian excursion (see, for instance, [BDSG24, Lemma 2.1]) ensure the existence of a random subset $\mathcal{R}_{\mathfrak{e}} \subset [0,1]$ such that:

- Almost surely, the complement of $\mathcal{R}_{\mathfrak{e}}$ has Hausdorff dimension $1/2$.

- For every $x, y \in \mathcal{R}_{\mathfrak{e}}$ with $x < y$, the minimum $\min_{[x,y]} \mathfrak{e}$ is reached at a unique point which is a strict local minimum.

In particular, the restriction of the relation $\lhd_{\mathfrak{e},\mathfrak{s},p}$ to $\mathcal{R}_{\mathfrak{e}}$ is a total order. Introducing the function

$$\forall t \in [0,1], \quad \psi_{\mathfrak{e},\mathfrak{s},p}(t) := \mathrm{Leb}\big(\{x \in [0,1] \mid x \lhd_{\mathfrak{e},\mathfrak{s},p} t\}\big), \tag{3.2}$$

then the (biased) **Brownian separable permuton** of parameter $p \in (0,1)$ is the push-forward of the Lebesgue measure on $[0,1]$ via the mapping $(\mathbb{I}, \psi_{\mathfrak{e},\mathfrak{s},p})$, where $\mathbb{I}$ denotes the identity. That is,

$$\boldsymbol{\mu}_p(\cdot) := (\mathbb{I}, \psi_{\mathfrak{e},\mathfrak{s},p})_* \mathrm{Leb}(\cdot) = \mathrm{Leb}(\{t \in [0,1] \mid (t, \psi_{\mathfrak{e},\mathfrak{s},p}(t)) \in \cdot\}). \tag{3.3}$$

Heuristically speaking, $\psi_{\mathfrak{e},\mathfrak{s},p}$ is the *continuum permutation* of the elements in the interval $[0,1]$ induced by the order $\lhd_{\mathfrak{e},\mathfrak{s},p}$ and $\boldsymbol{\mu}_p$ is the diagram of $\psi_{\mathfrak{e},\mathfrak{s},p}$. We highlight that $\boldsymbol{\mu}_p$ is a *random permuton*.

Remark 3.2. Recall from the discussion around (1.1) that, given $\boldsymbol{\mu}_p$, one can sample a corresponding permutation $\sigma_n = \sigma_n(p) = \mathrm{Perm}[\boldsymbol{\mu}_p, n]$ of size n. We also described the law of σ_n in terms of a p-signed uniform binary tree $T_n = T_n(p)$. We now explain how to obtain T_n directly from $\boldsymbol{\mu}_p$, or more precisely, how one can couple T_n with the signed excursion $(\mathfrak{e}, \mathfrak{s}, p)$ used in the construction of $\boldsymbol{\mu}_p$. As a corollary, we will see that the trees T_n are coupled for different values of n exactly as in the Rémy algorithm from Section 2.1.

For this, let $(U_i)_{i \geq 1}$ be a sequence of i.i.d. uniform random variables in $[0,1]$. For each $n \geq 1$, consider the first n points $(U_i)_{i \leq n}$ in $[0,1]$ and define the tree $t(\mathfrak{e}, \mathfrak{s}, p; (U_i)_{i \leq n})$ recursively as follows. Let $(\overline{U}_i)_{i \leq n}$ be the order statistic of $(U_i)_{i \leq n}$ and let m_i be the minimum value of $\mathfrak{e}$ on $[\overline{U}_i, \overline{U}_{i+1}]$ for all $i < n$. If $n = 1$, then $t(\mathfrak{e}, \mathfrak{s}, p; U_1)$ is a leaf. Otherwise, let $m = \min_i\{m_i\}$ and denote by j the index such that $m_j = m$. Then $t(\mathfrak{e}, \mathfrak{s}, p; (U_i)_{i \leq n})$ is the tree whose root is decorated by $\mathfrak{s}(m)$ and whose children are $t(\mathfrak{e}, \mathfrak{s}, p; (\overline{U}_i)_{i \leq j})$ and $t(\mathfrak{e}, \mathfrak{s}, p; (\overline{U}_i)_{j+1 \leq i \leq n})$. Then $t(\mathfrak{e}, \mathfrak{s}, p; (U_i)_{i \leq n}) \overset{\mathrm{d}}{=} T_n(p)$ and the trees $t(\mathfrak{e}, \mathfrak{s}, p; (U_i)_{i \leq n})$ are coupled for different values of n exactly as in the Rémy algorithm – see for instance [Le 05, Section 2.6] or [Pit06, Exercise 7.4.11].

We now turn to the Brownian cographon. Given the signed excursion $(\mathfrak{e}, \mathfrak{s}, p)$, the **Brownian cographon** $\boldsymbol{W}_p$ of parameter $p \in [0,1]$ is defined (following [BBF+22a]) as the graphon equivalence class of the random function

$$\begin{aligned} \boldsymbol{W}_p : [0,1]^2 &\to \{0,1\}, \\ (x,y) &\mapsto \mathbb{1}_{\mathfrak{s}(m_{x,y}) = \oplus}, \end{aligned} \tag{3.4}$$

23

where, as in (3.1), if $(x,y) \in \mathcal{R}_{\mathfrak{e}}^2$ and $x < y$, we denote by $m_{x,y} \in [x,y]$ the unique strict local minimum $\min_{[x,y]} \mathfrak{e}$. If x or y is not in $\mathcal{R}_{\mathfrak{e}}$ then we arbitrarily set $\mathfrak{s}(m_{x,y}) = \oplus$. This choice does not change the law of $\boldsymbol{W}_p$ because a.s. $\mathrm{Leb}([0,1]^2 \setminus \mathcal{R}_{\mathfrak{e}}^2) = 0$. We highlight that $\boldsymbol{W}_p$ is a *random* graphon.

3.3 A simple estimate for critical binary Bienaymé–Galton–Watson trees

Recall that T denotes a critical binary Bienaymé–Galton–Watson tree with i.i.d. Bernoulli(p) signs on its nodes, so that T conditioned to have size n is distributed as T_n. We will often use the following simple estimate on the size of T.

Lemma 3.3. *For all $n \geq 1$, set $r(n) := \mathbb{P}(|T| = n)$. Then*

$$r(n) = \frac{1}{2^{2n-1}} \cdot \frac{1}{n}\binom{2n-2}{n-1} = \frac{1}{2\sqrt{\pi}}n^{-3/2} + O\left(n^{-5/2}\right). \tag{3.5}$$

In particular, $r(n)$ is decreasing in n.

Proof. There are $\frac{1}{n}\binom{2n-2}{n-1}$ binary trees with n leaves, *i.e.* the $(n-1)$-th Catalan number. Moreover, each binary tree with n leaves has $n-1$ nodes. Since T is a critical binary Bienaymé–Galton–Watson tree, we immediately get the first equality in (3.5). The second one is then deduced by a standard application of Stirling's formula. The fact that $r(n)$ is decreasing in n follows by a simple direct calculation. $\square$

3.4 Decomposing p-signed uniform binary trees

Recall from Section 2.1 that given a subset $\mathcal{L}$ of the set of leaves of a sign-decorated binary tree t, we denote by $t|_{\mathcal{L}}$ the sign-decorated binary subtree with $\#\mathcal{L}$ leaves obtained from t by removing all vertices that do not have a descendant in $\mathcal{L}$ and contracting all vertices of degree 2 that may appear.

Lemma 3.4. *Let $n \geq m \geq 1$ and let T_n be a p-signed uniform binary tree with n leaves. Conditionally on T_n, let $\mathcal{L}_m = \{L_1, \ldots, L_m\}$ be a uniform random subset of the set of leaves of T_n with cardinality m. Then $T_n|_{\mathcal{L}_m}$ has the same law as T_m.*

Proof. This is an immediate consequence, for instance, of Remark 3.2. $\square$

We also note that $T_n|_{\mathcal{L}_m}$ has $m-1$ nodes which can be seen as nodes of the larger tree T_n. There is a natural correspondence between, on the one hand, the connected components of the complement of those vertices in T_n and, on the other hand, the edges of $T_n|_{\mathcal{L}_m}$. For any edge e of $T_n|_{\mathcal{L}_m}$, we denote by T_n^e the corresponding connected component in T_n; *cf.* Figure 8.

Let $e_1, \ldots, e_{2m-1}$ be the edges of $T_n|_{\mathcal{L}_m}$, labelled so that $e_1, \ldots, e_m$ are the edges incident to the leaves $L_1, \ldots, L_m$. For all $i \in \{1, \ldots, 2m-1\}$, we set

$$N_i := \text{the number of leaves of } T_n \text{ that belong to the region } T_n^{e_i},$$

and

$$\begin{aligned} \widetilde{N}_i &:= N_i - 1 && \text{if } i \in \{1, \ldots, m\}, \\ \widetilde{N}_i &:= N_i && \text{if } i \in \{m+1, \ldots, 2m-1\}. \end{aligned}$$

That is, $\widetilde{N}_i$ counts the number of leaves of T_n in each region, excluding the leaves $L_1, \ldots, L_m$.

We recall the definition of the *Dirichlet* distribution and of the *Dirichlet-multinomial* distribution, which will describe the law of the vector $(\widetilde{N}_i)_{1 \leq i \leq 2m-1}$.
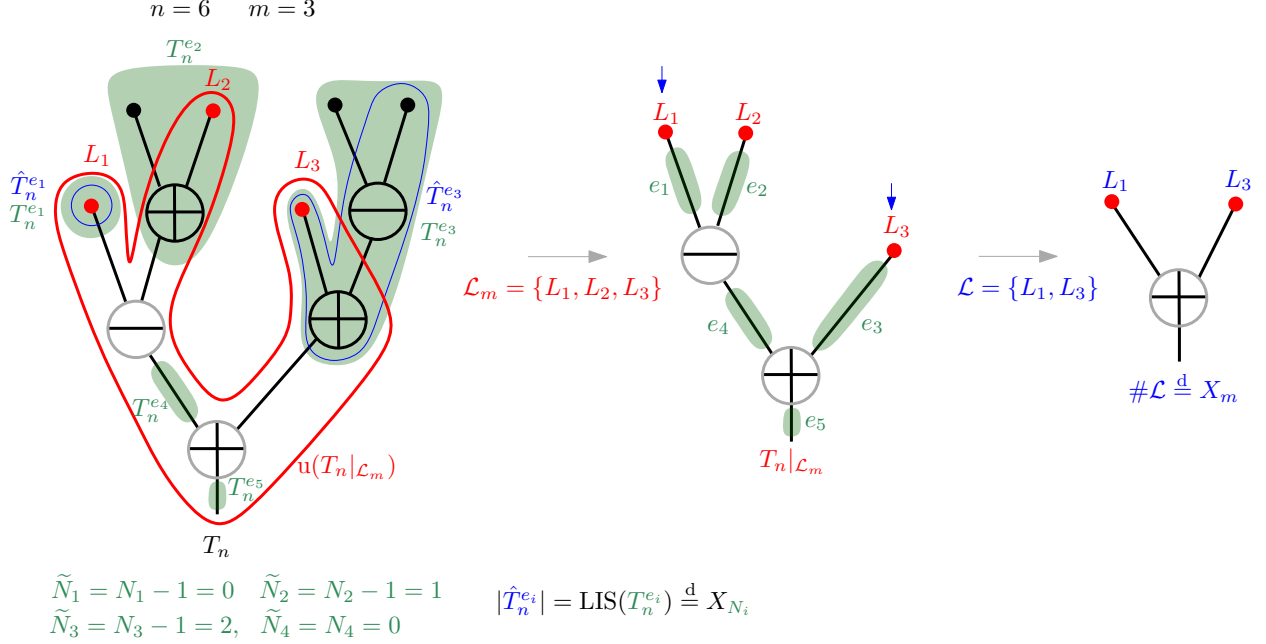
Figure 8: A diagram for the notation introduced in Section 3.4. The blue notation refers to the proof of Lemma 4.2. **Left:** The tree T_n, from which we select a subset of leaves $\mathcal{L}_m = \{L_1, L_2, L_3\}$ and form the uncontracted tree $\mathrm{u}(T_n|_{\mathcal{L}_m})$ highlighted in red. **Middle:** The tree $T_n|_{\mathcal{L}_m}$ with the two leaves in the set $\mathcal{L}$ corresponding to a maximal positive subtree marked by two blue arrows. To each edge e of $T_n|_{\mathcal{L}_m}$ corresponds a region T_n^e highlighted in green on the left part. We select a positive maximal subtree $\hat{T}_n^e$ in any such region adjacent to the subset $\mathcal{L}$. **Right:** The corresponding maximal positive subtree.

Definition 3.5. Let $n \geq 0, k \geq 1$ be integers, let $\alpha_1, \ldots, \alpha_k > 0$ and $\alpha_0 = \sum_{i=1}^{k} \alpha_i$.

- The Dirichlet distribution $\mathrm{Dir}(\alpha_1, \ldots, \alpha_k)$ is the law on k-tuples of non-negative real numbers $(x_1, \ldots, x_k)$ which sum up to 1 given by the probability density function

$$\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_k)} \prod_{i=1}^{k} x_i^{\alpha_i - 1}.$$

- The Dirichlet-multinomial distribution $\mathrm{DirMult}(n; \alpha_1, \ldots, \alpha_k)$ is the law on k-tuples of non-negative integers $(n_1, \ldots, n_k)$ which sum up to n given by the probability mass function

$$\frac{\Gamma(\alpha_0)\Gamma(n+1)}{\Gamma(n+\alpha_0)} \prod_{i=1}^{k} \frac{\Gamma(n_i + \alpha_i)}{\Gamma(\alpha_i)\Gamma(n_i + 1)}.$$

In other words, a $\mathrm{DirMult}(n; \alpha_1, \ldots, \alpha_k)$-distributed random variable $\mathbf{N} = (N_1, \ldots, N_k)$ can be obtained as follows: let $\mathbf{D} = (D_1, \ldots, D_k)$ follow the Dirichlet distribution $\mathrm{Dir}(\alpha_1, \ldots, \alpha_k)$; then, conditionally on $\mathbf{D}$, sample the vector $\mathbf{N}$ as a multinomial vector with parameters $(n; \mathbf{D})$.

Note that, by the law of large numbers, we immediately obtain the following result, stated in a convenient form for later use.

Lemma 3.6. *Fix an integer $m \geq 1$. Let $\mathbf{N} = (N_i)_{i \leq 2m-1}$ be a $\mathrm{DirMult}(n; \alpha_1, \ldots, \alpha_{2m-1})$-distributed random vector. Then*

$$\left(\frac{N_i}{n}\right)_{i \leq 2m-1} \xrightarrow[n \to \infty]{\mathrm{d}} (D_i^m)_{i \leq 2m-1},$$

where $(D_i^m)_{i \leq 2m-1}$ is a Dirichlet random vector with parameter $(\alpha_1, \ldots, \alpha_{2m-1})$.

We are making explicit the dependence on m of the random variable D_i^m because later on we will use this result in the case when $(D_i^m)_{i \leq 2m-1}$ is a Dirichlet random vector with parameter $(\frac{1}{2}, \ldots, \frac{1}{2})$ and the dependence on m will play a key role.

We also record a simple result about Dirichlet random variables.

Lemma 3.7. *Let $(D_1, \ldots D_n) \sim \mathrm{Dir}(1/2, \ldots, 1/2)$. Then*

$$\left(\frac{n}{2} D_1, \frac{n}{2} D_2\right) \xrightarrow[n \to \infty]{\mathrm{d}} (G_1, G_2),$$

where (G_1, G_2) is a pair of independent of $\mathrm{Gamma}\left(\frac{1}{2}\right)$-distributed random variables.

Proof. Let $(G_i)_{i \geq 1}$ be independent $\mathrm{Gamma}\left(\frac{1}{2}\right)$-distributed random variables. For any $n \geq 1$, we have the equality in distribution [Dev06, Chapter XI.4]

$$\frac{1}{\sum_{i=1}^{n} G_i} \cdot (G_1, G_2, \ldots, G_n) \stackrel{\mathrm{d}}{=} (D_1, \ldots D_n).$$

Since $\frac{n/2}{\sum_{i=1}^{n} G_i}$ converges to 1 in probability by the law of large numbers, the claimed result follows from Slutzky's lemma. $\square$

We finally state our last preliminary result on the decomposition of p-signed uniform binary trees.

Lemma 3.8. *Conditionally on $T_n|_{\mathcal{L}_m}$, the vector $(\widetilde{N}_i)_{1 \leq i \leq 2m-1}$ follows the Dirichlet-multinomial distribution $\mathrm{DirMult}\left(n - m; \frac{1}{2}, \ldots, \frac{1}{2}\right)$. Moreover, conditionally on $T_n|_{\mathcal{L}_m}$ and $(\widetilde{N}_i)_{1 \leq i \leq 2m-1}$, for any i with $1 \leq i \leq m$, the region $T_n^{e_i}$ corresponding to edge e_i (adjacent to the leaf L_i) has the same distribution as T_{N_i}, and the regions $(T_n^{e_i})_{1 \leq i \leq m}$ are jointly independent.*

In particular, since the laws of $(\widetilde{N}_i)_{1 \leq i \leq 2m-1}$ and $(T_n^{e_i})_{1 \leq i \leq m}$ conditionally on $T_n|_{\mathcal{L}_m}$ do not depend on $T_n|_{\mathcal{L}_m}$, then $(\widetilde{N}_i)_{1 \leq i \leq 2m-1}$ and $(T_n^{e_i})_{1 \leq i \leq m}$ are independent of $T_n|_{\mathcal{L}_m}$.

Proof. The result follows, for instance, by combining the two facts stated in [Pit06, Exercises 2.2.2 and 7.4.13]. $\square$

4 Existence of the exponent α via supermultiplicativity

Recall that $X_n = \mathrm{LIS}(T_n)$ denotes the maximal size (number of leaves) of a positive subtree of the p-signed uniform binary tree T_n with n leaves. The main goal of this section is to prove Theorem 2.1, *i.e.* the existence of the exponent α.

The proof of Theorem 2.1 consists of a supermultiplicativity argument. We first state a result of Dekking and Grimmett [DG88] which, although not directly applicable to our setting, has a similar flavor.

Proposition 4.1 ([DG88, Theorem 1]). *Let $(Y_n)_{n\geq 1}$ be a sequence of random variables with values in the set of positive integers. We assume that for any $m, n \geq 1$, we have the stochastic domination*

$$Y_{m+n} \succeq \sum_{i=1}^{Y_m} Y_n^{(i)}, \tag{4.1}$$

where the $Y_n^{(i)}$ are i.i.d. copies of Y_n and are independent of Y_m. Let $\lambda = \sup_{n\geq 1} \mathbb{E}[Y_n]^{1/n}$. If $\lambda < +\infty$, then we have the convergence in probability

$$Y_n^{1/n} \xrightarrow[n\to+\infty]{\mathbb{P}} \lambda.$$

In [DG88], processes satisfying (4.1) are called **superbranching processes**, by analogy with the branching property enjoyed by e.g. Bienaymé–Galton–Watson processes. Since we are interested in a quantity X_n with polynomial growth, it would be natural to try and obtain a superbranching equation of the form

$$X_{m\cdot n} \succeq \sum_{i=1}^{X_m} X_n^{(i)}. \tag{4.2}$$

This would mean breaking a tree of size $m \cdot n$ into m subtrees of size n. However, the most natural ways to decompose uniform binary trees into several regions result in regions with *random* and *different* sizes (recall Lemma 3.8 in Section 3.4), which is why our supermultiplicativity equation (4.3) will be more complicated than (4.2).

Recall the definition of the *Dirichlet-multinomial* distribution from Definition 3.5 in Section 3.4. We now state the superbranching-like equation that will be useful to us.

Lemma 4.2 (Supermultiplicativity equation for X_n). *For all $n \geq m \geq 1$, we have the stochastic domination*

$$X_n \succeq \sum_{i=1}^{X_m} X_{N_i}^{(i)}, \tag{4.3}$$

where:

- *The vector $(N_i - \mathbb{1}_{i\leq m})_{1\leq i\leq 2m-1}$ follows the $\mathrm{DirMult}\left(n-m; \frac{1}{2}, \ldots, \frac{1}{2}\right)$ distribution and is independent of X_m.*

- *Conditionally on $(N_i)_{1\leq i\leq 2m-1}$, the random variable $X_{N_i}^{(i)}$ has the law of X_{N_i} for all $i \in [\![1, m]\!]$ and the $(X_{N_i}^{(i)})_{1\leq i\leq m}$ are jointly independent and are independent of X_m.*

Remark 4.3. Note that since $X_m \leq m$, the random variables N_i involved in (4.3) are all among the first m coordinates of the vector $(N_i - \mathbb{1}_{i\leq m})_{1\leq i\leq 2m-1}$. That is, when the indicator function is equal to 1.

Proof. Let $n \geq m \geq 1$. We will build a subtree of T_n with only $\oplus$ nodes, whose total number of leaves will have the same distribution as the right-hand side of (4.3).

Conditionally on T_n, let $\mathcal{L}_m = \{L_1, \ldots, L_m\}$ be a uniform random subset of the set of leaves of T_n with size m. We consider the subtree $T_n|_{\mathcal{L}_m}$ as in Lemma 3.4. Then $T_n|_{\mathcal{L}_m}$ has the same law as T_m.

Recall the notation introduced above Lemma 3.8 and in Figure 8. Then, thanks to Lemma 3.8, the vector $(\widetilde{N}_i)_{1\leq i\leq 2m-1} = (N_i - \mathbb{1}_{i\leq m})_{1\leq i\leq 2m-1}$ follows the $\mathrm{DirMult}\left(n-m; \frac{1}{2}, \ldots, \frac{1}{2}\right)$ distribution

and is independent of $T_n|_{\mathcal{L}_m}$. Moreover, conditionally on $(N_i)_{1 \leq i \leq 2m-1}$, the regions $T_n^{e_i}$ for $1 \leq i \leq m$, corresponding to the edges e_i, are independent copies of T_{N_i} and are independent of $T_n|_{\mathcal{L}_m}$.

Now consider the largest positive subtree of $T_n|_{\mathcal{L}_m}$ (if there is more than one such maximal subtree, pick one arbitrarily), and denote by $\mathcal{L} \subset \mathcal{L}_m = \{L_1, \ldots, L_m\}$ the set of its leaves as on Figure 8. Since $T_n|_{\mathcal{L}_m}$ has the same law as T_m, the size of $\mathcal{L}$ has the same law as $X_m = \mathrm{LIS}(T_m)$. Moreover, for each $L_i \in \mathcal{L}$, consider a maximal positive subtree $\widehat{T}_n^{e_i}$ of $T_n^{e_i}$. Since $T_n^{e_i}$ has the same law as T_{N_i} by Lemma 3.8, the size $\mathrm{LIS}(T_n^{e_i})$ of $\widehat{T}_n^{e_i}$ has the same law as X_{N_i}.

The subtree of T_n induced by the leaves of $\bigcup_{L_i \in \mathcal{L}} \widehat{T}_n^{e_i}$ has only $\oplus$ nodes and, thanks to the previous paragraph, its number of leaves has the same distribution as $\sum_{i=1}^{X_m} X_{N_i}^{(i)}$, where N_i and the $X_{N_i}^{(i)}$ are as in the statement of the Lemma. This concludes the proof. $\qquad\square$

With Lemma 4.2 in our hands, we can turn to the proof of Theorem 2.1.

Proof of Theorem 2.1. For $n \geq 1$, we set $\alpha_n := \frac{\log \mathbb{E}[X_n]}{\log n}$ so that

$$\mathbb{E}[X_n] = n^{\alpha_n}.$$

Let $\alpha = \limsup_{n \to \infty} \alpha_n$ as in (2.2). By Markov's inequality, it is immediate that for any $\varepsilon > 0$, the probability that $X_n \geq n^{\alpha+\varepsilon}$ goes to 0 as $n \to \infty$. Therefore, to prove (2.1), we only need to show that for any $\varepsilon > 0$,

$$\mathbb{P}(X_n \leq n^{\alpha-\varepsilon}) \to 0 \quad \text{as } n \to \infty. \tag{4.4}$$

Once we have this, the fact that $\alpha \in (1/2, 1)$ is an immediate consequence of Theorem 1.2. We divide the proof of (4.4) into five main steps.

Step 1: Turning (4.3) into an equation closer to (4.2). We first turn (4.3) into an equation closer to (4.2) in order to compare X_n with a branching process. This goal is reached in (4.7). We fix $m \leq n$. The precise choice of m will be decided towards the end of the proof in Step 4.

Roughly speaking, Lemma 3.6 ensures that the random variables N_i for $1 \leq i \leq 2m-1$ appearing in (4.3) behave like $n \cdot D_i^m$, where $(D_1^m, D_2^m, \ldots, D_{2m-1}^m) \sim \mathrm{Dir}\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)$, so that they are of order $\frac{n}{m}$ and are random at that scale. In order to obtain a version of (4.3) with i.i.d. terms on the right-hand side, we will keep only the values of i for which $N_i > \frac{n}{m}$.

With the notation of Lemma 4.2, we can write the stochastic domination

$$X_n \succeq \sum_{i=1}^{X_m} X_{N_i}^{(i)} \succeq \sum_{i=1}^{X_m} X_{N_i}^{(i)} \mathbb{1}_{\{N_i \geq \frac{n}{m}\}} \succeq \sum_{i=1}^{\widetilde{Y}_m} X_{\lfloor \frac{n}{m} \rfloor}^{(i)}, \tag{4.5}$$

where

$$\widetilde{Y}_m = \widetilde{Y}_m(n) \overset{\mathrm{d}}{=} \sum_{i=1}^{X_m} \mathbb{1}_{\{N_i \geq \frac{n}{m}\}},$$

and the random variables $X_{\lfloor \frac{n}{m} \rfloor}^{(i)}$ are independent copies of $X_{\lfloor \frac{n}{m} \rfloor}$, also independent of $\widetilde{Y}_m$. Note that the last domination in (4.5) is possible because Lemma 4.2 guarantees that, conditional on (N_i), the $X_{N_i}^{(i)}$ are all independent and are also independent of X_m.

The domination in (4.5) is closer to (4.2) but we also need to replace $\widetilde{Y}_m(n)$ with a random variable Y_m that does not depend on n. By Lemma 3.6, we have the convergence in distribution

$$\sum_{i=1}^{X_m} \mathbb{1}_{\{N_i \geq \frac{n}{m}\}} \xrightarrow[n \to \infty]{\mathrm{d}} \sum_{i=1}^{X_m} \mathbb{1}_{\{D_i^m \geq \frac{1}{m}\}},$$

28

where $(D_1^m, \ldots D_{2m-1}^m) \sim \mathrm{Dir}(1/2, \ldots, 1/2)$ and is independent of X_m. We then introduce[8]

$$Y_m := \sum_{i=1}^{X_m} \mathbb{1}_{\{D_i^m \geq \frac{2}{m}\}}, \tag{4.6}$$

so that for any $k \in [\![1, m]\!]$, we have

$$\mathbb{P}\Big(\widetilde{Y}_m(n) \geq k\Big) = \mathbb{P}\Bigg(\sum_{i=1}^{X_m} \mathbb{1}_{\{N_i \geq \frac{n}{m}\}} \geq k\Bigg) \xrightarrow[n \to \infty]{} \mathbb{P}\Bigg(\sum_{i=1}^{X_m} \mathbb{1}_{\{D_i^m \geq \frac{1}{m}\}} \geq k\Bigg) > \mathbb{P}(Y_m \geq k).$$

Now, because of the strict inequality in the last display, the probability $\mathbb{P}\Big(\widetilde{Y}_m(n) \geq k\Big)$ is eventually larger than $\mathbb{P}(Y_m \geq k)$ for all $k \in [\![1, m]\!]$. Hence, for all $m \geq 1$, there exists $n_0 = n_0(m) > 0$ such that $\widetilde{Y}_m(n) \succeq Y_m$ for $n \geq n_0$. Summarizing,

$$X_n \succeq \sum_{i=1}^{Y_m} X_{\lfloor \frac{n}{m} \rfloor}^{(i)} \qquad \text{for all } m \geq 1 \text{ and } n \geq n_0(m), \tag{4.7}$$

where Y_m is as in (4.6). This domination is now closer to (4.2) as we want.

Step 2: Lower bounding the expectation of Y_m. Note that Y_m has expectation

$$\mathbb{E}[Y_m] = \mathbb{E}[X_m] \cdot \mathbb{P}\Big(D_1^m \geq \frac{2}{m}\Big) = m^{\alpha_m} \cdot \mathbb{P}\Big(D_1^m \geq \frac{2}{m}\Big), \tag{4.8}$$

since the vector $(D_i^m)_{1 \leq i \leq 2m-1}$ is independent of X_m. Using Lemma 3.7, we can set

$$b := \inf_{m \geq 1} \left\{ \mathbb{P}\Big(D_1^m \geq \frac{2}{m}\Big) \right\} > 0, \tag{4.9}$$

which gives

$$\mathbb{E}[Y_m] \geq m^{\alpha_m} b \qquad \text{for all } m \geq 1. \tag{4.10}$$

Step 3: Lower bounding X_n by comparing it with an appropriate branching process. Let $n_0 = n_0(m)$ be as in (4.7). By iterating (4.7), an induction on k shows that the process $(X_{n_0 m^k})_{k \geq 0}$ stochastically dominates a Bienaymé–Galton–Watson process $(Z_k)_{k \geq 0}$ started with one particle and with offspring distribution given by the law of Y_m. Therefore, setting

$$k = k(m, n) := \left\lfloor \frac{\log(n/n_0)}{\log m} \right\rfloor, \tag{4.11}$$

so that $n_0 m^{k+1} > n \geq n_0 m^k$, we get $X_n \succeq X_{n_0 \cdot m^k} \succeq Z_k$.

We now claim that there exists a threshold $m_0 > 0$, such that for every $m \geq m_0$, there exist $c = c(m) > 0$ and $n_0' = n_0'(m) > 0$ such that for all $n \geq n_0'(m)$:

$$\mathbb{P}\Big(X_n \geq c \cdot (m^{\alpha_m} b)^k\Big) \geq c, \tag{4.12}$$

where $k = k(m, n)$ is as in (4.11) (and so its value is fixed by m, n). Indeed, since $X_n \succeq Z_k$, for any choice of the constant $c > 0$, we can write

$$\mathbb{P}\Big(X_n \geq c \cdot (m^{\alpha_m} b)^k\Big) \geq \mathbb{P}\Big(Z_k \geq c \cdot (m^{\alpha_m} b)^k\Big) \geq \mathbb{P}\Big((\mathbb{E}[Y_m])^{-k} Z_k \geq c\Big),$$

---

[8]Note the extra factor 2 compared to the previous display.

where in the last equality we used that the offspring distribution Y_m of the branching process $(Z_i)_{i \geq 1}$ has expectation at least $m^{\alpha_m} b$ thanks to (4.10). By standard results (see for instance [KS66]), $(\mathbb{E}[Y_m])^{-k} Z_k$ converges almost surely as $k \to \infty$ to some $Z_\infty(m)$, which is positive on the event of non-extinction. Now, since $\mathbb{E}[Y_m] > 1$ for m large enough, there exists $m_0 > 0$ such that the event of non-extinction has positive probability (which depends on m) for $m > m_0$. Therefore, we can fix $c = c(m) > 0$ such that $\mathbb{P}(Z_\infty(m) \geq c) \geq 2c$. This ensures that $\mathbb{P}\big((\mathbb{E}[Y_m])^{-k} Z_k \geq c\big) \geq c$ for large enough k (*i.e.* for $n \geq n_0'(m)$), establishing (4.12).

In addition, when the event in (4.12) occurs, we have that

$$
\begin{aligned}
X_n &\geq c \cdot \exp(k(\alpha_m \log m + \log b)) \\
&\overset{(4.11)}{\geq} c \cdot \min\left\{ \exp\left(\left(\frac{\log(n/n_0)}{\log m} - 1\right)(\alpha_m \log m + \log b)\right), \exp\left(\frac{\log(n/n_0)}{\log m}(\alpha_m \log m + \log b)\right)\right\},
\end{aligned}
$$

where the minimum accounts for the fact that $\alpha_m \log m + \log b$ could *a priori* be negative. In any case, this implies that when the event in (4.12) occurs, we have

$$
\frac{\log X_n}{\log n} \geq \alpha_m + \frac{\log b}{\log m} + O_m\left(\frac{1}{\log n}\right). \tag{4.13}
$$

Step 4: Getting the right lower bound for X_n with positive probability. We will now specify the value of m. Let $\varepsilon > 0$. Recalling that $b \in (0,1)$ and $\alpha = \limsup_{m \to \infty} \alpha_m$, we choose m large enough so that $m \geq m_0$ and

$$
\frac{\log b}{\log m} \geq -\varepsilon/3 \qquad \text{and} \qquad \alpha_m \geq \alpha - \varepsilon/3. \tag{4.14}
$$

Now that m is fixed, (4.12) and (4.13) imply that there is a constant $c > 0$ such that for n large enough, we have

$$
\mathbb{P}\left(\frac{\log X_n}{\log n} \geq \alpha - \varepsilon\right) = \mathbb{P}(X_n \geq n^{\alpha - \varepsilon}) \geq c. \tag{4.15}
$$

In order to prove (4.4), it remains to turn the probability in (4.15) from c to $1 - \varepsilon$. This is the goal of the next (and last) step.

Step 5: Getting the right lower bound for X_n with high probability. The idea will be to use (4.7) again and argue that if each term on the right-hand side of (4.7) has a positive probability to be large (thanks to (4.15)), then the sum is very likely to be large, provided the number Y_m of terms is high enough. For this, we first claim that we can fix $m' > 0$ large enough so that

$$
\mathbb{P}\left(Y_{m'} \leq \frac{1}{\varepsilon c}\right) \leq \varepsilon, \tag{4.16}
$$

where c is as in (4.15). Indeed, recalling that $Y_m = \sum_{i=1}^{X_m} \mathbb{1}_{\{D_i^m \geq \frac{2}{m}\}}$, we have

$$
\mathbb{P}\left(Y_m \leq \frac{1}{\varepsilon c}\right) \leq \mathbb{P}(X_m \leq \sqrt{m}) + \mathbb{P}\left(\sum_{i=1}^{\sqrt{m}} \mathbb{1}_{\{D_i^m \geq \frac{2}{m}\}} \leq \frac{1}{\varepsilon c}\right).
$$

The first term on the right-hand side tends to zero as m tends to infinity thanks to Theorem 1.2, so we focus on the second term. Setting $S_m := \sum_{i=1}^{\sqrt{m}} \mathbb{1}_{\{D_i^m \geq \frac{2}{m}\}}$ and noting that $\mathbb{E}[S_m] \geq \lfloor \sqrt{m} \rfloor b$ by definition of b in (4.9), we get that for m large enough,

$$
\mathbb{P}\left(S_m > \frac{1}{\varepsilon c}\right) \geq \mathbb{P}\left(S_m > \frac{1}{\varepsilon c \lfloor \sqrt{m} \rfloor b} \mathbb{E}[S_m]\right) \geq \left(1 - \frac{1}{\varepsilon c \lfloor \sqrt{m} \rfloor b}\right)^2 \frac{\mathbb{E}[S_m]^2}{\mathbb{E}[S_m^2]},
$$

where the last inequality follows from the Paley–Zygmund inequality. The first factor on the right-hand side tends to one as m tends to infinity, so we are left to show that $\frac{\mathbb{E}[S_m]^2}{\mathbb{E}[S_m^2]} \xrightarrow[m \to \infty]{} 1$. Note that, computing the two expectations, we have that

$$\frac{\mathbb{E}[S_m]^2}{\mathbb{E}[S_m^2]} \geq \frac{\lfloor\sqrt{m}\rfloor^2 \mathbb{P}\left(D_1^m \geq \frac{2}{m}\right)^2}{\sqrt{m}\mathbb{P}\left(D_1^m \geq \frac{2}{m}\right) + m\mathbb{P}\left(D_1^m \geq \frac{2}{m}, D_2^m \geq \frac{2}{m}\right)}.$$

Thanks to Lemma 3.7, we know that $(mD_1^m, mD_2^m) \xrightarrow[m \to \infty]{\mathrm{d}} (G_1, G_2)$, where (G_1, G_2) is a pair of independent Gamma random variables of parameter $1/2$. This implies that the right-hand side of the above equation tends to one as we wanted, completing the proof of (4.16).

With (4.16) in our hands we can now return to (4.7). For n large enough, we have

$$\mathbb{P}\left(X_n < \left(\frac{n}{m'}\right)^{\alpha-\varepsilon}\right) \leq \mathbb{P}\left(Y_{m'} \leq \frac{1}{\varepsilon c}\right) + \mathbb{P}\left(\forall 1 \leq i \leq \frac{1}{\varepsilon c}, \ X^{(i)}_{\lfloor\frac{n}{m'}\rfloor} < \left(\frac{n}{m'}\right)^{\alpha-\varepsilon}\right)$$

$$\leq \varepsilon + (1-c)^{1/\varepsilon c}$$

$$\leq \varepsilon + \exp(-\varepsilon^{-1})$$

$$\leq 2\varepsilon,$$

where in the second inequality we used (4.16) and (4.15) with $\frac{n}{m'}$ in place of n. This proves (4.4). $\square$

5   Rough regularity estimates for the sequences q and Q

Recall the definition of the sequences q and Q from (2.5), and note from Section 2.4 that one of the main objectives of this paper is to establish the good regularity estimates (2.12) for these sequences. This goal will be accomplished in Theorem 7.1. As explained in Section 2.4 (see also the diagram in Figure 6), the proof of Theorem 7.1 relies crucially on the local convergence result established in Section 6. To obtain this result, we now prove the following *a priori* rough regularity estimates for the sequences q and Q.

Proposition 5.1 (Rough regularity estimates). *We have*

$$q(k) = k^{-\frac{1}{2\alpha}-1+o(1)} \qquad and \qquad Q(k) = k^{-\frac{1}{2\alpha}+o(1)} \qquad as\ k \to \infty,$$

where $\alpha \in (1/2, 1)$ is as in Theorem 2.1. Moreover, there is a constant $C > 0$ such that for any real $x \geq 1$, we have

$$(1 + o(1))\,x^{-C} \leq \frac{Q(xk)}{Q(k)} \leq 1 \qquad as\ k \to \infty, \tag{5.1}$$

and, for any $k \geq 1$,

$$C\,x^{-C} \leq \frac{q(xk)}{q(k)} \leq C, \tag{5.2}$$

where we are using the conventions $Q(z) = Q(\lceil z \rceil)$ and $q(z) = q(\lceil z \rceil)$ for all $z \in \mathbb{R}_{>0}$ as in Section 3.1.

The proof relies on the coupling of the decorated trees T_n between various values of n obtained through Rémy's algorithm introduced in Section 2.1. Recall that one can run Rémy's algorithm both forwards and backwards. For $n \geq 0$, we denote by $\mathcal{F}_n$ the σ-algebra generated by $(T_i)_{1 \leq i \leq n}$ and by $\overleftarrow{\mathcal{F}}_n$ the σ-algebra generated by $(T_i)_{i \geq n}$.

We recall that we set $X_n = \mathrm{LIS}(T_n)$ for all $n \geq 1$. Note that under this coupling, for all $n \geq 1$, we have $X_{n+1} - 1 \leq X_n \leq X_{n+1}$, and Theorem 1.2 guarantees that $X_n \to +\infty$ in probability. Therefore, for all $k \geq 1$, the stopping time

$$\tau_k := \min\{n \geq 1 \mid X_n = k\} \tag{5.3}$$

is almost surely finite (note that we also have the trivial a.s. bound $\tau_k \geq k$). In particular, let α be given by Theorem 2.1. Since X_n is non-decreasing in n, for all $\varepsilon > 0$ and $k \geq 1$, we have

$$\mathbb{P}\Big(\tau_k > k^{1/\alpha+\varepsilon}\Big) = \mathbb{P}\Big(X_{\lfloor k^{1/\alpha+\varepsilon}\rfloor} < k\Big) \quad \text{and} \quad \mathbb{P}\Big(\tau_k \leq k^{1/\alpha-\varepsilon}\Big) = \mathbb{P}\Big(X_{\lfloor k^{1/\alpha-\varepsilon}\rfloor} \geq k\Big). \tag{5.4}$$

Hence, Theorem 2.1 implies $\tau_k = k^{1/\alpha+o(1)}$ in probability. We start by making this statement a bit more precise, which will be useful later. We will rely on the description of the exponent α as a limsup, as we established in (2.2) of Theorem 2.1.

Lemma 5.2 (Control on the negative moments of τ_k). *Let $0 < \xi < \alpha$. Then we have*

$$\mathbb{E}\Big[\tau_k^{-\xi}\Big] = k^{-\frac{\xi}{\alpha}+o(1)} \qquad \text{as } k \to \infty. \tag{5.5}$$

In particular, we have

$$\mathbb{E}\Big[\tau_k^{-1/2}\Big] = k^{-\frac{1}{2\alpha}+o(1)} \qquad \text{and} \qquad \mathbb{E}\Big[\tau_k^{-3/2}\Big] = o\Big(\frac{1}{k}\mathbb{E}\Big[\tau_k^{-1/2}\Big]\Big) \qquad \text{as } k \to \infty. \tag{5.6}$$

Proof. The lower bound for (5.5) is immediate using $\tau_k = k^{1/\alpha+o(1)}$ in probability: for all $\varepsilon > 0$,

$$\mathbb{E}\Big[\tau_k^{-\xi}\Big] \geq \mathbb{E}\Big[\tau_k^{-\xi} \mathbb{1}_{\tau_k \leq k^{1/\alpha+\varepsilon}}\Big] \geq k^{-\frac{\xi}{\alpha}-\varepsilon\xi}\, \mathbb{P}\Big[\tau_k \leq k^{1/\alpha+\varepsilon}\Big] \sim k^{-\frac{\xi}{\alpha}-\varepsilon\xi} \quad \text{as } k \to \infty.$$

For the upper bound, let $\varepsilon > 0$ and let $k \geq 1$. We note that $\tau_k \geq k$, so $\tau_k^{-\xi} \leq k^{-\xi}$. For any $0 < x \leq k^{-\xi}$, using Markov's inequality and the description (2.2) of α as a limsup, we have

$$\mathbb{P}\Big(\tau_k^{-\xi} \geq x\Big) = \mathbb{P}\Big(\tau_k \leq x^{-1/\xi}\Big) = \mathbb{P}\Big(X_{\lfloor x^{-1/\xi}\rfloor} \geq k\Big) \leq \frac{1}{k}\mathbb{E}\Big[X_{\lfloor x^{-1/\xi}\rfloor}\Big] \leq \frac{1}{k}\lfloor x^{-1/\xi}\rfloor^{\alpha+\varepsilon}$$

if k is large enough (which implies that $x^{-1/\xi}$ is large enough). Note that this last inequality still holds if $x > k^{-\xi}$, because then the left-hand side is 0. By integrating over x, since $\frac{\alpha+\varepsilon}{\xi} > 1$, we find

$$\mathbb{E}\Big[\tau_k^{-\xi}\Big] = \int_0^\infty \mathbb{P}\Big(\tau_k^{-\xi} \geq x\Big)\mathrm{d}x \leq \int_0^\infty 1 \wedge \Big(\frac{1}{k}x^{-\frac{\alpha+\varepsilon}{\xi}}\Big)\mathrm{d}x \leq \int_0^{k^{-\frac{\xi}{\alpha+\varepsilon}}} 1 \, \mathrm{d}x + \frac{1}{k}\int_{k^{-\frac{\xi}{\alpha+\varepsilon}}}^\infty x^{-\frac{\alpha+\varepsilon}{\xi}}\mathrm{d}x$$
$$= O\Big(k^{-\frac{\xi}{\alpha+\varepsilon}}\Big).$$

Since this is true for any $\varepsilon > 0$, this proves (5.5) and the first claim in (5.6).

For the second claim in (5.6), we fix $\varepsilon > 0$ small enough to have $\frac{1}{2} + \varepsilon < \alpha$, which is possible because $\alpha > 1/2$ by Theorem 2.1. We write, using the fact that $\tau_k \geq k$ deterministically,

$$\mathbb{E}\Big[\tau_k^{-3/2}\Big] = \frac{1}{k}\mathbb{E}\Big[\frac{k}{\tau_k} \cdot \tau_k^{-1/2}\Big] \leq \frac{1}{k}\mathbb{E}\Big[\Big(\frac{k}{\tau_k}\Big)^\varepsilon \cdot \tau_k^{-1/2}\Big] = k^{\varepsilon-1}\mathbb{E}\Big[\tau_k^{-1/2-\varepsilon}\Big].$$

Using (5.5) with $\xi = \frac{1}{2} + \varepsilon$, we get $\mathbb{E}\Big[\tau_k^{-3/2}\Big] \leq k^{-\frac{1}{2\alpha}-1-(\frac{1}{\alpha}-1)\varepsilon+o(1)}$, whereas $\frac{1}{k}\mathbb{E}\Big[\tau_k^{-1/2}\Big] = k^{-\frac{1}{2\alpha}-1+o(1)}$ again by (5.5). This proves (5.6). $\qquad\square$

Proposition 5.1 will be an easy consequence of the estimates in the last Lemma 5.2 and the following result, which relates the sequences q and Q to the stopping times τ_k in order to compare them.

Lemma 5.3 (Relation between q, Q and the stopping times τ_k). *We have*

$$Q(k) = \left(1 + o\left(\frac{1}{k}\right)\right) \frac{1}{\sqrt{\pi}} \mathbb{E}\left[\tau_k^{-1/2}\right] \qquad as \ k \to \infty, \tag{5.7}$$

and

$$(1 + o(1)) \frac{1}{2\sqrt{\pi}} \frac{1}{k} \mathbb{E}\left[\tau_{k+1}^{-1/2}\right] \le q(k) \le (1 + o(1)) \frac{1}{2p\sqrt{\pi}} \frac{1}{k} \mathbb{E}\left[\tau_k^{-1/2}\right] \qquad as \ k \to \infty. \tag{5.8}$$

Moreover, there exists a constant $C > 0$ such that for all $k \ge 1$,

$$\frac{1}{C} \frac{Q(k)}{k} \le q(k) \le C \frac{Q(k)}{k}. \tag{5.9}$$

Proof. We divide the proof into five main steps.

Step 1: Proof of (5.7). For this, recall that we set $r(n) = \mathbb{P}(|T| = n)$ for all $n \ge 1$ and, from Lemma 3.3, we know that

$$r(n) = \frac{1}{2\sqrt{\pi}} n^{-3/2} \left(1 + O\left(\frac{1}{n}\right)\right) \quad as \ n \to \infty, \tag{5.10}$$

and $r(n)$ is decreasing in n. Therefore, summing over all possible values of $|T|$ and noting that $\mathbb{P}(X_n \ge k) = \mathbb{P}(\tau_k \le n)$, we can write as $k \to \infty$:

$$Q(k) = \mathbb{P}(\mathrm{LIS}(T) \ge k) = \sum_{n \ge 1} r(n)\mathbb{P}(X_n \ge k) = \mathbb{E}\left[\sum_{n \ge \tau_k} r(n)\right] = \frac{1}{2\sqrt{\pi}}\mathbb{E}\left[\sum_{n \ge \tau_k}\left(1 + O\left(\frac{1}{n}\right)\right)n^{-3/2}\right]$$

$$= \frac{1}{\sqrt{\pi}}\mathbb{E}\left[\tau_k^{-1/2}\right] + O\left(\mathbb{E}\left[\tau_k^{-3/2}\right]\right) = \frac{1}{\sqrt{\pi}}\mathbb{E}\left[\tau_k^{-1/2}\right] + o\left(\frac{1}{k}\mathbb{E}\left[\tau_k^{-1/2}\right]\right),$$

where the last equality follows from the second item of (5.6) in Lemma 5.2. Thus (5.7) holds.

Step 2: Preparation for the proof of (5.8). The key idea will be to use Rémy's algorithm forwards for the upper bound and backwards for the lower bound. As above, we first write

$$q(k) = \sum_{n \ge 1} r(n)\mathbb{P}(X_n = k) = \mathbb{E}\left[\sum_{n = \tau_k}^{\tau_{k+1}-1} r(n)\right] = \frac{1}{2\sqrt{\pi}}\left(1 + O\left(\frac{1}{k}\right)\right)\mathbb{E}\left[\sum_{n = \tau_k}^{\tau_{k+1}-1} n^{-3/2}\right], \tag{5.11}$$

where for the last equality we used (5.10) and that $\tau_k \ge k$. Bounding every term of the sum by either its largest or smallest term yields

$$\frac{1}{2\sqrt{\pi}}\left(1 + O\left(\frac{1}{k}\right)\right)\mathbb{E}\left[\frac{\tau_{k+1} - \tau_k}{\tau_{k+1}^{3/2}}\right] \le q(k) \le \frac{1}{2\sqrt{\pi}}\left(1 + O\left(\frac{1}{k}\right)\right)\mathbb{E}\left[\frac{\tau_{k+1} - \tau_k}{\tau_k^{3/2}}\right]. \tag{5.12}$$

Step 3: Proof of the upper bound of (5.8). Let $n \ge k \ge 1$, and let us condition on $X_n = k$. Recalling the terminology introduced at the beginning of Section 2.1, let $T_n^{\max}$ be a positive uncontracted subtree of T_n with k leaves (if it is not unique, pick one arbitrarily). Let also E_n be the edge of T_n that is picked by the Rémy algorithm run forwards at step n. Note that $T_n^{\max}$ is a tree with k

leaves and vertex degrees 1, 2 or 3, so it has at least $2k-1$ edges (this would be an equality if we considered the contracted version of $T_n^{\max}$ since all the vertices would have degree 1 or 3). If the edge E_n picked by the Rémy algorithm belongs to $T_n^{\max}$ and $S_n = \oplus$, then the union of $T_n^{\max}$ and the new leaf L_{n+1} forms a positive subtree of size $n+1$, so $\tau_{k+1} = n+1$. It follows that on the event $\{X_n = k\}$,

$$\mathbb{P}(\tau_{k+1} = n+1 \mid \mathcal{F}_n) \geq p\frac{2k-1}{2n-1} \geq p\frac{k-1}{n}, \tag{5.13}$$

since $\mathbb{P}(S_n = \oplus) = p$.

From here, we will obtain the upper bound of (5.8) by comparing the increment $\tau_{k+1} - \tau_k$ with a geometric random variable of parameter $p\frac{k}{\tau_k}$ conditionally on τ_k. More precisely, by an immediate induction, for all $j \geq 0$, we obtain from (5.13) that

$$\mathbb{P}(\tau_{k+1} > \tau_k + j \mid \tau_k) \leq \prod_{i=0}^{j-1}\left(1 - p\frac{k-1}{\tau_k + i}\right) \leq \exp\left(-p(k-1)\sum_{i=0}^{j-1}\frac{1}{\tau_k + i}\right),$$

where, when $j = 0$, we interpret the empty product as 1. Moreover, since

$$\sum_{i=0}^{j-1}\frac{1}{\tau_k + i} \geq \int_{\tau_k}^{\tau_k+j}\frac{dx}{x} = \log\left(1 + \frac{j}{\tau_k}\right),$$

we deduce that

$$\mathbb{P}(\tau_{k+1} > \tau_k + j \mid \tau_k) \leq \exp\left(-p(k-1)\log\left(1 + \frac{j}{\tau_k}\right)\right) = \left(1 + \frac{j}{\tau_k}\right)^{-p(k-1)}. \tag{5.14}$$

Therefore, we have for k large enough so that $p(k-1) > 1$,

$$\mathbb{E}[\tau_{k+1} - \tau_k \mid \tau_k] = \sum_{j\geq 0}\mathbb{P}(\tau_{k+1} > \tau_k + j \mid \tau_k) \leq \sum_{j\geq 0}\left(1 + \frac{j}{\tau_k}\right)^{-p(k-1)} \leq 1 + \int_0^\infty\left(1 + \frac{t}{\tau_k}\right)^{-p(k-1)}dt$$

$$= \left(1 + O\left(\frac{1}{k}\right)\right)\frac{\tau_k}{pk} + 1,$$

where the constant in the O notation is deterministic. This entails that

$$\mathbb{E}\left[\frac{\tau_{k+1} - \tau_k}{\tau_k^{3/2}}\right] = \left(1 + O\left(\frac{1}{k}\right)\right)\frac{1}{pk}\mathbb{E}\left[\tau_k^{-1/2}\right] + \mathbb{E}\left[\tau_k^{-3/2}\right]. \tag{5.15}$$

Plugging this back into (5.12), we obtain

$$q(k) \leq \frac{1}{2p\sqrt{\pi}}\left(1 + O\left(\frac{1}{k}\right)\right)\frac{1}{k}\mathbb{E}\left[\tau_k^{-1/2}\right] + O\left(\mathbb{E}\left[\tau_k^{-3/2}\right]\right).$$

Finally, the last error term is negligible by the right-hand side of (5.6) in Lemma 5.2, which proves the upper bound in (5.8).

Step 4: Proof of the lower bound of (5.8). The argument will be similar to the upper bound, but this time we will rely on the lower bound of (5.12) and use the Rémy algorithm run backwards. More precisely, let $n \geq k \geq 2$, and let us condition on $X_n = k$, so that $\tau_k \leq n$. We recall from Section 2.1 that, conditionally on $\overleftarrow{\mathcal{F}}_n$, the decorated tree T_{n-1} is obtained by removing a uniform leaf L_n of T_n. As before, let $T_n^{\max}$ be a positive uncontracted subtree of T_n with maximal size, *i.e.*

34

with k leaves. If $L_n \notin T_n^{\max}$ (which occurs with probability $1 - \frac{k}{n}$), then $T_n^{\max} \subset T_{n-1}$, which implies $X_{n-1} = X_n = k$, and so $\tau_k < n$. Therefore, on the event $\{X_n = k\}$, we have that

$$\mathbb{P}\Big(\tau_k < n \mid \overleftarrow{\mathcal{F}}_n\Big) \geq \mathbb{P}\Big(L_n \notin T_n^{\max} \mid \overleftarrow{\mathcal{F}}_n\Big) = 1 - \frac{k}{n}. \tag{5.16}$$

We refer to the end of the present section for a discussion on why this last inequality is not an equality, and why this matters a lot for the rest of the paper. For all $0 \leq j \leq \tau_{k+1} - k$, we get by induction from (5.16) that

$$\mathbb{P}(\tau_k < \tau_{k+1} - j \mid \tau_{k+1}) \geq \prod_{i=1}^{j}\Big(1 - \frac{k}{\tau_{k+1} - i}\Big) \geq \Big(1 - \frac{k}{\tau_{k+1} - j}\Big)^j, \tag{5.17}$$

where, when $j = 0$, we interpret the empty product as 1. Our upper bound argument (see (5.14)) suggests that the right order of magnitude for j is $\frac{\tau_{k+1}}{k}$. Therefore, we now fix a (large) constant $K > 0$, and assume $k > 2K$. By summing over j, we find that on the event $\{K\tau_{k+1}/k \leq \tau_{k+1} - k\}$, we have

$$\mathbb{E}[\tau_{k+1} - \tau_k \mid \tau_{k+1}] = \sum_{j \geq 0} \mathbb{P}(\tau_k < \tau_{k+1} - j \mid \tau_{k+1})$$

$$\geq \sum_{j=0}^{\lfloor K\tau_{k+1}/k \rfloor}\Big(1 - \frac{k}{\tau_{k+1} - \frac{K}{k}\tau_{k+1}}\Big)^j$$

$$= \Big(1 - \frac{K}{k}\Big)\frac{\tau_{k+1}}{k}\Big(1 - \Big(1 - \frac{k}{(1 - K/k)\tau_{k+1}}\Big)^{\lfloor K\tau_{k+1}/k \rfloor + 1}\Big)$$

$$\geq \Big(1 - \frac{K}{k}\Big)\frac{\tau_{k+1}}{k}\Big(1 - e^{-K}\Big),$$

where the last line just used that $1 - K/k \leq 1$ and $\lfloor K\tau_{k+1}/k \rfloor + 1 \geq K\tau_{k+1}/k$. It follows that

$$\mathbb{E}\left[\frac{\tau_{k+1} - \tau_k}{\tau_{k+1}^{3/2}}\right] \geq \mathbb{E}\Big[\mathbb{E}[\tau_{k+1} - \tau_k \mid \tau_{k+1}]\, \tau_{k+1}^{-3/2}\, \mathbb{1}_{\{K\tau_{k+1}/k \leq \tau_{k+1} - k\}}\Big]$$

$$\geq \Big(1 - \frac{K}{k}\Big)\Big(1 - e^{-K}\Big)\frac{1}{k}\mathbb{E}\Big[\tau_{k+1}^{-1/2}\mathbb{1}_{\{K\tau_{k+1}/k \leq \tau_{k+1} - k\}}\Big]. \tag{5.18}$$

To conclude, we argue that removing the indicator only costs a factor $1 + o(1)$. If the event in the indicator does not occur, then $\tau_{k+1} < \frac{k}{1 - K/k} \leq 2k$. Hence, using this with the trivial bound $\tau_{k+1} \geq k + 1$, and then Markov's inequality, we have

$$\mathbb{E}\Big[\tau_{k+1}^{-1/2}\mathbb{1}_{\{K\tau_{k+1}/k > \tau_{k+1} - k\}}\Big] \leq (k+1)^{-\frac{1}{2}} \cdot \mathbb{P}(\tau_{k+1} < 2k)$$

$$= (k+1)^{-\frac{1}{2}} \cdot \mathbb{P}(X_{2k-1} \geq k+1)$$

$$\leq (k+1)^{-\frac{3}{2}}\mathbb{E}[X_{2k-1}] = k^{-\frac{3}{2} + \alpha + o(1)},$$

by using (2.2) for the last inequality. On the other hand, by Lemma 5.2, we have $\mathbb{E}\Big[\tau_{k+1}^{-1/2}\Big] = k^{-\frac{1}{2\alpha} + o(1)}$. Moreover, by Theorem 2.1, we have $\frac{1}{2} < \alpha < 1$, which implies $-\frac{3}{2} + \alpha < -\frac{1}{2\alpha}$. Therefore, we have

$$\mathbb{E}\Big[\tau_{k+1}^{-1/2}\mathbb{1}_{\{K\tau_{k+1}/k \leq \tau_{k+1} - k\}}\Big] = (1 + o(1))\mathbb{E}\Big[\tau_{k+1}^{-1/2}\Big],$$

so (5.18) becomes

$$\mathbb{E}\left[\frac{\tau_{k+1} - \tau_k}{\tau_{k+1}^{3/2}}\right] \geq \left(1 - \frac{K}{k}\right)\left(1 - \mathrm{e}^{-K}\right)(1 + o(1))\frac{1}{k}\mathbb{E}\left[\tau_{k+1}^{-1/2}\right]$$

as $k \to \infty$. Since the above inequality is true for any constant $K > 0$, combined with the lower bound of (5.12), this proves the lower bound of (5.8).

Step 5: Proof of (5.9). The upper bound on q from (5.8) implies $q(k) = o(Q(k))$. Therefore $Q(k+1) = Q(k) - q(k) = (1 + o(1))Q(k)$ and again by (5.8),

$$\mathbb{E}\left[\tau_{k+1}^{-1/2}\right] = (1 + o(1))\mathbb{E}\left[\tau_k^{-1/2}\right] \qquad \text{as } k \to \infty. \tag{5.19}$$

Hence, we may replace $\mathbb{E}\left[\tau_{k+1}^{-1/2}\right]$ by $\mathbb{E}\left[\tau_k^{-1/2}\right]$ in the lower bound on q of Lemma 5.3. Therefore, there exists a constant $C > 0$ such that $\frac{1}{C}\frac{Q(k)}{k} \leq q(k) \leq C\frac{Q(k)}{k}$ for all k, as we wanted. □

We can now complete the proof of Proposition 5.1.

Proof of Proposition 5.1. The estimates $q(k) = k^{-\frac{1}{2\alpha}-1+o(1)}$ and $Q(k) = k^{-\frac{1}{2\alpha}+o(1)}$ are immediate by combining Lemma 5.3 with Lemma 5.2. Moreover, by Lemma 5.3, there exists a constant $C > 0$ such that for any k, we have

$$\frac{Q(k+1)}{Q(k)} = 1 - \frac{q(k)}{Q(k)} \geq 1 - \frac{C}{k}. \tag{5.20}$$

The estimate on Q then follows by writing a telescopic product: for all $x \geq 1$,

$$\frac{Q(xk)}{Q(k)} \geq \prod_{i=k}^{xk-1}\left(1 - \frac{C}{i}\right) = \exp\left(-\sum_{i=k}^{xk-1}\left(\frac{C}{i} + O\left(\frac{1}{i^2}\right)\right)\right)$$

$$\geq \exp\left(-C\log\frac{xk}{k} + O\left(\frac{1}{k}\right)\right) = (1 + o(1))\,x^{-C}.$$

Finally, it remains to prove the estimate (5.2) on q. By (5.9) in Lemma 5.3, there exists a constant $C > 0$ such that $\frac{1}{C}\frac{Q(k)}{k} \leq q(k) \leq C\frac{Q(k)}{k}$ for all k, so we can deduce the estimate on q from that on Q. □

What is still missing? As can be seen in this last proof, understanding the polynomial decay of Q and q is closely related to understanding the limit of the ratio $\frac{kq(k)}{Q(k)}$, which will be an important part of the rest of the paper. Note that Lemma 5.3 provides both an upper and a lower bound for this ratio, but neither is tight. Our strategy will be to turn the proof of the *lower* bound of Lemma 5.3 (Step 4 of the proof) into a more precise argument. We first notice that the "greedy" part of the argument is the discussion leading to (5.16), where we pick *one* specific maximal positive subtree T_n^{max}. To avoid such a waste, we can notice that $\tau_k = n$ if and only if the leaf L_n selected from T_n by the backward Rémy's algorithm belongs to *all* the maximal positive subtrees of T_n. Therefore, on the event $\{X_n = k\}$, we may write

$$\mathbb{P}\left(\tau_k < n \mid \overleftarrow{\mathcal{F}}_n\right) = 1 - \frac{\#\mathcal{L}_{\cap}^{\mathrm{max}}(T_n)}{n}, \tag{5.21}$$

or equivalently,

$$\mathbb{P}\left(\tau_k = n \mid \overleftarrow{\mathcal{F}}_n\right) = \frac{\#\mathcal{L}_{\cap}^{\mathrm{max}}(T_n)}{n}\mathbb{1}_{X_n=k}, \tag{5.22}$$

36

where we recall that $\mathcal{L}_{\cap}^{\max}(t)$ stands for the set of leaves in the intersection of *all* positive subtrees of t with maximal size. In particular, Section 7.1 will be devoted to proving that the ratio $\frac{\#\mathcal{L}_{\cap}^{\max}(T_n)}{X_n}$ is concentrated around a certain – non-explicit for now – deterministic constant λ (this is the law of large numbers in Theorem 7.2). The constant λ is exactly the multiplicative constant by which the lower bound in (5.8) of Lemma 5.3 is off. The proof of the law of large numbers will rely on the local convergence of a maximal positive subtree seen from a uniformly chosen leaf, resulting from Section 6. Finally, in Section 7.2, equipped with this concentration result, we will run again the proof of Proposition 5.1 in a more precise way to prove Theorem 7.1. We stress that the proof of the law of large numbers will not provide by itself an explicit value for the constant λ. However, once Theorem 7.1 is proved, we will be able to argue *a posteriori* that $\lambda = \alpha$; see Remark 7.9.

Note. The proofs of the local convergence and the law of large numbers presented in Section 6 and Section 7.1 are both intricate and rich in details. A reader encountering these sections for the first time may prefer to initially skip ahead, assuming the validity of the law of large numbers stated in Theorem 7.2, and proceed directly to the proof of Theorem 7.1 in Section 7.2. Following this approach also has the additional benefit of immediately seeing how the proof of Proposition 5.1 lays the ground for the more refined proof of Theorem 7.1. The reader is then encouraged to visit Section 6 and Section 7.1, as these sections offer deep insights into the landscape around a uniformly chosen leaf in a maximal positive subtree.

6   The local convergence of the leftmost maximal positive subtree of T^k around a uniformly chosen leaf

In the previous section, we proved the rough regularity estimates for the sequences q and Q stated in Proposition 5.1. Along the way, we realized that the problem of finding sharper estimates was closely related to the existence of the local limit of a maximal positive subtree. We have all freedom to decide which maximal positive subtree to consider: to fix ideas, we will look at the *leftmost* maximal subtree, which is defined by going left at each negative node carrying two subtrees with equal LIS (see Section 6.1.1 for a precise definition). The main goal of this section is to establish this local convergence (Theorem 6.2) for such leftmost maximal positive subtree around a uniformly chosen leaf. It turns out that the most natural model to state this local convergence result is the p-signed critical binary Bienaymé–Galton–Watson model T conditioned on $\mathrm{LIS}(T) = k$, denoted by T^k, in the regime where $k \to \infty$. This result will be later used in Section 7 to improve the rough regularity estimates for the sequences q and Q; completing the proof of Theorem 7.1. The proof of the local convergence hinges upon a key coupling argument which forms the bulk of this section, and essentially enables us to compare the situations when $\mathrm{LIS}(T) = k$ and when $\mathrm{LIS}(T) = k'$.

This section is structured as follows. In Section 6.1 we build the ground to state our local convergence result (Theorem 6.2) by introducing a Markov chain that records the behavior of T^k and its leftmost maximal positive subtree along the spine from the root to the chosen leaf. Section 6.2 is devoted to certain preliminary computations on this Markov chain that will be later used in Section 6.3 to run the main coupling argument.

Note. All the subtrees considered in Section 6 are uncontracted (see Section 2.1).

6.1   Preliminary results and statement of the local convergence

The main goal of this section is to state our local convergence result (Theorem 6.2) in Section 6.1.2. We first introduce some preliminary notation and results in Section 6.1.1.

6.1.1 Markovian description of the tree T^k around a uniformly chosen leaf of the leftmost maximal positive subtree

For any finite sign-decorated binary tree t, we denote by t^{lmax} the **leftmost maximal positive subtree** of t. More precisely, t^{lmax} can be constructed recursively in the following way (*cf.* Figure 9):

- If t consists of a single leaf, then $t^{\mathrm{lmax}} = t$.

- If t consists of a root vertex connected to two trees $t_\leftarrow$ on the left and $t_\rightarrow$ on the right, then:

  - If the root of t is decorated with a $\oplus$ sign, then t^{lmax} consists of a $\oplus$ node connected to the two trees $t_\leftarrow^{\mathrm{lmax}}$ and $t_\rightarrow^{\mathrm{lmax}}$.
  - If the root of t is decorated with a $\ominus$ sign and $\mathrm{LIS}(t_\leftarrow) \geq \mathrm{LIS}(t_\rightarrow)$, then $t^{\mathrm{lmax}} = t_\leftarrow^{\mathrm{lmax}}$.
  - If the root of t is decorated with a $\ominus$ sign and $\mathrm{LIS}(t_\leftarrow) < \mathrm{LIS}(t_\rightarrow)$, then $t^{\mathrm{lmax}} = t_\rightarrow^{\mathrm{lmax}}$.

Recall that T^k denotes the p-signed critical binary Bienaymé–Galton–Watson tree T conditioned on the event $\{\mathrm{LIS}(T) = k\}$. Conditionally on T^k, let L^k be a leaf chosen uniformly at random from the k leaves of the leftmost maximal positive subtree

$$T^{k,\mathrm{lmax}} := (T^k)^{\mathrm{lmax}}. \tag{6.1}$$

We first prove that the exploration of T^k along the spine directed towards L^k satisfies a nice Markov property. We denote by ρ_k the root of T^k and by $T_\leftarrow^k$ and $T_\rightarrow^k$ the left and right subtrees of ρ_k in T^k, respectively. We also denote by $D \in \{\leftarrow, \rightarrow\}$ the direction such that $L^k \in T_D^k$. For the unconditioned p-signed critical binary Bienaymé–Galton–Watson tree T, we analogously define the root ρ and the left and right subtrees $T_\leftarrow$ and $T_\rightarrow$ when $\deg(\rho) = 3$.

Lemma 6.1 (Decomposition of T^k at the root). *Let $k \geq 1$. Then the law of $\left(\mathrm{LIS}(T_\leftarrow^k), \mathrm{LIS}(T_\rightarrow^k), D\right)$ is described as follows:*

1. *With probability $\frac{1}{2q(1)} \mathbb{1}_{k=1}$, the tree T^k consists of a single vertex, which is a leaf.*

2. *For all $1 \leq i < k$, with probability $\frac{1}{2}(1-p)q(i)$, the root ρ_k is negative, $\mathrm{LIS}(T_\leftarrow^k) = i$ and $\mathrm{LIS}(T_\rightarrow^k) = k$. Moreover, in this case $D = \rightarrow$.*

3. *For all $1 \leq i \leq k$, with probability $\frac{1}{2}(1-p)q(i)$, the root ρ_k is negative, $\mathrm{LIS}(T_\leftarrow^k) = k$ and $\mathrm{LIS}(T_\rightarrow^k) = i$. Moreover, in this case $D = \leftarrow$.*

4. *For all $1 \leq i \leq k-1$, with probability $\frac{p}{2} \frac{q(i)q(k-i)}{q(k)}$, the root ρ_k is positive, $\mathrm{LIS}(T_\leftarrow^k) = i$ and $\mathrm{LIS}(T_\rightarrow^k) = k - i$. Moreover, conditionally on this, the direction D is $\leftarrow$ with probability $\frac{i}{k}$ and $\rightarrow$ with probability $\frac{k-i}{k}$.*

Moreover, conditionally on $\{(\mathrm{LIS}(T_\leftarrow^k), \mathrm{LIS}(T_\rightarrow^k), D) = (k_1, k_2, d)\}$, the trees $T_\leftarrow^k$ and $T_\rightarrow^k$ are independent trees with respective laws T^{k_1} and T^{k_2}, and L^k is a uniform leaf of $(T_d^k)^{\mathrm{lmax}}$.

Proof. The first claim simply follows from

$$\mathbb{P}\left(T^k = \{\rho_k\}\right) = \frac{\mathbb{P}(T = \{\rho\}, \mathrm{LIS}(T) = k)}{\mathbb{P}(\mathrm{LIS}(T) = k)} = \frac{\mathbb{P}(T = \{\rho\})}{\mathbb{P}(\mathrm{LIS}(T) = k)} \mathbb{1}_{k=1} = \frac{1}{2q(1)} \mathbb{1}_{k=1}.$$

For the second claim, fix $1 \leq i < k$. We can rewrite $\mathbb{P}\Big(s(\rho_k) = \ominus, \mathrm{LIS}(T_\leftarrow^k) = i, \mathrm{LIS}(T_\rightarrow^k) = k\Big)$ as

$$\frac{\mathbb{P}(\deg(\rho) = 2, s(\rho) = \ominus, \mathrm{LIS}(T_\leftarrow) = i, \mathrm{LIS}(T_\rightarrow) = k, \mathrm{LIS}(T) = k)}{\mathbb{P}(\mathrm{LIS}(T) = k)}$$

$$= \frac{\mathbb{P}(\deg(\rho) = 2, s(\rho) = \ominus, \mathrm{LIS}(T_\leftarrow) = i, \mathrm{LIS}(T_\rightarrow) = k)}{q(k)}.$$

By the branching property at the root vertex ρ of T, we conclude that

$$\mathbb{P}\Big(s(\rho_k) = \ominus, \mathrm{LIS}(T_\leftarrow^k) = i, \mathrm{LIS}(T_\rightarrow^k) = k\Big) = \frac{1}{2}(1-p)q(i).$$

Moreover, on the above event, we have $D = \rightarrow$ since $\mathrm{LIS}(T_\rightarrow^k) > \mathrm{LIS}(T_\leftarrow^k)$, so L^k must be in $T_\rightarrow$.

The proof of the third claim is the same as the second one. Note that in this case $D = \leftarrow$ since $\mathrm{LIS}(T_\leftarrow^k) \geq \mathrm{LIS}(T_\rightarrow^k)$ and we are considering the leftmost maximal subtree as defined at the beginning of Section 6.1.1.

For the fourth claim, fix $1 \leq i \leq k-1$ so that in particular $k-1 \geq 1$. We rewrite the probability $\mathbb{P}\Big(s(\rho_k) = \oplus, \mathrm{LIS}(T_\leftarrow^k) = i, \mathrm{LIS}(T_\rightarrow^k) = k-i\Big)$ as

$$\frac{\mathbb{P}(\deg(\rho) = 2, s(\rho) = \oplus, \mathrm{LIS}(T_\leftarrow) = i, \mathrm{LIS}(T_\rightarrow) = k-i, \mathrm{LIS}(T) = k)}{\mathbb{P}(\mathrm{LIS}(T) = k)} = \frac{p}{2}\frac{q(i)q(k-i)}{q(k)},$$

where for the equality we used that $\mathrm{LIS}(T) = \mathrm{LIS}(T_\leftarrow) + \mathrm{LIS}(T_\rightarrow) = k$ is implied by all the other conditions. The claim about the conditional distribution of D immediately follows.

The very last claim about the conditional independence and the conditional distribution of L^k is immediate. □

By iterating this Markov property, we can introduce a Markov chain that describes the law of the exploration of T^k and its leftmost maximal positive subtree along the spine directed towards L^k.
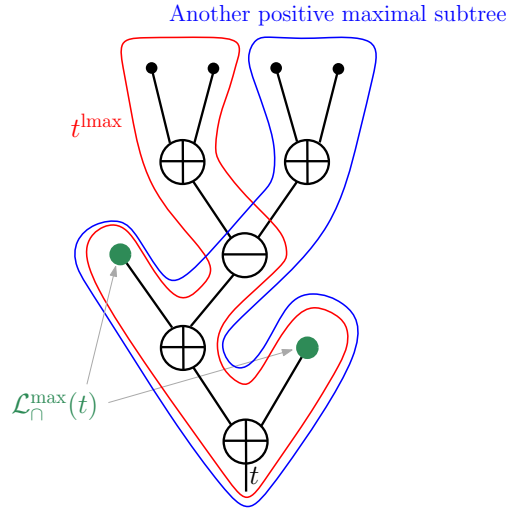

Figure 9: A sign-decorated binary tree t with two maximal positive subtrees highlighted in red and blue. The leftmost maximal positive subtree t^{lmax} is the red tree. The leaves in $\mathcal{L}_\cap^{\max}(t)$ are highlighted in green.

6.1.2 Description of the Markov chain and statement of the local convergence

For every $h \geq 0$, consider the vertex v_h in the tree T^k which is the ancestor of the marked leaf L^k at height $h \geq 0$; *cf.* Figure 10. In particular, $v_0 = \rho_k$ is the root of the tree T^k and, denoting by $\eta - 1$ the height of L^k in T^k, we have that $v_{\eta-1} = L^k$ and we can interpret v_η as a fictitious vertex lying above L^k playing the role of a *cemetery*. Importantly, we consider $v_{\eta-1}$ as a leaf, but not v_η.

Given the sign-decorated binary tree T^k, we introduce[9] the process Z^k,

$$(Z_h^k)_{h \geq 0} := (S_h, D_h, M_h^*, W_h)_{h \geq 0}, \tag{6.2}$$

with values in $\{\ominus, \oplus\} \times \{\leftarrow, \rightarrow\} \times \mathbb{N} \times \mathbb{N}$, defined as follows (and illustrated on Figure 10):

1. Let $(S_0, D_0, M_0^*, W_0) = (\ominus, \leftarrow, k, 0)$, by convention.

2. For any $h \in [\![1, \eta - 1]\!]$:

   - S_h is the **sign** in $\{\ominus, \oplus\}$ of v_{h-1}.
   - D_h is the **direction** in $\{\leftarrow, \rightarrow\}$ indicating whether v_h is the left or the right child of v_{h-1}.
   - M_h^* is the number of leaves of $T^{k,\text{lmax}}$ *weakly*[10] above v_h (the size of the **main** tree, which contains the marked leaf).
   - W_h is the number of leaves in a maximal positive subtree of the subtree rooted at the sibling of v_h, *i.e.* the child of v_{h-1} which is not v_h (we use the letter W since it is M upside down). Note that if $S_h = \oplus$, then W_h is the number of leaves of $T^{k,\text{lmax}}$ *weakly* above the sibling of v_h.

3. We use the convention that the chain is sent at time η to some **cemetery state**

$$\dagger := (\oplus, \leftarrow, 0, 0) \tag{6.3}$$

   and stays there afterwards. Note that defining the chain in this way allows us to write

$$\eta = \inf\{h \geq 0 \mid M_h^* = 0\}.$$

We will denote by S the process $(S_h)_{h \geq 0}$ in the first component of Z^k. Similarly, we denote by D, M^*, and W the processes in the other components of Z^k. Since Z^k is sent to the absorbing cemetery state $\dagger$ from time η onward, we shall restrict the study of Z^k to that of $(Z_h^k)_{0 \leq h \leq \eta}$.

It follows from Lemma 6.1 that Z^k is a Markov chain, whose transition probabilities will be given in Section 6.2. Recall that $\mathcal{L}_\cap^{\max}(t)$ denotes the set of leaves of t contained in all the maximal positive subtrees of t. Note that the only way for T^k to have maximal positive subtrees that do not contain L^k is to find some $h \geq 1$ such that the node v_{h-1} is negative and $M_h^* = W_h$. Hence, we can express the event that L^k belongs to $\mathcal{L}_\cap^{\max}(T^k)$ as a function of the trajectory of the chain, that is,

$$\left\{ L^k \in \mathcal{L}_\cap^{\max}(T^k) \right\} = \left\{ \sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h\} = 0 \right\}. \tag{6.4}$$

---

[9]There are two reasons why the third component M_h^* is marked with a star. First, it is the most important component of the process Z^k since the transition probabilities of Z^k will only depend on the value of M_h^*, as we will see in Section 6.2.1. Second, starting from (6.9), we will mainly focus on an accelerated version of M^* which we will simply denote by M.

[10]Note that the *weakly* condition is relevant only when $h = \eta - 1$ so that $v_{\eta-1}$ is a leaf, *i.e.* at the very end of the process.

$k=5,\ \eta=5$

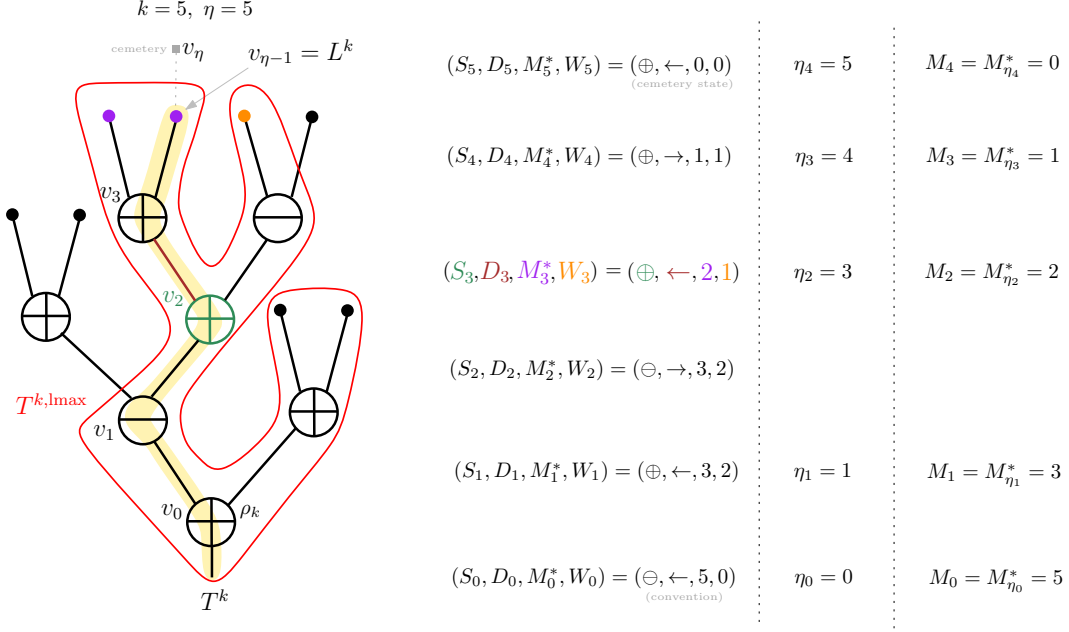|		
$(S_5, D_5, M_5^*, W_5) = (\oplus, \leftarrow, 0, 0)$ (cemetery state)	$\eta_4 = 5$	$M_4 = M_{\eta_4}^* = 0$
$(S_4, D_4, M_4^*, W_4) = (\oplus, \rightarrow, 1, 1)$	$\eta_3 = 4$	$M_3 = M_{\eta_3}^* = 1$
$(S_3, D_3, M_3^*, W_3) = (\oplus, \leftarrow, 2, 1)$	$\eta_2 = 3$	$M_2 = M_{\eta_2}^* = 2$
$(S_2, D_2, M_2^*, W_2) = (\ominus, \rightarrow, 3, 2)$		
$(S_1, D_1, M_1^*, W_1) = (\oplus, \leftarrow, 3, 2)$	$\eta_1 = 1$	$M_1 = M_{\eta_1}^* = 3$
$(S_0, D_0, M_0^*, W_0) = (\ominus, \leftarrow, 5, 0)$ (convention)	$\eta_0 = 0$	$M_0 = M_{\eta_0}^* = 5$

Figure 10: A schematic representation of the notation introduced for the process $(Z_h^k)_{h \geq 0} = (S_h, D_h, M_h^*, W_h)_{h \geq 0}$. Here $k = 5$ and $\eta = 5$. The spine from the root to the marked leaf L^k is highlighted in yellow and $T^{k,\text{lmax}}$ is marked in red. We used the other colors to highlight the elements of the tree T^k which determine the values of (S_3, D_3, M_3^*, W_3). The stopping times η_t and the decreasing Markov chain $M_t = M_{\eta_t}^*$ on the right will be introduced in (6.8) and (6.9).

This chain will be used later in Theorem 7.2 to establish, through (6.4) and the local convergence in Theorem 6.2, that

$$\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k} \to \lambda > 0$$

in probability as k tends to infinity. This is why we consider this four-component Markov chain.

We want to prove the convergence of the process Z^k defined in (6.2) "seen from the end", *i.e.* of the time-reversal of Z^k, because our interest lies in the behavior of Z^k near the chosen leaf L^k.

Theorem 6.2 (Local convergence of the leftmost maximal positive subtree around a uniformly chosen leaf)**.** *The process $(Z_{\eta-h}^k)_{0 \leq h \leq \eta}$ converges in distribution as $k \to \infty$ to a limiting process $(\overleftarrow{Z}_h)_{h \geq 0}$, in the sense of finite-dimensional distributions.*

The proof of the theorem above will mainly rely on a coupling argument; see Proposition 6.6 and Corollary 6.7.

We conclude this section by explaining why Theorem 6.2 can be interpreted as a local convergence result. Indeed, one can deduce[11] from Theorem 6.2 that the triplet $(T^k, T^{k,\text{lmax}}, L^k)$, viewed as a tree T^k marked at the leaf L^k and equipped with the distinguished subtree $T^{k,\text{lmax}}$, converges in the Benjamini–Schramm sense to a triplet

$$\left(T^\infty, T^{\infty,\text{lmax}}, L^\infty\right).$$

---

[11]This proof is standard and omitted, as it is not strictly necessary for our purposes.

This limit consists of an infinite tree T^∞ with a marked leaf L^∞ and a distinguished infinite "leftmost maximal" positive subtree $T^{\infty,\mathrm{lmax}}$. In particular, the triplet $\left(T^\infty, T^{\infty,\mathrm{lmax}}, L^\infty\right)$ is constructed as follows:

- Start by letting L^∞ be the marked leaf, with an infinite spine below it. Set $u_0 = L^\infty$ and for all $h \geq 1$, let u_h be the h-th vertex below L^∞ along this infinite spine. The infinite tree T^∞ should be regarded as rooted at the bottom of this spine, in line with our usual convention of rooting trees at their bottom vertex.

- Then, for all $h \geq 1$, conditioning on the value $\overleftarrow{Z}_h = (\overleftarrow{S}_h, \overleftarrow{D}_h, \overleftarrow{M}_h^*, \overleftarrow{W}_h)$:

  – Assign the sign $\overleftarrow{S}_h$ to u_h.

  – Attach an independent tree $T^{\overleftarrow{W}_h}$ to u_h on the left[12] (resp. right) side of the infinite spine if $\overleftarrow{D}_h = \rightarrow$ (resp. if $\overleftarrow{D}_h = \leftarrow$).

- Let T^∞ be the tree obtained through the procedure above regarded as rooted at the bottom of this spine and with the marked leaf L^∞.

- Finally, $T^{\infty,\mathrm{lmax}}$ is the union of the spine $\{u_h, h \geq 0\}$ and of the trees $T^{\overleftarrow{W}_h,\mathrm{lmax}}$ for $h \geq 0$ such that $\overleftarrow{S}_h = \oplus$. This is also an infinite tree regarded as rooted at the bottom of the spine.

6.2 Transition probabilities and the associated decreasing Markov chain

The first step in the proof of Theorem 6.2 is to compute the transition probabilities of the Markov chain $(Z_h^k)_{0 \leq h \leq \eta}$ introduced in (6.2). This is done in Section 6.2.1. Then, in Section 6.2.2 we will introduce a simpler (single-component) decreasing Markov chain M that is a reduced and time-changed version of Z^k. The heart of Theorem 6.2 will then be to produce a coupling of M and M' started from two values k and k', in which the two chains merge fast enough as $k, k' \to \infty$. This goal is achieved in Section 6.3.

6.2.1 Transition probabilities of the Markov chain

Recall that the process $(Z_h^k)_{0 \leq h \leq \eta}$ defined in (6.2) is a Markov chain. Let

$$\pi\big((s, d, m, w); (s', d', m', w')\big)$$

denote its transition probabilities. We start by noticing that by the very end of Lemma 6.1, these transition probabilities do not depend on s, d, w.

We first determine these transition probabilities when $m \geq 2$. From Items 2 and 3 of Lemma 6.1, we have

$$\pi\big((*, *, m, *); (\ominus, d, m, w')\big) = \frac{1-p}{2} q(w') \tag{6.5}$$

whenever either $d = \leftarrow$, $m \geq 2$, $1 \leq w' \leq m$, or $d = \rightarrow$, $m \geq 2$, $1 \leq w' < m$. Additionally, from Item 4 of Lemma 6.1, we have

$$\pi\big((*, *, m, *); (\oplus, d, m', m - m')\big) = \frac{p}{2} \frac{m'}{m} \frac{q(m')q(m - m')}{q(m)} \quad \text{if } d \in \{\leftarrow, \rightarrow\},\ m \geq 2,\ 1 \leq m' \leq m - 1. \tag{6.6}$$

---

[12] Here left and right refer to the usual notions for the plane where the infinite spine is drawn.

Note that this indeed defines a Markov kernel, since

$$\sum_{(s',d',m',w')} \pi((*,*,m,*);(s',d',m',w'))$$

$$= \sum_{m'=1}^{m-1} \frac{1-p}{2} q(m') + \sum_{m'=1}^{m} \frac{1-p}{2} q(m') + \sum_{m'=1}^{m-1} \frac{p}{2} \frac{m'}{m} \frac{q(m')q(m-m')}{q(m)} + \sum_{m'=1}^{m-1} \frac{p}{2} \frac{m'}{m} \frac{q(m')q(m-m')}{q(m)}$$

$$= \frac{1}{q(m)} \cdot \left((1-p)q(m) \sum_{m'=1}^{m-1} q(m') + \frac{1}{2}(1-p)q(m)^2 + \frac{p}{2} \sum_{m'=1}^{m-1} \frac{m'+m-m'}{m} q(m')q(m-m') \right)$$

$$= \frac{1}{q(m)} \cdot q(m) = 1, \tag{6.7}$$

where for the second-to-last equality we used the recursive relation for $q(m)$ introduced in Lemma 2.4.

For $m = 1$, note that $\pi((*,*,1,*);(\oplus,*,*,*)) = 0$. Moreover, from Item 1 of Lemma 6.1, we have probability

$$\pi((*,*,1,*);(\ominus,\leftarrow,0,0)) = \frac{1}{2q(1)}$$

to jump to the cemetery state $(\oplus,\leftarrow,0,0)$, and complementary probability

$$\pi((*,*,1,*);(\ominus,\leftarrow,1,1)) = 1 - \frac{1}{2q(1)}$$

to jump to $(\ominus,\leftarrow,1,1)$. Since the transition probabilities only depend on the third coordinate m, most of the rest of this section will be devoted to the study of the Markov chain $M^* = (M_h^*)_{0 \le h \le \eta}$.

6.2.2 The associated decreasing Markov chain

The Markov chain M^* is non-increasing, but it will be convenient for us to make it decreasing. For this, we consider the stopping times $(\eta_t)_{t \ge 0}$, where $\eta_0 = 0$ and for $t \ge 1$,

$$\eta_t := \inf\{h > \eta_{t-1} \mid M_h^* < M_{h-1}^*\} \tag{6.8}$$

is the t-th decreasing ladder epoch; see the second column on the right-hand side of Figure 10 for an example. This means that either $S_{\eta_t} = \oplus$, so that $v_{\eta_t - 1}$ is the t-th vertex along the spine from the root to L^k that carries a $\oplus$ sign, or that we reached the fictitious vertex v_η and that $M_{\eta_t}^* = 0$. We let

$$\mathtt{T} := \inf\left\{t \ge 0 \mid M_{\eta_t}^* = 0\right\}$$

be the total number of times that the chain M^* decreases during the process. Note that $\eta_\mathtt{T} = \eta$.

From now on, we focus our attention on the decreasing Markov chain M defined by

$$(M_t)_{0 \le t \le \mathtt{T}} := \left(M_{\eta_t}^* \right)_{0 \le t \le \mathtt{T}}, \tag{6.9}$$

see the third column on the right-hand side of Figure 10 for an example. Indeed, from the transitions of the chain, we can easily check that, conditionally on the trajectory $(M_t)_{0 \le t \le \mathtt{T}}$, the distribution of the trajectory of the whole process Z^k is given as follows:

- The waiting times $(\eta_t - \eta_{t-1})_{1 \leq t \leq \mathtt{T}-1}$ are independent geometric random variables with respective parameters $\theta(M_{t-1})$, where for $m \geq 2$ (note that $M_{t-1} \geq 2$ when $1 \leq t \leq \mathtt{T}-1$),

$$
\begin{aligned}
\theta(m) = \mathbb{P}(S_1 = \oplus \mid M_0^* = m) &= \sum_{m'=1}^{m-1} \sum_{d' \in \{\leftarrow, \rightarrow\}} \pi((*, *, m, *); (\oplus, d', m', m - m')) \\
&= 1 - (1 - p)\left(\sum_{m'=1}^{m-1} q(m') + \frac{q(m)}{2} \right) \\
&= 1 - (1 - p)\left(1 - Q(m) + \frac{Q(m) - Q(m+1)}{2} \right) \\
&= p + (1 - p)\frac{Q(m) + Q(m+1)}{2}, \qquad (6.10)
\end{aligned}
$$

where in the third equality we used (6.6) and the fact that we know that the second line of (6.7) equals one. Furthermore, the final waiting time $\eta_{\mathtt{T}} - \eta_{\mathtt{T}-1}$ is another independent geometric random variable, but with parameter $\theta(1) = \frac{1}{2q(1)}$.

- For any $t \in [\![0, \mathtt{T}-1]\!]$ and $h \in [\![1, \eta_{t+1} - \eta_t - 1]\!]$, we have

$$
M_{\eta_t + h}^* = M_t \qquad \text{and} \qquad S_{\eta_t + h} = \ominus. \qquad (6.11)
$$

Moreover, conditionally on $(M_t)_{0 \leq t \leq \mathtt{T}}$ and $(\eta_t)_{0 \leq t \leq \mathtt{T}}$, the random variables $(D_{\eta_t + h}, W_{\eta_t + h})$ are independent for all $t \in [\![0, \mathtt{T}-1]\!]$ and $h \in [\![1, \eta_{t+1} - \eta_t - 1]\!]$. Their conditional distribution is given as follows: for $d = \leftarrow, 1 \leq w \leq m$, or for $d = \rightarrow, 1 \leq w < m$,

$$
\begin{aligned}
\mathbb{P}\Big(D_{\eta_t + h} = d, W_{\eta_t + h} = w \mid M_t = m\Big) &= \mathbb{P}(D_1 = d, W_1 = w \mid S_1 = \ominus, M_0^* = m) \\
&= \frac{\pi((*, *, m, *); (\ominus, d, m, w))}{\mathbb{P}(S_1 = \ominus \mid M_0^* = m)} \\
&= \frac{1}{2} \frac{q(w)}{\left(1 - \frac{Q(m) + Q(m+1)}{2} \right)}, \qquad (6.12)
\end{aligned}
$$

where for the last equality we used (6.5) and (6.10).

- Finally, for all $t \in [\![0, \mathtt{T}-1]\!]$, we have $S_{\eta_t} = \oplus$,

$$
W_{\eta_t} = M_{\eta_t - 1}^* - M_{\eta_t}^* \overset{(6.11)}{=} M_{t-1} - M_t,
$$

and

$$
\mathbb{P}\Big(D_{\eta_t} = \leftarrow \mid S, M^*, W\Big) = \mathbb{P}\Big(D_{\eta_t} = \rightarrow \mid S, M^*, W\Big) = \frac{1}{2}. \qquad (6.13)
$$

We now compute the transition probabilities of the decreasing Markov chain M. By the branching property, we have

$$
\pi(m; m') := \mathbb{P}(M_{t+1} = m' \mid M_t = m) = \mathbb{P}\Big(M_{\eta_1}^* = m' \mid M_0^* = m\Big),
$$

so M is a time-homogeneous Markov chain. Moreover, the transition probabilities $\pi(\cdot\,;\cdot)$ can be expressed, for all $m \geq 2$ and $1 \leq m' < m$, as

$$
\begin{aligned}
\pi(m;m') &= \sum_{s=0}^{\infty} \mathbb{P}\big(\eta_1 \geq s, M_{s+1}^* = m' \mid M_0^* = m\big) \\
&= \sum_{s=0}^{\infty} \left((1-p)\left(\sum_{j=1}^{m-1} q(j) + \frac{q(m)}{2}\right)\right)^s \cdot \mathbb{P}(M_{s+1}^* = m' \mid M_0^* = m, \eta_1 \geq s) \qquad \text{by (6.10)} \\
&= \sum_{s=0}^{\infty} \left((1-p)\left(\sum_{j=1}^{m-1} q(j) + \frac{q(m)}{2}\right)\right)^s \cdot \mathbb{P}(M_1^* = m' \mid M_0^* = m) \qquad \text{by the Markov property} \\
&= \sum_{s=0}^{\infty} \left((1-p)\left(\sum_{j=1}^{m-1} q(j) + \frac{q(m)}{2}\right)\right)^s \cdot p \cdot \frac{m'}{m} \cdot \frac{q(m')q(m-m')}{q(m)} \qquad \text{by (6.6)} \\
&= \frac{1}{1 + \frac{(1-p)}{p}\left(Q(m) - \frac{q(m)}{2}\right)} \cdot \frac{m'}{m} \cdot \frac{q(m')q(m-m')}{q(m)}.
\end{aligned}
\tag{6.14}
$$

Moreover $\pi(1;0) = 1$ and the chain is killed at 0.

Important discussion. It can also be useful to write the transition probabilities in the following way:

$$
\begin{aligned}
\pi(m\,; m-w) &= \mathbb{P}(M_t - M_{t+1} = w \mid M_t = m) \\
&= \frac{1}{1 + \frac{(1-p)}{p}\left(Q(m) - \frac{q(m)}{2}\right)} \cdot \frac{m-w}{m} \cdot \frac{q(m-w)}{q(m)} \cdot q(w),
\end{aligned}
\tag{6.15}
$$

where it is reasonable to expect that all the factors except $q(w)$ are close to 1 when w is fixed and m is large.[13] This suggests comparing the jumps of M with those of a decreasing random walk with step distribution $\mathcal{Q}$ given by $(q(i))_{i \geq 1}$. Note from the rough regularity estimates in Proposition 5.1 that $q(i) = i^{-\gamma + o(1)}$ as $i \to \infty$, with $\gamma \in (3/2, 2)$ as in (2.7). In particular, such a random walk has jumps with infinite mean.

Remark 6.3. The decreasing process $(M_t)_{0 \leq t \leq \mathtt{T}}$ was defined in (6.9) as a time-change of $(M_h^*)_{0 \leq h \leq \eta}$, which in turn is the third coordinate of the process $(Z_h^k)_{0 \leq h \leq \eta}$ defined in (6.2). We will use the notation $\mathbb{P}_k$ and $\mathbb{E}_k$ to denote the corresponding probability measures and expectations of the process Z^k. In particular, by construction, under $\mathbb{P}_k$, the processes $(M_t)_{0 \leq t \leq \mathtt{T}}$ and $(M_h^*)_{0 \leq h \leq \eta}$ are both started at k because $Z_0^k = (S_0, D_0, M_0^*, W_0) = (\ominus, \leftarrow, k, 0)$, by convention.

Remark 6.4. It can be useful to think of the set of values taken by the Markov chain M under the measure $\mathbb{P}_k$ as a point process $\mathscr{P}^k = \{M_t,\ t \geq 0\}$ on $\mathbb{N}$. Indeed, Theorem 6.2 restricted to M may be recast as the (vague) convergence of $\mathscr{P}^k$, as $k \to \infty$, to a limiting point process $\mathscr{P}$. A nice feature of $\mathscr{P}^k$ is that conditioning on $\mathscr{P}^k \cap [m, \infty)$ allows for more flexibility than conditioning on the first steps of M, as it allows to condition on partial information about one of the jumps. In particular, this "spatial" Markov property explains why we can expect independence across scales, even though one single step of M can in theory jump over several scales (see the "spatial" filtration introduced in (6.29)). However, we will not explicitly use this point process representation.

---

[13]Equation (5.2) in Proposition 5.1 proves that $\frac{q(m-w)}{q(m)}$ is bounded from above and below by positive constants, but not that it is close to 1.

6.3 The coupling argument

Let M and M' be two versions of the decreasing Markov chain introduced in (6.9), started from two large values k and k'. The goal of this section is to construct a coupled version of those chains (see Definition 6.5) such that, with high probability as k and k' tend to infinity, the last finite number of steps of the two chains before extinction are the same (Proposition 6.6). We will then quickly explain how to upgrade this coupling to the full Markov chain Z^k (Corollary 6.7), and how to deduce Theorem 6.2. The bulk of this section is devoted to the proof of Proposition 6.6, which will be completed in Section 6.3.5, building on the coupling estimates of Section 6.3.4.

6.3.1 Introducing the coupling: the black and golden instructions

We assume for the moment that there exists an infinite deterministic set of good scales, denoted by GoodScales $\subset \mathbb{N}$, that satisfy certain regularity conditions; see Definition 6.10 for a precise definition and Lemma 6.9 for the existence of such good scales. In particular, these good scales are *well-separated*, in the sense that if $\ell, \ell' \in$ GoodScales with $\ell' > \ell$ then $\ell' > 7\ell$.

Then in Definition 6.5, we construct a coupling $\mathbb{P}_{k,k'}$ of the two Markov chains M and M' in (6.9), with marginals $\mathbb{P}_k$ and $\mathbb{P}_{k'}$ respectively, by using a set of rules for determining what step the chains M and M' take at every site of $\mathbb{N}$. The sort of coupling that we have in mind is what Aldous and Thorisson coined as *shift-couplings* [AT93], since the two chains will *a priori* merge at different times in their own evolutions.

First, we introduce the **black instructions** $(B_m)_{m \geq 1}$ as independent random variables defined so that

$$\forall m' < m, \qquad \mathbb{P}(B_m = m') = \pi(m; m'), \tag{6.16}$$

where $\pi(m; m')$ is as in (6.14). Then, for every good scale $\ell \in$ GoodScales, we introduce the **golden instructions** $(G_m)_{5\ell \leq m \leq 7\ell}$ that are coupled together for a given good scale $\ell \in$ GoodScales, but independent between different good scales and independent of the black instructions $(B_m)_{m \geq 1}$, with marginal law

$$\forall m' \in [\![\ell, 4\ell]\!], \quad \mathbb{P}(G_m = m') = \mathbb{P}(B_m = m' \mid B_m \in [\![\ell, 4\ell]\!]). \tag{6.17}$$

The coupling of $(G_m)_{5\ell \leq m \leq 7\ell}$ will be chosen so that, with uniformly positive probability over all $\ell \in$ GoodScales, we have

$$\forall a, a' \in [\![5\ell, 7\ell]\!], \qquad G_a = G_{a'}.$$

The existence of such coupling of the golden instructions will be shown later in Item 3 of Lemma 6.14.

Definition 6.5 (The black-golden coupling). We say that the decreasing chains M and M' have the **black-golden coupling** when the dynamics of M (and M') are described for all $t \geq 0$ as follows:[14]

$$M_{t+1} = \begin{cases} G_{M_t} & \text{if } M_t \in [\![5\ell, 7\ell]\!] \text{ for some good scale } \ell \text{ and } B_{M_t} \in [\![\ell, 4\ell]\!]; \\ B_{M_t} & \text{otherwise.} \end{cases} \tag{6.18}$$

Moreover, when two chains M and M' are started from k and k' respectively, we denote their joint law under this coupling by $\mathbb{P}_{k,k'}$.

---

[14]We stress that here we are using a *unique* set of black and golden instructions for *both* chains M and M'.

In more intuitive terms, the chains M and M' typically follow the black instructions, except in a specific situation: when a chain is instructed to jump – via a black instruction – from a position in $[\![5\ell, 7\ell]\!]$ to somewhere inside $[\![\ell, 4\ell]\!]$, it then follows the corresponding golden instruction instead. It is clear from this definition that the marginals M and M' under $\mathbb{P}_{k,k'}$ are indeed Markov chains with transition probabilities π as in (6.14).

For all $m \in \mathbb{N}$, the black instruction B_m should be thought of as the instruction to take a *small step* from m to B_m, while for all $m \in [\![5\ell, 7\ell]\!]$ the golden instruction G_m should be thought of as the instruction to take a *large step* from $m \geq 5\ell$ to $G_m \in [\![\ell, 4\ell]\!]$. The reason why we consider golden instructions pointing towards $[\![\ell, 4\ell]\!]$ rather than $[\![0, 4\ell]\!]$ is that we will run a multiscale argument that requires the chains not to jump across several good scales at once. Furthermore, the reason for the terminology for colors is that when taking a golden instruction, the chains will merge with positive probability and then stick together,[15] ending up in a win, as precisely described in the next result.

We will write V_{mer} for the merging value of M and M', *i.e.* the largest common value taken by the chains M and M'. Let also $\mathsf{t}_{\mathrm{mer}}$ (resp. $\mathsf{t}'_{\mathrm{mer}}$) be the time when M (resp. M') visits V_{mer}. Note that by definition of the black-golden coupling,

$$(M_{\mathsf{t}_{\mathrm{mer}}+s})_{s \geq 0} = (M'_{\mathsf{t}'_{\mathrm{mer}}+s})_{s \geq 0}.$$

Finally, for $k \geq A \geq 1$, we denote by $s(A, k)$ the number of good scales $\ell \in [\![A, k]\!]$ such that $7\ell < k$.

Proposition 6.6 (Fast merging of M and M'). *Let M and M' be defined as the decreasing chains started from k and k', respectively, under the black-golden coupling $\mathbb{P}_{k,k'}$ of Definition 6.5. Then, the following assertions hold:*

1. *There exists a constant $c \in (0, 1)$ such that for all $k' \geq k \geq 1$ and all $A \geq 1$,*

$$\mathbb{P}_{k,k'}(V_{\mathrm{mer}} < A) \leq (1 - c)^{s(A,k)}.$$

2. *Furthermore, the total times $\mathsf{T} - \mathsf{t}_{\mathrm{mer}}$ and $\mathsf{T}' - \mathsf{t}'_{\mathrm{mer}}$, representing the durations between merging and dying, tend to infinity in probability as $k, k' \to \infty$.*

The proof of Proposition 6.6 will occupy most of the remainder of Section 6. Once this is proved, we easily obtain a similar result for the whole Markov chain Z^k. We note that with the notation of (6.8), the time where M^* (resp. $(M^*)'$) first hits V_{mer} is given by $\eta_{\mathsf{t}_{\mathrm{mer}}}$ (resp. $\eta'_{\mathsf{t}'_{\mathrm{mer}}}$). We also recall that $\eta = \eta_{\mathsf{T}}$ (resp. $\eta' = \eta'_{\mathsf{T}'}$) is the hitting time of 0 by M^* (resp. $(M^*)'$).

Corollary 6.7. *For any $k' \geq k \geq 1$, there exists a coupling $\widetilde{\mathbb{P}}_{k,k'}$ of the full Markov chains Z^k and $Z^{k'}$ started from k and k' respectively, under which:*

- *The pair (M, M') follows the black-golden coupling $\mathbb{P}_{k,k'}$ of Definition 6.5.*

- *For any $h \geq 1$, we have $Z^k_{\eta_{\mathsf{t}_{\mathrm{mer}}}+h} = Z^{k'}_{\eta'_{\mathsf{t}'_{\mathrm{mer}}}+h}$.*

In particular:

1. *For all $A \geq 1$, we have*

$$\widetilde{\mathbb{P}}_{k,k'}(V_{\mathrm{mer}} < A) \leq (1 - c)^{s(A,k)}.$$

---

[15]We emphasize that this is not the only mechanism by which the two chains can merge. In fact, it is totally possible for them to merge through two black instructions that happen to land at the same value, causing the chains to stick together from that point onward. However, since our aim is to lower bound the probability that the two chains will merge fast enough, it suffices for us to focus on the case involving the golden instructions.

2. For all $A \geq 1$, we have
$$\widetilde{\mathbb{P}}_{k,k'}(V_{\mathrm{mer}} \geq A) \to 1 \quad as\ k, k' \to \infty.$$

3. The durations $\eta - \eta_{\mathsf{t}_{\mathrm{mer}}}$ and $\eta' - \eta'_{\mathsf{t}'_{\mathrm{mer}}}$ go to infinity in probability as $k, k' \to \infty$.

Proof of Corollary 6.7 given Proposition 6.6. It is explained in Section 6.2.2 how to reconstruct Z^k from M using independent decorations η, S, D, W, and similarly for $Z^{k'}$. Since M and M' merge at respective times $\mathsf{t}_{\mathrm{mer}}$ and $\mathsf{t}'_{\mathrm{mer}}$, we can choose these decorations such that for all $t \geq 0$ and all $h \geq 1$,

$$\eta_{\mathsf{t}_{\mathrm{mer}}+t+1} - \eta_{\mathsf{t}_{\mathrm{mer}}+t} = \eta'_{\mathsf{t}'_{\mathrm{mer}}+t+1} - \eta'_{\mathsf{t}'_{\mathrm{mer}}+t} \quad \text{and} \quad (S, D, W)_{\eta_{\mathsf{t}_{\mathrm{mer}}}+h} = (S', D', W')_{\eta'_{\mathsf{t}'_{\mathrm{mer}}}+h}.$$

This implies that Z^k after time $\eta_{\mathsf{t}_{\mathrm{mer}}} + 1$ and $Z^{k'}$ after time $\eta'_{\mathsf{t}'_{\mathrm{mer}}} + 1$ coincide, which shows the existence of $\widetilde{\mathbb{P}}_{k,k'}$. Item 1 then follows immediately from Item 1 of Proposition 6.6, and implies Item 2 because our set of good scales is infinite. Finally, Item 3 follows from Item 2 of Proposition 6.6 since $\eta - \eta_{\mathsf{t}_{\mathrm{mer}}} \geq \mathsf{T} - \mathsf{t}_{\mathrm{mer}}$ and $\eta' - \eta'_{\mathsf{t}'_{\mathrm{mer}}} \geq \mathsf{T}' - \mathsf{t}'_{\mathrm{mer}}$. $\qquad \square$

Proof of Theorem 6.2 given Corollary 6.7. We need to prove that there exists a process $(\overleftarrow{Z}_t)_{t \geq 0}$ such that, for all $r \geq 1$,
$$d_{\mathrm{TV}}\left((Z^k_{\eta-t})_{t \leq r}, (\overleftarrow{Z}_t)_{t \leq r}\right) \to 0 \quad as\ k \to \infty, \tag{6.19}$$

where we take the convention that $Z^k_{\eta-t} = (\ominus, \leftarrow, k, 0)$ when $t > \eta$. We prove (6.19) by Cauchy characterization. Consider the two decreasing Markov chains Z^k and $Z^{k'}$ started at k and k' respectively and having the coupling $\widetilde{\mathbb{P}}_{k,k'}$ of Corollary 6.7. By standard coupling techniques (see e.g. [LP17, Proposition 4.7]), we can write

$$d_{\mathrm{TV}}\left((Z^k_{\eta-t})_{t \leq r}, (Z^{k'}_{\eta'-t})_{t \leq r}\right) \leq \widetilde{\mathbb{P}}_{k,k'}\left(Z^k_{\eta-t} \neq Z^{k'}_{\eta'-t} \text{ for some } t \leq r\right) \leq \widetilde{\mathbb{P}}_{k,k'}\left(\eta - \eta_{\mathsf{t}_{\mathrm{mer}}} \leq r\right)$$

since, under the coupling, the two chains Z^k and $Z^{k'}$ stick together after $\eta_{\mathsf{t}_{\mathrm{mer}}}$ and $\eta'_{\mathsf{t}'_{\mathrm{mer}}}$. Then, Item 3 of Corollary 6.7 guarantees that the probability on the right-hand side of the last display goes to 0 as $k, k' \to \infty$, which proves the desired Cauchy characterization.

We conclude that for any fixed r, the random vector $\left(Z^k_{\eta-t}\right)_{t \leq r}$ started from k has a limit as $k \to \infty$ that we write $(\overleftarrow{Z}_t)_{t \leq r}$. Finally, an application of Kolmogorov's extension theorem shows that the limit defines a valid stochastic process $\overleftarrow{Z}$. $\qquad \square$

All we have left to do to prove Theorem 6.2 is therefore to prove Proposition 6.6, which will be completed in Section 6.3.5. As it forms the most technical part of the paper, the reader may on a first read directly jump to Section 7.

The rest of this section is organized as follows: in Section 6.3.2 we will give an informal description of the strategy for the proof of Proposition 6.6. Then, Section 6.3.3 is devoted to the description of the good scales, Section 6.3.4 to the proof of the main technical estimates, and finally Section 6.3.5 to the conclusion of the proof of Proposition 6.6.

6.3.2 Description of the strategy for the coupling argument

We now describe the strategy for the proof of Item 1 of Proposition 6.6; the second item will follow as a simple consequence. We will show that there exists a constant $c \in (0, 1)$ such that:

1. From any starting point k, for each good scale ℓ such that $7\ell \leq k$, the chain M visits $[\![6\ell, 7\ell]\!]$ with probability at least c (this will be a consequence of Item 1 in Lemma 6.14).

2. If M is ever in $[\![6\ell, 7\ell]\!]$, it will try to jump below 4ℓ from somewhere in $[\![5\ell, 7\ell]\!]$ with probability at least c, meaning it will use a golden instruction with probability at least c (this is a consequence of Item 2 in Lemma 6.14). Moreover, conditional on this, the same facts hold for M' with probability at least $c/2$ (this is a consequence of Lemma 6.15).

3. For all good scales ℓ, we can couple the $(G_m)_{5\ell \leq m \leq 7\ell}$ so that with probability at least c they are all equal (this is Item 3 in Lemma 6.14).

4. Since, for any good scale, Item 1 in Lemma 6.14 holds conditionally on all the information above 7ℓ, we can iterate the argument at each scale.

We recall from the discussion around (6.15) that M can be compared to a decreasing random walk with jump distribution $(q(i))_{i \geq 1}$, with $q(i) = i^{-\gamma + o(1)}$ as $i \to \infty$, and $\gamma \in (3/2, 2)$ as in (2.7). Informally speaking, the *good scales* that we need in Steps 1, 2 and 3 are scales where the behavior of this walk is reasonably close to that of a stable subordinator.[16]

If all the above is true (and using the fact that M and M' are essentially independent up until merging), there is a uniformly positive probability at every good scale that both M and M' follow two golden instructions landing at the same height. Since we can iterate this attempt to land at the same height by successive conditionings above each good scale (thanks to Step 4 above), with probability tending to one as the starting points approach infinity, the two chains will eventually follow two golden instructions landing at the same height. From this, we can conclude by construction of the black-golden coupling that the future trajectories of M and M' will stick together beyond this point, completing the proof of Item 1 of Proposition 6.6.

6.3.3 Definition and existence of good scales and additional preliminary estimates

In this whole section, the value of the constants c and C may change from line to line. Recall that, in Section 5, we proved the rough regularity estimates for the sequences q and Q stated in Proposition 5.1, which have the following immediate consequences.

Corollary 6.8 (Bounds on q and Q). *Fix $r, R > 0$. Then there exist constants $c, C > 0$ such that:*

1. For all $k \geq 1$,

$$c \leq \frac{kq(k)}{Q(k)} \leq C. \tag{6.20}$$

2. For all $1 \leq k' \leq k$,

$$\frac{q(k')}{q(k)} \geq c. \tag{6.21}$$

---

[16]The convergence of such a walk to a $(\gamma - 1)$-stable subordinator would be immediate given Theorem 7.1, but unfortunately proving Theorem 7.1 is precisely our long-term goal.

3. For all $k, k' \geq 1$ with $r \leq \frac{k'}{k} \leq R$,

$$c \leq \frac{q(k')}{q(k)} \leq C \qquad and \qquad c \leq \frac{Q(k')}{Q(k)} \leq C. \tag{6.22}$$

Proof. The bounds in (6.20) are exactly the ones in Equation (5.9) in Lemma 5.3. The upper bound in (5.2) of Proposition 5.1 gives that $\frac{q(k)}{q(k')} \leq C$ for all $1 \leq k' \leq k$, which immediately implies (6.21). Finally, Proposition 5.1 ensures that as long as the ratio $\frac{k'}{k}$ is bounded above and below by some fixed constants, so are the ratios $\frac{q(k')}{q(k)}$ and $\frac{Q(k')}{Q(k)}$. This gives the bounds in (6.22). □

In order to show the existence of large jumps, we will need an additional regularity result holding (at least) for an infinite number of integers ℓ.

Lemma 6.9 (Good scale condition)**.** *There exists a constant $C > 0$ such that there are infinitely many integers $\ell \geq 1$ satisfying the inequality*

$$\sum_{i=1}^{\ell-1} iq(i) \leq C\ell Q(\ell). \tag{6.23}$$

Proof. We prove that there exists a constant $c > 0$ such that for infinitely many values of ℓ we have

$$\ell Q(\ell) \geq c \sum_{i=1}^{\ell-1} Q(i). \tag{6.24}$$

By the bounds in (6.20), this would imply our claim. For all $\ell \geq 1$, we write $\mathbf{Q}(\ell) = \sum_{i=1}^{\ell-1} Q(i)$. By Proposition 5.1, we know that $Q(i) = i^{-\frac{1}{2\alpha}+o(1)}$ where $\alpha \in (1/2, 1)$. This entails that

$$\mathbf{Q}(\ell) = \ell^{1-\frac{1}{2\alpha}+o(1)}. \tag{6.25}$$

Now, assume for a contradiction that for any choice of the constant $c > 0$, (6.24) fails for all large enough ℓ. Then, for all large enough ℓ, we have $\mathbf{Q}(\ell+1) - \mathbf{Q}(\ell) = Q(\ell) < \frac{c}{\ell}\mathbf{Q}(\ell)$, so that $\mathbf{Q}(\ell+1) < (1 + \frac{c}{\ell})\mathbf{Q}(\ell)$. Hence, for all ℓ large enough,

$$\mathbf{Q}(\ell) \leq \ell^{c+o(1)}.$$

We get a contradiction with (6.25) if $c < 1 - \frac{1}{2\alpha}$. Since $1 - \frac{1}{2\alpha} > 0$ as $\alpha \in (1/2, 1)$, we conclude that there exists a constant $c > 0$ such that (6.24) holds for infinitely many values of ℓ. □

We can now introduce the good scales that we mentioned at the beginning of Section 6.3.1.

Definition 6.10 (Good scales)**.** Fix $C > 0$ as in Lemma 6.9. We say that an infinite set GoodScales $\subset \mathbb{N}$ is a collection of **good scales** if the following two conditions hold:

1. Every $\ell \in$ GoodScales satisfies the condition $\sum_{i=1}^{\ell-1} iq(i) \leq C\ell Q(\ell)$.

2. For all $\ell, \ell' \in$ GoodScales with $\ell' > \ell$, we have that $\ell' > 7\ell$.

If $\ell \in$ GoodScales, we say that the interval $[\![\ell, 7\ell]\!]$ is a **good interval**.

Lemma 6.9 guarantees the existence of these collections of good scales and from now on we fix GoodScales to be one arbitrary, but fixed, such collection. Note that good intervals are disjoint by the second condition in the definition.

Remark 6.11. Note that, assuming the good regularity estimates in Theorem 7.1, there exists $C > 0$ such that any integer $\ell \geq 1$ satisfies the good scale condition in Item 1 of Definition 6.10. Indeed, if $q(i) = i^{-\frac{1}{2\alpha}-1}\varphi(i)$ as in Theorem 7.1, then by Karamata's Theorem (see [BGT89, Theorem 1.5.11]), we have $\sum_{i=1}^{\ell-1} iq(i) \sim \frac{\ell^2 q(\ell)}{1-\frac{1}{2\alpha}}$ as $\ell \to \infty$, and $\ell^2 q(\ell) \leq C\ell Q(\ell)$ by the bounds in (6.20). Therefore, provided we knew Theorem 7.1, we could take GoodScales to be any collection of well-separated integers. This will be important later for the scaling limit result in Section 9.

The bounds in Corollary 6.8 and Lemma 6.9 help us control the transition probabilities $\pi(m; m')$ from (6.14) of the decreasing Markov chain M, as shown in the next lemma.

Lemma 6.12 (Bounds for the transition probabilities of M)**.** *Fix two constants $c_1, c_2 > 0$ with $c_1 < c_2$. Then there exist $c > 0$ and $C > 0$ such that the following is true:*

1. *If $c_1 < 1$, then for all $j \geq 1$,*

$$\forall i < j, \quad \pi(j; i) \geq c\frac{i}{j}q(i) \qquad and \qquad \forall i \leq c_1 j, \quad \pi(j; i) \leq C\frac{i}{j}q(i).$$

2. *For all $\ell \geq 1$, and all $j \geq 1$ and $x, y \geq 1$ such that $c_1 \ell \leq x \leq y \leq c_2 \ell < j$,*

$$\mathbb{P}(M_{t+1} \in [\![x, y]\!] \mid M_t = j) \geq c\frac{y-x}{j}Q(\ell).$$

3. *If $c_1 < 1$, then for all $\ell \in \mathrm{GoodScales}$, and all $j \geq 1$ and $y \geq 1$ such that $y \leq (c_1 j) \wedge (c_2 \ell)$,*

$$\mathbb{P}(M_{t+1} \in [\![1, y]\!] \mid M_t = j) \leq C\frac{\ell}{j}Q(\ell).$$

4. *If $c_1 > 1$, then for all $\ell \in \mathrm{GoodScales}$ and $j \in [\![c_1 \ell, c_2 \ell]\!]$, writing $\Delta M_1 = M_0 - M_1$,*

$$\mathbb{E}_j\big[\Delta M_1 \mathbb{1}_{\Delta M_1 \leq \ell}\big] \leq C\ell Q(\ell).$$

Proof. The bounds in Item 1 follow from Corollary 6.8. Indeed, by (6.14), we have

$$\forall i < j, \quad \pi(j; i) = \frac{1}{1 + \frac{(1-p)}{p}(Q(j) - \frac{q(j)}{2})} \cdot \frac{i}{j} \cdot \frac{q(i)q(j-i)}{q(j)}. \tag{6.26}$$

Now for all $j \geq 1$, the ratio $\frac{1}{1+\frac{(1-p)}{p}(Q(j)-\frac{q(j)}{2})}$ is bounded above and below by constants. Besides, on the one hand, by (6.21),

$$\forall i \leq j, \qquad \frac{q(j-i)}{q(j)} \geq c.$$

On the other hand, by the first item of (6.22), if $c_1 < 1$,

$$\forall i \leq c_1 j, \qquad \frac{q(j-i)}{q(j)} \leq C.$$

since $1 - c_1 \leq 1 - \frac{i}{j} = \frac{j-i}{j} \leq 1$. The bounds in Item 1 immediately follow.

We now move on to Item 2. We recall from Section 3.1 the notation for sums between real numbers, which will be extensively used in the rest of the proof. Let $\ell \geq 1$, and $j \geq 1$ and $x, y \geq 1$,

such that $c_1\ell \leq x \leq y \leq c_2\ell < j$. Allowing the value of the constant c to change from line to line, we get by Item 1 that

$$\mathbb{P}(M_{t+1} \in [\![x\,,y]\!] \mid M_t = j) = \sum_{i=x}^{y} \pi(j;i) \geq c \sum_{i=x}^{y} \frac{i}{j} \cdot q(i)$$

$$\geq c \cdot \frac{\ell}{j} \sum_{i=x}^{y} q(i) \geq c\frac{\ell}{j}(y-x)q(\ell), \qquad (6.27)$$

where in the last inequality we used the first bound in (6.22), which holds since we are summing over $i \in [\![x\,,y]\!] \subseteq [\![c_1\ell\,,c_2\ell]\!]$. By the bounds in (6.20), this establishes Item 2.

We now assume that $c_1 < 1$ and prove Item 3. Fix $\ell \in$ GoodScales and consider $j \geq 1$ and $y \geq 1$ such that $y \leq (c_1 j) \wedge (c_2\ell)$. Using the fact that we sum over $i \leq y \leq c_1 j$, together with the second bound in Item 1, we get

$$\mathbb{P}(M_{t+1} \in [\![1\,,y]\!] \mid M_t = j) = \sum_{i=1}^{y} \pi(j;i) \leq C \sum_{i=1}^{y} \frac{i}{j} q(i) \leq \frac{C}{j} \left(\sum_{i=1}^{\ell-1} iq(i) + \sum_{i=\ell}^{y} iq(i) \right).$$

Since $y \leq c_2\ell$, the latter sum can be handled again using first the bounds in (6.22) and then in (6.20), yielding

$$\sum_{i=\ell}^{y} iq(i) \leq C\ell^2 q(\ell) \leq C\ell Q(\ell).$$

The first one is also bounded by $C\ell Q(\ell)$ by the good scale condition of Definition 6.10, which proves Item 3.

It only remains to prove Item 4 for which we assume that $c_1 > 1$. Taking $\ell \in$ GoodScales and $j \in [\![c_1\ell\,,c_2\ell]\!]$, we have

$$\mathbb{E}_j\big[\Delta M_1 \mathbb{1}_{\Delta M_1 \leq \ell}\big] = \sum_{i=j-\ell}^{j-1} (j-i)\pi(j;i).$$

Any i in this sum satisfies $(c_1 - 1)\ell \leq i < j \leq c_2\ell$ (with $c_1 - 1 > 0$). Therefore, we have

$$\pi(j;i) \overset{(6.26)}{\leq} C\frac{i}{j} \cdot \frac{q(i)q(j-i)}{q(j)} \overset{(6.22)}{\leq} Cq(j-i).$$

Recalling that ℓ is a good scale, we finally get by the good scale condition (Definition 6.10):

$$\mathbb{E}_j\big[\Delta M_1 \mathbb{1}_{\Delta M_1 \leq \ell}\big] \leq C \sum_{i=j-\ell}^{j-1} (j-i)q(j-i) = C \sum_{i=1}^{\ell} iq(i) \leq C\ell Q(\ell).$$

This concludes the proof of Item 4 and therefore establishes Lemma 6.12. $\qquad\qquad\square$

6.3.4 Statements and proofs of the three main ingredients for the coupling

In this section, we consider two chains M and M' under the black-golden coupling (see Definition 6.5). Recall the notation $\mathbb{P}_{k,k'}$ for the probability measure of this coupling under which M starts at k and M' at k'. Whenever we work with events that only depend on M, we use $\mathbb{P}_k$ for shorter expressions.

We write H_i for the first hitting time of $[\![0\,,i]\!]$ by the chain M, that is

$$H_i := \inf\{t \geq 1 : M_t \in [\![0\,,i]\!]\}. \qquad (6.28)$$

We define H_i' similarly for M'. We will often write $M(H_i)$ instead of M_{H_i} to avoid notational clutter. Note that for any $h \in [\![0\,,k]\!]$ we have

$$M(H_h) = \sup(\{M_t,\ t \geq 0\} \cap [0\,,h]),$$
$$M(H_h - 1) = \inf(\{M_t,\ t \geq 0\} \cap [h+1\,,\infty)).$$

Define for all $\ell \in \mathrm{GoodScales}$ the σ-algebra (*cf.* Remark 6.4)

$$\mathcal{G}_\ell := \sigma\left(G_m : m \in \bigcup_{\ell' > \ell} [\![5\ell'\,,7\ell']\!]\right) \vee \sigma\left([7\ell+1\,,+\infty) \cap \{M_t, t \geq 0\}, [7\ell+1\,,+\infty) \cap \{M_t', t \geq 0\}\right).$$

(6.29)

This σ-algebra records the values of the golden instructions for all golden intervals strictly above 7ℓ, as well as the trajectories of M and M' from their starting point up to the last time that they take a value strictly larger than 7ℓ, *i.e.* $(M_t)_{0 \leq t \leq H_{7\ell}-1}$ and $(M_t')_{0 \leq t \leq H_{7\ell}'-1}$.

Lemma 6.13 (Spatial Markov property at level 7ℓ). *Fix $\ell \in \mathrm{GoodScales}$ and $k, k' \geq 7\ell + 1$. Let M and M' be defined under the black-golden coupling $\mathbb{P}_{k,k'}$ of Definition 6.5. Then the following assertions are true:*

1. *For $m \in [\![7\ell+1\,,k]\!]$ and $m' \in [\![7\ell+1\,,k']\!]$, the $\mathcal{G}_\ell$-conditional distribution of $(M(H_{7\ell}), M'(H_{7\ell}'))$ on the event $\{M(H_{7\ell} - 1) = m, M'(H_{7\ell}' - 1) = m'\}$ is given by*

$$\mathrm{Law}\big((M(H_{7\ell}), M'(H_{7\ell}')) \mid \mathcal{G}_\ell\big) = \mathrm{Law}((B_m, B_{m'}) \mid B_m \leq 7\ell, B_m' \leq 7\ell).$$ (6.30)

2. *Conditionally on $\sigma(\mathcal{G}_\ell, (M(H_{7\ell}), M'(H_{7\ell}'))$, the chains $(M(H_{7\ell} + t))_{t \geq 0}$ and $(M'(H_{7\ell}' + t))_{t \geq 0}$ have the distribution of M and M' under the black-golden coupling $\mathbb{P}_{M(H_{7\ell}), M'(H_{7\ell}')}$.*

Proof. We first prove Item 1. Fix any $m, m' \geq 7\ell + 1$ and let us reason on the event $\{M(H_{7\ell} - 1) = m, M'(H_{7\ell}' - 1) = m'\}$. We know by definition of $H_{7\ell}$ and $H_{7\ell}'$ that $M(H_{7\ell}) \leq 7\ell$ and $M'(H_{7\ell}') \leq 7\ell$. From the definition of the black-golden coupling, it is clear that we have $M(H_{7\ell}) = B_m$, because a golden instruction G_m (if taken at all) would land in some interval $[\![\ell'\,,4\ell']\!]$ with $\ell' > 7\ell$, which would contradict the fact that $M(H_{7\ell}) \leq 7\ell$. We similarly have $M'(H_{7\ell}') = B_{m'}'$. Now we have

$$\mathrm{Law}((M(H_{7\ell}), M'(H_{7\ell}')) \mid \mathcal{G}_\ell) = \mathrm{Law}((B_m, B_{m'}) \mid \mathcal{G}_\ell) = \mathrm{Law}((B_m, B_{m'}) \mid B_m \leq 7\ell, B_m' \leq 7\ell),$$

which proves Item 1.

We now turn to Item 2. Just note that by definition of the black-golden coupling, the future trajectories $(M(H_{7\ell} + t))_{t \geq 0}$ and $(M'(H_{7\ell}' + t))_{t \geq 0}$ are entirely determined by the positions $(M(H_{7\ell}), M'(H_{7\ell}'))$ and the values of the black and golden instructions below 7ℓ, namely $(B_m)_{1 \leq m \leq 7\ell}$ and $(G_m)_{5\ell' \leq m \leq 7\ell', \ell' \leq \ell}$. Note also that these random variables are independent of $\mathcal{G}_\ell, M(H_{7\ell}), M'(H_{7\ell}')$, because those are determined from the value of the black and golden instructions *above* $7\ell + 1$. This ensures that Item 2 holds. $\square$

The next lemma presents the three main ingredients for the proof of Proposition 6.6, following the strategy outlined in Section 6.3.2. Then, in Lemma 6.15, we will extend the second item to incorporate the coupling with M'.

Lemma 6.14. *There exists a constant $c \in (0,1)$ such that the following three properties hold:*

1. *For any $\ell \in \mathrm{GoodScales}$ and $m > 7\ell$,*

$$\mathbb{P}(B_m \in [\![6\ell\,,7\ell]\!] \mid B_m \leq 7\ell) \geq c.$$

53

2. *For any $\ell \in$ GoodScales and $k \in [\![6\ell, 7\ell]\!]$,*

$$\mathbb{P}_k(M(H_{4\ell} - 1) \in [\![5\ell, 7\ell]\!], \ M(H_{4\ell}) \in [\![\ell, 4\ell]\!]) \geq c.$$

3. *For any $\ell \in$ GoodScales, we can couple the golden instructions $(G_m)_{5\ell \leq m \leq 7\ell}$, with law given by* (6.17), *in such a way that*

$$\mathbb{P}(\forall a, a' \in [\![5\ell, 7\ell]\!], \ G_a = G_{a'}) \geq c.$$

See Figure 11 for a schematic representation of the quantities involved in Lemma 6.14 and some further explanations on the role of the three items.
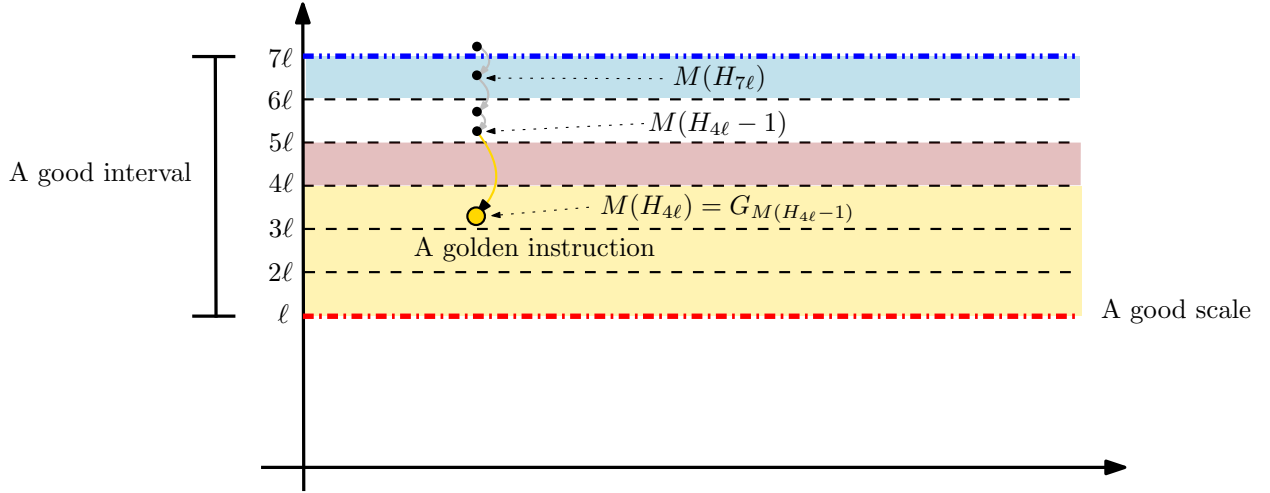

Figure 11: A schematic representation for the statement of Lemma 6.14. A good scale ℓ is shown as a dashed-dotted red line and the corresponding value of 7ℓ is shown as a dashed-dotted blue line. The good interval $[\![\ell, 7\ell]\!]$ is divided into six sub-intervals by the dashed black lines, with the sub-intervals $[\![6\ell, 7\ell]\!]$, $[\![4\ell, 5\ell]\!]$ and $[\![\ell, 4\ell]\!]$ highlighted in blue, brown and gold, respectively. Items 1 and 2 ensure that there is a uniformly positive probability among all good scales ℓ that M enters the blue sub-interval and then (possibly after a few steps) jumps across the brown sub-interval; this trajectory is schematically represented by the black dots. The jump across the brown good sub-interval corresponds to a golden instruction landing at the golden dot at height $G_{M(H_{4\ell}-1)}$ inside the golden interval $[\![\ell, 4\ell]\!]$. Since Item 3 allows us to couple all the golden instructions in such a way that they land at *the same* place with positive probability, this (with the additional help of Lemma 6.15) will allow us to argue that there is a uniformly positive probability over all good scales ℓ that M and M' meet. Since Item 1 guarantees that this actually holds conditionally on the scales above, we will be able to conclude that M and M' meet with probability tending to one as their starting points tend to infinity.

We prove the three items of Lemma 6.14 sequentially.

Proof of Item 1 of Lemma 6.14. Let ℓ be a good scale and $m > 7\ell$. We write

$$\mathbb{P}(B_m \in [\![6\ell, 7\ell]\!] \mid B_m \leq 7\ell) = \frac{\mathbb{P}(B_m \in [\![6\ell, 7\ell]\!])}{\mathbb{P}(B_m \leq 7\ell)} = \frac{\mathbb{P}(M_{t+1} \in [\![6\ell, 7\ell]\!] \mid M_t = m)}{\mathbb{P}(M_{t+1} \in [\![1, 7\ell]\!] \mid M_t = m)}, \qquad (6.31)$$

for any $t \geq 0$. Now by Item 2 in Lemma 6.12 (with $c_1 = 6$, $c_2 = 7$, $x = 6\ell$ and $y = 7\ell$) there is a constant $c > 0$ (independent of ℓ) such that for all $m > 7\ell$,

$$\mathbb{P}(M_{t+1} \in [\![6\ell, 7\ell]\!] \mid M_t = m) \geq c \frac{\ell}{m} Q(\ell).$$

Likewise, by Item 3 in Lemma 6.12 (with $c_1 = 6/7$, $c_2 = 6$, and $y = 6\ell - 1$), there is a constant $C > 0$ (independent of ℓ) such that for all $m > 7\ell$,

$$\mathbb{P}(M_{t+1} \in [\![1, 6\ell - 1]\!] \mid M_t = m) \leq C \frac{\ell}{m} Q(\ell).$$

Combining the previous two displays, we find that for all $m > 7\ell$,

$$\mathbb{P}(B_m \in [\![6\ell, 7\ell]\!] \mid B_m \leq 7\ell) \geq \frac{c}{C + c}.$$

This proves Item 1 of Lemma 6.14. $\qquad\square$

Proof of Item 2 of Lemma 6.14. We divide the proof into five steps.

<u>*Step 1: Introducing a good event E_ℓ.*</u> For $a > 0$ and $\ell \geq 1$, let $b = b(\ell) := \left\lceil \frac{a}{Q(\ell)} \right\rceil$, where $Q(\cdot)$ is as in (2.5). Define the following event E_ℓ:

  (i) M takes at most b jumps smaller than ℓ, and then a big jump of size between 3ℓ and 4ℓ.

  (ii) The sum of the above small jumps is actually smaller than ℓ.

Note that the event E_ℓ depends implicitly on b, and therefore on a. If the chain M starts from $k \in [\![6\ell, 7\ell]\!]$ and satisfies (i) and (ii) above, it first travels less than distance ℓ in less than b units of time (so it stays in $[\![5\ell, 7\ell]\!]$), and then lands in $[\![\ell, 4\ell]\!]$ by a big jump, jumping over $[\![4\ell, 5\ell]\!]$. Therefore,

$$\forall k \in [\![6\ell, 7\ell]\!], \quad \mathbb{P}_k(M(H_{4\ell} - 1) \in [\![5\ell, 7\ell]\!], M(H_{4\ell}) \in [\![\ell, 4\ell]\!]) \geq \mathbb{P}_k(E_\ell). \tag{6.32}$$

We will prove in the next steps that we can find $a > 0$ such that for some constant $c > 0$ and all $\ell \in \text{GoodScales}$,

$$\forall k \in [\![6\ell, 7\ell]\!], \quad \mathbb{P}_k(E_\ell) \geq c. \tag{6.33}$$

Assuming (6.33), we note that (6.32) yields the claim in Item 2 of Lemma 6.14. It remains to prove (6.33).

<u>*Step 2: Splitting the event E_ℓ.*</u> We treat $a > 0$ as a constant that will be determined at the very end of the proof (the constants c, C appearing in the estimates will not depend on a). For convenience, set for the decreasing Markov chain M,

$$\Delta M_t := M_{t-1} - M_t, \qquad t \geq 1.$$

Note that $\Delta M_t \geq 1$ for all $t \geq 1$. Let $\ell \in \text{GoodScales}$ and $k \in [\![6\ell, 7\ell]\!]$. We first partition E_ℓ as

$$\mathbb{P}_k(E_\ell) = \mathbb{P}_k\big(\exists b' \leq b : \max\{\Delta M_1, \ldots, \Delta M_{b'}\} \leq \ell, \Delta M_{b'+1} \in [\![3\ell, 4\ell]\!], k - M_{b'} \leq \ell\big)$$

$$= \sum_{b'=0}^{b} \mathbb{P}_k(\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \leq \ell, \Delta M_{b'+1} \in [\![3\ell, 4\ell]\!], k - M_{b'} \leq \ell). \tag{6.34}$$

Note that even if the event $\{k - M_{b'} \le \ell\}$ is contained in $\{\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell\}$, it will be convenient for us to keep both events explicit. Fix $b' \in [\![0, b]\!]$. Using the Markov property,

$$\mathbb{P}_k(\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell, \Delta M_{b'+1} \in [\![3\ell, 4\ell]\!], k - M_{b'} \le \ell)$$
$$= \mathbb{E}_k\Big[\mathbb{1}\{\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell\}\mathbb{1}\{k - M_{b'} \le \ell\}\mathbb{P}_{M_{b'}}(\Delta M_1 \in [\![3\ell, 4\ell]\!])\Big].$$

Recalling that $k \ge 6\ell$, on the event $\{k - M_{b'} \le \ell\}$, we have $M_{b'} \ge 5\ell$. Moreover, $M_{b'} \le k \le 7\ell$. Therefore, we can use Item 2 in Lemma 6.12 (with $c_1 = 1$, $c_2 = 4$, $j \in [\![5\ell, 7\ell]\!]$, $x = j - 4\ell$ and $y = j - 3\ell$) to bound, on the event $\{k - M_{b'} \le \ell\}$, the probability $\mathbb{P}_{M_{b'}}(\Delta M_1 \in [\![3\ell, 4\ell]\!])$ from below:

$$\mathbb{P}_{M_{b'}}(\Delta M_1 \in [\![3\ell, 4\ell]\!]) = \mathbb{P}_{M_{b'}}(M_1 \in [\![M_{b'} - 4\ell, M_{b'} - 3\ell]\!]) \ge c\frac{\ell}{M_{b'}}Q(\ell) \ge c\,Q(\ell).$$

This yields

$$\mathbb{P}_k(\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell, \Delta M_{b'+1} \in [\![3\ell, 4\ell]\!], k - M_{b'} \le \ell)$$
$$\ge c\,Q(\ell) \cdot \mathbb{P}_k(\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell, k - M_{b'} \le \ell). \quad (6.35)$$

For future purposes, we note that, assuming $M_0 = k$, we can rewrite the event $\{k - M_{b'} \le \ell\}$ as

$$\{k - M_{b'} \le \ell\} = \Big\{\sum_{i=1}^{b'} \Delta M_i \mathbb{1}_{M_{i-1} \ge k-\ell} \le \ell\Big\}.$$

Likewise, we will need to introduce a cutoff in order to later apply the estimates of Lemma 6.12. Therefore, we rewrite the event

$$\{\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell, k - M_{b'} \le \ell\}$$
$$= \{\max\{\Delta M_1 \mathbb{1}_{M_0 \ge k-\ell}, \ldots, \Delta M_{b'} \mathbb{1}_{M_{b'-1} \ge k-\ell}\} \le \ell, \ \sum_{i=1}^{b'} \Delta M_i \mathbb{1}_{M_{i-1} \ge k-\ell} \le \ell\}$$
$$= \{\max\{\Delta M_1 \mathbb{1}_{M_0 \ge k-\ell}, \ldots, \Delta M_{b'} \mathbb{1}_{M_{b'-1} \ge k-\ell}\} \le \ell, \ \sum_{i=1}^{b'} \Delta M_i \mathbb{1}_{\Delta M_i \le \ell}\mathbb{1}_{M_{i-1} \ge k-\ell} \le \ell\}.$$

Hence, the probability on the right-hand side of (6.35) is

$$\mathbb{P}_k(\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \le \ell, k - M_{b'} \le \ell)$$
$$= \mathbb{P}_k\Bigg(\max\{\Delta M_1 \mathbb{1}_{M_0 \ge k-\ell}, \ldots, \Delta M_{b'} \mathbb{1}_{M_{b'-1} \ge k-\ell}\} \le \ell, \ \sum_{i=1}^{b'} \Delta M_i \mathbb{1}_{\Delta M_i \le \ell}\mathbb{1}_{M_{i-1} \ge k-\ell} \le \ell\Bigg)$$
$$\ge \mathbb{P}_k\Big(\max\{\Delta M_1 \mathbb{1}_{M_0 \ge k-\ell}, \ldots, \Delta M_{b'} \mathbb{1}_{M_{b'-1} \ge k-\ell}\} \le \ell\Big) - \mathbb{P}_k\Bigg(\sum_{i=1}^{b'} \Delta M_i \mathbb{1}_{\Delta M_i \le \ell}\mathbb{1}_{M_{i-1} \ge k-\ell} > \ell\Bigg).$$
$$(6.36)$$

Step 3: Lower bound for the first term in (6.36). We deal with the first term in (6.36), bringing it down by successive applications of the Markov property to

$$\mathbb{P}_k\Big(\max\{\Delta M_1 \mathbb{1}_{M_0 \ge k-\ell}, \ldots, \Delta M_{b'} \mathbb{1}_{M_{b'-1} \ge k-\ell}\} \le \ell\Big) \ge (1 - CQ(\ell))^{b'} \quad (6.37)$$

for some constant $C > 0$. Indeed, by the Markov property,

$$\mathbb{P}_k\Big(\max\{\Delta M_1\mathbb{1}_{M_0 \geq k-\ell}, \ldots, \Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq k-\ell}\} \leq \ell\Big)$$
$$= \mathbb{E}_k\Big[\mathbb{1}\big\{\max\{\Delta M_1\mathbb{1}_{M_0 \geq k-\ell}, \ldots, \Delta M_{b'-1}\mathbb{1}_{M_{b'-2} \geq k-\ell}\} \leq \ell\big\}\mathbb{P}\big(\Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq k-\ell} \leq \ell \mid M_{b'-1}\big)\Big].$$
$$(6.38)$$

Now fix $j \leq 7\ell$. One can write

$$\mathbb{P}\Big(\Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq k-\ell} \leq \ell \mid M_{b'-1} = j\Big) = 1 - \mathbb{P}\Big(\Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq k-\ell} > \ell \mid M_{b'-1} = j\Big).$$

Since $k \geq 6\ell$, the last term is smaller than $\mathbb{P}\Big(\Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq 5\ell} > \ell \mid M_{b'-1} = j\Big)$. Besides, since $\ell \in \text{GoodScales}$, we can use Item 3 in Lemma 6.12 (with $c_1 = 6/7$, $c_2 = 6$ and $y = j - \ell - 1$) to see that, for all $j \in [\![5\ell, 7\ell]\!]$,

$$\mathbb{P}\Big(\Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq 5\ell} > \ell \mid M_{b'-1} = j\Big) = \mathbb{P}(M_1 \in [\![1, j - \ell - 1]\!] \mid M_0 = j) \leq CQ(\ell).$$

The last bound actually holds for all $j \leq 7\ell$, since the probability on the left is zero if $j < 5\ell$, so we proved that

$$\mathbb{P}\Big(\Delta M_{b'}\mathbb{1}_{M_{b'-1} \geq k-\ell} \leq \ell \mid M_{b'-1}\Big) \geq 1 - CQ(\ell).$$

Plugging this back into (6.38), we obtain (6.37) by induction on b'.

Step 4: Upper bound for the second term in (6.36). We now deal with the second term of (6.36). By Markov's inequality,

$$\mathbb{P}_k\left(\sum_{i=1}^{b'} \Delta M_i\mathbb{1}_{\Delta M_i \leq \ell}\mathbb{1}_{M_{i-1} \geq k-\ell} > \ell\right) \leq \frac{1}{\ell} \sum_{i=1}^{b'} \mathbb{E}_k\big[\Delta M_i\mathbb{1}_{\Delta M_i \leq \ell}\mathbb{1}_{M_{i-1} \geq k-\ell}\big].$$

By the Markov property, each expectation is

$$\mathbb{E}_k\big[\Delta M_i\mathbb{1}_{\Delta M_i \leq \ell}\mathbb{1}_{M_{i-1} \geq k-\ell}\big] = \mathbb{E}_k\big[\mathbb{1}_{M_{i-1} \geq k-\ell}\mathbb{E}_{M_{i-1}}\big[\Delta M_1\mathbb{1}_{\Delta M_1 \leq \ell}\big]\big].$$

But now, on the event $\{M_{i-1} \geq k - \ell\}$, we have $7\ell \geq k \geq M_{i-1} \geq k - \ell \geq 5\ell$, i.e. $M_{i-1} \in [\![5\ell, 7\ell]\!]$. Hence, the inner expectation can be bounded using Item 4 in Lemma 6.12 as

$$\mathbb{E}_{M_{i-1}}\big[\Delta M_1\mathbb{1}_{\Delta M_1 \leq \ell}\big] \leq C\ell Q(\ell).$$

We deduce

$$\mathbb{P}_k\left(\sum_{i=1}^{b'} \Delta M_i\mathbb{1}_{\Delta M_i \leq \ell}\mathbb{1}_{M_{i-1} \geq k-\ell} > \ell\right) \leq Cb'Q(\ell).$$

Step 5: Completing the proof of (6.33). Going back to (6.36), the bounds from Steps 3 and 4 give

$$\mathbb{P}_k(\max\{\Delta M_1, \ldots, \Delta M_{b'}\} \leq \ell, k - M_{b'} \leq \ell) \geq (1 - CQ(\ell))^{b'} - Cb'Q(\ell).$$

Then (6.34)–(6.35) provide, for some constants $c, C > 0$,

$$\mathbb{P}_k(E_\ell) \geq c \cdot Q(\ell) \cdot \left(\frac{1 - (1 - CQ(\ell))^{b+1}}{CQ(\ell)} - Cb^2 Q(\ell)\right) \geq \frac{c}{C} \cdot \left(1 - (1 - CQ(\ell))^b - C^2(bQ(\ell))^2\right).$$

We recall that $b = b(\ell) := \left\lceil \frac{a}{Q(\ell)} \right\rceil$. To conclude, it remains to see that one can choose $a > 0$ (small enough) so that the right-hand side is bounded from below by a positive constant, uniformly over ℓ. For the first term, we note that

$$1 - (1 - CQ(\ell))^b \geq 1 - \mathrm{e}^{-CbQ(\ell)} \geq 1 - \mathrm{e}^{-Ca}.$$

On the other hand, we can bound the second term by $C^2(a + Q(\ell))^2 \leq 2C^2 a^2$ for ℓ large enough (where the threshold only depends on a). Overall we have proven the existence of two constants $c > 0$ and $C > 0$ such that for all $\ell \in \text{GoodScales}$ large enough and $k \in [\![6\ell, 7\ell]\!]$,

$$\mathbb{P}_k(E_\ell) \geq c(1 - \mathrm{e}^{-Ca} - 2C^2 a^2).$$

We can now take a small enough so that this is positive, and remove the condition that ℓ is large enough by using $\mathbb{P}_k(E_\ell) > 0$ for all $\ell \in \text{GoodScales}$ and $k \in [\![6\ell, 7\ell]\!]$. This concludes the proof of (6.33). $\qquad \square$

Proof of Item 3 of Lemma 6.14. Let $\ell \in \text{GoodScales}$. Recalling that $\mathbb{P}(B_m = m') = \pi(m; m')$ and using Item 1 in Lemma 6.12, we get that for all $m \in [\![5\ell, 7\ell]\!]$,

$$\forall m' \leq 4\ell, \quad \frac{c}{m} \cdot m' \cdot q(m') \leq \mathbb{P}(B_m = m') \leq \frac{C}{m} \cdot m' \cdot q(m'),$$

and so

$$\frac{c}{m} \sum_{m'=\ell}^{4\ell} m' \cdot q(m') \leq \mathbb{P}(B_m \in [\![\ell, 4\ell]\!]) \leq \frac{C}{m} \sum_{m'=\ell}^{4\ell} m' \cdot q(m').$$

As a consequence, there exists a constant $C > 0$ such that

$$\forall m \in [\![5\ell, 7\ell]\!], \quad \forall m' \in [\![\ell, 4\ell]\!], \quad \frac{1}{C} \cdot \mathbb{P}(B_{5\ell} = m') \leq \mathbb{P}(B_m = m') \leq C \cdot \mathbb{P}(B_{5\ell} = m'), \quad (6.39)$$

and

$$\forall m \in [\![5\ell, 7\ell]\!], \quad \frac{1}{C} \cdot \mathbb{P}(B_{5\ell} \in [\![\ell, 4\ell]\!]) \leq \mathbb{P}(B_m \in [\![\ell, 4\ell]\!]) \leq C \cdot \mathbb{P}(B_{5\ell} \in [\![\ell, 4\ell]\!]). \quad (6.40)$$

Combining (6.39) and (6.40), we get that for all $m \in [\![5\ell, 7\ell]\!]$,

$$\forall m' \in [\![\ell, 4\ell]\!], \quad \mathbb{P}(G_m = m') \overset{(6.17)}{=} \frac{\mathbb{P}(B_m = m')}{\mathbb{P}(B_m \in [\![\ell, 4\ell]\!])} \geq \frac{1}{C^2} \cdot \frac{\mathbb{P}(B_{5\ell} = m')}{\mathbb{P}(B_{5\ell} \in [\![\ell, 4\ell]\!])} = \frac{1}{C^2} \cdot \mathbb{P}(G_{5\ell} = m').$$

This ensures that

$$\sum_{m'=\ell}^{4\ell} \left(\inf_{m \in [\![5\ell, 7\ell]\!]} \mathbb{P}(G_m = m') \right) \geq c \sum_{m'=\ell}^{4\ell} \mathbb{P}(G_{5\ell} = m') = c, \quad (6.41)$$

where in the last equality we used that the sum is one by definition (6.17) of $G_{5\ell}$. By standard arguments (see for example [Pre22]), this allows us to construct a coupling G_m, $m \in [\![5\ell, 7\ell]\!]$, of the golden instructions such that[17]

$$\mathbb{P}(\forall a, a' \in [\![5\ell, 7\ell]\!], \ G_a = G_{a'}) \geq c.$$

This proves our claim. $\qquad \square$

_____

[17]Even if not strictly needed, we note that here we can take the same constant c as in (6.41).

As explained in Section 6.3.2, we also need to show that the facts stated in Item 2 of Lemma 6.14 for M also hold for M', conditional on them holding for M, with uniformly positive probability across all good scales.

Lemma 6.15. *Let M and M' be defined as the decreasing chains started from respectively k and k' under the black-golden coupling $\mathbb{P}_{k,k'}$ of Definition 6.5. Let $c \in (0,1)$ be the constant in Lemma 6.14. Then, for any $\ell \in \mathrm{GoodScales}$ and for all $k, k' \in [\![6\ell, 7\ell]\!]$, we have that*

$$\mathbb{P}_{k,k'}\big(M'(H'_{4\ell} - 1) \in [\![5\ell, 7\ell]\!], M'(H'_{4\ell}) \in [\![\ell, 4\ell]\!] \mid M(H_{4\ell} - 1) \in [\![5\ell, 7\ell]\!], M(H_{4\ell}) \in [\![\ell, 4\ell]\!]\big) \geq \frac{c}{2}.$$

Proof. Fix $\ell \in \mathrm{GoodScales}$ and $k, k' \in [\![6\ell, 7\ell]\!]$, and define the finite set of paths

$$\mathcal{M}_{k,\ell} := \Big\{\vec{m} = (m_i)_{i \in [\![0, h-1]\!]} : h > 0, \ k = m_0 > m_1 > \ldots > m_{h-1}, \ m_{h-1} \in [\![5\ell, 7\ell]\!]\Big\}.$$

We then have

$$\{M_0 = k, M(H_{4\ell} - 1) \in [\![5\ell, 7\ell]\!], M(H_{4\ell}) \in [\![\ell, 4\ell]\!])\}$$
$$= \bigsqcup_{\vec{m} \in \mathcal{M}_{k,\ell}} \Big\{(M_i)_{i \in [\![0, H_{4\ell} - 1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\Big\},$$

where we recall that $B_{m_{h-1}}$ is the black instruction at height m_{h-1}. We now fix a path $\vec{m} = (m_i)_{i \in [\![0, h-1]\!]} \in \mathcal{M}_{k,\ell}$ and set $\mathcal{H}'_\ell := \{M'(H'_{4\ell} - 1) \in [\![5\ell, 7\ell]\!], M'(H'_{4\ell}) \in [\![\ell, 4\ell]\!]\}$. We will show that

$$\mathbb{P}_{k,k'}\Big(\mathcal{H}'_\ell \mid (M_i)_{i \in [\![0, H_{4\ell} - 1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\Big) \geq \frac{c}{2}. \tag{6.42}$$

Note that this immediately implies the lemma statement. We need to introduce some new notation. Define, for all $t' > 0$, the finite set of paths

$$\mathcal{M}'_{k'}(\vec{m}, t') := \Big\{\vec{m}' = (m'_i)_{i \in [\![0, t']\!]} : k' = m'_0 > m'_1 > \ldots > m'_{t'},$$
$$m'_{t'} = m_t \text{ for some } t \leq h - 1, \ m'_i \neq m_j \text{ for all } i < t' \text{ and } j < t\Big\},$$

that is, the set of all possible paths for M' up to t' that are merging into the path $\vec{m}$ at time t'. We also define the set

$$\mathcal{M}'^{\mathrm{c}}_{k'}(\vec{m}) := \Big\{\vec{m}' = (m'_i)_{i \in [\![0, h'-1]\!]} : h' > 0, \ k' = m'_0 > m'_1 > \ldots > m'_{h'-1},$$
$$m'_{h'-1} \in [\![5\ell, 7\ell]\!], \ m'_i \neq m_j \text{ for all } i \leq h' - 1 \text{ and } j \leq h - 1\Big\},$$

that is, the set of all possible paths for M' up to $H'_{4\ell} - 1$ on the event $\mathcal{H}'_\ell$ that are not merging into the path $\vec{m}$ up to time $h - 1$. We also introduce the events

$$\mathcal{E}_{k',\vec{m}} := \bigsqcup_{t' > 0} \Big\{(M'_i)_{i \in [\![0, t']\!]} \in \mathcal{M}'_{k'}(\vec{m}, t')\Big\},$$
$$\mathcal{E}^{\mathrm{c}}_{k',\vec{m}} := \Big\{(M'_i)_{i \in [\![0, H'_{4\ell} - 1]\!]} \in \mathcal{M}'^{\mathrm{c}}_{k'}(\vec{m}), \ B_{M'(H'_{4\ell} - 1)} \in [\![\ell, 4\ell]\!]\Big\}.$$

Note that $\mathcal{E}^{\mathrm{c}}_{k',\vec{m}} \subset \mathcal{H}'_\ell$, and on the event $\Big\{(M_i)_{i \in [\![0, H_{4\ell} - 1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\Big\}$,

$$\mathcal{H}'_\ell = \mathcal{E}_{k',\vec{m}} \sqcup \mathcal{E}^{\mathrm{c}}_{k',\vec{m}}. \tag{6.43}$$

59

We divide our analysis of (6.42) into two cases.

Case 1: We first assume that $\mathbb{P}_{k'}(\mathcal{E}_{k',\vec{m}}) \geq \frac{c}{2}$. Note that thanks to (6.43),

$$\mathbb{P}_{k,k'}\left(\mathcal{H}'_\ell \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right) \geq \mathbb{P}_{k,k'}\left(\mathcal{E}_{k',\vec{m}} \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}\in[\![\ell,4\ell]\!]}\right)$$

$$= \sum_{t'>0} \mathbb{P}_{k,k'}\left((M'_i)_{i\in[\![0,t']\!]} \in \mathcal{M}'_{k'}(\vec{m},t') \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}\in[\![\ell,4\ell]\!]}\right). \quad (6.44)$$

Now, fix $t' > 0$. For any path $\vec{m}' = (m'_i)_{i\in[\![0,t']\!]} \in \mathcal{M}'_{k'}(\vec{m},t')$, the event $\left\{(M'_i)_{i\in[\![0,t']\!]} = \vec{m}'\right\}$ is measurable w.r.t. the black instructions $(B_{m'_i})_{i\in[\![0,t'-1]\!]}$, while $\left\{(M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right\}$ is measurable w.r.t. the black instructions $(B_{m_i})_{i\in[\![0,h-1]\!]}$. But, by the definition of $\mathcal{M}'_{k'}(\vec{m},t')$, we have that $m'_i \neq m_j$ for all $i < t'$ and $j < t$ (and so also for all $j \leq h - 1$ since both paths are decreasing and $m'_{t'} = m_t$). Therefore, these two events are independent by the independence of the black instructions. As a result, we have

$$\mathbb{P}_{k,k'}\left(\mathcal{E}_{k',\vec{m}} \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right) = \mathbb{P}_{k,k'}(\mathcal{E}_{k',\vec{m}}) \geq \frac{c}{2} \quad (6.45)$$

by our assumption that $\mathbb{P}_{k'}(\mathcal{E}_{k',\vec{m}}) \geq \frac{c}{2}$. Combining (6.44) and (6.45), we get (6.42).

Case 2: We now assume that $\mathbb{P}_{k'}(\mathcal{E}_{k',\vec{m}}) < \frac{c}{2}$. Because of Equation (6.43), we have the bound

$$\mathbb{P}_{k,k'}\left(\mathcal{H}'_\ell \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right)$$

$$\geq \mathbb{P}_{k,k'}\left(\mathcal{E}^{\mathrm{c}}_{k',\vec{m}} \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right). \quad (6.46)$$

Now, for any $\vec{m}' = (m'_i)_{i\in[\![0,h'-1]\!]} \in \mathcal{M}'^{\mathrm{c}}_{k'}(\vec{m})$, the event $\left\{(M'_i)_{i\in[\![0,H'_{4\ell}-1]\!]} = \vec{m}', B_{m'_{h'-1}} \in [\![\ell, 4\ell]\!]\right\}$ is measurable w.r.t. the black instructions $(B_{m'_i})_{i\in[\![0,h'-1]\!]}$, while $\left\{(M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right\}$ is measurable w.r.t. the black instructions $(B_{m_i})_{i\in[\![0,h-1]\!]}$. But, by definition of $\mathcal{M}'^{\mathrm{c}}_{k'}(\vec{m})$, we have that $m'_i \neq m_j$ for all $i \leq h' - 1$ and $j \leq h$ and thus these two events are independent. Combined with (6.46), this shows that

$$\mathbb{P}_{k,k'}\left(\mathcal{H}'_\ell \,\middle|\, (M_i)_{i\in[\![0,H_{4\ell}-1]\!]} = \vec{m}, B_{m_{h-1}} \in [\![\ell, 4\ell]\!]\right) \geq \mathbb{P}_{k,k'}\left(\mathcal{E}^{\mathrm{c}}_{k',\vec{m}}\right). \quad (6.47)$$

Now, since $\mathcal{E}^{\mathrm{c}}_{k',\vec{m}} \subset \mathcal{H}'_\ell \subset \mathcal{E}^{\mathrm{c}}_{k',\vec{m}} \cup \mathcal{E}_{k',\vec{m}}$, we can write

$$\mathbb{P}_{k,k'}(\mathcal{H}'_\ell) = \mathbb{P}_{k,k'}\left(\mathcal{E}^{\mathrm{c}}_{k',\vec{m}}\right) + \mathbb{P}_{k,k'}(\mathcal{H}'_\ell \mid \mathcal{E}_{k',\vec{m}}) \cdot \mathbb{P}_{k,k'}(\mathcal{E}_{k',\vec{m}}).$$

Thanks to Item 2 in Lemma 6.14, we have that $\mathbb{P}_{k,k'}(\mathcal{H}'_\ell) \geq c$. Since $\mathbb{P}_{k,k'}(\mathcal{E}_{k',\vec{m}}) < \frac{c}{2}$ by assumption, we immediately get that

$$\mathbb{P}_{k,k'}\left(\mathcal{E}^{\mathrm{c}}_{k',\vec{m}}\right) > \frac{c}{2}. \quad (6.48)$$

Thus, we get (6.42) as a consequence of (6.47) and (6.48).

This concludes the proof of the lemma. $\qquad\square$

6.3.5 Completing the coupling argument: fast merging of the two walks

We are now ready to complete the proof of Proposition 6.6.

Proof of Proposition 6.6. Most of the proof will be devoted to establishing Item 1. The gist of the proof is the following: by Items 1 and 2 of Lemma 6.14 (combined with Lemma 6.15), at each good scale, there is positive probability that both M and M' follow a golden instruction, and therefore positive probability to merge thanks to Item 3 of Lemma 6.14. Moreover, by the specific conditioning in Item 1 of Lemma 6.14, we can iterate this argument at each good scale, so this will happen at a fair amount of scales. Since there are infinitely many good scales, it will happen with probability going to 1 as the starting points k and k' of M and M' go to infinity.

We now make this argument rigorous. Without loss of generality, we assume that $k' \geq k$. Fix $A \geq 1$. For $k \geq A$, we recall that $s = s(A, k)$ is the number of good scales ℓ in $[\![A, k]\!]$ such that $7\ell < k$. We denote, for all $k \geq 1$,

$$\ell^1 = \ell^1(k) := \max\{\ell' \in \text{GoodScales} : 7\ell' < k\}, \tag{6.49}$$

with the convention $\ell^1(k) = 0$ if no such scale exists, and then write $\ell^i = \ell^i(k)$ for the i-th good scale down from $\ell^1(k)$.

We start by noting that M and M' will meet above height A if all the following events occur simultaneously at some scale ℓ^i with $1 \leq i \leq s$ (recall Figure 11):

- The chains M and M', at times $H_{7\ell^i}$ and $H'_{7\ell^i}$ respectively, land in the sub-interval $[\![6\ell^i, 7\ell^i]\!]$ (highlighted in blue in Figure 11).

- The chains M and M' both exit the interval $[\![5\ell^i, 7\ell^i]\!]$ by following a golden instruction, *i.e.* landing in the golden sub-interval in Figure 11. This means that $M(H_{4\ell^i} - 1) \in [\![5\ell^i, 7\ell^i]\!]$ and the black instruction $B_{M(H_{4\ell^i}-1)}$ is inside the interval $[\![\ell^i, 4\ell^i]\!]$, and that the same holds for M'.

- The golden instructions $G_{M(H_{4\ell^i}-1)}$ and $G_{M'(H'_{4\ell^i}-1)}$ are equal.

Therefore, for any $i \in [\![1, s]\!]$ we introduce the event

$$\mathcal{E}_i := \mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \cap \mathcal{E}_i^2 \cap (\mathcal{E}_i^2)' \cap \mathcal{E}_i^3, \tag{6.50}$$

where

$$\begin{cases} \mathcal{E}_i^1 = \left\{ M(H_{7\ell^i}) \in [\![6\ell^i, 7\ell^i]\!] \right\}, \\ \mathcal{E}_i^2 = \left\{ M(H_{4\ell^i} - 1) \in [\![5\ell^i, 7\ell^i]\!], B_{M(H_{4\ell^i}-1)} \in [\![\ell^i, 4\ell^i]\!] \right\}, \\ \mathcal{E}_i^3 = \left\{ \forall a, a' \in [\![5\ell^i, 7\ell^i]\!], \ G_a = G_{a'} \right\}, \end{cases}$$

and $(\mathcal{E}_i^1)'$ and $(\mathcal{E}_i^2)'$ are constructed similarly to $\mathcal{E}_i^1$ and $\mathcal{E}_i^2$ but for the chain M'. Recall the σ-field $\mathcal{G}_{\ell^i}$ of (6.29). Note that for any $j < i$, the event $\mathcal{E}_j$ is $\mathcal{G}_{\ell^i}$-measurable.

The preceding discussion leads to

$$\mathbb{P}_{k,k'}(V_{\text{mer}} < A) \leq \mathbb{P}_{k,k'}\left(\bigcap_{i=1}^s \mathcal{E}_i^{\text{c}} \right). \tag{6.51}$$

We start by proving that for any $i \in [\![1, s]\!]$ we almost surely have

$$\mathbb{P}_{k,k'}(\mathcal{E}_i \mid \mathcal{G}_{\ell^i}) \geq c^5/2, \tag{6.52}$$

where $c \in (0,1)$ is as in Lemma 6.14. We consider the events sequentially.

Bound on $\mathbb{P}_{k,k'}(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \mid \mathcal{G}_{\ell i})$: By Item 1 of Lemma 6.13 together with Item 1 of Lemma 6.14, we get that on the event $\{M(H_{7\ell i} - 1) = m, M'(H'_{7\ell i} - 1) = m'\}$ with $m \in [\![7\ell^i + 1, k]\!]$ and $m' \in [\![7\ell^i + 1, k']\!]$ such that $m \neq m'$, we have

$$
\begin{aligned}
\mathbb{P}_{k,k'}\Big(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \ \Big| \ \mathcal{G}_{\ell i}\Big) &= \mathbb{P}(B_m, B_{m'} \in [\![6\ell, 7\ell]\!] \mid B_m \leq 7\ell, B_{m'} \leq 7\ell) \\
&= \mathbb{P}(B_m \in [\![6\ell, 7\ell]\!] \mid B_m \leq 7\ell) \cdot \mathbb{P}(B_{m'} \in [\![6\ell, 7\ell]\!] \mid B_{m'} \leq 7\ell) \\
&\geq c^2.
\end{aligned}
$$

Moreover, if $m = m'$, the same computation gives a lower bound $c \geq c^2$, so the last inequality is true for every possible value (m, m') that can be taken by the pair $(M(H_{7\ell i} - 1), M'(H'_{7\ell i} - 1))$. Therefore, we almost surely have

$$
\mathbb{P}_{k,k'}\Big(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \ \Big| \ \mathcal{G}_{\ell i}\Big) \geq c^2. \tag{6.53}
$$

Bound on $\mathbb{P}_{k,k'}(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \cap \mathcal{E}_i^2 \cap (\mathcal{E}_i^2)' \mid \mathcal{G}_{\ell i})$: Let $j, j' \in [\![6\ell^i, 7\ell^i]\!]$. Conditioning on $\mathcal{E}_i^2$, we have

$$
\mathbb{P}_{j,j'}\Big(\mathcal{E}_i^2 \cap (\mathcal{E}_i^2)'\Big) = \mathbb{E}_j\Big[\mathbb{1}_{\mathcal{E}_i^2} \cdot \mathbb{P}_{j,j'}\Big((\mathcal{E}_i^2)' \ \Big| \ \mathcal{E}_i^2\Big)\Big].
$$

From there, we use Lemma 6.15 to get $\mathbb{P}_{j,j'}((\mathcal{E}_i^2)' \mid \mathcal{E}_i^2) \geq c/2$ and then Item 2 in Lemma 6.14 to obtain

$$
\mathbb{P}_{j,j'}\Big(\mathcal{E}_i^2 \cap (\mathcal{E}_i^2)'\Big) \geq c^2/2. \tag{6.54}
$$

Note that the events $\mathcal{E}_i^2$ and $(\mathcal{E}_i^2)'$ are realized for the chains M and M' if and only if they are realized for the chains $(M(H_{7\ell i} + t))_{t \geq 0}$ and $(M'(H'_{7\ell i} + t))_{t \geq 0}$. By Item 2 of Lemma 6.13, this ensures that

$$
\begin{aligned}
\mathbb{P}_{k,k'}\Big(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \cap \mathcal{E}_i^2 \cap (\mathcal{E}_i^2)' \ \Big| \ \mathcal{G}_{\ell i}\Big) &= \mathbb{E}_{k,k'}\Big[\mathbb{1}_{\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)'} \mathbb{1}_{\mathcal{E}_i^2 \cap (\mathcal{E}_i^2)'} \ \Big| \ \mathcal{G}_{\ell i}\Big] \\
&= \mathbb{E}_{k,k'}\Big[\mathbb{1}_{\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)'} \mathbb{P}_{M(H_{7\ell i}), M'(H'_{7\ell i})}\Big(\mathcal{E}_i^2 \cap (\mathcal{E}_i^2)'\Big) \ \Big| \ \mathcal{G}_{\ell i}\Big] \\
&\overset{(6.54)}{\geq} \frac{c^2}{2} \mathbb{E}_{k,k'}\Big[\mathbb{1}_{\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)'} \ \Big| \ \mathcal{G}_{\ell i}\Big] \overset{(6.53)}{\geq} \frac{c^4}{2}
\end{aligned}
$$

almost surely.

Bound on $\mathbb{P}_{k,k'}(\mathcal{E}_i \mid \mathcal{G}_{\ell i})$: Finally, by independence of black and golden instructions, and independence of golden instructions between different scales, the event $\mathcal{E}_i^3$ is independent of the chains up to the times $H_{4\ell i} - 1$ and $H'_{4\ell i} - 1$ and of $B_{M(H_{4\ell i}-1)}$ and $B_{M'(H'_{4\ell i}-1)}$, so it is independent of $\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \cap \mathcal{E}_i^2 \cap (\mathcal{E}_i^2)'$. Since it is also independent of $\mathcal{G}_{\ell i}$, we have, using Item 3 of Lemma 6.14,

$$
\mathbb{P}_{k,k'}\Big(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \cap \mathcal{E}_i^2 \cap (\mathcal{E}_i^2)' \cap \mathcal{E}_i^3 \ \Big| \ \mathcal{G}_{\ell i}\Big) = \mathbb{P}\Big(\mathcal{E}_i^3\Big)\mathbb{P}_{k,k'}\Big(\mathcal{E}_i^1 \cap (\mathcal{E}_i^1)' \cap \mathcal{E}_i^2 \cap (\mathcal{E}_i^2)' \ \Big| \ \mathcal{G}_{\ell i}\Big) \geq c \cdot \frac{c^4}{2} \geq \frac{c^5}{2}
$$

almost surely, which proves (6.52).

Now, using in cascade that $\mathcal{E}_1, \dots, \mathcal{E}_i$ are $\mathcal{G}_{\ell i+1}$-measurable together with the fact (6.52) we get

$$
\begin{aligned}
\mathbb{P}_{k,k'}\left(\bigcap_{i=1}^s \mathcal{E}_i^{\mathrm{c}}\right) &= \mathbb{E}_{k,k'}\Big[\mathbb{1}_{\mathcal{E}_1^{\mathrm{c}}} \dots \mathbb{1}_{\mathcal{E}_{s-1}^{\mathrm{c}}} \mathbb{P}(\mathcal{E}_s^{\mathrm{c}} \mid \mathcal{G}_{\ell s})\Big] \\
&\leq \left(1 - \frac{c^5}{2}\right) \cdot \mathbb{E}_{k,k'}\Big[\mathbb{1}_{\mathcal{E}_1^{\mathrm{c}}} \dots \mathbb{1}_{\mathcal{E}_{s-1}^{\mathrm{c}}}\Big] \\
&\leq \left(1 - \frac{c^5}{2}\right)^s = \left(1 - \frac{c^5}{2}\right)^{s(A,k)}.
\end{aligned}
$$

This proves Item 1 of Proposition 6.6.

We now prove Item 2, *i.e.* the one about $\mathsf{T} - \mathsf{t}_{\mathrm{mer}}$ and $\mathsf{T}' - \mathsf{t}'_{\mathrm{mer}}$, which will be a consequence of the first one. Recall from the statement of Proposition 6.6 that $\mathsf{t}_{\mathrm{mer}}$ and $\mathsf{t}'_{\mathrm{mer}}$ are the respective first times when the two chains M and M' visit V_{mer}. By symmetry, we may focus on the chain M, say, and prove that $\mathsf{T} - \mathsf{t}_{\mathrm{mer}}$ goes to infinity in probability. For all $t > 0$, since the chain M is decreasing, we have

$$\mathbb{P}_{k,k'}(\mathsf{T} - \mathsf{t}_{\mathrm{mer}} < t) = \mathbb{P}_{k,k'}(\mathsf{t}_{\mathrm{mer}} > \mathsf{T} - t) = \mathbb{P}_{k,k'}(V_{\mathrm{mer}} < M_{\mathsf{T}-t}).$$

We now introduce another parameter $A > 0$, that we will tune later, and split the above display as follows:

$$\mathbb{P}_{k,k'}(\mathsf{T} - \mathsf{t}_{\mathrm{mer}} < t) \leq \mathbb{P}_{k,k'}(V_{\mathrm{mer}} \leq A) + \mathbb{P}_k(M_{\mathsf{T}-t} > A) \leq \mathbb{P}_{k,k'}(V_{\mathrm{mer}} \leq A) + \sup_{\mathsf{k}} \mathbb{P}_{\mathsf{k}}(M_{\mathsf{T}-t} > A). \tag{6.55}$$

By Item 1 of Proposition 6.6, for any fixed $A > 0$, the first term on the right-hand side goes to 0 as $k, k' \to \infty$. It remains to prove that $\sup_{\mathsf{k}} \mathbb{P}_{\mathsf{k}}(M_{\mathsf{T}-t} > A) \to 0$ as $A \to \infty$. To see this, we use a decomposition into blocks and Item 1 of Lemma 6.14. Let $s \geq 1$, and take A large enough so that there are at least $s \cdot (t+1)$ good scales ℓ such that $\ell < A$. We see $t+1$ as the number of blocks, each containing s consecutive good scales. Then we note that, on the event $\{M_{\mathsf{T}-t} > A\}$, the chain M only has at most t units of time to visit $(t+1)$ blocks of good scales. As a consequence, on that event, at least one block of s good scales is not visited by M. By a union bound and iterating the bound of Item 1 in Lemma 6.14, this gives that for all $\mathsf{k} > A$,

$$\mathbb{P}_{\mathsf{k}}(M_{\mathsf{T}-t} > A) \leq (t+1)(1-c)^s.$$

The above estimate is actually uniform over $\mathsf{k} \geq 1$ since the probability is 0 if $\mathsf{k} \leq A$. This entails that $\sup_{\mathsf{k}} \mathbb{P}_{\mathsf{k}}(M_{\mathsf{T}-t} > A) \to 0$ as $A \to \infty$, providing the missing estimate in (6.55), and hence Item 2. $\qquad\square$

7 Good regularity estimates for the sequences q and Q

The main goal of this section is to prove the desired good regularity estimates.

Theorem 7.1 (Good regularity estimates: q and Q are regularly varying)**.** *There exist two slowly varying functions φ, Φ such that for all $k \geq 1$,*

$$q(k) = k^{-\frac{1}{2\alpha}-1}\varphi(k) \qquad and \qquad Q(k) = k^{-\frac{1}{2\alpha}}\Phi(k),$$

where $\alpha \in (1/2, 1)$ is as in Theorem 2.1.

Recall that T^k denotes the p-signed critical binary Bienaymé–Galton–Watson tree T conditioned on the event that $\mathrm{LIS}(T) = k$. A key step is to first prove in Section 7.1 the following law of large numbers for the number $\#\mathcal{L}_{\cap}^{\max}(T^k)$ of leaves contained in the intersection of all maximal positive subtrees of T^k. This will rely on the local convergence in Theorem 6.2 (or, rather, on the good merging properties of the coupling in Corollary 6.7).

Theorem 7.2 (Law of large numbers for $\#\mathcal{L}_{\cap}^{\max}(T^k)$)**.** *We have the convergence in probability*

$$\frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k} \xrightarrow{\mathbb{P}} \lambda \quad as \ k \to \infty,$$

where λ is a constant in $(0,1)$.

Building on this law of large numbers, we will then obtain in Section 7.2 the good regularity estimates for the sequences q and Q, proving Theorem 7.1.

7.1 The size of the intersection of all maximal positive subtrees

The goal of this section is to prove the law of large numbers in Theorem 7.2 using first and second moment arguments. Controlling the latter moments is the purpose of the next two subsections. We will aim for rather precise estimates, as such refined control will prove useful later in Section 9 for establishing the full scaling limit result of Theorem 2.3.

7.1.1 The first moment

We begin with the first moment. Recall the Markov chain $(Z_h^k)_{h\geq 0} = (S_h, D_h, M_h^*, W_h)_{h\geq 0}$ introduced in (6.2). From (6.4), for a leaf L^k chosen uniformly at random among the k leaves of the leftmost maximal positive subtree $T^{k,\mathrm{lmax}}$, we have

$$\left\{ L^k \in \mathcal{L}_\cap^{\max}(T^k) \right\} = \left\{ \sum_{h\geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h\} = 0 \right\}. \tag{7.1}$$

We also note that

$$\mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k} \right] = \mathbb{P}\left(L^k \in \mathcal{L}_\cap^{\max}(T^k) \right). \tag{7.2}$$

Therefore, proving the convergence of the first moment of $\mathcal{L}_\cap^{\max}(T^k)$ will crucially rely on controlling the number of **forks**

$$F := \sum_{m\geq 1} \sum_{h\geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\} \tag{7.3}$$

along the spine towards the marked leaf L^k. We also introduce the version of F truncated below A:

$$F_A := \sum_{m\geq A} \sum_{h\geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\}. \tag{7.4}$$

The following claim says that F is *local*, in the sense that F_A is small when A is large, and so F is well-approximated by a random variable that only depends on the finite number of values taken by the non-increasing chain M^* before extinction. Recall from Remark 6.3 that under $\mathbb{P}_k$, the chain M^* starts from k.

Lemma 7.3 (The number of forks is local)**.** *There exists a constant $C > 0$ such that, for all $A \geq 1$,*

$$\sup_{k\geq 1} \mathbb{E}_k[F_A] \leq C \cdot Q(A), \tag{7.5}$$

where $Q(\cdot)$ is as in (2.5). Moreover, for all $k \geq 1$,

$$\mathbb{P}_k(F = 0) \leq \frac{1}{2q(1)} < 1. \tag{7.6}$$

Proof. We have for all $A \geq 2$ and $k \geq 1$,

$$\mathbb{E}_k[F_A] = \sum_{m=A}^{\infty} \sum_{h\geq 1} \mathbb{P}_k(S_h = \ominus \text{ and } M_h^* = W_h = m)$$

$$= \sum_{m=A}^{\infty} \sum_{h\geq 1} \mathbb{P}_k(M_{h-1}^* = m) \cdot \mathbb{P}_k(S_h = \ominus, M_h^* = W_h = m \mid M_{h-1}^* = m)$$

since, on the event $\{S_h = \ominus, M_h^* = W_h = m\}$, we must have that $M_{h-1}^* = m$. By the Markov property and the expression of the transition probabilities (6.5), the above display boils down to

$$\mathbb{E}_k[F_A] = \sum_{m=A}^{\infty} \mathbb{E}_k\left[\sum_{h \geq 0} \mathbb{1}\{M_h^* = m\}\right] \cdot \frac{1}{2}(1-p)q(m).$$

Now for fixed $m \geq 2$, conditionally on M ever hitting the value m, the random variable $\sum_{h \geq 0} \mathbb{1}\{M_h^* = m\}$ is geometric with parameter $\theta(m) = p + (1-p)\frac{Q(m)+Q(m+1)}{2}$ as shown in (6.10), and is 0 if M does not hit m. Hence, for all $m \geq 2$, we get the upper bound

$$\mathbb{E}_k\left[\sum_{h \geq 0} \mathbb{1}\{M_h^* = m\}\right] \leq \left(p + (1-p)\frac{Q(m) + Q(m+1)}{2}\right)^{-1} \leq C,$$

where C is a constant that is independent of k and m. The same argument carries over to the case $m = 1$ using that $\theta(1) = \frac{1}{2q(1)}$ (see the discussion following (6.10)). As a result,

$$\mathbb{E}_k[F_A] \leq \frac{C}{2}(1-p)\sum_{m=A}^{\infty} q(m) = \frac{C}{2}(1-p)Q(A).$$

This bound is uniform in k, and we deduce (7.5).

To prove (7.6), we follow similar ideas. Fix $k \geq 1$. By a crude bound, we may only look at the term $m = 1$ in (7.3):

$$\mathbb{P}_k(F = 0) \leq \mathbb{P}_k\left(\sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = 1\} = 0\right).$$

Note that if $\eta_{\mathrm{T}-1} < h < \eta_{\mathrm{T}}$, then $M_h^* = W_h = 1$ and $S_h = \ominus$ (recall (6.11)). Thus we have the inclusion of events:

$$\left\{\sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = 1\} = 0\right\} \subset \{\eta_{\mathrm{T}} - \eta_{\mathrm{T}-1} = 1\}.$$

But by the discussion following (6.10), we know that $\eta_{\mathrm{T}} - \eta_{\mathrm{T}-1}$ is geometric with parameter $\theta(1) = \frac{1}{2q(1)}$. This provides (7.6). $\qquad \square$

From Lemma 7.3 and the fast merging property of Corollary 6.7, we deduce the convergence of the first moment. Recall the notation V_{mer} introduced in Proposition 6.6.

Lemma 7.4 (The first moment converges)**.** *The following assertions are true:*

1. For $k' \geq k$ and for all $A > 0$,

$$\left|\mathbb{E}\left[\frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k}\right] - \mathbb{E}\left[\frac{\#\mathcal{L}_{\cap}^{\max}(T^{k'})}{k'}\right]\right| \leq 2\sup_{k \geq 1}\mathbb{E}_k[F_A] + \widetilde{\mathbb{P}}_{k,k'}(V_{\mathrm{mer}} \leq A). \tag{7.7}$$

2. The first moment converges:

$$\mathbb{E}\left[\frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k}\right] \to \lambda \quad \text{as } k \to \infty.$$

65

3. *The limit λ is in $(0,1)$.*

The quantitative bound in Item 1 will also be used later in the proof of Theorem 9.1.

Proof. We first prove the bound in Item 1 using Corollary 6.7.

Let now F and F' be the number of forks along the trajectories of (Z_h^k) and $(Z_h^{k'})$ respectively, as in (7.3), and F_A and F'_A their truncated version, as in (7.4). Under the coupling $\widetilde{\mathbb{P}}_{k,k'}$ of Corollary 6.7, we can write $F - F'$ as

$$F - F' = \sum_{m \geq 1} \sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\} - \sum_{m \geq 1} \sum_{h \geq 1} \mathbb{1}\{S'_h = \ominus \text{ and } (M_h^*)' = W'_h = m\}$$

$$= \sum_{m \geq V_{\mathrm{mer}}} \sum_{h \geq 1} (\mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\} - \mathbb{1}\{S'_h = \ominus \text{ and } (M_h^*)' = W'_h = m\}), \quad (7.8)$$

where in the last equality we used that Z^k and $Z^{k'}$ are identical as soon as M^* and $(M^*)'$ are both strictly smaller than V_{mer}. To conclude, note from (7.1)–(7.3) that

$$\left| \mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k} \right] - \mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^{k'})}{k'} \right] \right| = |\mathbb{P}_k(F = 0) - \mathbb{P}_{k'}(F' = 0)|,$$

and

$$\begin{aligned}
|\mathbb{P}_k(F = 0) - \mathbb{P}_{k'}(F' = 0)| &\leq \widetilde{\mathbb{E}}_{k,k'}[|\mathbb{1}_{F=0} - \mathbb{1}_{F'=0}|] \\
&\leq \widetilde{\mathbb{E}}_{k,k'}[1 \wedge |F - F'|] \\
&= \widetilde{\mathbb{E}}_{k,k'}[(1 \wedge |F - F'|)\mathbb{1}_{V_{\mathrm{mer}} \geq A}] + \widetilde{\mathbb{E}}_{k,k'}[(1 \wedge |F - F'|)\mathbb{1}_{V_{\mathrm{mer}} < A}] \\
&\leq 2 \sup_{k \geq 1} \mathbb{E}_k[F_A] + \widetilde{\mathbb{P}}_{k,k'}(V_{\mathrm{mer}} < A),
\end{aligned}$$

where for the last inequality we used (7.8). This completes the proof of Item 1.

We now prove the convergence in Item 2 by Cauchy characterization. The first term on the right-hand side of (7.7) is small as $A \to \infty$ because of (7.5) in Lemma 7.3, while the second term is small for fixed A and $k, k' \to \infty$ as a result of Item 2 in Corollary 6.7. This proves the Cauchy characterization and so Item 2.

It remains to check Item 3, *i.e.* that the limit λ is in $(0,1)$. The fact that λ is not 1 is a straightforward consequence of (7.6). We now check that it is non-zero. We fix $A \geq 1$ and write

$$F = \sum_{m \geq 1} \sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\}$$

$$= F_{A+1} + \sum_{m=1}^{A} \sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\}.$$

We recall from (6.28) the notation $H_A = \inf\{t \geq 1 \mid M_t \leq A\}$. Using the strong Markov property of Z together with the fact that the transitions only depend on M (Section 6.2.1), we have

$$\mathbb{P}_k(F = 0) = \mathbb{E}_k\left[\mathbb{1}_{F_{A+1}=0} \cdot \mathbb{P}_{M(H_A)}(F = 0)\right] \geq \inf_{1 \leq k \leq A} \mathbb{P}_k(F = 0) \cdot \mathbb{P}_k(F_{A+1} = 0),$$

where for the inequality we used that $1 \leq M(H_A) \leq A$. But by Markov's inequality

$$\mathbb{P}_k(F_{A+1} = 0) = 1 - \mathbb{P}_k(F_{A+1} \geq 1) \geq 1 - \sup_{k \geq 1} \mathbb{E}_k[F_{A+1}],$$

which is non-zero if A is large enough thanks to (7.5). Combining the last two displayed equations, we get a uniform lower bound for $\mathbb{P}_k(F = 0)$, as we wanted. □

7.1.2 The second moment

We now turn to the second moment. First, let us recall some notation and introduce some more. Recall from the beginning of Section 6.1.1 and (6.2) that conditionally on T^k, we let L^k be a leaf chosen uniformly at random among the k leaves of the leftmost maximal positive subtree $T^{k,\mathrm{lmax}}$ and denote, for $h \geq 0$, by v_h the ancestor (if it exists) of L^k at height h; see Figure 10. The process $Z^k = (Z^k_h)_{h \geq 0} = (S_h, D_h, M^*_h, W_h)_{h \geq 0}$ defined in (6.2) records some useful information about the subtrees above each of the vertices $(v_h)_{h \geq 0}$. Now, pick $\widetilde{L}^k$ a second leaf with the same distribution as L^k, independently of L^k conditionally on T^k. We introduce the process $\widetilde{Z}^k = (\widetilde{Z}^k_h)_{h \geq 0}$ similarly to the process Z^k, recording information along the ancestral line of $\widetilde{L}^k$; and with † as in (6.3), we let

$$\mathfrak{h}_{\mathrm{div}} := \inf\Big\{ h \geq 0 \ \Big| \ Z^k_h \neq \widetilde{Z}^k_h \text{ or } Z^k_h = \widetilde{Z}^k_h = \dagger \Big\}, \tag{7.9}$$

so that $v_{\mathfrak{h}_{\mathrm{div}}-1}$ is the highest common ancestor of L^k and $\widetilde{L}^k$. Note that as soon as $L^k \neq \widetilde{L}^k$, then $\mathfrak{h}_{\mathrm{div}} < \min(\eta, \widetilde{\eta})$.

We record some useful facts in a preliminary lemma. The proof would follow straightforwardly from the arguments presented in Section 6.2 so we omit it.

Lemma 7.5 (Following the paths leading to the two marked leaves). *The following properties hold:*

1. *The two processes Z^k and $\widetilde{Z}^k$ have the same law.*

2. *The process $(Z^k_h, \widetilde{Z}^k_h)_{h \geq 0}$ is a Markov chain.*

3. *The divergence height $\mathfrak{h}_{\mathrm{div}}$ is a stopping time for the Markov chain $(Z^k_h, \widetilde{Z}^k_h)_{h \geq 0}$.*

4. *Whenever $L^k \neq \widetilde{L}^k$, the two leaves L^k and $\widetilde{L}^k$ belong respectively to one and the other subtree above $v_{\mathfrak{h}_{\mathrm{div}}-1}$ so that*

$$(M^*_{\mathfrak{h}_{\mathrm{div}}}, W_{\mathfrak{h}_{\mathrm{div}}}) = (\widetilde{W}_{\mathfrak{h}_{\mathrm{div}}}, \widetilde{M}^*_{\mathfrak{h}_{\mathrm{div}}}).$$

5. *Conditionally on $(Z^k_h, \widetilde{Z}^k_h)_{0 \leq h \leq \mathfrak{h}_{\mathrm{div}}}$, the two processes $(Z^k_{\mathfrak{h}_{\mathrm{div}}+h})_{h \geq 0}$ and $(\widetilde{Z}^k_{\mathfrak{h}_{\mathrm{div}}+h})_{h \geq 0}$ are independent.*

We then prove a first estimate, which roughly says that the explorations towards L_k and $\widetilde{L}_k$ do not stick together for too long.

Lemma 7.6 (Divergence occurs quickly). *There exists a constant $C > 0$ such that for any $k \geq A > 1$,*

$$\mathbb{P}_k\Big(M^*_{\mathfrak{h}_{\mathrm{div}}} < A \Big) \leq \frac{A^2}{k} + \frac{C}{A}.$$

Proof. Let $k \geq A > 1$. We write

$$\mathbb{P}_k\Big(M^*_{\mathfrak{h}_{\mathrm{div}}} < A \Big) = \mathbb{P}_k\Big(M^*_{\mathfrak{h}_{\mathrm{div}}} < A, M^*_{\mathfrak{h}_{\mathrm{div}}} + W_{\mathfrak{h}_{\mathrm{div}}} \leq A^2 \Big) + \mathbb{P}_k\Big(M^*_{\mathfrak{h}_{\mathrm{div}}} < A, M^*_{\mathfrak{h}_{\mathrm{div}}} + W_{\mathfrak{h}_{\mathrm{div}}} > A^2 \Big).$$

We show below that the first term is small because it would require picking the second leaf $\widetilde{L}^k$ quite close to L^k. For the second term, we show that such an event can only happen if the process M has an unlikely trajectory.

Assume that $L^k \neq \widetilde{L}^k$, meaning $\mathfrak{h}_{\mathrm{div}} < \eta$. Note that $S_{\mathfrak{h}_{\mathrm{div}}} = \oplus$ because L^k and $\widetilde{L}^k$ are leaves of the same maximal positive subtree. Thus we have $M^*_{\mathfrak{h}_{\mathrm{div}}} + W_{\mathfrak{h}_{\mathrm{div}}} = M^*_{\mathfrak{h}_{\mathrm{div}}-1}$. Recall from (6.9) the definition of the process M as a time-changed version of M^*, so that by the remark above

$M^*_{\mathfrak{h}_{\text{div}}} + W_{\mathfrak{h}_{\text{div}}}$ and $M^*_{\mathfrak{h}_{\text{div}}}$ are two consecutive values along the trajectory $\{M_t \mid t \geq 0\}$. Recall from (6.28), the random variable H_m which is such that $M(H_m) = \sup(\llbracket 1, m \rrbracket \cap \{M_t \mid t \geq 0\})$. For any $m \geq 1$, we can associate a node v^m in the original tree T^k, which corresponds to the lowest vertex on the spine from the root to the leaf L^k such that $T^{k,\text{lmax}}$ has at most m leaves (weakly) above v^m. We denote by $\mathcal{L}^{k,\text{lmax}}(v^m)$ the set of leaves of $T^{k,\text{lmax}}$ above vertex v^m, which has cardinality $M(H_m)$, that is at most m.

Consider the event $\{M^*_{\mathfrak{h}_{\text{div}}} < A, \, M^*_{\mathfrak{h}_{\text{div}}} + W_{\mathfrak{h}_{\text{div}}} \leq A^2\}$. By the above discussion, if this event is realized, then necessarily the leaf $\widetilde{L}^k$ belongs to $\mathcal{L}^{k,\text{lmax}}(v^{A^2})$ (note that this is obvious when $L^k = \widetilde{L}^k$). Since by definition this tree has $M(H_{A^2}) \leq A^2$ leaves, by reasoning conditionally on Z^k, we get

$$\mathbb{P}_k\left(M^*_{\mathfrak{h}_{\text{div}}} < A, \, M^*_{\mathfrak{h}_{\text{div}}} + W_{\mathfrak{h}_{\text{div}}} \leq A^2\right) \leq \mathbb{P}\left(\widetilde{L}^k \in \mathcal{L}^{k,\text{lmax}}(v^{A^2})\right) = \mathbb{E}\left[\frac{M(H_{A^2})}{k}\right] \leq \frac{A^2}{k}.$$

We now handle the second term. By a similar discussion, we have

$$\mathbb{P}_k\left(M^*_{\mathfrak{h}_{\text{div}}} < A, \, M^*_{\mathfrak{h}_{\text{div}}} + W_{\mathfrak{h}_{\text{div}}} > A^2\right) \leq \mathbb{P}_k(M(H_{A^2}) < A) = \mathbb{E}\left[\sum_{i=1}^{A-1} \pi(M(H_{A^2} - 1), i)\right]$$

$$\leq \sum_{i=1}^{A-1} C \frac{A}{A^2} q(i) \leq \frac{C}{A}$$

for some constant C, by the second bound in Item 1 of Lemma 6.12. $\qquad\square$

Recalling the first moment convergence provided by Lemma 7.4, we now show that the second moment converges.

Lemma 7.7 (The second moment converges). *The following assertions are true:*

1. *For any $k \geq A > 1$ we have*

$$\left|\mathbb{E}\left[\left(\frac{\#\mathcal{L}^{\text{max}}_{\cap}(T^k)}{k}\right)^2\right] - \lambda^2\right| \leq \mathbb{E}_k[F_A] + 3 \sup_{k \geq A}\left|\mathbb{E}\left[\frac{\#\mathcal{L}^{\text{max}}_{\cap}(T^k)}{k}\right] - \lambda\right| + 7\left(\frac{A^2}{k} + \frac{C}{A}\right).$$

2. *We have the convergence*

$$\mathbb{E}\left[\left(\frac{\#\mathcal{L}^{\text{max}}_{\cap}(T^k)}{k}\right)^2\right] \to \lambda^2 \qquad \text{as } k \to \infty.$$

The precise form of the first item will be used to obtain quantitative bounds in Theorem 9.1.

Proof. Conditionally on T^k, let L^k and $\widetilde{L}^k$ be two (possibly equal) independent leaves chosen uniformly at random among the k leaves of the leftmost maximal positive subtree $T^{k,\text{lmax}}$. Recall from (7.3) the random variable

$$F = \sum_{m \geq 1} \sum_{h \geq 1} \mathbb{1}\{S_h = \ominus \text{ and } M^*_h = W_h = m\},$$

counting the number of forks along the spine towards the marked leaf L^k. Denote by $\widetilde{F}$ the analogous quantity for the marked leaf $\widetilde{L}^k$. We have

$$\mathbb{E}\left[\left(\frac{\#\mathcal{L}^{\text{max}}_{\cap}(T^k)}{k}\right)^2\right] = \mathbb{P}_k\left(F = 0, \widetilde{F} = 0\right).$$

Recall that $\mathfrak{h}_{\mathrm{div}}$ is one plus the height of the highest common ancestor of L^k and $\widetilde{L}^k$ and set

$$F_{\leq \mathfrak{h}_{\mathrm{div}}} := \sum_{m \geq 1} \sum_{h=1}^{\mathfrak{h}_{\mathrm{div}}} \mathbb{1}\{S_h = \ominus \text{ and } M_h^* = W_h = m\},$$

$$F_{>\mathfrak{h}_{\mathrm{div}}} := F - F_{\leq \mathfrak{h}_{\mathrm{div}}} \quad \text{and} \quad \widetilde{F}_{>\mathfrak{h}_{\mathrm{div}}} := \widetilde{F} - F_{\leq \mathfrak{h}_{\mathrm{div}}},$$

where we note that $F_{\leq \mathfrak{h}_{\mathrm{div}}}$ is a function of $(Z_h^k, \widetilde{Z}_h^k)_{0 \leq h \leq \mathfrak{h}_{\mathrm{div}}}$ and $F_{>\mathfrak{h}_{\mathrm{div}}}$ and $\widetilde{F}_{>\mathfrak{h}_{\mathrm{div}}}$ are a function of $(Z_h^k, \widetilde{Z}_h^k)_{h \geq \mathfrak{h}_{\mathrm{div}}+1}$. With this notation, we get

$$\mathbb{E}\left[\left(\frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k}\right)^2\right] = \mathbb{P}_k\left(F_{\leq \mathfrak{h}_{\mathrm{div}}} = 0, F_{>\mathfrak{h}_{\mathrm{div}}} = 0, \widetilde{F}_{>\mathfrak{h}_{\mathrm{div}}} = 0\right). \tag{7.10}$$

By Lemma 7.5, we get that conditionally on $(Z_h^k, \widetilde{Z}_h^k)_{0 \leq h \leq \mathfrak{h}_{\mathrm{div}}}$, the processes $(Z_h^k)_{h \geq \mathfrak{h}_{\mathrm{div}}+1}$ and $(\widetilde{Z}_h^k)_{h \geq \mathfrak{h}_{\mathrm{div}}+1}$ are distributed as two independent versions of $(Z_h)_{h \geq 1}$ under the respective measures $\mathbb{P}_{M_{\mathfrak{h}_{\mathrm{div}}}^*}$ and $\mathbb{P}_{W_{\mathfrak{h}_{\mathrm{div}}}}$.[18] Introduce the notation

$$\epsilon(k) := \mathbb{P}_k(F = 0) - \lambda = \mathbb{E}\left[\frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k}\right] - \lambda.$$

This ensures that we have

$$\mathbb{P}_k\left(F_{>\mathfrak{h}_{\mathrm{div}}} = 0, \widetilde{F}_{>\mathfrak{h}_{\mathrm{div}}} = 0 \mid (Z_h^k, \widetilde{Z}_h^k)_{0 \leq h \leq \mathfrak{h}_{\mathrm{div}}}\right) = (\lambda + \epsilon(M_{\mathfrak{h}_{\mathrm{div}}}^*)) \cdot (\lambda + \epsilon(W_{\mathfrak{h}_{\mathrm{div}}})).$$

Using the last display allows us to rewrite the right-hand side of (7.10) as

$$\mathbb{P}_k\left(F_{\leq \mathfrak{h}_{\mathrm{div}}} = 0, F_{>\mathfrak{h}_{\mathrm{div}}} = 0, \widetilde{F}_{>\mathfrak{h}_{\mathrm{div}}} = 0\right)$$

$$= \mathbb{E}_k\left[\mathbb{1}_{F_{\leq \mathfrak{h}_{\mathrm{div}}}=0} \cdot (\lambda + \epsilon(M_{\mathfrak{h}_{\mathrm{div}}}^*)) \cdot (\lambda + \epsilon(W_{\mathfrak{h}_{\mathrm{div}}}))\right]$$

$$= \lambda^2 - \lambda^2 \mathbb{P}_k(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1) + \mathbb{E}_k\left[\mathbb{1}_{F_{\leq \mathfrak{h}_{\mathrm{div}}}=0} \cdot (\lambda \epsilon(M_{\mathfrak{h}_{\mathrm{div}}}^*) + \lambda \epsilon(W_{\mathfrak{h}_{\mathrm{div}}}) + \epsilon(M_{\mathfrak{h}_{\mathrm{div}}}^*)\epsilon(W_{\mathfrak{h}_{\mathrm{div}}}))\right]. \tag{7.11}$$

Let us first handle the term $\mathbb{P}_k(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1)$ in (7.11). For all $A > 1$, we can write

$$\mathbb{P}_k(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1) = \mathbb{P}_k\left(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1, M_{\mathfrak{h}_{\mathrm{div}}}^* \geq A\right) + \mathbb{P}_k\left(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1, M_{\mathfrak{h}_{\mathrm{div}}}^* < A\right).$$

Now on the event $\{M_{\mathfrak{h}_{\mathrm{div}}}^* \geq A\}$, we have $F_{\leq \mathfrak{h}_{\mathrm{div}}} \leq F_A$, where we recall that F_A was introduced in (7.4). Hence, by Markov's inequality,

$$\mathbb{P}_k\left(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1, M_{\mathfrak{h}_{\mathrm{div}}}^* \geq A\right) \leq \mathbb{P}_k(F_A \geq 1) \leq \mathbb{E}_k[F_A].$$

We conclude that

$$\mathbb{P}_k(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1) \leq \mathbb{E}_k[F_A] + \mathbb{P}_k\left(F_{\leq \mathfrak{h}_{\mathrm{div}}} \geq 1, M_{\mathfrak{h}_{\mathrm{div}}}^* < A\right) \leq \mathbb{E}_k[F_A] + \mathbb{P}_k\left(M_{\mathfrak{h}_{\mathrm{div}}}^* < A\right).$$

---

[18]Note that here we are using that the Markov property of $(Z_h^k)_h$ holds when conditioning either on the full state $Z_h^k = (S_h, D_h, M_h^*, W_h)$ or just on M_h^*, since the transitions depend only on this third coordinate; see Section 6.2.1.

Now let us control the last term in (7.11). We have

$$
\left| \mathbb{E}_k \left[\mathbb{1}_{F_{\leq \mathfrak{h}_{\mathrm{div}}}=0} \cdot (\lambda \epsilon(M^*_{\mathfrak{h}_{\mathrm{div}}}) + \lambda \epsilon(W_{\mathfrak{h}_{\mathrm{div}}}) + \epsilon(M^*_{\mathfrak{h}_{\mathrm{div}}}) \epsilon(W_{\mathfrak{h}_{\mathrm{div}}})) \right] \right|
$$

$$
\leq \mathbb{E}_k \left[\lambda |\epsilon(M^*_{\mathfrak{h}_{\mathrm{div}}})| + \lambda |\epsilon(W_{\mathfrak{h}_{\mathrm{div}}})| + |\epsilon(M^*_{\mathfrak{h}_{\mathrm{div}}}) \epsilon(W_{\mathfrak{h}_{\mathrm{div}}})| \right]
$$

$$
\leq \mathbb{E}_k \left[\mathbb{1}_{\{M^*_{\mathfrak{h}_{\mathrm{div}}} \geq A, W_{\mathfrak{h}_{\mathrm{div}}} \geq A\}} \cdot (\lambda |\epsilon(M^*_{\mathfrak{h}_{\mathrm{div}}})| + \lambda |\epsilon(W_{\mathfrak{h}_{\mathrm{div}}})| + |\epsilon(M^*_{\mathfrak{h}_{\mathrm{div}}}) \epsilon(W_{\mathfrak{h}_{\mathrm{div}}})|) \right]
$$

$$
+ 3 \mathbb{P}_k \left(\{ M^*_{\mathfrak{h}_{\mathrm{div}}} \geq A, W_{\mathfrak{h}_{\mathrm{div}}} \geq A \}^{\mathrm{c}} \right)
$$

$$
\leq 3 \sup_{m \geq A} |\epsilon(m)| + 3 \mathbb{P}_k \left(\{ M^*_{\mathfrak{h}_{\mathrm{div}}} \geq A, W_{\mathfrak{h}_{\mathrm{div}}} \geq A \}^{\mathrm{c}} \right),
$$

where we used the two bounds $\lambda \leq 1$ and $|\epsilon(\cdot)| \leq 1$. Finally, using a union bound and the fact that $M^*_{\mathfrak{h}_{\mathrm{div}}}$ and $W_{\mathfrak{h}_{\mathrm{div}}} = \widetilde{M}^*_{\mathfrak{h}_{\mathrm{div}}}$ have the same distribution by symmetry of L^k and $\widetilde{L}^k$, we can write

$$
\mathbb{P}_k \left(\{ M^*_{\mathfrak{h}_{\mathrm{div}}} \geq A, W_{\mathfrak{h}_{\mathrm{div}}} \geq A \}^{\mathrm{c}} \right) \leq 2 \mathbb{P} \left(M^*_{\mathfrak{h}_{\mathrm{div}}} < A \right).
$$

Putting everything together and using Lemma 7.6, we get

$$
\left| \mathbb{E} \left[\left(\frac{\# \mathcal{L}^{\max}_{\cap}(T^k)}{k} \right)^2 \right] - \lambda^2 \right| \leq \mathbb{E}_k[F_A] + 3 \sup_{m \geq A} |\epsilon(m)| + 7 \mathbb{P} \left(M^*_{\mathfrak{h}_{\mathrm{div}}} < A \right)
$$

$$
\leq \mathbb{E}_k[F_A] + 3 \sup_{k \geq A} \left| \mathbb{E} \left[\frac{\# \mathcal{L}^{\max}_{\cap}(T^k)}{k} \right] - \lambda \right| + 7 \left(\frac{A^2}{k} + \frac{C}{A} \right).
$$

Hence, Item 1 is proved.

The convergence in Item 2 is a consequence of Item 1, together with Lemma 7.3 and Lemma 7.4, by taking A going to infinity with k, with $A = o(\sqrt{k})$. $\qquad\square$

Lemma 7.4 and Lemma 7.7 entail our law of large numbers in Theorem 7.2.

Proof of Theorem 7.2. The result follows from Chebyshev's inequality since, for any $\delta > 0$,

$$
\mathbb{P} \left(\left| \frac{\# \mathcal{L}^{\max}_{\cap}(T^k)}{k} - \lambda \right| \geq \delta \right) \leq \frac{1}{\delta^2} \mathbb{E} \left[\left(\frac{\# \mathcal{L}^{\max}_{\cap}(T^k)}{k} - \lambda \right)^2 \right] \to 0 \qquad \text{as } k \to \infty,
$$

by Lemma 7.4 and Lemma 7.7. $\qquad\square$

7.2   Good regularity estimates for the sequences q and Q

With the law of large numbers of Theorem 7.2 in hand, we can now show the good regularity estimates for the sequences q and Q stated in Theorem 7.1.

Recall, from (5.3), the a.s. finite stopping times

$$
\tau_k := \min \{ n \geq 1 \mid X_n = k \}, \qquad k \geq 1.
$$

We first state a preliminary lemma, which improves the estimates in (5.8) of Lemma 5.3 and constitutes the main input towards Theorem 7.1. The constant λ was introduced in Theorem 7.2.

Lemma 7.8 (Asymptotics for the sequence $q(k)$). *As $k \to \infty$, we have*

$$
q(k) \sim \frac{1}{2\sqrt{\pi} \lambda k} \mathbb{E} \left[\tau_k^{-1/2} \right].
$$

Before proving the lemma, we show how it implies Theorem 7.1.

Proof of Theorem 7.1 given Lemma 7.8. The proof is basically the same as the end of the proof of Proposition 5.1 on page 36. By (5.7) in Lemma 5.3 and Lemma 7.8,

$$\frac{Q(k+1)}{Q(k)} = 1 - \frac{q(k)}{Q(k)} = 1 - \frac{1}{2\lambda k} + o\left(\frac{1}{k}\right) = \exp\left(-\frac{1}{2\lambda k} + o\left(\frac{1}{k}\right)\right) \quad \text{as } k \to \infty. \tag{7.12}$$

We first show the claim on Q. Let $x \geq 1$. Multiplying the estimates (7.12) over a telescopic product, we have as $k \to \infty$,

$$\frac{Q(xk)}{Q(k)} = \exp\left(\sum_{i=k}^{xk-1}\left(-\frac{1}{2\lambda i} + o\left(\frac{1}{i}\right)\right)\right) = \exp\left(-\frac{1}{2\lambda}\sum_{i=k}^{xk-1}\frac{1}{i} + o(1)\right) \to x^{-\frac{1}{2\lambda}}. \tag{7.13}$$

By Karamata's characterization theorem [BGT89, Theorem 1.4.1], we obtain the existence of a slowly varying function Φ such that $Q(k) = k^{-\frac{1}{2\lambda}}\Phi(k)$. In particular, Proposition 5.1 implies $\lambda = \alpha$, which proves the part of Theorem 7.1 on Q.

Finally, by comparing Lemma 7.8 with (5.7) in Lemma 5.3, we have

$$q(k) \sim \frac{1}{2\lambda k}Q(k) = \frac{1}{2\alpha}\Phi(k)k^{-1-\frac{1}{2\alpha}} \quad \text{as } k \to \infty,$$

which concludes the proof. □

The rest of this section is devoted to the proof of Lemma 7.8.

Proof of Lemma 7.8. We divide the proof into five steps. The core of the proof is Step 3, where we refine our proof of the lower bound on $q(k)$ in Lemma 5.3. The improvement will come from the precise asymptotics for $\#\mathcal{L}_\cap^{\max}(T^k)$ obtained in Theorem 7.2.

Step 1: Truncation. We fix a constant $K \geq 1$ (to be thought of as a large constant to be tuned at the end of the proof in Step 5) and for the rest of the proof we assume that $k > K$. The proof starts in the same way as the proof of Lemma 5.3: by the second equality in (5.11), for all $k \geq 1$, we have

$$q(k) = \mathbb{E}\left[\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)\right].$$

In view of Lemma 5.3 and (5.15) say, we expect $\tau_{k+1} - \tau_k$ to be typically of order $\frac{\tau_{k+1}}{k}$. Therefore, we split the sum in two, according as $n \geq \tau_{k+1} - \lfloor K\frac{\tau_{k+1}}{k}\rfloor$ or not. If $n \geq \tau_{k+1} - \lfloor K\frac{\tau_{k+1}}{k}\rfloor$, note that by the asymptotics in Lemma 3.3,

$$r(n) = \left(1 + O_K\left(\frac{1}{k}\right)\right)r(\tau_{k+1}),$$

that is, there is a constant C_K depending on K (but not on k) such that $|r(n)-r(\tau_{k+1})| \leq \frac{C_K}{k}r(\tau_{k+1})$. This implies

$$q(k) = \left(1 + O_K\left(\frac{1}{k}\right)\right)\mathbb{E}\left[r(\tau_{k+1})\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right)\right] + \mathbb{E}\left[\sum_{n\geq\tau_k}\mathbb{1}_{n<\tau_{k+1}-\lfloor K\frac{\tau_{k+1}}{k}\rfloor}r(n)\right].$$
$$\tag{7.14}$$

Roughly speaking, we will estimate the first term by applying the strategy of the lower bound of Lemma 5.3 in a more accurate way, where the additional precision will come from Theorem 7.2. The second term will be bounded crudely by the same proof as the upper bound of Lemma 5.3. Our main goal will be to show that all the error terms that will appear are small relative to $q(k)$.

Step 2: Neglecting the second term of Equation (7.14). Since it is easier, let us start with the second term. We note that if $n < \tau_{k+1} - \lfloor K\frac{\tau_{k+1}}{k} \rfloor$, then $\tau_{k+1} > n + \lfloor K\frac{\tau_k}{k} \rfloor$. Using the monotonicity of $r(\cdot)$ from Lemma 3.3, we can bound $r(n)$ by $r(\tau_k)$ and condition on τ_k to get

$$\mathbb{E}\left[\sum_{n \geq \tau_k} \mathbb{1}_{n < \tau_{k+1} - \lfloor K\frac{\tau_{k+1}}{k} \rfloor} r(n) \right] \leq \mathbb{E}\left[r(\tau_k) \sum_{n \geq \tau_k} \mathbb{P}\left(\tau_{k+1} > n + \left\lfloor K\frac{\tau_k}{k} \right\rfloor \ \middle| \ \tau_k \right) \right]. \qquad (7.15)$$

Using the change of variables $j = n - \tau_k + \lfloor K\frac{\tau_k}{k} \rfloor$, we can now use (5.14) to obtain, almost surely,

$$\sum_{n \geq \tau_k} \mathbb{P}\left(\tau_{k+1} > n + \left\lfloor K\frac{\tau_k}{k} \right\rfloor \ \middle| \ \tau_k \right) = \sum_{j \geq \lfloor K\frac{\tau_k}{k} \rfloor} \mathbb{P}(\tau_{k+1} > \tau_k + j \mid \tau_k)$$

$$\overset{(5.14)}{\leq} \sum_{j \geq \lfloor K\frac{\tau_k}{k} \rfloor} \left(1 + \frac{j}{\tau_k} \right)^{-p(k-1)}$$

$$\leq \int_{\lfloor K\tau_k/k \rfloor - 1}^{+\infty} \left(1 + \frac{t}{\tau_k} \right)^{-p(k-1)} \mathrm{d}t.$$

Now for k large enough so that the integral is finite, we find that

$$\int_{\lfloor K\tau_k/k \rfloor - 1}^{+\infty} \left(1 + \frac{t}{\tau_k} \right)^{-p(k-1)} \mathrm{d}t = \frac{\tau_k}{p(k-1) - 1} \left(1 + \frac{\lfloor K\frac{\tau_k}{k} \rfloor - 1}{\tau_k} \right)^{1-p(k-1)}$$

$$\leq \frac{\tau_k}{p(k-1) - 1} \exp\left((-pk + 1 + p) \log\left(1 + \frac{K-2}{k} \right) \right),$$

where for the bound in the log we used that $\tau_k \geq k$. We conclude that there exists a constant $C > 0$ (independent of K) and a threshold A_K such that for $k \geq A_K$,

$$\sum_{n \geq \tau_k} \mathbb{P}\left(\tau_{k+1} > n + \left\lfloor K\frac{\tau_k}{k} \right\rfloor \ \middle| \ \tau_k \right) \leq C\frac{\tau_k}{k} \mathrm{e}^{-pK}.$$

By integrating in τ_k, we finally obtain that for $k \geq A_K$,

$$\mathbb{E}\left[\sum_{n \geq \tau_k} \mathbb{1}_{n < \tau_{k+1} - \lfloor K\frac{\tau_{k+1}}{k} \rfloor} r(n) \right] \overset{(7.15)}{\leq} C\frac{\mathrm{e}^{-pK}}{k} \mathbb{E}[r(\tau_k)\tau_k] \overset{(3.5)}{\leq} C\frac{\mathrm{e}^{-pK}}{k} \mathbb{E}\left[\tau_k^{-1/2} \right] \overset{(5.8)}{\leq} C\mathrm{e}^{-pK}q(k), \quad (7.16)$$

for some other constant C. This bounds the second term of (7.14).

Step 3: Estimating the first expectation in Equation (7.14). For the first expectation in (7.14), we will be more precise and rely on Theorem 7.2. Recall the notation $\mathcal{L}_\cap^{\max}(t)$ for the set of leaves contained in all maximal positive subtrees of t. Let $\lambda \in (0, 1)$ be the constant appearing in Theorem 7.2. We now take $\varepsilon \in (0, 1)$ (to be thought of as a small constant to be tuned at the end of the proof in Step 5) and denote by $\mathcal{T}_\varepsilon$ the set of finite decorated trees t such that

$$\left| \frac{\#\mathcal{L}_\cap^{\max}(t)}{\mathrm{LIS}(t)} - \lambda \right| \geq \varepsilon.$$

In particular, Theorem 7.2 means that $\mathbb{P}(T \in \mathcal{T}_\varepsilon \mid \mathrm{LIS}(T) = k) \to 0$ as $k \to \infty$. For future reference, observe that by definition of τ_k and since T conditioned to have size n is distributed as T_n, this implies that as $k \to \infty$,

$$\mathbb{E}\left[\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)\mathbb{1}_{T_n \in \mathcal{T}_\varepsilon}\right] = \sum_{n \geq 1} r(n)\mathbb{P}(T_n \in \mathcal{T}_\varepsilon, X_n = k) = \mathbb{P}(T \in \mathcal{T}_\varepsilon, \mathrm{LIS}(T) = k) = o(q(k)). \quad (7.17)$$

Since our goal is to estimate the term

$$\mathbb{E}\left[r(\tau_{k+1}) \min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right)\right] = \mathbb{E}\left[\mathbb{E}\left[r(\tau_{k+1}) \min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right) \,\Big|\, \tau_{k+1}\right]\right]$$

in Equation (7.14), we will first be interested in estimating

$$\mathbb{P}(\tau_k < n - m \mid \tau_{k+1} = n), \qquad (7.18)$$

which will be achieved in (7.23). For this, we mimic (in a more precise way) the proof of the lower bound on $q(k)$ in Lemma 5.3.

For the rest of this step of the proof we fix $k \geq 2K/\varepsilon$. We start by looking at the event we condition on in (7.18). Let $n > k$. We recall from Section 2.1 that conditionally on $\overleftarrow{\mathcal{F}}_n$, the decorated tree T_{n-1} is obtained by removing a uniformly chosen leaf L_n of T_n. Therefore, if $X_n = k$, we have $\tau_k = n$ if and only if L_n belongs to the intersection of all maximal positive subtrees of T_n, i.e. almost surely,

$$\mathbb{P}\left(\tau_k = n \,\Big|\, \overleftarrow{\mathcal{F}}_n\right) = \frac{\#\mathcal{L}_\cap^{\max}(T_n)}{n}\mathbb{1}_{X_n=k}.$$

By distinguishing whether $T_n \in \mathcal{T}_\varepsilon$ or not, and in the case $T_n \in \mathcal{T}_\varepsilon$, using the crude bound $0 \leq \#\mathcal{L}_\cap^{\max}(T_n) \leq X_n$ and that $\lambda \in (0,1)$, we obtain for all $n > k$ and all $\varepsilon \in (0,1)$, that

$$\left|\mathbb{P}\left(\tau_k = n \,\Big|\, \overleftarrow{\mathcal{F}}_n\right) - \lambda\frac{k}{n}\mathbb{1}_{X_n=k}\right| \leq \frac{k}{n}(\varepsilon + \mathbb{1}_{T_n \in \mathcal{T}_\varepsilon}). \qquad (7.19)$$

We now turn to (7.18). For all $i, n \geq 1$ such that $i \leq n - 2$ and $n - i > k$, the events $\{\tau_{k+1} = n\}$ and $\{\tau_k \leq n - i\}$ belong to $\overleftarrow{\mathcal{F}}_{n-i}$ and have positive probability, so

$$\mathbb{P}(\tau_k < n - i \mid \tau_{k+1} = n, \tau_k \leq n - i) = 1 - \mathbb{E}\left[\mathbb{P}\left(\tau_k = n - i \,\Big|\, \overleftarrow{\mathcal{F}}_{n-i}\right) \,\Big|\, \tau_{k+1} = n, \tau_k \leq n - i\right].$$

In addition, note that on the event $\{\tau_{k+1} = n, \tau_k \leq n - i\}$, we have that $X_{n-i} = k$. Therefore, using the bound in (7.19) with $n - i$ in place of n, for all $i, n \geq 1$ such that $i \leq n - 2$ and $n - i > k$,

$$\left|\mathbb{P}(\tau_k < n - i \mid \tau_{k+1} = n, \tau_k \leq n - i) - 1 + \lambda\frac{k}{n-i}\right|$$
$$\leq \frac{k}{n-i}\left(\varepsilon + \mathbb{P}(T_{n-i} \in \mathcal{T}_\varepsilon \mid \tau_{k+1} = n, \tau_k \leq n - i)\right).$$

We now also assume $i \leq K\frac{n}{k}$ and define, for all $i, n \geq 1$ such that $i \leq K\frac{n}{k}$ and $n - i > k$,

$$x_i^{k,n} := \mathbb{P}(\tau_k < n - i \mid \tau_{k+1} = n, \tau_k \leq n - i) - 1 + \lambda\frac{k}{n}.$$

73

Recalling that we fixed $k \geq 2K/\varepsilon$, our last bound and the triangle inequality entail that if $1 \leq i \leq K\frac{n}{k}$ and $n - i > k$,

$$
\begin{aligned}
|x_i^{k,n}| &\leq \frac{k}{n-i}\left(\varepsilon + \mathbb{P}(T_{n-i} \in \mathcal{T}_\varepsilon \mid \tau_{k+1} = n, \tau_k \leq n - i)\right) + \lambda\left|\frac{k}{n-i} - \frac{k}{n}\right| \\
&\leq \frac{3k}{n}\left(\varepsilon + \mathbb{P}(T_{n-i} \in \mathcal{T}_\varepsilon \mid \tau_{k+1} = n, \tau_k \leq n - i)\right), \quad (7.20)
\end{aligned}
$$

where for the last inequality we used, for the first term, that $n - i \geq n(1 - K/k) \geq n(1 - \varepsilon/2) \geq n/2$ since $\varepsilon \in (0,1)$, and for the second term, that $\frac{k}{n-i} - \frac{k}{n} = \frac{k}{n}\frac{i}{n-i} \leq \frac{k}{n}\varepsilon$ since $i \leq K\frac{n}{k} \leq \frac{\varepsilon}{2}n$. By writing a telescopic product, for $j \geq 0$ and $n - j > k$, we have

$$
\mathbb{P}(\tau_k < n - j \mid \tau_{k+1} = n) = \prod_{i=1}^{j}\left(1 - \lambda\frac{k}{n} + x_i^{k,n}\right) = \left(1 - \lambda\frac{k}{n}\right)^j \exp\left(\sum_{i=1}^{j}\log\left(1 + \frac{x_i^{k,n}}{1 - \lambda k/n}\right)\right),
$$
$$(7.21)$$

where, when $j = 0$, we interpret the empty product as 1 and the empty sum as 0. Therefore, for $0 \leq j \leq K\frac{n}{k}$ and $n - j > k$, we write

$$
y_j^{k,n} := \sum_{i=1}^{j}\frac{|x_i^{k,n}|}{1 - \lambda k/n}. \quad (7.22)
$$

For later convenience, we also set $y_j^{k,n} := 0$ for all $0 \leq j \leq K\frac{n}{k}$ with $n - j \leq k$.

Since $n > n - j > k$, the denominators $1 - \lambda k/n$ are bounded away from 0 (recall that $\lambda < 1$). We deduce from (7.20) that for $0 \leq j \leq K\frac{n}{k}$ and $n - j > k$, the sum $y_j^{k,n}$ is bounded by a constant that depends only on K, and thus so is the sum

$$
\left|\sum_{i=1}^{j}\log\left(1 + \frac{x_i^{k,n}}{1 - \lambda k/n}\right)\right| \leq y_j^{k,n}.
$$

Therefore, by the mean value inequality, we can find a constant $C_K > 0$ that depends only on K such that for all $0 \leq j \leq K\frac{n}{k}$ and $n - j > k$,

$$
\left|\exp\left(\sum_{i=1}^{j}\log\left(1 + \frac{x_i^{k,n}}{1 - \lambda k/n}\right)\right) - 1\right| \leq C_K y_j^{k,n} \leq C_K y_{\lfloor Kn/k\rfloor}^{k,n}.
$$

We conclude that for all $0 \leq j \leq K\frac{n}{k}$,

$$
\mathbb{P}(\tau_k < n - j \mid \tau_{k+1} = n) = \mathbb{1}_{n-j>k}\left(1 - \lambda\frac{k}{n}\right)^j\left(1 + O_K\left(y_{\lfloor Kn/k\rfloor}^{k,n}\right)\right), \quad (7.23)
$$

since the probability on the left-hand side is zero when $n - j \leq k$. The latter estimate gives us that for $k \geq 2K/\varepsilon$, almost surely

$$
\begin{aligned}
\mathbb{E}\left[\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right) \,\middle|\, \tau_{k+1}\right] &= \sum_{j=0}^{\infty}\mathbb{P}\left(\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right) > j \,\middle|\, \tau_{k+1}\right) \\
&= \sum_{j=0}^{\lfloor K\frac{\tau_{k+1}}{k}\rfloor - 1}\mathbb{P}(\tau_k < \tau_{k+1} - j \mid \tau_{k+1}) \\
&= \sum_{j=0}^{J_K - 1}\left(\left(1 - \lambda\frac{k}{\tau_{k+1}}\right)^j\left(1 + O_K\left(y_{\lfloor K\tau_{k+1}/k\rfloor}^{k,\tau_{k+1}}\right)\right)\right),
\end{aligned}
$$

74

with $J_K := \min\left(\left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor, \tau_{k+1} - k\right)$. This is a geometric sum, so we end up with

$$\mathbb{E}\left[\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right) \,\Big|\, \tau_{k+1}\right] = \frac{\tau_{k+1}}{\lambda k}\left(1 - \left(1 - \lambda\frac{k}{\tau_{k+1}}\right)^{J_K}\right)\left(1 + O_K\left(y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\right)\right).$$

We now simplify the exponent J_K above. Notice that if $\tau_{k+1} - k < K\frac{\tau_{k+1}}{k}$, then we must have that $\tau_{k+1} < \frac{k}{1-\frac{K}{k}}$. Since $k \geq 2K/\varepsilon$ and $\varepsilon \in (0,1)$, this would imply $\tau_{k+1} < 2k$. Hence, if $\tau_{k+1} \geq 2k$, then $J_K = \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor$. Besides, if $J_K = \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor$, then $\left(1 - \lambda\frac{k}{\tau_{k+1}}\right)^{J_K} = O(e^{-\lambda K})$. Therefore, we can combine these two observations into

$$\mathbb{E}\left[\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right) \,\Big|\, \tau_{k+1}\right]\mathbb{1}_{\tau_{k+1}\geq 2k}$$
$$= \frac{\tau_{k+1}}{\lambda k}\left(1 - O\left(e^{-\lambda K}\right)\right)\left(1 + O_K\left(y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\right)\right)\mathbb{1}_{\tau_{k+1}\geq 2k}.$$

Integrating against $r(\tau_{k+1})$, we conclude that there is an absolute constant $C > 0$ (that may change from one inequality to the next) and a constant $C_K > 0$ depending on K such that for all $k \geq 2K/\varepsilon$,

$$\left|\mathbb{E}\left[r(\tau_{k+1})\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right)\mathbb{1}_{\tau_{k+1}\geq 2k}\right] - \mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}\mathbb{1}_{\tau_{k+1}\geq 2k}\right]\right|$$
$$\leq Ce^{-\lambda K}\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}\mathbb{1}_{\tau_{k+1}\geq 2k}\right] + C_K\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\mathbb{1}_{\tau_{k+1}\geq 2k}\right]$$
$$\overset{(3.5)}{\leq} C\frac{e^{-\lambda K}}{\lambda k}\mathbb{E}\left[\tau_{k+1}^{-1/2}\right] + C_K\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\mathbb{1}_{\tau_{k+1}\geq 2k}\right]$$
$$\overset{(5.8)}{\leq} Ce^{-\lambda K}q(k) + C_K\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\mathbb{1}_{\tau_{k+1}\geq 2k}\right]. \quad (7.24)$$

Although the indicator $\mathbb{1}_{\tau_{k+1}\geq 2k}$ is 1 with high probability as $k \to \infty$, we decide to keep it, as it will be useful in the next step of the proof.

We now estimate the first term on the right-hand side of (7.14) using (7.24): there is an absolute constant $C > 0$ and a (possibly different) constant $C_K > 0$ depending on K such that for all $k \geq 2K/\varepsilon$,

$$\left|\mathbb{E}\left[r(\tau_{k+1})\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right)\right] - \mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}\right]\right|$$
$$\leq Ce^{-\lambda K}q(k) + C_K\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\mathbb{1}_{\tau_{k+1}\geq 2k}\right] + C_K\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{k}\mathbb{1}_{\tau_{k+1}<2k}\right]. \quad (7.25)$$

Finally, we may further bound the last term as follows:

$$\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{k}\mathbb{1}_{\tau_{k+1}<2k}\right] \overset{(3.5)}{\leq} C\mathbb{E}\left[\frac{\tau_{k+1}^{-1/2}}{k}\mathbb{1}_{\tau_{k+1}<2k}\right] \leq C\mathbb{E}\left[\frac{\tau_{k+1}^{-1/2}}{k}\frac{2k}{\tau_{k+1}}\right] \leq 2C\mathbb{E}\left[\tau_{k+1}^{-3/2}\right].$$

By (5.6) and (5.8), we deduce that the last term of (7.25) is $o_K(q(k))$ as $k \to \infty$. We have reached the following bound: for all $k \geq 2K/\varepsilon$,

$$\left|\mathbb{E}\left[r(\tau_{k+1})\min\left(\tau_{k+1} - \tau_k, \left\lfloor K\frac{\tau_{k+1}}{k}\right\rfloor\right)\right] - \mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}\right]\right|$$
$$\leq Ce^{-\lambda K}q(k) + o_K(q(k)) + C_K\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k\rfloor}\mathbb{1}_{\tau_{k+1}\geq 2k}\right]. \quad (7.26)$$

The next step will be to bound the final error term in the above expression.

Step 4: Estimating the final error term in (7.26). Most of the remaining work is to show that the last term above is small compared to $q(k)$. By definition of $y_j^{k,n}$ in (7.22) and the bound (7.20), this ultimately boils down to bounding the probability that $T_n \in \mathcal{T}_\varepsilon$ for various values of n. Therefore, we will rewrite this error term in such a way that we can apply (7.17).

We start by fixing $k \geq 2K/\varepsilon$. By definition of $y_j^{k,n}$ in (7.22), recalling in particular that $y_j^{k,n} = 0$ for all $0 \leq j \leq K\frac{n}{k}$ and $n - j \leq k$, the last term of (7.26) can be written

$$\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y_{\lfloor K\tau_{k+1}/k\rfloor}^{k,\tau_{k+1}}\mathbb{1}_{\tau_{k+1}\geq 2k}\right] = \sum_{n=2k}^{\infty}\mathbb{P}(\tau_{k+1}=n)r(n)\frac{n}{\lambda k}\frac{1}{1-\lambda k/n}\mathbb{1}_{n-\lfloor Kn/k\rfloor>k}\sum_{i=1}^{\lfloor Kn/k\rfloor}|x_i^{k,n}|$$

$$\leq \frac{1}{\lambda(1-\lambda)}\sum_{n=2k}^{\infty}\mathbb{P}(\tau_{k+1}=n)r(n)\frac{n}{k}\mathbb{1}_{n-\lfloor Kn/k\rfloor>k}\sum_{i=1}^{\lfloor Kn/k\rfloor}|x_i^{k,n}|.$$

First, we note that the above indicator is always 1. Indeed, using $K/k \leq \varepsilon/2 < 1/2$ and $n \geq 2k$ we get $n - \lfloor Kn/k\rfloor > k$. Now, using the upper bound (7.20), we get that

$$\mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}y_{\lfloor K\tau_{k+1}/k\rfloor}^{k,\tau_{k+1}}\mathbb{1}_{\tau_{k+1}\geq 2k}\right] \leq \frac{3\varepsilon K}{\lambda(1-\lambda)}\sum_{n=2k}^{\infty}\mathbb{P}(\tau_{k+1}=n)r(n)\frac{n}{k}$$

$$+ \frac{3}{\lambda(1-\lambda)}\sum_{n=2k}^{\infty}\mathbb{P}(\tau_{k+1}=n)r(n)\sum_{i=1}^{\lfloor Kn/k\rfloor}\mathbb{P}(T_{n-i}\in\mathcal{T}_\varepsilon\mid\tau_{k+1}=n,\tau_k\leq n-i). \quad (7.27)$$

We now take care of the two sums separately. For the first sum we have that

$$\sum_{n=2k}^{\infty}\mathbb{P}(\tau_{k+1}=n)r(n)\frac{n}{k} = \mathbb{E}\left[r(\tau_{k+1})\frac{\tau_{k+1}}{k}\mathbb{1}_{\tau_{k+1}\geq 2k}\right] \overset{(3.5)}{\leq} \frac{C}{k}\mathbb{E}\left[\tau_{k+1}^{-1/2}\right] \overset{(5.8)}{\leq} Cq(k) \quad (7.28)$$

for some absolute constant $C > 0$ that may change from one inequality to the next.

We now turn to the second sum in (7.27). We write it as

$$\sum_{n=2k}^{\infty}\mathbb{P}(\tau_{k+1}=n)r(n)\sum_{i=1}^{\lfloor Kn/k\rfloor}\mathbb{P}(T_{n-i}\in\mathcal{T}_\varepsilon\mid\tau_{k+1}=n,\tau_k\leq n-i)$$

$$= \sum_{n=2k}^{\infty}r(n)\sum_{i=1}^{\lfloor Kn/k\rfloor}\frac{\mathbb{P}(T_{n-i}\in\mathcal{T}_\varepsilon,\tau_k\leq n-i,\tau_{k+1}=n)}{\mathbb{P}(\tau_k\leq n-i\mid\tau_{k+1}=n)}. \quad (7.29)$$

Let us now bound the denominators from below when $n \geq 2k$ and $1 \leq i \leq \lfloor Kn/k\rfloor$, recalling that we fixed $k \geq 2K/\varepsilon$. These bounds guarantee that $\frac{k}{n-i} \leq \frac{1}{1-\frac{\varepsilon}{2}}\frac{k}{n} < 1$. Therefore, we can use the crude bound (5.17) to obtain

$$\mathbb{P}(\tau_k\leq n-i\mid\tau_{k+1}=n)\geq\left(1-\frac{k}{n-i}\right)^i\geq\left(1-\frac{1}{1-\frac{\varepsilon}{2}}\frac{k}{n}\right)^i.$$

Using $\log(1-x)\geq-\frac{x}{1-x}$ for $0 < x < 1$, we obtain

$$\mathbb{P}(\tau_k\leq n-i\mid\tau_{k+1}=n)\geq\exp\left(-\frac{i\frac{k}{n}}{(1-\frac{\varepsilon}{2})-\frac{k}{n}}\right)\geq\exp\left(-\frac{2K}{1-\varepsilon}\right),$$

where, in the last inequality, we used that $i \le \lfloor Kn/k \rfloor$ and $n \ge 2k$. Then, setting $C_{K,\varepsilon} := \exp\left(\frac{2K}{1-\varepsilon}\right)$, (7.29) yields

$$\sum_{n=2k}^{\infty} \mathbb{P}(\tau_{k+1} = n) r(n) \sum_{i=1}^{\lfloor Kn/k \rfloor} \mathbb{P}(T_{n-i} \in \mathcal{T}_\varepsilon \mid \tau_{k+1} = n, \tau_k \le n - i)$$

$$\le C_{K,\varepsilon} \cdot \sum_{n=2k}^{\infty} r(n) \sum_{i=1}^{\lfloor Kn/k \rfloor} \mathbb{P}(T_{n-i} \in \mathcal{T}_\varepsilon, \tau_k \le n - i, \tau_{k+1} = n)$$

$$\le C_{K,\varepsilon} \cdot \mathbb{E}\left[\sum_{n=k+1}^{\infty} \mathbb{1}_{\tau_{k+1}=n}\, r(n) \sum_{i=1}^{\lfloor Kn/k \rfloor} \mathbb{1}_{\tau_k \le n-i} \mathbb{1}_{T_{n-i} \in \mathcal{T}_\varepsilon} \right]$$

$$= C_{K,\varepsilon} \cdot \mathbb{E}\left[r(\tau_{k+1}) \sum_{i=1}^{\lfloor K\tau_{k+1}/k \rfloor} \mathbb{1}_{\tau_k \le \tau_{k+1}-i} \mathbb{1}_{T_{\tau_{k+1}-i} \in \mathcal{T}_\varepsilon} \right]$$

$$\le C_{K,\varepsilon} \cdot \mathbb{E}\left[r(\tau_{k+1}) \sum_{j=\tau_k}^{\tau_{k+1}-1} \mathbb{1}_{T_j \in \mathcal{T}_\varepsilon} \right].$$

Since r is decreasing by Lemma 3.3, the last expectation is $o_\varepsilon(q(k))$ as $k \to \infty$ by (7.17). Hence, the second sum in (7.27) is $o_{K,\varepsilon}(q(k))$ as $k \to \infty$. Since the first sum in (7.27) was bounded by $Cq(k)$ in (7.28), we conclude (taking into account the additional constants in (7.27)) that

$$\mathbb{E}\left[r(\tau_{k+1}) \frac{\tau_{k+1}}{\lambda k} y^{k,\tau_{k+1}}_{\lfloor K\tau_{k+1}/k \rfloor} \mathbb{1}_{\tau_{k+1} \ge 2k} \right] \le \varepsilon C_K q(k) + o_{K,\varepsilon}(q(k)) \quad \text{as } k \to \infty. \tag{7.30}$$

Step 5: Conclusion. Let us come back to the formula (7.14) for $q(k)$. Recall the parameters $K \ge 1$ and $\varepsilon \in (0,1)$ that we have introduced before. We bound the second term on the right-hand side of (7.14) using (7.16), and estimate the first one with (7.26) (with its last error term bounded by (7.30)). We obtain for k large enough (where "large enough" depends on K and ε):

$$q(k) = \left(1 + O_K\left(\frac{1}{k}\right) \right) \mathbb{E}\left[r(\tau_{k+1}) \frac{\tau_{k+1}}{\lambda k} \right] + \left(1 + O_K\left(\frac{1}{k}\right) \right) E(\varepsilon, K, k), \tag{7.31}$$

where the error term $E(\varepsilon, K, k)$ is bounded by

$$|E(\varepsilon, K, k)| \le (Ce^{-CK} + \varepsilon C_K) q(k) + o_{K,\varepsilon}(q(k)) \tag{7.32}$$

for some absolute constant $C > 0$ and some constant $C_K > 0$ depending on K. We stress that the estimate (7.31) and the bound (7.32) hold for all $\varepsilon \in (0,1)$ and $K \ge 1$. We now tune the parameters ε and K. Let $\delta > 0$. There exist $K = K(\delta)$ (large enough) and $\varepsilon = \varepsilon(\delta) \in (0,1)$ (small enough) such that the first term on the right-hand side of (7.32) is bounded by $\delta q(k)$. The values of K and ε are now fixed (and only depend on δ). These values being fixed, we can now take $k_0 = k_0(\delta)$ large enough so that by (7.32), for all $k \ge k_0$,

$$|E(\varepsilon, K, k)| \le 2\delta q(k).$$

This estimate holds for all $\delta > 0$. Together with (7.31), this implies that

$$q(k) = \mathbb{E}\left[r(\tau_{k+1}) \frac{\tau_{k+1}}{\lambda k} \right] + o(q(k)) \quad \text{as } k \to \infty,$$

which means that $q(k) \sim \mathbb{E}\big[r(\tau_{k+1})\frac{\tau_{k+1}}{\lambda k}\big]$ as $k \to \infty$. We conclude, using once more Lemma 3.3 and (5.19), that

$$q(k) \sim \frac{1}{2\sqrt{\pi}\lambda k}\mathbb{E}\big[\tau_k^{-1/2}\big] \quad \text{as } k \to \infty,$$

which establishes Lemma 7.8. □

Remark 7.9. From now on, we know that $\lambda = \alpha$ in Theorem 7.2, as identified in the proof of Theorem 7.1 on page 71; see in particular the discussion below (7.13).

8 The value of the exponent α from the good regularity estimates for q and Q

The goal of this section is to complete the proof of Theorem 2.2, determining the exact value of the exponent α. We stress that so far α is only defined through Theorem 2.1, but we also know that the good regularity estimates in Theorem 7.1 hold for *the same* exponent α.

The latter estimates ensure that there are two slowly varying functions φ, Φ and an exponent

$$\gamma \overset{(2.7)}{=} \frac{1}{2\alpha} + 1 \in \left(\frac{3}{2}, 2\right) \tag{8.1}$$

such that

$$q(k) = k^{-\gamma}\varphi(k) \qquad \text{and} \qquad Q(k) = k^{1-\gamma}\Phi(k). \tag{8.2}$$

We will use these estimates to deduce the relation of Theorem 2.2 between the values of α and p.

8.1 The exact value of the exponent α

To identify the exact value of α, we first rewrite the recursive relation in Lemma 2.4 as

$$q(k) = \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i)q(k-i) + \frac{1-p}{2p}q(k)^2 + \mathbb{1}_{k \text{ even}}\frac{1}{2}q(k/2)^2\right). \tag{8.3}$$

Because we are not able to tackle this equation on the sequence q directly, we take interest in some related *inequations* that can be satisfied for a test sequence $\hat{q}$, namely

$$\hat{q}(k) \geq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i)\hat{q}(k-i) + \frac{1-p}{2p}q(k)^2 + \mathbb{1}_{k \text{ even}}\frac{1}{2}q(k/2)^2\right), \tag{8.4}$$

and

$$\hat{q}(k) \leq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i)\hat{q}(k-i)\right). \tag{8.5}$$

We will deduce Theorem 2.2 from the following two key ingredients:

1. A comparison lemma for test sequences (Lemma 8.1) which ensures that if we can check that one of the inequalities (8.4) or (8.5) holds for k large enough for some sequence $\hat{q}$, then we can compare the values of q and $\hat{q}$. This ingredient is partially inspired by [AC17].

2. The good regularity estimates in Theorem 7.1 (see (8.2) for a closer reference) and their consequences in Lemma 8.2 below.

We start by stating our main tool to compare the values of q and $\hat{q}$.

Lemma 8.1 (Comparison lemma for test sequences). *Let $(\hat{q}(k))_{k\geq 1}$ be a sequence of positive numbers.*

1. *If $\hat{q}(k)$ satisfies (8.4) for k large enough, then there exists a constant $C > 0$ such that*

$$\forall k \geq 1, \qquad q(k) \leq C\hat{q}(k).$$

2. *If $\hat{q}(k)$ satisfies (8.5) for k large enough, then there exists a constant $c > 0$ such that*

$$\forall k \geq 1, \qquad c\hat{q}(k) \leq q(k).$$

We proceed by stating the consequences of the good regularity estimates that we need here. Recall the exponent $\gamma \in (3/2, 2)$ related to α through (8.1), and the slowly varying function Φ appearing in (8.2).

Lemma 8.2 (Asymptotic expansion lemma for test sequences). *For $\hat{q}(k) = k^{-\beta}$ with $\beta \in (\frac{3}{2}, 2)$, we have as $k \to \infty$,*

- $$\left(\sum_{i=1}^{(k-1)/2} q(i)\hat{q}(k-i)\right) - \hat{q}(k) = k^{1-\gamma-\beta}\Phi(k)\left(-2^{\gamma+\beta-1} + \beta \int_0^{1/2} t^{1-\gamma}(1-t)^{-1-\beta}\,\mathrm{d}t + o(1)\right).$$

- $$\frac{1-p}{p}Q(k)\hat{q}(k) \sim \frac{1-p}{p}k^{1-\gamma-\beta}\Phi(k).$$

- $$\frac{1-p}{2p}q(k)^2 + \mathbb{1}_{k\,\text{even}}\frac{1}{2}q(k/2)^2 = o(k^{1-\gamma-\beta}\Phi(k)).$$

We will also need the following simple technical fact.

Lemma 8.3 (Monotonicity). *The function $f : \gamma \mapsto \frac{4^{\gamma-1}\sqrt{\pi}\,\Gamma(2-\gamma)}{\Gamma(\frac{3}{2}-\gamma)}$ is (strictly) decreasing on $(\frac{3}{2}, 2)$. Moreover, for any $x \in (3/2, 2)$, we have $\frac{4^{x-1}\sqrt{\pi}\,\Gamma(2-x)}{\Gamma(\frac{3}{2}-x)} = 2^{2x-1} - x \cdot \int_0^{1/2} t^{1-x}(1-t)^{-1-x}\,\mathrm{d}t.$*

We now prove Theorem 2.2 from Lemmas 8.1, 8.2 and 8.3.

Proof of Theorem 2.2 (assuming Lemmas 8.1, 8.2 and 8.3). For $\beta, \nu \in (\frac{3}{2}, 2)$ we write

$$g(\beta, \nu) = -2^{\nu+\beta-1} + \beta \int_0^{1/2} t^{1-\nu}(1-t)^{-1-\beta}\,\mathrm{d}t.$$

Note that $g(\beta, \beta) = -f(\beta)$ for f defined in Lemma 8.3. Recall the exponent $\gamma \in (3/2, 2)$ in (8.1) and suppose, for the sake of contradiction, that $g(\gamma, \gamma) > \frac{1-p}{p}$. Then by continuity, we also have $g(\beta, \gamma) > \frac{1-p}{p}$ for some $\beta \in (3/2, \gamma)$. By the asymptotics of Lemma 8.2, taking $\hat{q}(k) = k^{-\beta}$ with this choice of β we then have, for k large enough,

$$\frac{1-p}{p}Q(k)\hat{q}(k) \leq \left(\sum_{i=1}^{(k-1)/2} q(i)\hat{q}(k-i)\right) - \hat{q}(k),$$

so that (8.5) is satisfied.

Thanks to Item 2 in Lemma 8.1, there exists $c > 0$ such that $ck^{-\beta} = c\hat{q}(k) \leq q(k)$ for all $k \geq 1$. Yet by Theorem 7.1, $q(k) = k^{-\gamma}\varphi(k)$. This is absurd because $\beta < \gamma$.

Symmetrically, if $g(\gamma, \gamma) = -f(\gamma) < \frac{1-p}{p}$, then by continuity, we also have $g(\beta, \gamma) < \frac{1-p}{p}$ for some $\beta > \gamma$. The same argument as above shows that (8.4) is satisfied for k large enough if $\hat{q}(k) = k^{-\beta}$, so that by Lemma 8.1, we have $Ck^{-\beta} = C\hat{q}(k) \geq q(k)$ for all $k \geq 1$, for some constant $C > 1$. Because $\beta > \gamma$, this again contradicts the behavior of $q(k) = k^{-\gamma}\varphi(k)$ established in Theorem 7.1.

In the end, we proved that necessarily $g(\gamma, \gamma) = -f(\gamma) = \frac{1-p}{p}$. Since by Lemma 8.3 the function f is strictly decreasing on $(\frac{3}{2}, 2)$, this ensures that this equality determines γ uniquely as we wanted. Substituting $\alpha \overset{(2.7)}{=} \frac{1}{2(\gamma-1)}$, we get that the equivalent equation $1/f(\gamma) = \frac{p}{p-1}$ is exactly the equation in (1.3) for α. $\qquad\square$

In the following subsections, we give the missing proofs of Lemmas 8.1, 8.2 and 8.3.

8.2 The comparison lemma for test sequences

We give here the missing proof of Lemma 8.1.

Proof of Lemma 8.1. We start with the proof of Item 1. Assume that a sequence of positive numbers $(\hat{q}(k))_{k \geq 1}$ satisfies (8.4) for k larger than some value $k_0 \geq 1$. Let $C > 1$ be a constant chosen large enough so that

$$C\hat{q}(k) \geq q(k) \tag{8.6}$$

for all $1 \leq k \leq k_0$, which exists because we assumed $\hat{q}$ to be positive. We now prove by induction that this inequality also holds for all values of $k \geq 1$. Let $k \geq k_0 + 1$, and assume the bound $C\hat{q}(j) \geq q(j)$ holds for all $1 \leq j \leq k - 1$. Using sequentially the inequality (8.4) that holds for any $k > k_0$, the induction hypothesis, and the recursion equation (8.3) for q, we get

$$C\hat{q}(k) \geq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i) \cdot C\hat{q}(k-i) + C\frac{1-p}{2p}q(k)^2 + C\mathbb{1}_{k \text{ even}}\frac{1}{2}q(k/2)^2 \right)$$

$$\geq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i)q(k-i) + \frac{1-p}{2p}q(k)^2 + \mathbb{1}_{k \text{ even}}\frac{1}{2}q(k/2)^2 \right) = q(k).$$

Hence, the induction step is proved, so (8.6) holds for all $k \geq 1$, and Item 1 is proved.

The proof of Item 2 follows the exact same steps: Assume that a sequence of positive numbers $(\hat{q}(k))_{k \geq 1}$ satisfies (8.5) for k larger than some value $k_0 \geq 1$. Let $c > 0$ be a constant chosen small enough so that $c\hat{q}(k) \leq q(k)$ for all $1 \leq k \leq k_0$, which exists because q is positive. We now prove by induction that this inequality in fact holds for all values of $k \geq 1$. Let $k > k_0$, and assume the bound $c\hat{q}(j) \leq q(j)$ holds for all $1 \leq j \leq k - 1$. Using sequentially the inequality (8.5) that holds

for any $k \geq k_0$, the induction hypothesis, and the recursion equation (8.3) for q, we get

$$c\hat{q}(k) \leq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i) \cdot c\hat{q}(k-i) \right)$$

$$\leq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i) \cdot c\hat{q}(k-i) + \frac{1-p}{2p}q(k)^2 + \mathbb{1}_{k \text{ even}}\frac{1}{2}q(k/2)^2 \right)$$

$$\leq \frac{1}{1 + \frac{1-p}{p}Q(k)} \cdot \left(\sum_{i=1}^{(k-1)/2} q(i)q(k-i) + \frac{1-p}{2p}q(k)^2 + \mathbb{1}_{k \text{ even}}\frac{1}{2}q(k/2)^2 \right) = q(k).$$

This proves Item 2. $\qquad\square$

8.3 The asymptotic expansion lemma for test sequences

We give here the missing proof of Lemma 8.2.

Proof of Lemma 8.2. The second estimate is in fact an equality. The third one is immediate from the equality $q(k)^2 = k^{-2\gamma}\varphi(k)^2$ shown in Theorem 7.1, and the fact that $\gamma > \beta - 1$ (since $\beta, \gamma \in (3/2, 2)$). Therefore, the rest of the proof is devoted to the proof of the first estimate.

By Abel summation, the sum in the first estimate is

$$\sum_{i=1}^{(k-1)/2} q(i)\hat{q}(k-i)$$

$$= \sum_{i=1}^{(k-1)/2} Q(i)\hat{q}(k-i) - \sum_{i=2}^{(k+1)/2} Q(i)\hat{q}(k-i+1)$$

$$= Q(1)\hat{q}(k-1) - Q\left(\left\lfloor \frac{k+1}{2} \right\rfloor\right)\hat{q}\left(\left\lceil \frac{k+1}{2} \right\rceil\right) + \sum_{i=2}^{(k-1)/2} Q(i)(\hat{q}(k-i) - \hat{q}(k-i+1)).$$

Since $\hat{q}(k) = k^{-\beta}$, we have

$$\hat{q}(k-i) - \hat{q}(k-i+1) = \beta(k-i)^{-1-\beta} + O((k-i)^{-2-\beta}).$$

Using this estimate, together with the fact that $Q(1) = 1$ and $Q\left(\left\lfloor \frac{k+1}{2} \right\rfloor\right) \sim \left(\frac{k}{2}\right)^{1-\gamma}\Phi(k)$ thanks to Theorem 7.1, we can rewrite the above sum as

$$\sum_{i=1}^{(k-1)/2} q(i)\hat{q}(k-i) = k^{-\beta}(1 + O(k^{-1})) - \Phi(k)\left(\frac{k}{2}\right)^{1-\gamma-\beta}(1 + o(1))$$

$$+ \beta\Phi(k)k^{1-\gamma-\beta}\left(\frac{1}{k}\sum_{i=2}^{(k-1)/2} \frac{\Phi(i)}{\Phi(k)}\left(\frac{i}{k}\right)^{1-\gamma}\left(1 - \frac{i}{k}\right)^{-1-\beta} \right) + O(k^{-1-\beta}), \quad (8.7)$$

where for the last error term we used that

$$\sum_{i=2}^{(k-1)/2} Q(i)(k-i)^{-2-\beta} \leq \sum_{i=2}^{(k-1)/2} (k-i)^{-2-\beta} = O(k^{-1-\beta}).$$

81

We now focus on the sum on the right-hand side of (8.7). We split it into two parts: one for small values of i and one for large values of i. Fix $\varepsilon > 0$. For the values of $i \geq \varepsilon k$, by the uniform convergence theorem recalled in Item 1 of Lemma 3.1, we have that $\frac{\Phi(i)}{\Phi(k)} = 1 + o(1)$ as $k \to \infty$, uniformly for all $i \in [\![\varepsilon k, (k-1)/2]\!]$. Hence, by Riemann sum approximation,

$$\frac{1}{k} \sum_{i=\lfloor \varepsilon k \rfloor}^{(k-1)/2} \frac{\Phi(i)}{\Phi(k)} \left(\frac{i}{k}\right)^{1-\gamma} \left(1 - \frac{i}{k}\right)^{-1-\beta} = (1+o(1)) \frac{1}{k} \sum_{i=\lfloor \varepsilon k \rfloor}^{(k-1)/2} \left(\frac{i}{k}\right)^{1-\gamma} \left(1 - \frac{i}{k}\right)^{-1-\beta}$$

$$= (1+o(1)) \int_{\varepsilon}^{1/2} t^{1-\gamma} (1-t)^{-1-\beta} \, \mathrm{d}t.$$

We now focus on the values of $i < \varepsilon k$. Let $\delta > 0$ be a small parameter to be fixed later. Using the Potter bounds recalled in Item 2 of Lemma 3.1, there exists a constant $C > 1$ such that

$$\frac{\Phi(i)}{\Phi(k)} \leq C \left(\frac{i}{k}\right)^{-\delta} \qquad \text{for all } i, k \geq 1 \text{ such that } i < \varepsilon k.$$

Therefore, again by Riemann sum approximation,

$$\frac{1}{k} \sum_{i=2}^{\lfloor \varepsilon k \rfloor - 1} \frac{\Phi(i)}{\Phi(k)} \left(\frac{i}{k}\right)^{1-\gamma} \left(1 - \frac{i}{k}\right)^{-1-\beta} \leq C \frac{1}{k} \sum_{i=2}^{\lfloor \varepsilon k \rfloor - 1} \left(\frac{i}{k}\right)^{1-\gamma-\delta} \left(1 - \frac{i}{k}\right)^{-1-\beta}$$

$$= C(1+o(1)) \int_{0}^{\varepsilon} t^{1-\gamma-\delta} (1-t)^{-1-\beta} \, \mathrm{d}t.$$

Since $\gamma < 2$, we can always choose $\delta > 0$ small enough so that $1 - \gamma - \delta > -1$, ensuring that the integral above is finite.

Putting together the estimates for small and large values of i and sending ε to zero, we get

$$\frac{1}{k} \sum_{i=2}^{(k-1)/2} \frac{\Phi(i)}{\Phi(k)} \left(\frac{i}{k}\right)^{1-\gamma} \left(1 - \frac{i}{k}\right)^{-1-\beta} = (1+o(1)) \int_{0}^{1/2} t^{1-\gamma} (1-t)^{-1-\beta} \, \mathrm{d}t.$$

Substituting this last estimate in (8.7), and noting that $O(k^{-1-\beta}) = o(k^{1-\gamma-\beta} \Phi(k))$ since $\gamma < 2$ and $k^{\gamma-2} = o(\Phi(k))$, we obtain the first estimate of Lemma 8.2. $\qquad \square$

8.4 The monotonicity lemma

We finally give the missing proof of Lemma 8.3.

Proof of Lemma 8.3. We first prove that the function $f : \gamma \mapsto \frac{4^{\gamma-1}\sqrt{\pi}\, \Gamma(2-\gamma)}{\Gamma(\frac{3}{2}-\gamma)}$ is strictly decreasing on $(\frac{3}{2}, 2)$. Since $\frac{4^{\gamma-1}\sqrt{\pi}\, \Gamma(2-\gamma)}{\Gamma(3/2-\gamma)} < 0$ for all $\gamma \in (3/2, 2)$, it is enough to show that

$$\frac{\partial}{\partial \gamma} \log(-f(\gamma)) > 0 \qquad \text{for all } \gamma \in (3/2, 2). \tag{8.8}$$

We have that

$$\frac{\partial}{\partial \gamma} \log(-f(\gamma)) = \frac{\partial}{\partial \gamma} \left(\log(4^{\gamma-1}) + \log(\sqrt{\pi}) + \log(\Gamma(2-\gamma)) - \log(-\Gamma(3/2-\gamma)) \right)$$

$$= \log(4) - \psi(2-\gamma) + \psi(3/2-\gamma),$$

where $\psi(\cdot)$ denotes the Digamma function $\psi = \Gamma'/\Gamma$. Using that $\psi(x) = \psi(x+1) - 1/x$ and that for all $x > 0$, $\log(x) - 1/x \leq \psi(x) \leq \log(x) - 1/(2x)$ (see e.g. [BE53, Chap. 1.7.1] and [Alz97] respectively), we get the bound

$$\frac{\partial}{\partial\gamma}\log(-f(\gamma)) \geq \log\left(\frac{4(5/2 - \gamma)}{2 - \gamma}\right) + \frac{1}{4 - 2\gamma} - \frac{1}{3/2 - \gamma} - \frac{1}{5/2 - \gamma}.$$

Then a standard analysis of the function on the right-hand side of the above equation shows that (8.8) holds.

Finally, the second part of the lemma statement can be directly proved using basic properties of incomplete beta functions. More precisely, the integral that we want to express is nothing but the incomplete beta function

$$B_{1/2}(2 - x, -x) := \int_0^{1/2} t^{1-x}(1 - t)^{-1-x}\mathrm{d}t, \quad x < 2.$$

The integration by parts relations for incomplete beta functions (see e.g. [BE53, Chap. 2.5.3]) give

$$I_{1/2}(a, b) = I_{1/2}(a, b + 1) - \frac{x^a(1 - x)^b}{bB(a, b)},$$

where B denote the usual beta function and $I_{1/2}(a, b) := \frac{B_{1/2}(a,b)}{B(a,b)}$. Using twice these relations, we are able to express $B_{1/2}(2 - x, -x)$ in terms of (usual) beta functions and $B_{1/2}(2 - x, 2 - x)$. But by symmetry $B_{1/2}(a, a) = \frac{1}{2}B(a, a)$, so we are able to express $B_{1/2}(2 - x, -x)$ in terms of beta functions, and hence ultimately of gamma functions. The formula claimed in Lemma 8.3 then comes from standard bookkeeping. □

9   Scaling limit results

The main goal of this section is to establish the scaling limit result for $X_n = \mathrm{LIS}(T_n)$ as $n \to \infty$, stated in Theorem 2.3. This will be carried out in Sections 9.2, 9.3 and 9.4. It is worth emphasizing that this result does not require any knowledge of the precise value of the exponent α, which was determined in the previous section. Then, in Section 9.5, we present an argument that completes the proof of Theorem 1.1, and in Section 9.6, we establish a scaling limit result for the tree $T^{k,\mathrm{lmax}}$ as $k \to \infty$, proving Theorem 2.6.

9.1   The scaling limit of the LIS: proof outline

We start with an outline of the proof of Theorem 2.3. Recall the notation $\overleftarrow{\mathcal{F}}_n = \sigma(T_m, \ m \geq n)$ for the σ-algebra generated by the tree process after time n, as defined by Rémy's algorithm. We recall that for every $n \geq 1$, the tree T_n is obtained by removing a uniform leaf L_{n+1} from T_{n+1}. We write

$$X_n = X_{n+1} - \mathbb{1}_{L_{n+1} \in \mathcal{L}_\cap^{\max}(T_{n+1})},$$

so that for $n \geq 1$,

$$\mathbb{E}\left[X_n \mid \overleftarrow{\mathcal{F}}_{n+1}\right] = X_{n+1} - \frac{\#\mathcal{L}_\cap^{\max}(T_{n+1})}{n + 1}$$

$$= X_{n+1}\left(1 - \frac{\alpha}{n + 1}\right) + \frac{1}{n + 1}(\alpha X_{n+1} - \#\mathcal{L}_\cap^{\max}(T_{n+1})),$$

with the idea that the second term on the right-hand side of the last display should be negligible in some sense, as suggested by Theorem 7.2. From there, it is quite natural to introduce the sequence $(c_n)_{n \geq 1}$ defined as

$$c_n := \prod_{i=1}^{n} \left(1 - \frac{\alpha}{i} \right) = \prod_{i=1}^{n} \frac{i - \alpha}{i} = \frac{\Gamma(n + 1 - \alpha)}{\Gamma(n+1)\Gamma(1-\alpha)}, \tag{9.1}$$

so that the rescaled process

$$\mathcal{X}_n := c_n X_n \tag{9.2}$$

satisfies

$$\mathbb{E}\left[\mathcal{X}_n \mid \overleftarrow{\mathcal{F}}_{n+1} \right] = \mathcal{X}_{n+1} + R_{n+1}, \quad \text{with} \quad R_{n+1} := \frac{c_n}{n+1}(\alpha X_{n+1} - \#\mathcal{L}_{\cap}^{\max}(T_{n+1})). \tag{9.3}$$

Now, if the remainder terms R_n were identically zero, then the process $(\mathcal{X}_n)_{n \geq 1}$ would be a backward martingale and would thus converge almost surely as $n \to \infty$. Since, by properties of the Gamma function, it is easy to check that

$$c_n \underset{n \to \infty}{\sim} \frac{1}{\Gamma(1-\alpha)} n^{-\alpha}, \tag{9.4}$$

this would show that $n^{-\alpha} X_n$ converges almost surely, as stated in Theorem 2.3.

Following this intuition, the strategy of our proof consists of showing that the remainder terms R_n are small in some quantifiable way. For this, we rely on the quantitative version of Theorem 7.2 stated below, where we recall that T^k is a p-signed critical binary Bienaymé–Galton–Watson tree T conditioned on the event $\{\mathrm{LIS}(T) = k\}$. We recall that we identified at this point the constant λ in Theorem 7.2 with the exponent α in Theorem 2.1; see Remark 7.9.

Theorem 9.1 (Quantitative version of Theorem 7.2). *There exists some $\varepsilon > 0$ such that*

$$\mathbb{E}\left[\left\| \frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k} - \alpha \right\| \right] = O(k^{-\varepsilon}). \tag{9.5}$$

This theorem ensures that for a p-signed critical binary Bienaymé–Galton–Watson tree T conditioned on having $\mathrm{LIS}(T) = k$, the number of leaves $\#\mathcal{L}_{\cap}^{\max}(T)$ in the intersection of all its maximal positive subtrees is close to $\alpha \, \mathrm{LIS}(T) = \alpha k$. In our case, we can see from (9.3) that R_n is rather defined as (a constant times) the corresponding quantity for T_n, where we recall that T_n has the distribution of T conditioned on $\{|T| = n\}$. Nevertheless, with some additional work, we manage to translate the information of Theorem 9.1 under this different conditioning into some usable result on the remainders R_n.

Below, we first prove Theorem 9.1 in Section 9.2. We then turn it into Lemma 9.2, which is a quantitative result about the remainders R_n, through a series of technical lemmas (Lemmas 9.3, 9.4, 9.5 and 9.6) in Section 9.3. Based on that, we will be able in Section 9.4 to make rigorous sense of the martingale arguments sketched above, and to conclude the proof of Theorem 2.3.

9.2   A quantitative version of the law of large numbers

This section is devoted to the proof of Theorem 9.1. Recall the definition of good scales from Definition 6.10. The proof of Theorem 9.1 relies on the fact that the good regularity estimates from Theorem 7.1 guarantee (as explained in Remark 6.11) the existence of a constant $C > 0$ such that every integer satisfies the good scale condition from Item 1 in Definition 6.10.

Thus, in what follows, we set

$$\text{GoodScales} := \{10^m \mid m \geq 0\}$$

to be our new set of good scales.

Proof of Theorem 9.1. From Lemma 7.4 and Lemma 7.7 we know that the first and second moments of $\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}$ converge towards respectively α and α^2 (recall that $\lambda = \alpha$). We hence write

$$\mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}\right] = \alpha + \delta_1(k) \qquad \text{and} \qquad \mathbb{E}\left[\left(\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}\right)^2\right] = \alpha^2 + \delta_2(k), \tag{9.6}$$

with $\delta_1(k), \delta_2(k) \to 0$ as $k \to \infty$. This ensures that

$$\mathbb{E}\left[\left(\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k} - \alpha\right)^2\right] = \mathbb{E}\left[\left(\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}\right)^2\right] - 2\alpha\mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}\right] + \alpha^2$$
$$= \delta_2(k) - 2\alpha\delta_1(k),$$

and hence by Jensen's inequality, we get

$$\mathbb{E}\left[\left|\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k} - \alpha\right|\right] \leq \sqrt{\mathbb{E}\left[\left(\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k} - \alpha\right)^2\right]} = \sqrt{\delta_2(k) - 2\alpha\delta_1(k)}. \tag{9.7}$$

From there, it is clear that if we obtain that $\delta_1(k), \delta_2(k) = O(k^{-\varepsilon})$ for some $\varepsilon > 0$, *i.e.* that the convergence (9.6) of the first and second moments happens at least with polynomial speed, then we get the result claimed in the theorem. We fix $k \geq 1$ and take care of the two terms separately.

First, we bound $\delta_1(k)$. Item 1 in Lemma 7.4 ensures that for any $k' \geq k$ and any $A > 0$, we have

$$\left|\mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}\right] - \mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^{k'})}{k'}\right]\right| \leq 2\sup_{k \geq 1}\mathbb{E}_k[F_A] + \widetilde{\mathbb{P}}_{k,k'}(V_{\text{mer}} \leq A). \tag{9.8}$$

We fix $A = k^{1/4}$ and estimate the two terms on the right-hand side of (9.8) in turn. First, by Lemma 7.3, for some constant $C > 0$ we have

$$\sup_{k \geq 1}\mathbb{E}_k[F_A] \leq C \cdot Q(A) = C \cdot Q(k^{1/4}) = k^{-\frac{1}{8\alpha}+o(1)}, \tag{9.9}$$

where the last equality comes from the rough estimate on Q in Proposition 5.1. Second, the number of good intervals included in $[\![k^{1/4}, k]\!]$ is $s(k^{1/4}, k) = \frac{3\log k}{4\log 10}(1 + o(1))$. Therefore, by Item 1 of Corollary 6.7, there exists a constant $c > 0$ such that

$$\widetilde{\mathbb{P}}_{k,k'}\left(V_{\text{mer}} \leq k^{1/4}\right) \leq (1-c)^{s(k^{1/4},k)} = (1-c)^{\frac{3\log k}{4\log 10}(1+o(1))} \leq k^{\frac{3\log(1-c)}{4\log 10}(1+o(1))} \leq k^{-\varepsilon+o(1)}$$

as $k \to \infty$, for some $\varepsilon > 0$. Since we have upper bounded the two terms on the right-hand side of (9.8) by quantities that do not depend on k', we can then take $k' \to \infty$ in that inequality and get, for $\varepsilon' > 0$ small enough,

$$|\delta_1(k)| = \left|\mathbb{E}\left[\frac{\#\mathcal{L}_\cap^{\max}(T^k)}{k}\right] - \alpha\right| \leq k^{-\frac{1}{8\alpha}+o(1)} + k^{-\varepsilon+o(1)} = O(k^{-\varepsilon'}). \tag{9.10}$$

Second, we bound $\delta_2(k)$. We use Item 1 of Lemma 7.7: for any $k \geq A > 1$, we have

$$|\delta_2(k)| = \left| \mathbb{E}\left[\left(\frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k} \right)^2 \right] - \lambda^2 \right| \leq \mathbb{E}_k[F_A] + 3 \sup_{\mathsf{k} \geq A} |\delta_1(\mathsf{k})| + 7\left(\frac{A^2}{k} + \frac{C}{A} \right).$$

We again take $A = k^{1/4}$ and use (9.9) and (9.10) to obtain

$$|\delta_2(k)| \leq k^{-\frac{1}{8\alpha}+o(1)} + O(k^{-\frac{\varepsilon}{8}}) + O(k^{-\frac{1}{4}}) = O(k^{-\varepsilon'}) \tag{9.11}$$

for some $\varepsilon' > 0$ small enough.

Finally, substituting the bounds (9.10) and (9.11) into (9.7) yields the claimed result. $\qquad\square$

9.3 Control of the remainders

The goal of this subsection is to prove the following lemma, which allows us to control the sum of the remainder terms R_n introduced in (9.3).

Lemma 9.2 (Control on the remainders). *The random sum $\sum_{n=2}^{\infty} |R_n|$ is almost surely finite. Furthermore, there exists $\varepsilon > 0$ such that almost surely, for n_0 large enough,*

$$\sum_{n=n_0}^{\infty} |R_n| \leq n_0^{-\varepsilon}.$$

We point out that in this section, as in Lemma 9.2, whenever we write "almost surely, for n large enough," the threshold for "large enough" may be random. The proof will rely on four intermediate lemmas, which we will state and prove as we proceed. The proof of Lemma 9.2 will then be provided at the end of this section. Recall from Section 5 the notation τ_k for the hitting time of k by X_n introduced in (5.3), as well as the probability $r(n) = \mathbb{P}(|T| = n)$ from Lemma 3.3. For clearer expressions, we introduce the following notation: for a finite sign-decorated binary tree t, we write

$$f(t) := |\alpha \operatorname{LIS}(t) - \#\mathcal{L}_{\cap}^{\max}(t)|.$$

Note that since $|R_n| = \frac{c_{n-1}}{n}|\alpha X_n - \#\mathcal{L}_{\cap}^{\max}(T_n)| = \frac{c_{n-1}}{n}f(T_n)$, we can write for any $n_0 \geq 2$,

$$\sum_{n=n_0}^{\infty} |R_n| = \sum_{n=n_0}^{\infty} \frac{c_{n-1}}{n}f(T_n). \tag{9.12}$$

Looking at the form of the expression (9.12), we can already see some similarity with the quantities appearing (in expectation) in the following rephrasing of Theorem 9.1.

Lemma 9.3 (Rephrasing of Theorem 9.1). *There is $\varepsilon > 0$ such that we have*

$$\mathbb{E}\left[\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n) \cdot f(T_n) \right] = O(k^{1-\varepsilon}q(k)) \quad \text{as } k \to \infty. \tag{9.13}$$

Proof. First, Theorem 9.1 tells us that there exists $\varepsilon > 0$ such that

$$\mathbb{E}\left[f(T^k) \right] = \mathbb{E}\left[|\alpha \operatorname{LIS}(T^k) - \#\mathcal{L}_{\cap}^{\max}(T^k)| \right] = k\mathbb{E}\left[\left| \frac{\#\mathcal{L}_{\cap}^{\max}(T^k)}{k} - \alpha \right| \right] = O(k^{1-\varepsilon}).$$

Then, recalling that $r(n) = \mathbb{P}(|T| = n)$, we can rewrite the left-hand side as

$$\mathbb{E}\Big[f(T^k)\Big] = \frac{1}{\mathbb{P}(|T| = k)} \cdot \mathbb{E}\Big[f(T) \cdot \mathbb{1}_{\mathrm{LIS}(T)=k}\Big] = \frac{1}{q(k)} \sum_{n=1}^{\infty} r(n) \cdot \mathbb{E}\Big[f(T_n)\mathbb{1}_{\mathrm{LIS}(T_n)=k}\Big]$$

$$= \frac{1}{q(k)} \mathbb{E}\left[\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n) \cdot f(T_n)\right].$$

Combining the last two displays finishes the proof. □

The next lemma aims to upper bound the sum of the remainders R_n in terms of the quantities controlled (in expectation) by the previous result.

Lemma 9.4 (Upper bound on the sum of remainders). *There exists a constant $C > 0$ such that, almost surely, for any $n_0 \geq 2$, we have*

$$\sum_{n=n_0}^{\infty} |R_n| \leq C \cdot \sum_{k=X_{n_0}}^{\infty} (\tau_k)^{\frac{1}{2}-\alpha} \left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right).$$

Proof. Using (9.12), we write

$$\sum_{n=n_0}^{\infty} |R_n| = \sum_{n=n_0}^{\infty} \frac{c_{n-1}}{n} f(T_n) \leq \sum_{n=\tau_{X_{n_0}}}^{\infty} \frac{c_{n-1}}{n} f(T_n)$$

$$= \sum_{k=X_{n_0}}^{\infty} \left(\sum_{n=\tau_k}^{\tau_{k+1}-1} \frac{c_{n-1}}{n} f(T_n)\right). \tag{9.14}$$

Since $c_n \sim \frac{1}{\Gamma(1-\alpha)} n^{-\alpha}$ by (9.4) and $r(n) \sim c n^{-3/2}$ thanks to Lemma 3.3, we bound the term appearing in front of $f(T_n)$ in the sum as

$$\frac{c_{n-1}}{n} \leq C n^{\frac{1}{2}-\alpha} r(n),$$

for some constant $C > 0$. We substitute this into (9.14) and additionally use the fact that if $n \geq \tau_k$ then $n^{\frac{1}{2}-\alpha} \leq (\tau_k)^{\frac{1}{2}-\alpha}$ because $\frac{1}{2} - \alpha < 0$ (recall Theorem 1.2). We get

$$\sum_{n=n_0}^{\infty} \frac{c_{n-1}}{n} f(T_n) \leq C \cdot \sum_{k=X_{n_0}}^{\infty} (\tau_k)^{\frac{1}{2}-\alpha} \left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right),$$

which finishes the proof. □

In the next lemma, we use the estimates from Lemma 9.3 to further control the upper bound in Lemma 9.4.

Lemma 9.5 (Polynomial tail control on the random sum). *The random sum*

$$\sum_{k=1}^{\infty} (\tau_k)^{\frac{1}{2}-\alpha} \left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right)$$

is almost surely finite. Furthermore, there exists $\varepsilon > 0$ such that almost surely, for k_0 large enough,

$$\sum_{k=k_0}^{\infty} (\tau_k)^{\frac{1}{2}-\alpha} \left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right) \leq k_0^{-\varepsilon}.$$

Proof of Lemma 9.5. Let $\delta > 0$ be a small parameter to be fixed later. By (2.2) in Theorem 2.1, there is a constant $C > 0$ such that, for all $n \geq 1$, $\mathbb{E}[X_n] \leq Cn^{\alpha+\delta/2}$. Using Markov's inequality for X_{2^i}, for all $i \geq 1$, we have

$$\mathbb{P}\Big(X_{2^i} \geq (2^{i-1})^{\alpha+\delta}\Big) \leq (2^{i-1})^{-\alpha-\delta}\mathbb{E}[X_{2^i}] \leq C2^{-i\delta/2}$$

for some (other) constant $C > 0$. Therefore

$$\sum_{i\geq 1}\mathbb{P}\Big(X_{2^i} \geq (2^{i-1})^{\alpha+\delta}\Big) < \infty.$$

By the Borel–Cantelli lemma and the fact that X_n is non-decreasing, we conclude that, almost surely, $X_n \leq n^{\alpha+\delta}$ for all n large enough, so that $\tau_k \geq k^{\frac{1}{\alpha+\delta}}$ for k large enough. We let

$$K = 1 + \sup\Big\{k \geq 1 \,\Big|\, \tau_k < k^{\frac{1}{\alpha+\delta}}\Big\},$$

and note that the argument above proves that $K < \infty$ almost surely.

The definition of K ensures that on the event $\{k \geq K\}$ we have $\tau_k \geq k^{\frac{1}{\alpha+\delta}}$ so that, for any $k \geq 1$, noting that $1/2 - \alpha < 0$ by Theorem 1.2,

$$\mathbb{E}\left[\mathbb{1}_{k\geq K}(\tau_k)^{\frac{1}{2}-\alpha}\left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right)\right] \leq \mathbb{E}\left[\mathbb{1}_{k\geq K}k^{\frac{1-2\alpha}{2\alpha+2\delta}}\left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right)\right]$$

$$\leq k^{\frac{1-2\alpha}{2\alpha+2\delta}}\mathbb{E}\left[\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right].$$

By Lemma 9.3 this can be bounded as

$$\mathbb{E}\left[\mathbb{1}_{k\geq K}(\tau_k)^{\frac{1}{2}-\alpha}\left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right)\right] \leq Ck^{\frac{1-2\alpha}{2\alpha+2\delta}} \cdot q(k) \cdot k^{1-\varepsilon}$$

for some constants $C, \varepsilon > 0$. Now, using the fact that $q(k) = k^{-\frac{1}{2\alpha}-1+o(1)}$, thanks to Proposition 5.1, we find that by taking δ small enough, the last expression is smaller than $Ck^{-1-\varepsilon/2}$. Summing over $k \geq 2^i$ we get that

$$\mathbb{E}\left[\sum_{k=2^i}^{\infty}\mathbb{1}_{k\geq K}(\tau_k)^{\frac{1}{2}-\alpha}\left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right)\right] = O(2^{-i\varepsilon/2}) \quad \text{as } i \to \infty,$$

which is summable. Hence, by Markov's inequality,

$$\sum_{i\geq 1}\mathbb{P}\left(\sum_{k=2^i}^{\infty}\mathbb{1}_{k\geq K}(\tau_k)^{\frac{1}{2}-\alpha}\left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right) > 2^{-i\varepsilon/4}\right) < \infty.$$

Therefore, by the Borel–Cantelli lemma and monotonicity, with probability one we have

$$\sum_{k=k_0}^{\infty}\mathbb{1}_{k\geq K}(\tau_k)^{\frac{1}{2}-\alpha}\left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n)f(T_n)\right) \leq 2k_0^{-\varepsilon/4}$$

for k_0 large enough. Finally, by taking k_0 large enough, we can assume that $k_0 \geq K$, which yields the claim of Lemma 9.5. $\qquad\square$

Lastly, because the sum appearing on the right-hand side of Lemma 9.4 starts from the random variable X_{n_0}, which can, in principle, be small, we state the next lemma to ensure that this quantity grows at least polynomially fast in n_0 almost surely.

Lemma 9.6 (Almost sure polynomial lower bound on X_n). *There exists an almost surely positive random variable M such that for all $n \geq 1$,*

$$X_n \geq M n^{\frac{p}{2}} \quad \text{almost surely.}$$

Proof of Lemma 9.6. We use a greedy construction to get a positive subtree $\underline{T}_n$ of T_n. The way this is done is reminiscent of the proof of Proposition 5.1. We let $\underline{T}_1 = T_1$. Then at every step $n \geq 1$ of the forward Rémy's algorithm (Section 2.1):

  (i) If the edge E_{n+1} that is picked is adjacent to one of the leaves of $\underline{T}_n$ *and* the created vertex is a $\oplus$, then we add the newly added leaf to $\underline{T}_n$ to get $\underline{T}_{n+1}$.

  (ii) Otherwise we let $\underline{T}_{n+1} := \underline{T}_n$.

Now let us consider the dynamics of the number Y_n of leaves in $\underline{T}_n$. Conditionally on $Y_1, \dots, Y_n$ we have for all $n \geq 1$,

$$Y_{n+1} = \begin{cases} Y_n + 1 & \text{with probability } \dfrac{pY_n}{2n-1}, \\ Y_n & \text{with probability } \dfrac{2n-1-pY_n}{2n-1}. \end{cases}$$

From these transitions, we note that the sequence of vectors $(pY_n, 2n-1-pY_n)_{n \geq 1}$ follows the same dynamics as the (non-necessarily integer) amount of black and white balls in a generalized Pólya urn where:

  • At time 1, the urn contains p black balls and $1 - p$ white balls.

  • At every step, a color is drawn at random from the urn with a distribution corresponding to its current composition. Then if the sampled color is black, we add p black balls and $2 - p$ white balls to the urn. Otherwise, we add 2 white balls to the urn.

Such an urn is called *balanced* because the number of balls added at each step is constant, and *triangular* because the corresponding replacement matrix is triangular. From general results on triangular balanced Pólya urns (e.g. [Agu09, Theorem 2]), we get that $n^{-\frac{p}{2}} \cdot Y_n \to L$ almost surely as $n \to \infty$, where $L > 0$ almost surely. Hence, for $M := \min_{n \geq 1} Y_n n^{-\frac{p}{2}}$ we get the result stated in the lemma, because by definition of X_n we have $X_n \geq Y_n$ for all $n \geq 1$. □

We can now prove Lemma 9.2 by combining all of the above.

Proof of Lemma 9.2. Using consecutively the results of Lemmas 9.4, 9.5 and 9.6, we get

$$\sum_{n=n_0}^{\infty} |R_n| \leq C \cdot \sum_{k=X_{n_0}}^{\infty} (\tau_k)^{\frac{1}{2}-\alpha} \left(\sum_{n=\tau_k}^{\tau_{k+1}-1} r(n) f(T_n) \right) \leq C X_{n_0}^{-\varepsilon} \leq C M^{-\varepsilon} n_0^{-p\varepsilon/2}.$$

Taking $n_0 = 2$ shows the first claim of the lemma. Then taking $\varepsilon' = \frac{p\varepsilon}{4}$ ensures that for n_0 large enough (the threshold being random), we have

$$\sum_{n=n_0}^{\infty} |R_n| \leq n_0^{-\varepsilon'}.$$

This concludes the proof of Lemma 9.2. □

9.4   The scaling limit result

We can now establish the scaling limit result of Theorem 2.3 for the sequence $(X_n)_{n \geq 1}$.

Proof of Theorem 2.3. Recall the notation $\overleftarrow{\mathcal{F}}_n = \sigma(T_m, \ m \geq n)$ and denote $\overleftarrow{\mathcal{F}}_\infty = \bigcap_{n \geq 1} \overleftarrow{\mathcal{F}}_n$. Recall from Section 9.1 that we defined, for all $n \geq 1$,

$$c_n = \prod_{i=1}^{n}\left(1 - \frac{\alpha}{i}\right) \quad \text{and} \quad \mathcal{X}_n = c_n X_n \quad \text{and} \quad R_{n+1} = \frac{c_n}{n+1}(\alpha X_{n+1} - \#\mathcal{L}_\cap^{\max}(T_{n+1})).$$

We can now introduce the auxiliary $(\overleftarrow{\mathcal{F}}_n)$-adapted process $(\mathscr{X}_n)_{n \geq 1}$ as

$$\mathscr{X}_n := \mathcal{X}_n - \sum_{i=n+1}^{\infty} R_i, \quad n \geq 1, \tag{9.15}$$

which is almost surely finite thanks to Lemma 9.2.

Step 1: The process $(\mathscr{X}_n)_{n \geq 1}$ has martingale increments. Using $X_n = X_{n+1} - \mathbb{1}_{L_{n+1} \in \mathcal{L}_\cap^{\max}(T_{n+1})}$ and $c_{n+1} = \left(1 - \frac{\alpha}{n+1}\right)c_n$ along with the expression $R_{n+1} = \frac{c_n}{n+1}(\alpha X_{n+1} - \#\mathcal{L}_\cap^{\max}(T_{n+1}))$, we can write the increment of the auxiliary process as

$$\mathscr{X}_n - \mathscr{X}_{n+1} = (\mathcal{X}_n - \mathcal{X}_{n+1}) - R_{n+1} = c_n\left(\frac{\#\mathcal{L}_\cap^{\max}(T_{n+1})}{n+1} - \mathbb{1}_{L_{n+1} \in \mathcal{L}_\cap^{\max}(T_{n+1})}\right).$$

From this expression, it is clear that

$$\mathbb{E}\left[\mathscr{X}_n - \mathscr{X}_{n+1} \mid \overleftarrow{\mathcal{F}}_{n+1}\right] = 0, \tag{9.16}$$

and

$$\mathbb{E}\left[(\mathscr{X}_n - \mathscr{X}_{n+1})^2 \mid \overleftarrow{\mathcal{F}}_{n+1}\right] = (c_n)^2 \cdot \frac{\#\mathcal{L}_\cap^{\max}(T_{n+1})}{n+1} \cdot \left(1 - \frac{\#\mathcal{L}_\cap^{\max}(T_{n+1})}{n+1}\right)$$
$$\leq (c_n)^2 = O(n^{-2\alpha}), \tag{9.17}$$

where, in the inequality, we applied the deterministic bound $\#\mathcal{L}_\cap^{\max}(T_{n+1}) \leq n+1$, and then we used the asymptotics (9.4) for the sequence $(c_n)_{n \geq 1}$.

We have just proved that the *increments* of the process $(\mathscr{X}_n)_{n \geq 1}$ are those of an L^2 backward martingale. Unfortunately, we actually do not know at this point whether $\mathscr{X}_n$ is in L^1 or not, so we cannot claim that it is a backward martingale.

Step 2: Introducing related backward supermartingales. We introduce

$$\tau := \sup\left\{n \geq 2 \ \middle| \ \sum_{i \geq n} |R_i| > n^{-\varepsilon}\right\}, \tag{9.18}$$

where $\varepsilon > 0$ is as in Lemma 9.2 and we take the convention that $\tau = +\infty$ if no such n exists. By the second part of Lemma 9.2, we have $\tau < +\infty$ almost surely. Then the negative part $\mathscr{X}_{n \vee \tau}^- := -\min(\mathscr{X}_{n \vee \tau}, 0)$ is bounded for all $n \geq 1$, since

$$\mathscr{X}_{n \vee \tau}^- \leq \sum_{i > n \vee \tau} |R_i| \leq (n \vee \tau)^{-\varepsilon} \leq n^{-\varepsilon} \leq 1 \qquad \text{almost surely.} \tag{9.19}$$

Moreover, using that $\{\tau \leq n\}$ is $\overleftarrow{\mathcal{F}}_{n+1}$-measurable, we have

$$\mathbb{E}\Big[\mathscr{X}_{n\vee\tau} - \mathscr{X}_{(n+1)\vee\tau} \,\Big|\, \overleftarrow{\mathcal{F}}_{n+1}\Big] = \mathbb{E}\Big[\mathbb{1}_{\tau\leq n}(\mathscr{X}_{n\vee\tau} - \mathscr{X}_{(n+1)\vee\tau}) \,\Big|\, \overleftarrow{\mathcal{F}}_{n+1}\Big]$$
$$= \mathbb{1}_{\tau\leq n}\mathbb{E}\Big[\mathscr{X}_n - \mathscr{X}_{n+1} \,\Big|\, \overleftarrow{\mathcal{F}}_{n+1}\Big] = 0.$$

Hence, the increments of $(\mathscr{X}_{n\vee\tau})_{n\geq 1}$ have vanishing (conditional) expectation. Now for any $K > 0$, the process $(\mathscr{X}_{n\vee\tau} \wedge K)_{n\geq 1}$ is bounded between -1 and K so it is in L^1. Noting that for any $x \in \mathbb{R}$, the function $y \mapsto (x + y) \wedge K$ is concave, we use Jensen's inequality to get

$$\mathbb{E}\Big[\mathscr{X}_{n\vee\tau} \wedge K \,\Big|\, \overleftarrow{\mathcal{F}}_{n+1}\Big] = \mathbb{E}\Big[(\mathscr{X}_{(n+1)\vee\tau} + (\mathscr{X}_{n\vee\tau} - \mathscr{X}_{(n+1)\vee\tau})) \wedge K \,\Big|\, \overleftarrow{\mathcal{F}}_{n+1}\Big]$$
$$\leq \Big(\mathscr{X}_{(n+1)\vee\tau} + \mathbb{E}\Big[\mathscr{X}_{n\vee\tau} - \mathscr{X}_{(n+1)\vee\tau} \,\Big|\, \overleftarrow{\mathcal{F}}_{n+1}\Big]\Big) \wedge K$$
$$= \mathscr{X}_{(n+1)\vee\tau} \wedge K,$$

which ensures that the process $(\mathscr{X}_{n\vee\tau} \wedge K)_{n\geq 1}$ is a backward supermartingale.

Step 3: Proving the almost sure convergence of $(\mathscr{X}_n)_{n\geq 1}$ and hence of $(n^{-\alpha}X_n)_{n\geq 1}$.

For any $K > 0$, the process $(\mathscr{X}_{n\vee\tau} \wedge K)_{n\geq 1}$ is a bounded backward supermartingale. By standard results on backward supermartingales (see e.g. [Dur19, Section 5.6][19]), this implies that $(\mathscr{X}_{n\vee\tau} \wedge K)_{n\geq 1}$ converges almost surely as $n \to \infty$ towards a limit $\mathscr{X}_\infty^K$ and that for all $n \geq 1$,

$$\mathbb{E}\Big[\mathscr{X}_{n\vee\tau} \wedge K \,\Big|\, \overleftarrow{\mathcal{F}}_\infty\Big] \leq \mathscr{X}_\infty^K.$$

By monotonicity in K, we get that $(\mathscr{X}_{n\vee\tau})_{n\geq 1}$ converges almost surely towards a (possibly infinite) limit $\mathscr{X}_\infty = \lim_{K\to\infty} \mathscr{X}_\infty^K$. Now since for all $n \geq 1$ we have $\mathscr{X}_{n\vee\tau} \geq -1$, we can use the monotone convergence theorem when we pass to the limit $K \to \infty$ in the previous display and get

$$\mathbb{E}\Big[\mathscr{X}_{n\vee\tau} \,\Big|\, \overleftarrow{\mathcal{F}}_\infty\Big] = \lim_{K\to\infty} \mathbb{E}\Big[\mathscr{X}_{n\vee\tau} \wedge K \,\Big|\, \overleftarrow{\mathcal{F}}_\infty\Big] \leq \lim_{K\to\infty} \mathscr{X}_\infty^K = \mathscr{X}_\infty. \tag{9.20}$$

Since $\tau < \infty$ almost surely, we deduce that $(\mathscr{X}_n)_{n\geq 1}$ also converges almost surely to the same $\mathscr{X}_\infty$. Note that the definition (9.15) of $(\mathscr{X}_n)_{n\geq 1}$, the asymptotic $c_n \sim n^{-\alpha}/\Gamma(1 - \alpha)$ and its a.s. convergence towards $\mathscr{X}_\infty$ ensure that we have the almost sure convergence

$$n^{-\alpha}X_n \xrightarrow[n\to\infty]{} X \qquad \text{where} \qquad X := \Gamma(1 - \alpha)\mathscr{X}_\infty. \tag{9.21}$$

Step 4: Proving that $X < \infty$ almost surely. First, we prove that $X < \infty$ almost surely relying on L^2 estimates. By (9.17), the random variables $(\mathscr{X}_{n+1} - \mathscr{X}_n)_{n\geq 1}$ are all in L^2 and by the martingale property (9.16), they are uncorrelated, hence orthogonal in L^2. By the Pythagorean theorem, for any $1 \leq N_0 \leq N$ we have, now using the second moment inequality (9.17),

$$\mathbb{E}\Big[(\mathscr{X}_{N_0} - \mathscr{X}_N)^2\Big] = \sum_{n=N_0}^{N-1} \mathbb{E}\Big[(\mathscr{X}_n - \mathscr{X}_{n+1})^2\Big] \overset{(9.17)}{=} \sum_{n=N_0}^{N-1} O(n^{-2\alpha}) \leq \sum_{n=N_0}^{\infty} O(n^{-2\alpha}) = O((N_0)^{1-2\alpha}),$$

where we used that $\alpha > 1/2$, which is guaranteed by Theorem 1.2. This ensures that the sequence $(\mathscr{X}_n - \mathscr{X}_1)_{n\geq 1}$ is a Cauchy sequence in L^2 so it converges in L^2 (hence in probability) towards some almost surely finite random variable Y. Thus $\mathscr{X}_n \to X = \mathscr{X}_1 + Y$ in probability, which is then almost surely finite.

---

[19]Although technically the reference only deals with backward *martingales*, the argument readily extends to supermartingales since it only relies on Doob's upcrossing lemma.

Step 5: Proving that X is almost surely positive. Secondly, we argue that $\mathbb{E}[X] > 0$, or equivalently $\mathbb{E}[\mathscr{X}_\infty] > 0$. By (9.20),

$$\mathbb{E}[\mathscr{X}_\infty] \geq \mathbb{E}[\mathscr{X}_{n\vee\tau}] = \mathbb{E}[\mathcal{X}_{n\vee\tau}] - \mathbb{E}\left[\sum_{i>n\vee\tau} |R_i|\right] \geq \mathbb{E}[\mathcal{X}_{n\vee\tau}] - n^{-\varepsilon}, \tag{9.22}$$

where the last inequality follows by definition of τ in (9.18). By the convergence in probability in Theorem 2.1, we know that $\mathbb{P}\big(\mathcal{X}_n \geq n^{-\varepsilon/2}\big) \geq 1/2$ for n large enough. In addition, since τ is almost surely finite, we also have $\mathbb{P}(\tau > n) < 1/3$ for n large enough. This implies that for n large enough,

$$\mathbb{E}[\mathcal{X}_{n\vee\tau}] \geq \mathbb{E}\left[\mathcal{X}_{n\vee\tau} \mathbb{1}_{\mathcal{X}_n \geq n^{-\varepsilon/2}} \mathbb{1}_{\tau \leq n}\right] \geq n^{-\varepsilon/2}\mathbb{P}\big(\mathcal{X}_n \geq n^{-\varepsilon/2}, \tau \leq n\big)$$

$$\geq n^{-\varepsilon/2}\Big(\mathbb{P}\big(\mathcal{X}_n \geq n^{-\varepsilon/2}\big) - \mathbb{P}(\tau > n)\Big)$$

$$\geq \frac{1}{6}n^{-\varepsilon/2}.$$

Therefore we deduce from (9.22) that $\mathbb{E}[X] > 0$.

Now let us prove that $X > 0$ almost surely. Picking two independent uniform leaves $L_n^{(1)}$ and $L_n^{(2)}$ of T_n, we break the tree into three regions: $T_n^{(1)}$ and $T_n^{(2)}$ are the two subtrees above the highest common ancestor of $L_n^{(1)}$ and $L_n^{(2)}$ containing the leaves $L_n^{(1)}$ and $L_n^{(2)}$ respectively, and $T_n^{(3)}$ is the rest of the tree. According to Lemma 3.8 we have $(|T_n^{(1)}| - 1, |T_n^{(2)}| - 1, |T_n^{(3)}|) \sim$ DirMult$(n - 2; 1/2, 1/2, 1/2)$ and conditionally on those numbers, the trees $T_n^{(1)}$ and $T_n^{(2)}$ are independent p-signed uniform binary trees with these sizes. We then write

$$\frac{X_n}{n^\alpha} \geq \max\left(\left(\frac{|T_n^{(1)}|}{n}\right)^\alpha \cdot \frac{\text{LIS}(T_n^{(1)})}{|T_n^{(1)}|^\alpha}, \left(\frac{|T_n^{(2)}|}{n}\right)^\alpha \cdot \frac{\text{LIS}(T_n^{(2)})}{|T_n^{(2)}|^\alpha}\right). \tag{9.23}$$

By Lemma 3.6, $\frac{1}{n}(|T_n^{(1)}|, |T_n^{(2)}|, |T_n^{(3)}|) \xrightarrow{\text{d}} (D_1, D_2, D_3)$ where $(D_1, D_2, D_3) \sim$ Dir$(1/2, 1/2, 1/2)$. Besides, Step 3 above implies that, jointly with the last convergence, we have

$$\left(\frac{\text{LIS}(T_n^{(1)})}{|T_n^{(1)}|^\alpha}, \frac{\text{LIS}(T_n^{(2)})}{|T_n^{(2)}|^\alpha}\right) \xrightarrow{\text{d}} (X^{(1)}, X^{(2)}) \tag{9.24}$$

as $n \to \infty$, where $X^{(1)}, X^{(2)}$ are copies of X, independent from each other and from (D_1, D_2, D_3). Taking $n \to \infty$ in (9.23), we get that

$$X \succeq \max\Big(D_1^\alpha \cdot X^{(1)}, D_2^\alpha \cdot X^{(2)}\Big).$$

Thus,

$$\mathbb{P}(X = 0) \leq \mathbb{P}\Big(X^{(1)} = X^{(2)} = 0\Big) = \mathbb{P}(X = 0)^2,$$

so that $\mathbb{P}(X = 0) \in \{0, 1\}$. We know since $\mathbb{E}[X] > 0$ that this probability cannot be 1, so we must have $\mathbb{P}(X = 0) = 0$.

Step 6: Proving that X is not deterministic. Finally, we justify using similar arguments that X cannot be a finite constant almost surely. Assume by contradiction that $X = b \in (0, \infty)$ almost

surely. Then, considering the same decomposition as above, we write $S_n^{(1,2)}$ for the sign in $\{\ominus, \oplus\}$ of the highest common ancestor of the leaves $L_n^{(1)}$ and $L_n^{(2)}$. We have

$$\frac{\mathrm{LIS}(T_n)}{n^\alpha} \geq \mathbb{1}_{\{S_n^{(1,2)}=\oplus\}} \cdot \left(\left(\frac{|T_n^{(1)}|}{n}\right)^\alpha \cdot \frac{\mathrm{LIS}(T_n^{(1)})}{|T_n^{(1)}|^\alpha} + \left(\frac{|T_n^{(2)}|}{n}\right)^\alpha \cdot \frac{\mathrm{LIS}(T_n^{(2)})}{|T_n^{(2)}|^\alpha} \right). \tag{9.25}$$

Now, from our assumptions and using again Step 3 above, the left-hand side converges in law towards a constant b. The right-hand side converges by (9.24) towards a random variable

$$\mathrm{Ber}(p) \cdot b \cdot (D_1^\alpha + D_2^\alpha),$$

where $(D_1, D_2, D_3) \sim \mathrm{Dir}(1/2, 1/2, 1/2)$ and $\mathrm{Ber}(p)$ denotes a Bernoulli random variable with parameter p independent of everything else. Hence, $b \succeq \mathrm{Ber}(p) \cdot b \cdot (D_1^\alpha + D_2^\alpha)$. Since $\alpha < 1$, the random variable on the right takes values strictly greater than b with positive probability. This is a contradiction, so X is not deterministic. $\qquad\square$

9.5 The limiting random variable X is a measurable function of the Brownian separable permuton

We prove the only remaining claim of Theorem 1.1, *i.e.* that X is a deterministic measurable function of the Brownian separable permuton. The same explanation works for the companion remaining claim in the case of the Brownian cographon in Theorem 1.3.

Recall from Remark 3.2 that the permutations σ_n in Theorem 1.1 can be obtained from the p-signed uniform binary trees $\widetilde{T}_n = t(\mathfrak{e}, \mathfrak{s}, p; U_1, \dots, U_n)$, where $(U_i)_{i\geq 1}$ are i.i.d. uniform random variables in $[0,1]$ and $(\mathfrak{e}, \mathfrak{s}, p)$ is the signed Brownian excursion used to construct the Brownian separable permuton $\boldsymbol{\mu}_p$. In particular, the trees $(\widetilde{T}_n)_{n\geq 1}$ are coupled as in Rémy's algorithm.

Fix $m \geq 1$ and introduce, for all $n \geq 1$, the tree $\widetilde{T}_n^{(m)} = t(\mathfrak{e}, \mathfrak{s}, p; U_{m+1}, \dots U_{m+n})$. Note that, since the two collections of points $\{U_1, \dots, U_n\}$ and $\{U_{m+1}, \dots, U_{m+n}\}$ have all but $2m$ points in common, by definition of $\mathrm{LIS}(\cdot)$, we have for all $n \geq 1$ the deterministic bound

$$\left| \mathrm{LIS}(\widetilde{T}_n^{(m)}) - \mathrm{LIS}(\widetilde{T}_n) \right| \leq 2m.$$

As a consequence, from Theorem 2.3, we have the almost sure convergence

$$X = \lim_{n\to\infty} n^{-\alpha} \mathrm{LIS}(\widetilde{T}_n) = \lim_{n\to\infty} n^{-\alpha} \mathrm{LIS}(\widetilde{T}_n^{(m)}). \tag{9.26}$$

Hence, the limiting random variable X can be written as a deterministic measurable function of $((\mathfrak{e}, \mathfrak{s}, p), (U_i)_{i>m})$. Since this is true for any $m \geq 1$, we get that X is $\mathcal{F}$-measurable for $\mathcal{F} = \bigcap_{m\geq 1} \sigma((\mathfrak{e}, \mathfrak{s}, p), (U_i)_{i>m})$. Since in addition the random variables U_i for $i \geq 1$ are i.i.d., by Kolmogorov's 0-1 law, the σ-field $\mathcal{G} = \bigcap_{m>1} \sigma((U_i)_{i\geq m})$ is trivial – in the sense that it only contains events of probability 0 or 1. Since $\mathcal{F} = \sigma(\mathcal{G} \cup \sigma(\mathfrak{e}, \mathfrak{s}, p))$, this ensures that X is almost surely a deterministic measurable function of $(\mathfrak{e}, \mathfrak{s}, p)$, *i.e.* of the Brownian separable permuton $\boldsymbol{\mu}_p$.

9.6 Scaling limit of the leftmost maximal positive subtree of T^k

We prove in this final section the only remaining result, *i.e.* Theorem 2.6. Recall from (6.1) that we introduced the more compact notation $T^{k,\mathrm{lmax}} = (T^k)^{\mathrm{lmax}}$.

Proof of Theorem 2.6. We first notice that Lemma 6.1 provides a recursive construction of $T^{k,\mathrm{lmax}}$ as follows. Let $\mathcal{L}^{k,\mathrm{lmax}}$ be the set of leaves of $T^{k,\mathrm{lmax}}$. Let v_1, v_2 be the two children of the root of T^k and let K_1, K_2 be their respective number of descendants in $\mathcal{L}^{k,\mathrm{lmax}}$. We choose to order v_1 and v_2 in such a way that $K_1 \geq K_2$. We also set the convention $K_2 = 0$ if the root has only one child in $T^{k,\mathrm{lmax}}$, and $K_1 = K_2 = 0$ if it has none (which implies $k = 1$). Then conditionally on $(K_1, K_2) = (k_1, k_2)$, the two trees of descendants of v_1 and v_2 with leaves in $\mathcal{L}^{k,\mathrm{lmax}}$ are independent, each distributed as $T^{k_1,\mathrm{lmax}}$ and $T^{k_2,\mathrm{lmax}}$ (taking the convention that $T^{0,\mathrm{lmax}}$ is empty). Moreover, the law[20] q_k of (K_1, K_2) is given for $k \geq 1$ by

$$q_k((k_1, k_2)) = p\frac{q(k_1)q(k_2)}{q(k)} \qquad \text{if } k_1 + k_2 = k \text{ and } 1 \leq k_2 < \tfrac{k}{2},$$

$$q_k((k_1, k_2)) = \frac{1}{2}p\frac{q(k_1)q(k_2)}{q(k)} \qquad \text{if } k_1 = k_2 = \tfrac{k}{2},$$

$$q_k((k, 0)) = (1 - p)\left(\frac{q(k)}{2} + \sum_{i=1}^{k-1} q(i)\right),$$

$$q_k(0, 0) = \frac{1}{2q(k)}\mathbb{1}_{k=1}.$$

In the terminology of [HM12, Section 1.2.1], this means that $T^{k,\mathrm{lmax}}$ is a Markov branching tree with splitting distributions q_k (more precisely, we are in the case where the splitting is binary, *i.e.* q_k is supported on partitions of k with 1 or 2 parts). Therefore, in order to conclude by [HM12, Theorem 5], it is sufficient to check that the splitting distributions (q_k) satisfy the assumption (H) of [HM12], with the measure $\nu = \nu_\gamma$ on binary partitions of 1 given by (2.24).

Let us now state this assumption (H) explicitly. We first define the normalized version $\bar{q}_k$ of q_k as the measure on $[0, 1]^2$ such that, for any measurable function $f : [0, 1]^2 \to \mathbb{R}_{>0}$, we have

$$\int f \, \mathrm{d}\bar{q}_k = \sum_{k_1, k_2} q_k((k_1, k_2))f\left(\frac{k_1}{k}, \frac{k_2}{k}\right).$$

Then assumption (H) with the measure ν_γ of (2.24) states that there is a slowly varying function $\ell(\cdot)$ such that for any continuous function $f : [0, 1]^2 \to \mathbb{R}_{>0}$, we have

$$k^{\gamma-1}\ell(k)\int(1 - s_1)f(s_1, s_2)\,\bar{q}_k(\mathrm{d}s_1, \mathrm{d}s_2) \xrightarrow[k\to\infty]{} \int_0^{1/2} f(1 - x, x)x^{1-\gamma}(1 - x)^{-\gamma}\,\mathrm{d}x. \qquad (9.27)$$

Note that the parameter γ of [HM12] corresponds to our $\gamma - 1$. It turns out that we can choose ℓ to be $1/\varphi$ with φ as in Theorem 7.1, *i.e.* φ is slowly varying and $q(k) = k^{-\gamma}\varphi(k)$. Using the explicit probabilities q_k given above, the left-hand side of (9.27) then rewrites for $k \geq 2$ as

$$\frac{k^{\gamma-1}}{\varphi(k)}p\left(\sum_{i=1}^{\lfloor\frac{k-1}{2}\rfloor}\frac{q(k-i)q(i)}{q(k)}\frac{i}{k}f\left(\frac{k-i}{k}, \frac{i}{k}\right) + \mathbb{1}_{k \text{ even}}\frac{1}{2}\frac{q(k/2)^2}{q(k)}\cdot\frac{1}{2}f\left(\frac{1}{2}, \frac{1}{2}\right)\right).$$

Note first that $\frac{q(k/2)^2}{q(k)} = 4^\gamma k^{-\gamma}\varphi(k)(1 + o(1))$, so the second term is negligible. After replacing $q(k)$ by $k^{-\gamma}\varphi(k)$, the last equation becomes

$$p\frac{1}{k}\sum_{i=1}^{\lfloor\frac{k-1}{2}\rfloor}\left(1 - \frac{i}{k}\right)^{-\gamma}\left(\frac{i}{k}\right)^{1-\gamma}\frac{\varphi(k-i)\varphi(i)}{\varphi(k)^2}f\left(\frac{k-i}{k}, \frac{i}{k}\right) + o(1).$$

---

[20]The notation q_k is chosen to be consistent with [HM12].

94

We now fix $\varepsilon > 0$, and split the sum according to whether $i < \varepsilon k$ or not. First, by Item 1 of Lemma 3.1, the ratio $\frac{\varphi(k-i)\varphi(i)}{\varphi(k)^2}$ converges to 1 as $k \to \infty$ uniformly in $i \in \left[\varepsilon k, \frac{k}{2}\right]$. By Riemann sum approximation, this implies

$$p\frac{1}{k}\sum_{i=\varepsilon k}^{\lfloor\frac{k-1}{2}\rfloor}\left(1-\frac{i}{k}\right)^{-\gamma}\left(\frac{i}{k}\right)^{1-\gamma}\frac{\varphi(k-i)\varphi(i)}{\varphi(k)^2}f\left(\frac{k-i}{k},\frac{i}{k}\right)\xrightarrow[k\to\infty]{}p\int_{\varepsilon}^{1/2}f(1-x,x)x^{1-\gamma}(1-x)^{-\gamma}\,\mathrm{d}x.$$
(9.28)

Now let us now handle the terms where $i < \varepsilon k$. We fix $\delta > 0$ small enough so that $\gamma + \delta < 2$ (recall from (2.7) that $\gamma < 2$). Again, by Item 1 of Lemma 3.1, the ratio $\frac{\varphi(k-i)}{\varphi(k)}$ is bounded by an absolute constant which does not depend on ε. So are the values of f and the factors $\left(1-\frac{i}{k}\right)^{-\gamma}$. On the other hand, Item 2 of Lemma 3.1 shows $\frac{\varphi(i)}{\varphi(k)} = O\left(\left(\frac{i}{k}\right)^{-\delta}\right)$ uniformly in $i < k$, so we can write

$$p\frac{1}{k}\sum_{i=1}^{\varepsilon k}\left(1-\frac{i}{k}\right)^{-\gamma}\left(\frac{i}{k}\right)^{1-\gamma}\frac{\varphi(k-i)\varphi(i)}{\varphi(k)^2}f\left(\frac{k-i}{k},\frac{i}{k}\right)=O\left(\frac{1}{k}\sum_{i=1}^{\varepsilon k}\left(\frac{i}{k}\right)^{1-\gamma-\delta}\right)=O\left(\varepsilon^{2-\gamma-\delta}\right).$$

Combined with (9.28), this shows that (9.27) is satisfied and concludes the proof of Theorem 2.6. □

A   Numerical simulations for the limiting random variable X

We present in Figure 12 and Table 13 some numerical simulations for the limiting random variables $X(p)$ appearing in Theorem 1.1. Recall the notation $\sigma_n(p) \coloneqq \mathrm{Perm}[\boldsymbol{\mu}_p, n]$ from (1.1).

For the diagram on the left-hand side of Figure 12, we sampled $100\,000$ independent copies of $\sigma_n(p_i) \coloneqq \mathrm{Perm}[\boldsymbol{\mu}_{p_i}, n]$ of size $n = 1\,000\,000$, for each $p_i = \frac{i}{10}$ with $i \in [\![1\,,9]\!]$. We then estimated the density function of (recall Theorem 1.1)

$$\frac{\mathrm{LIS}(\sigma_n(p_i))}{n^{\alpha(p_i)}} \approx X(p_i) \qquad \text{for all } i \in [\![1\,,9]\!],$$

using the function SmoothKernelDistribution from Mathematica.

For the diagram on the right-hand side of Figure 12, we sampled $100\,000$ independent uniform separable permutations $\overline{\sigma}_n$ of the same size $n = 1\,000\,000$, and estimated the density function (recall Conjecture 1.4) of

$$\frac{\mathrm{LIS}(\overline{\sigma}_n)}{0.901 \cdot n^{\alpha(1/2)}} \approx X(1/2),$$

again using SmoothKernelDistribution. The constant $c = 0.901$ was calibrated so that the two estimated densities on the right-hand side of Figure 12 would be as close as possible. More precisely, we chose the constant c as the minimizer of

$$\int_0^{100} |f_c(x) - g(x)|\,\mathrm{d}x,$$

where $f_c(x)$ is the approximate density function of $\frac{\mathrm{LIS}(\overline{\sigma}_n)}{c \cdot n^{\alpha(1/2)}}$ and $g(x)$ is the approximate density function of $\frac{\mathrm{LIS}(\sigma_n(1/2))}{n^{\alpha(1/2)}}$. At $c = 0.901$, the above integral takes the value 0.055.

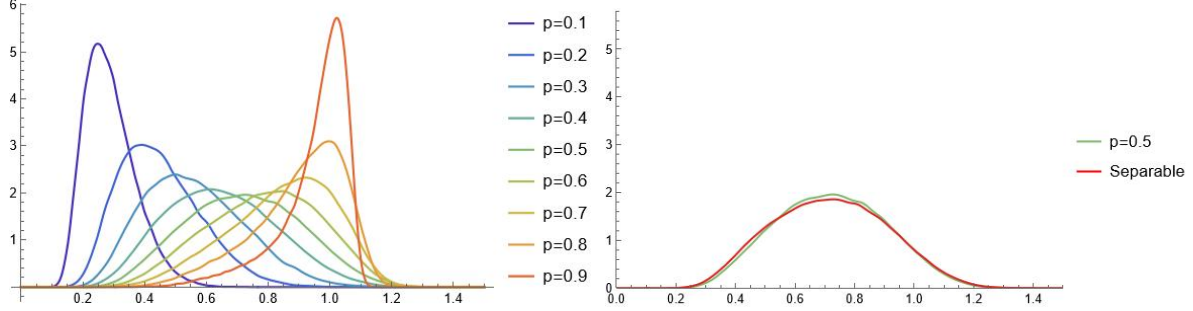Finally, in Table 13 we show the approximated values for the mean and variance of the random variables $X(p_i)$.

Figure 12: **Left:** Approximate densities of the random variables $X(p)$. **Right:** Comparison between the approximate density of the limiting random variable for $\frac{\text{LIS}(\sigma_n(1/2))}{n^{\alpha(1/2)}}$ when the permutations σ_n are sampled from the Brownian separable permuton with parameter $1/2$ and the approximate density of the limiting random variable for $\frac{\text{LIS}(\overline{\sigma}_n)}{0.901 \cdot n^{\alpha(1/2)}}$, when the permutations $\overline{\sigma}_n$ are exactly uniform separable permutations.

p	$\mathbb{E}[X(p)]$	$\text{Var}(X(p))$
0.1	0.288	0.007
0.2	0.449	0.017
0.3	0.560	0.026
0.4	0.648	0.032
0.5	0.722	0.034
0.6	0.787	0.033
0.7	0.846	0.029
0.8	0.901	0.022
0.9	0.952	0.012

Table 13: Approximate mean and variance of $X(p)$ for different values of p.

Index of notation

$\boldsymbol{\mu}_p$	Brownian separable permuton of parameter $p \in (0,1)$
σ_n	Permutation of size n sampled from $\boldsymbol{\mu}_p$
$\mathrm{LIS}(\sigma)$	Length of the longest increasing sequence of the permutation σ
$\boldsymbol{W}_p$	The Brownian cographon of parameter $p \in (0,1)$
G_n	Random graph on n vertices sampled from the Brownian cographon
$\mathrm{LCL}(G),\ \mathrm{LIN}(G)$	Size of the largest clique, independent set, of the graph G
$\alpha = \alpha(p)$	Explicit exponent (depending on the parameter p), see (1.3)
$\gamma = \gamma(p)$	Regular variation exponent related to α, see (2.7)
$t\vert_{\mathcal{L}}$	Subtree of t induced by the set of leaves $\mathcal{L}$, see Section 2.1
$\vert t \vert$	Size (*i.e.* number of leaves) of the tree t
$\oplus,\ \ominus$	Signs carried by the nodes of a sign-decorated tree
$\mathrm{LIS}(t)$	Maximal size of a positive subtree of t
$\mathcal{L}_{\cap}^{\max}(t)$	Set of leaves in the intersection of all maximal positive subtrees of t
E_n, S_n, L_n	Edge, sign and leaf at step n of Rémy's algorithm, see Section 2.1
T_n	p-signed uniform binary tree with n leaves, see Section 1.1
T	p-signed critical binary Bienaymé–Galton–Watson tree
T^k	T conditioned on $\mathrm{LIS}(T) = k$
$q,\ Q$	Law and tail function of $\mathrm{LIS}(T)$, see (2.5)
X_n	$X_n = \mathrm{LIS}(T_n)$, see (2.3)
$\varphi,\ \Phi$	Slowly varying functions in Theorem 7.1
$r(n)$	$r(n) = \mathbb{P}(\vert T \vert = n)$, see Lemma 3.3
$\mathcal{F}_n,\ \overleftarrow{\mathcal{F}}_n$	Filtrations associated with the forward, backward, Rémy's algorithm
τ_k	Stopping time $\tau_k = \min\{n \geq 1 \mid X_n = k\}$, see (5.3)
t^{lmax}	Leftmost maximal subtree of t, see Section 6.1.1
L^k	Uniformly chosen leaf of $T^{k,\mathrm{lmax}}$
$Z^k = (S, D, M^*, W)$	Four-component Markov chain, see (6.2)
η	Death time of M^*
$\mathbb{P}_k$	Law of Z^k started at k, see Remark 6.3
$\pi((*,*,*,*);(*,*,*,*))$	Transition probabilities of Z^k, see Section 6.2.1
$\pi(*;*)$	Transition probabilities of M, see (6.14)
η_t	Time-change in M^*, see (6.8)
M	Decreasing Markov chain, (η_t) time-change of M^*
$\mathtt{T}$	Death time of M
$\theta(m)$	Parameters of geometric random variable in (6.10)
$(B_m)_m$	Black instructions, see (6.16)
$(G_m)_m$	Golden instructions, see (6.17)
$\mathbb{P}_{k,k'}$	Law of the coupling of M, M' started from k, k', see Definition 6.5
$\widetilde{\mathbb{P}}_{k,k'}$	Law of the coupling of Z, Z' started from k, k', see Corollary 6.7
V_{mer}	Merging value of M and M', see Proposition 6.6
$\mathtt{t}_{\mathrm{mer}},\ \mathtt{t}'_{\mathrm{mer}}$	Merging times of M and M', see Proposition 6.6
$\mathrm{GoodScales}$	A set of good scales, as defined in Definition 6.10
H_i	Hitting time of $[\![0, i]\!]$ by M, see (6.28)
$\mathcal{G}_\ell$	σ-field recording information strictly above 7ℓ, see (6.29)
λ	Scaling limit of $\mathcal{L}_{\cap}^{\max}(T^k)$, see Theorem 7.2; later equated with α
$\mathfrak{h}_{\mathrm{div}}$	Divergence height between L^k and $\widetilde{L}^k$, see (7.9)

References

[ABS⁺25]   D. Amankwah, J. Björnberg, S. Ö. Stefánsson, B. Stufler, and J. Turunen. The scaling limit of random 2-connected series-parallel maps. *ArXiv e-prints*, 2025, 2503.19705.

[AC17]   A. Auffinger and D. Cable. Pemantle's min-plus binary tree. *ArXiv e-prints*, 2017, 1709.07849.

[Agu09]   R. Aguech. Limit theorems for random triangular urn schemes. *Journal of applied probability*, 46(3):827–843, 2009.

[Alz97]   H. Alzer. On some inequalities for the Gamma and Psi functions. *Mathematics of computation*, 66(217):373–389, 1997.

[AT93]   D. Aldous and H. Thorisson. Shift-coupling. *Stochastic Processes and Their Applications*, 44(1):1–14, 1993.

[BBD⁺22]   F. Bassino, M. Bouvel, M. Drmota, V. Féray, L. Gerin, M. Maazoun, and A. Pierrot. Linear-sized independent sets in random cographs and increasing subsequences in separable permutations. *Comb. Theory*, 2(3):35, 2022, 2104.07444. Id/No 15.

[BBF⁺18]   F. Bassino, M. Bouvel, V. Féray, L. Gerin, and A. Pierrot. The Brownian limit of separable permutations. *Ann. Probab.*, 46(4):2134–2189, 2018, 1602.04960. MR3813988

[BBF⁺20]   F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun, and A. Pierrot. Universal limits of substitution-closed permutation classes. *J. Eur. Math. Soc. (JEMS)*, 22(11):3565–3639, 2020, 1706.08333. MR4167015

[BBF⁺22a]   F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun, and A. Pierrot. Random cographs: Brownian graphon limit and asymptotic degree distribution. *Random Structures & Algorithms*, 60(2):166–200, 2022, 1907.08517.

[BBF⁺22b]   F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun, and A. Pierrot. Scaling limits of permutation classes with a finite specification: a dichotomy. *Advances in Mathematics*, 405:108513, 2022, 1903.07522.

[BBMF11]   N. Bonichon, M. Bousquet-Mélou, and É. Fusy. Baxter permutations and plane bipolar orientations. *Sém. Lothar. Combin.*, 61A:Art. B61Ah, 29, 2009/11, arXiv:0805.4180.

[BDSG24]   J. Borga, W. Da Silva, and E. Gwynne. Power-law bounds for increasing subsequences in Brownian separable permutons and homogeneous sets in Brownian cographons. *Advances in Mathematics*, 439:109480, 2024, 2303.17030.

[BE53]   H. Bateman and A. Erdélyi. Higher transcendental functions. *Bateman Manuscript Project Mc Graw-Hill Book Company*, Vol. I, 1953.

[BG24]   J. Borga and E. Gwynne. Reconstructing SLE-decorated Liouville quantum gravity surfaces from random permutons. *ArXiv e-prints*, 2024, 2409.15494.

[BGS25]   J. Borga, E. Gwynne, and X. Sun. Permutons, meanders, and SLE-decorated Liouville quantum gravity. *Journal of the European Mathematical Society*, to appear, 2025, 2207.02319.

[BGT89]   N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation.* Number 27. Cambridge university press, 1989.

[BHSY23] J. Borga, N. Holden, X. Sun, and P. Yu. Baxter permuton and Liouville quantum gravity. *Probab. Theory Relat. Fields*, 186(3-4):1225–1273, 2023, 2203.12176.

[BM22]   J. Borga and M. Maazoun. Scaling and local limits of Baxter permutations and bipolar orientations through coalescent-walk processes. *Ann. Probab.*, 50(4):1359–1417, 2022, 2008.09086. MR4420422

[Bor21]   J. Borga. Random Permutations – A geometric point of view. *ArXiv e-prints (PhD thesis)*, July 2021, 2107.09699.

[Bor23]   J. Borga. The skew Brownian permuton: A new universality class for random constrained permutations. *Proc. Lond. Math. Soc. (3)*, 126(6):1842–1883, 2023, 2112.00156.

[Dev06]   L. Devroye. Nonuniform random variate generation. *Handbooks in operations research and management science*, 13:83–121, 2006.

[DG88]   F. M. Dekking and G. R. Grimmett. Superbranching processes and projections of random cantor sets. *Probability Theory and Related Fields*, 78:335–355, 1988.

[Dur19]   R. Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[FK17]   F. Friedli and A. Karlsson. Spectral zeta functions of graphs and the Riemann zeta function in the critical strip. *Tohoku Mathematical Journal, Second Series*, 69(4):585–610, 2017, 1410.8010.

[FRL23]   V. Féray and K. Rivera-Lopez. The permuton limit of random recursive separable permutations. *Confluentes Mathematici*, 15:45–82, 2023, 2306.04278.

[Ghy16]   É. Ghys. A singular mathematical promenade. *ArXiv e-prints*, 2016, 1612.06373.

[HM12]   B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *Annals of Probability*, 40:2589–2666, 2012, 1003.3632.

[Kar20]   A. Karlsson. Spectral zeta functions. *Discrete and Continuous Models in the Theory of Networks*, pages 199–211, 2020, 1907.01832.

[KP22]   A. Karlsson and M. Pallich. Volumes of spheres and special values of zeta functions of $\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$. *ArXiv e-prints*, 2022, 2209.03590.

[KS66]   H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton-Watson processes. *The Annals of Mathematical Statistics*, 37(5):1211–1223, 1966.

[Le 05]   J.-F. Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005, math/0511515. MR2203728 (2007h:60078)

[Len23a] T. Lenoir. Graph classes with few P_4's: Universality and Brownian graphon limits. *arXiv e-prints*, 2023, 2301.13607.

[Len23b] T. Lenoir. Subgraph densities and scaling limits of random graphs with a prescribed modular decomposition. *ArXiv e-prints*, 2023, 2310.01318.

[Lov12] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[LP17] D. A. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

[Maa20] M. Maazoun. On the Brownian separable permuton. *Combin. Probab. Comput.*, 29(2):241–266, 2020, 1711.08986. MR4079636

[Pit06] J. Pitman. *Combinatorial stochastic processes: Ecole d'été de probabilités de Saint-Flour XXXII-2002.* Springer Science & Business Media, 2006.

[Pre22] V. V. Prelov. Coupling of several random variables. *Problems of Information Transmission*, 58(4):300–305, 2022.

[Rém85] J.-L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO. Informatique théorique*, 19(2):179–195, 1985.

[Stu20] B. Stufler. Limits of random tree-like discrete structures. *Probability Surveys*, 17:318–477, 2020, 1612.02580.

[Stu21] B. Stufler. Graphon convergence of random cographs. *Random Structures & Algorithms*, 59(3):464–491, 2021, 1906.10355.