Geometry of weighted recursive and affine preferential attachment trees

Delphin Sénizergues*

Abstract

We study two models of growing recursive trees. For both models, the tree initially contains a single vertex u_1 and at each time $n \geq 2$ a new vertex u_n is added to the tree and its parent is chosen randomly according to some rule. In the *weighted recursive tree*, we choose the parent u_k of u_n among $\{u_1, u_2, \ldots, u_{n-1}\}$ with probability proportional to w_k, where $(w_n)_{n \geq 1}$ is some deterministic sequence that we fix beforehand. In the *affine preferential attachment tree with fitnesses*, the probability of choosing any u_k is proportional to $a_k + \deg^+(u_k)$, where $\deg^+(u_k)$ denotes its current number of children, and the sequence of *fitnesses* $(a_n)_{n \geq 1}$ is deterministic and chosen as a parameter of the model.

We show that for any sequence $(a_n)_{n \geq 1}$, the corresponding preferential attachment tree has the same distribution as some weighted recursive tree with a *random* sequence of weights (with some explicit distribution). We then prove almost sure scaling limit convergences for some statistics associated with weighted recursive trees as time goes to infinity, such as degree sequence, height, profile and also the weak convergence of some measures carried on the tree. Thanks to the connection between the two models, these results also apply to affine preferential attachment trees.

1 Introduction

The uniform recursive tree has been introduced in the 70's as an example of random graph constructed by addition of vertices: starting from a tree containing a single vertex, the vertices arrive one by one and the n-th vertex picks its parent uniformly at random from the $n - 1$ already present vertices. Many properties of this tree were then investigated due to its particularly simple dynamics: number of leaves, profile, height,

---
*University of British Columbia, Canada. E-mail: senizergues@math.ubc.ca

degrees, size of subtrees and others. We refer to the survey [52] and the more recent book [15, Section 6] for an overview of the results obtained for this model.

We consider a generalization of the uniform recursive tree called the weighted recursive tree (WRT), which was introduced in [8] in 2006. In this model, each vertex is assigned a non-negative weight, constant in time. When a newcomer randomly picks its parent, it does so with probability proportional to those weights. Most of the contributions concerning WRTs are very recent, see e.g. [34, 24, 25, 20, 33]. In [34] those trees are studied because of their connection to a model of random walk with preferential relocation (a.k.a. "monkey walk"). The authors prove some limiting results for the distribution of the weight of vertices at different heights in the tree, for different assumptions on the weight sequence which cover a wide range of behaviours. The two papers [25, 20] study the properties of some more general models of increasing trees, and their results apply to WRTs when the sequence of weights is i.i.d. and identify in that case the asymptotic degree distribution in the tree. Finally [33] studies the behaviour of the maximal degree, also in the i.i.d. case, for different tail behaviours of the weight distribution.

In this paper, we prove asymptotic results for this model concerning the degree sequence, the height, the profile and the convergence of some probability measures carried on the tree, mainly under some assumptions that ensure that the sequence $(w_n)_{n \geq 1}$ describing the weights of the vertices in order of creation behaves roughly as a power of n. Our deepest result is the one that concerns the asymptotic behaviour of the profile of the tree, which is the function that maps each integer k to the number of vertices in the tree at height k. Both the statement and the proof of this result are inspired from the work carried out in the last 20 years for different models of logarithmic trees, see [9, 10, 53, 50, 29]. They rely on the study of the Laplace transform of the profile using tools that ultimately date back to Biggins [5] in the context of the branching random walk, together with a Fourier inversion argument, which in our case is handled by a very precise theorem of [29]. The rest of our results and proofs on WRTs are less involved and mostly rely on more elementary arguments, as well as a connection with Pemantle's time-dependent Pólya urns, introduced in [43].

We also consider another model of trees which we call the affine preferential attachment tree (PAT) with fitnesses. In this one, every vertex has a fixed fitness, and the probability of picking any vertex to be the parent of a newcomer is proportional to its fitness plus its current number of children.

The term "preferential attachment", coined by Barabási and Albert in [2], refers to the property that a vertex in the graph that has a high degree tends to increase its degree even more over time, also referred to as a "rich-get-richer" effect. Many different preferential attachment mechanisms have then been studied in the last two decades because the degree distribution that emerges from this type of construction shares some quantitative properties with real-world networks, see [54, 37] for good overviews of the vast literature on this subject.

One of the original motivations of this paper comes from the fact that PATs arise in the analysis of some growing random graphs, studied in the companion paper [51]. The class of models that we study there is designed to encompass Rémy's algorithm, described in [47], which creates a sequence of binary trees, and a lot of its natural generalizations, studied separately in [19, 35, 11, 22, 23, 48]. In particular, we show that the sequences of graphs obtained using these constructions, considered as metric spaces, almost surely converge in the so-called Gromov–Hausdorff–Prokhorov scaling limit towards a limiting random continuous metric space. This proof relies on a decomposition of our graphs along the structure of a tree, whose evolution is that of a PAT. Notably, a crucial result that is needed in this argument is a uniform control over the degrees of all the vertices

in the tree, which we prove in this paper.

Note that only a few contributions in the literature concern this particular version of the model, where the fitness can depend on the vertex. In the case where the fitnesses are i.i.d., the model is considered for the first time in [17] and the first rigorous mathematical result can be found in [4]. Recently, still in the case of i.i.d. fitnesses, it has been studied in more detail in [32] along with some other similar models. In [32] the authors study the asymptotic degree distribution and maximum degree in the tree and show that these can exhibit different behaviours depending on the tail of the fitness distribution, which the authors classify as weak, strong and extreme disorder. A result from [25] also applies in that case and also identifies the limiting degree distribution. Even more recently, the local weak limit of this model has been investigated in [31], and the limit is expressed as the so-called π-Pólya point tree. We also mention two models that do not fall in our setting but are somewhat related, studied in [13] and [7], in which the reinforcement is affine in the degree of the vertices but there is some inhomogeneity between vertices. Instead of coming from different fitnesses associated to vertices like in our model, the inhomogeneity is introduced using a random initial degree, or respectively a random time of creation.

Our approach for studying this model relies on the connection between the PAT and the WRT models (this was already known in the field in the case of constant fitnesses but stated in a slightly different form, see [7, 3]). Indeed we shall see that using a de Finetti-type argument, a PAT can be seen as a WRT with a random sequence of weights that almost surely decays like a power of n. This enables us to translate all of the results obtained for WRTs to corresponding results for PATs, and hence prove asymptotics for degrees, height and profile of the tree. In particular, we prove the almost sure scaling limit convergence of the sequence of degrees of the vertices in order of creation in an ℓ^p norm. For some regular sequences of fitnesses, we can explicitly describe the distribution of the limiting sequence using Beta, Gamma and Mittag-Leffler distributions. This relates in various ways to other results that can be found in the literature associated to preferential attachment trees or to urn models, contained in [36, 27, 41, 26, 42, 40, 39, 1].

1.1 Two related models of growing trees

Definitions. For any sequence of non-negative real numbers $(w_n)_{n \geq 1}$ with $w_1 > 0$, we define the distribution $\mathrm{WRT}((w_n)_{n \geq 1})$ on sequences of growing rooted labelled trees[1], which is called the *weighted recursive tree with weights* $(w_n)_{n \geq 1}$. We construct a sequence of rooted trees $(\mathrm{T}_n)_{n \geq 1}$ by starting from T_1 containing only one root-vertex u_1 with label 1 and letting it evolve in the following manner: the tree T_{n+1} is obtained from T_n by adding a vertex u_{n+1} with label $n + 1$. The parent of this new vertex is chosen to be the vertex with label K_{n+1}, where the random variable K_{n+1} is such that

$$\forall k \in \{1, \dots, n\}, \qquad \mathbb{P}\left(K_{n+1} = k \mid \mathrm{T}_1, \mathrm{T}_2, \dots \mathrm{T}_n\right) \propto w_k. \qquad (1.1)$$

Remark that this conditional distribution does not depend on the evolution $\mathrm{T}_1, \mathrm{T}_2, \dots \mathrm{T}_n$ up to time n, which ensures in particular that the random variables $K_2, K_3, \dots$ are independent. In this definition, we also allow sequences of weights $(\mathsf{w}_n)_{n \geq 1}$ that are random and in this case the distribution $\mathrm{WRT}((\mathsf{w}_n)_{n \geq 1})$ denotes the law of the random tree obtained by first sampling the random sequence $(\mathsf{w}_n)_{n \geq 1}$, and then running the above process conditionally on $(\mathsf{w}_n)_{n \geq 1}$.

Similarly, for any sequence $(a_n)_{n \geq 1}$ of real numbers, with $a_1 > -1$ and $a_n \geq 0$ for $n \geq 2$, we define another model of growing tree. The construction goes on as before: P_1

---

[1]In fact, in the rest of the paper we see them as plane trees, see Section 1.2.2.

contains only one root-vertex u_1 with label 1 and P_{n+1} is obtained from P_n by adding a vertex u_{n+1} with label $n+1$ and the parent of the newcomer is chosen to be the vertex with label J_{n+1}, where now

$$\forall k \in \{1, \ldots, n\}, \qquad \mathbb{P}\left(J_{n+1} = k \mid \mathsf{P}_1, \mathsf{P}_2, \ldots, \mathsf{P}_n\right) \propto \deg_{\mathsf{P}_n}^+(u_k) + a_k, \qquad (1.2)$$

where $\deg_{\mathsf{P}_n}^+(\cdot)$ denotes the number of children in the tree P_n. In the particular case where $n = 1$, the second vertex u_2 is always defined as a child of u_1, even in the case $-1 < a_1 \leq 0$ for which the last display does not make sense. We call this sequence of tree an *affine preferential attachment tree with fitnesses* $(a_n)_{n \geq 1}$ and its law is denoted by $\mathrm{PAT}((a_n)_{n \geq 1})$.

Notation. Here and in the rest of the paper, whenever we have any sequence of non-negative real numbers $(x_n)_{n \geq 1}$, we write $\boldsymbol{x} = (x_n)_{n \geq 1}$ in a bold font as a shorthand for the sequence itself, and $(X_n)_{n \geq 0}$ with a capital letter to denote the sequence of partial sums defined for all $n \geq 0$ as $X_n := \sum_{i=1}^n x_i$, where the empty sum is interpreted as 0. This ensures that we can write $(x_n)_{n \geq 1} = (X_n - X_{n-1})_{n \geq 1}$ and hence that defining $(X_n)_{n \geq 1}$ is equivalent to defining $(x_n)_{n \geq 1}$. In particular, we do so for sequences of fitnesses $\mathbf{a} = (a_n)_{n \geq 1}$, for deterministic sequences of weights $\boldsymbol{w} = (w_n)_{n \geq 1}$ and for random sequence of weights $\mathbf{w} = (\mathsf{w}_n)_{n \geq 1}$.

Representation result. The following result gives a connection between these two models of growing trees. It is an analogue of the so-called "Pólya urn-representation" result described in [3, Theorem 2.1] or [7, Section 1.2] for related models, which already cover the case of constant sequences $\mathbf{a}$.

For $a, b > 0$ the distribution $\mathrm{Beta}(a, b)$ has density $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot x^{a-1}(1-x)^{b-1} \cdot \mathbf{1}_{\{0 \leq x \leq 1\}}$ with respect to Lebesgue measure. If $b = 0$ and $a > 0$, we use the convention that the distribution $\mathrm{Beta}(a, b)$ is a Dirac mass at 1.

Theorem 1.1 (WRT-representation of PATs). *For any sequence* $\mathbf{a}$ *of fitnesses, we define the associated random sequence* $\mathbf{w^a} = (\mathsf{w}_n^{\mathbf{a}})_{n \geq 1}$ *through its partial sums* $(\mathsf{W}_n^{\mathbf{a}})_{n \geq 1}$ *as*

$$\mathsf{W}_1^{\mathbf{a}} = 1 \qquad \text{and} \qquad \forall n \geq 2, \quad \mathsf{W}_n^{\mathbf{a}} = \prod_{k=1}^{n-1} \beta_k^{-1},$$

where the $(\beta_k)_{k \geq 1}$ *are independent with respective distribution* $\mathrm{Beta}(A_k + k, a_{k+1})$. *Then, the distributions* $\mathrm{PAT}(\mathbf{a})$ *and* $\mathrm{WRT}(\mathbf{w^a})$ *coincide.*

The result of the theorem is obtained by studying the evolution of the degree of fixed vertices in the preferential attachment model $(\mathsf{P}_n)_{n \geq 1}$. The key argument lies in the fact that we can describe the whole process $(\mathsf{P}_n)_{n \geq 1}$ using a sequence of Pólya urns, related to the degrees of those vertices. The connection of the evolution of the degrees to Pólya urns in the context of preferential attachment models is well-know and was observed for the first time in [36]. It explains why Beta-distributed random variables appear in the limit. In our case, the theorem relies on applying the de Finetti theorem to this sequence of urns and on proving that those urns are jointly independent.

The result stated in the theorem can be made a bit more precise than an equality in distribution as soon as the sequence $\mathbf{a}$ is chosen in such way that almost surely the degree of the first vertex $\deg_{\mathsf{P}_n}^+(u_1)$ tends to infinity as $n \to \infty$. It is easy to check that the condition $A_n = O(n)$ is sufficient to ensure this behaviour, and in this case we can state the following corollary.

Corollary 1.2. *For a sequence* **a** *such that* $A_n = O(n)$, *we can construct the sequence* $\mathbf{w^a}$ *from* $(\mathrm{P}_n)_{n \geq 1}$ *in such a way that for all* $k \geq 1$:

$$\mathrm{w}_k^{\mathbf{a}} = \lim_{n \to \infty} \frac{\deg_{\mathrm{P}_n}^+(u_k)}{\deg_{\mathrm{P}_n}^+(u_1)} \quad almost\ surely. \tag{1.3}$$

The obtained sequence $\mathbf{w^a} = (\mathrm{w}_k^{\mathbf{a}})_{k \geq 1}$ *has the distribution described in Theorem 1.1 and conditionally on* $\mathbf{w^a}$, *the sequence of trees* $(\mathrm{P}_n)_{n \geq 1}$ *has distribution* $\mathrm{WRT}(\mathbf{w^a})$.

In fact, and this is the content of Proposition 1.3 below, if A_n grows linearly as some $c \cdot n$ for some $c > 0$, then the sequence $(\mathrm{W}_n^{\mathbf{a}})_{n \geq 1}$ almost surely grows as some power of n whose exponent depends on c. This is achieved through moment computations, using the explicit definition of $(\mathrm{W}_n^{\mathbf{a}})_{n \geq 1}$ given by the theorem. In the rest of the paper, we investigate several properties of WRTs under this type of assumptions for the sequence of weights, such as convergence of height, profile and measures carried on the tree. Thanks to this connection, our results then also hold for PATs under the assumption that A_n grows linearly.

Assumptions on the sequences. For two sequences (x_n) and (y_n) we say that

$$x_n \underset{n \to \infty}{\bowtie} y_n \qquad \text{if and only if} \qquad \exists \epsilon > 0,\ x_n \underset{n \to \infty}{=} y_n \cdot (1 + O(n^{-\epsilon})).$$

Our main assumption for sequences $\mathbf{a} = (a_n)_{n \geq 1}$ of fitnesses is the following (H_c), which is parametrised by some positive $c > 0$ and ensures that the fitness of vertices is c on average

$$A_n \underset{n \to \infty}{\bowtie} c \cdot n. \tag{H_c}$$

For sequences of weights $\boldsymbol{w} = (w_n)_{n \geq 1}$, we introduce the following hypothesis, which depends on a parameter $\gamma > 0$

$$W_n \underset{n \to \infty}{\bowtie} \mathrm{cst} \cdot n^\gamma, \tag{$\square_\gamma$}$$

where the symbol "cst" denotes an unspecified positive constant. The following proposition ensures in particular that our assumption on sequences of fitnesses **a** translates to a power-law behaviour for the random sequence of cumulated weights $(\mathrm{W}_n^{\mathbf{a}})_{n \geq 1}$ defined in Theorem 1.1.

Proposition 1.3. *Suppose that there exists* $c > 0$ *such that* **a** *satisfies* (H_c), *then the random sequence* $(\mathrm{w}_n^{\mathbf{a}})_{n \geq 1}$ *defined in Theorem 1.1 almost surely satisfies* $(\square_\gamma)$ *with*

$$\gamma = \frac{c}{c+1}.$$

If furthermore **a** *is such that* $a_n \leq (n+1)^{c'+o(1)}$ *for some* $c' \in [0, 1)$, *then almost surely* $\mathrm{w}_n^{\mathbf{a}} \leq (n+1)^{c' - \frac{1}{c+1} + o^\omega(1)}$, *where* $o^\omega(1)$ *denotes a random function of* n *that almost surely tends to* 0 *as* $n \to \infty$.

Convergence of degrees using the WRT representation. In the WRT with a deterministic sequence of weights $\boldsymbol{w}$ that satisfies

$$W_n \underset{n \to \infty}{\sim} C \cdot n^\gamma$$

for some $\gamma \in (0, 1)$, the degree of a fixed vertex evolves as a sum of independent Bernoulli random variables and it is possible to handle it with elementary methods and obtain that a.s. for all $k \geq 1$,

$$\deg_{\mathrm{T}_n}^+(u_k) \underset{n \to \infty}{\sim} \frac{w_k}{C(1-\gamma)} \cdot n^{1-\gamma}.$$

Further calculations allow us to improve this statement to an almost sure convergence

$$n^{-(1-\gamma)} \cdot (\deg_{\mathsf{T}_n}^+(u_1), \deg_{\mathsf{T}_n}^+(u_2), \dots) \to \frac{1}{C(1-\gamma)} \cdot (w_1, w_2, \dots) \tag{1.4}$$

in the space ℓ^p of p-th power summable sequences for some $p \geq 1$, for weight sequences $\boldsymbol{w}$ that satisfy some additional control. A precise version of this statement is given in Proposition 2.1.

Suppose that $\mathbf{a}$ satisfies (H_c) and consider $(\mathsf{P}_n)_{n\geq 1}$ which has distribution PAT($\mathbf{a}$). Then, according to Theorem 1.1 and Corollary 1.2, we know that conditionally on the sequence $\mathbf{w^a} = (\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$ obtained as in (1.3), the sequence $(\mathsf{P}_n)_{n\geq 1}$ has distribution WRT($\mathbf{w^a}$). Also, thanks to Proposition 1.3, we know that there exists some random variable Z such that $\mathsf{W}_n^{\mathbf{a}} \sim Z \cdot n^\gamma$ almost surely as $n \to \infty$, with $\gamma = \frac{c}{c+1}$. This motivates the introduction of the following random sequence

$$(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1} := \frac{1}{Z(1-\gamma)} \cdot (\mathsf{w}_n^{\mathbf{a}})_{n\geq 1} = \frac{c+1}{Z} \cdot (\mathsf{w}_n^{\mathbf{a}})_{n\geq 1} \quad \text{a.s..} \tag{1.5}$$

Applying the convergence (1.4) conditionally on the sequence $(\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$ (or equivalently conditionally on $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$) yields an almost sure convergence in the product topology on sequences, which can be improved to an ℓ^p convergence if $\mathbf{a}$ satisfies some additional control. This is stated below as a theorem.

Theorem 1.4. *Suppose that $\mathbf{a}$ satisfies (H_c). Then for a sequence $(\mathsf{P}_n)_{n\geq 1} \sim \mathrm{PAT}(\mathbf{a})$ we obtain the following almost sure convergence in the product topology*

$$n^{-\frac{1}{c+1}} \cdot (\deg_{\mathsf{P}_n}^+(u_1), \deg_{\mathsf{P}_n}^+(u_2), \dots) \xrightarrow[n\to\infty]{} (\mathsf{m}_1^{\mathbf{a}}, \mathsf{m}_2^{\mathbf{a}}, \dots). \tag{1.6}$$

Furthermore, if $a_n \leq (n+1)^{c'+o(1)}$, for some $0 \leq c' < \frac{1}{c+1}$, the previous convergence also takes place in the space ℓ^p of p-th power summable sequences, for all $p > \frac{c+1}{1-(c+1)c'}$.

Note that the function $\max : \ell^p \to \mathbb{R}$ that outputs the maximum of a sequence is a continuous function, so that the scaling limit of the maximal degree in the tree P_n is ensured by the theorem whenever the appropriate condition on the sequence $\mathbf{a}$ is satisfied. Convergence of the rescaled degree of fixed vertices in preferential attachment trees is a well-know phenomenon in the case a preferential attachment trees with constant fitnesses, as is the convergence of the maximum of that sequence, see [36]. However, to the best of the author's knowledge, Theorem 1.4 is the first result that ensures an almost sure convergence of the rescaled degrees as a sequence in such a topology. This improves the ℓ^p convergence proved in distribution in [40] for a related model, which we treat in Proposition 5.5.

The distribution of the limiting sequence $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$ can be characterized, and even has a reasonable description for certain regular sequences of fitnesses $\mathbf{a}$, as it is explained in the following paragraph. This result is actually related to the study of some urn models like the Pólya urns with immigration of [42] or the periodic Pólya urns of [1] and allows us to provide some alternative proofs and complete some of the known results about those processes. This is developed in Section 5.2.

Distribution of the limiting chain. Let us say a word on the properties of the non-decreasing sequence $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ that corresponds using our notation to the partial sums of the sequence $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$ defined in (1.5). Using the random variables $(\beta_n)_{n\geq 1}$ defined in Theorem 1.1, we can write for any $n \geq 1$,

$$\mathsf{M}_n^{\mathbf{a}} = \frac{c+1}{Z} \cdot \prod_{k=1}^{n-1} \beta_k^{-1},$$

but then, since the random variable Z depends on the whole sequence $(\beta_n)_{n \geq 1}$, the sequence $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ is not just an iterated product of independent random variables, as it was the case for $(\mathsf{W}_n^{\mathbf{a}})_{n \geq 1}$. Nevertheless, the sequence still has the nice property of being a time-inhomogeneous Markov chain with a simple *backward* transition, characterised by the equality

$$\mathsf{M}_n^{\mathbf{a}} = \beta_n \cdot \mathsf{M}_{n+1}^{\mathbf{a}},$$

where β_n is independent of $\mathsf{M}_{n+1}^{\mathbf{a}}$ and has distribution $\mathrm{Beta}(A_n + n, a_{n+1})$. This is the content of Proposition 4.2.

For some specific choices of sequences $\mathbf{a}$, the distribution of the chain $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ is explicit. Whenever $\mathbf{a}$ is of the form

$$\mathbf{a} = a, b, b, b, \ldots \qquad \text{with } a > -1 \text{ and } b > 0,$$

we retrieve Goldschmidt and Haas' Mittag-Leffler Markov chain family, introduced in [21] and also studied by James [26].

The other case where the chain is explicit is when $\mathbf{a}$ is periodic starting from the second term, of the form

$$\mathbf{a} = a, \underbrace{b_1, b_2, \ldots, b_\ell}, \underbrace{b_1, b_2, \ldots, b_\ell}, b_1, b_2 \ldots,$$

where $a > -1$ is a real number and $b_1, b_2, \ldots, b_\ell$ are non-negative integers such that at least one of them is non-zero. Then the sequence $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ has an explicit distribution defined using products of Gamma-distributed random variables. We discuss this in Section 5.1.2.

1.2 Other geometric properties of weighted random trees

We now state the convergence for other statistics of weighted random trees, namely profile, height and probability measures. Here $(\mathsf{T}_n)_{n \geq 1}$ denotes a sequence of trees evolving according to the distribution $\mathrm{WRT}(\boldsymbol{w})$ for some deterministic sequence $\boldsymbol{w}$ and state our results in this setting. Our results also apply to random sequences of weights $\mathbf{w}$ that satisfy the assumptions of the theorems almost surely, they hence apply to PATs with appropriate sequences of fitnesses, thanks to Theorem 1.1 and Proposition 1.3.

1.2.1 Height and profile of WRT

We define

$$\mathbb{L}_n(k) := \# \{1 \leq i \leq n \mid \mathrm{ht}(u_i) = k\}$$

the number of vertices of T_n at height k. The function $k \mapsto \mathbb{L}_n(k)$ is called the *profile* of the tree T_n. The height of the tree is the maximal distance of a vertex to the root, which we can also express as $\mathrm{ht}(\mathsf{T}_n) := \max \{k \geq 0 \mid \mathbb{L}_n(k) > 0\}$. We are interested in the asymptotic behaviour of $\mathbb{L}_n$ and $\mathrm{ht}(\mathsf{T}_n)$ as $n \to \infty$.

In order to express our results, we need to introduce some quantities. For $\gamma > 0$, we define the function $f_\gamma : \mathbb{R} \to \mathbb{R}$ as

$$f_\gamma : z \mapsto f_\gamma(z) := 1 + \gamma \left(e^z - 1 - z e^z\right).$$

This function is increasing on $(-\infty, 0]$ and decreasing on $[0, \infty)$ with $f_\gamma(-\infty) = 1 - \gamma$ and $f_\gamma(0) = 1$ and $f_\gamma(\infty) = -\infty$. We define z_+ and z_- as

$$z_+ := \sup \{z \in \mathbb{R} \mid f_\gamma(z) > 0\} \quad \text{and} \quad z_- := \begin{cases} -\infty & \text{if } \gamma \leq 1, \\ \log((\gamma - 1)/\gamma) & \text{if } \gamma > 1. \end{cases} \qquad (1.7)$$

We assume that we work with a sequence $\boldsymbol{w}$ that satisfies the following assumption ($\square_\gamma^p$) for some $\gamma > 0$ and $p \in (1, 2]$,

$$W_n \underset{n \to \infty}{\bowtie} \text{cst} \cdot n^\gamma \qquad \text{and} \qquad \sum_{i=n}^{2n} w_i^p \le n^{1+(\gamma-1)p+o(1)}. \qquad (\square_\gamma^p)$$

Thanks to Proposition 1.3, this property is almost surely satisfied for $\gamma = \frac{c}{c+1}$ by the random sequence $\boldsymbol{w}^{\mathbf{a}}$ for any sequence $\mathbf{a}$ of fitnesses satisfying $A_n \underset{n \to \infty}{\bowtie} c \cdot n$ and $a_n \le (n+1)^{o(1)}$.

Theorem 1.5. *Suppose that there exists $\gamma > 0$ and $p \in (1, 2]$ such that the sequence $\boldsymbol{w}$ satisfies ($\square_\gamma^p$). Then, for a sequence of random trees $(\mathrm{T}_n)_{n \ge 1} \sim \mathrm{WRT}(\boldsymbol{w})$, we have the almost sure asymptotics for the profile*

$$\mathbb{L}_n(k) \underset{n \to \infty}{=} \frac{n}{\sqrt{2\pi \log n}} \exp\left\{ -\frac{1}{2} \cdot \left(\frac{k - \gamma \log n}{\sqrt{\gamma \log n}} \right)^2 \right\} + O\left(\frac{n}{\log n} \right), \qquad (1.8)$$

where the error term is uniform in $k \ge 0$. Also for any compact $K \subset (z_-, z_+)$ we have almost surely for all $z \in K$

$$\mathbb{L}_n\left(\lfloor \gamma e^z \log n \rfloor \right) = n^{f_\gamma(z) - \frac{1}{2} \frac{\log \log n}{\log n} + O\left(\frac{1}{\log n} \right)}, \qquad (1.9)$$

where the error term is uniform in $z \in K$. Moreover, we have the almost sure convergence

$$\frac{\mathrm{ht}(\mathrm{T}_n)}{\log n} \underset{n \to \infty}{\longrightarrow} \gamma \cdot e^{z_+}. \qquad (1.10)$$

The proof of this result follows the path used for many similar results for trees with logarithmic growth (see [9, 10, 30]): we study the Laplace transform of the profile $z \mapsto \sum_{k=0}^n e^{zk} \mathbb{L}_n(k)$ on a domain of the complex plane and prove its convergence to some random analytic function when appropriately rescaled. Then, we apply [29, Theorem 2.1], which consists of a fine Fourier inversion argument and hence allows to obtain precise asymptotics for $\mathbb{L}_n$. The application of [29, Theorem 2.1] in its full generality proves a so-called Edgeworth expansion for $\mathbb{L}_n$, which we express here in a weaker form by equations (1.8) and (1.9). The convergence (1.8) expresses that the profile is asymptotically close to a Gaussian shape centred around $\gamma \log n$ and with variance $\gamma \log n$, so that a majority of vertices have a height of order $\gamma \log n$. The second equation (1.9) describes the behaviour of the number of vertices at a given height, for heights that are not necessarily close to $\gamma \log n$ (for which the preceding result ensure that there are of order $\frac{n}{\sqrt{\log n}}$ vertices per level). According to this result, at height $\lfloor \gamma e^z \log n \rfloor$ for any $z \in (z_-, z_+)$ there are of order $\frac{n^{f_\gamma(z)}}{\sqrt{\log n}}$ vertices.

Remark that the exponent $f_\gamma(z)$ is continuous in z and tends to 0 when $z \to z_+$. Although this does not directly prove the convergence (1.10), it already provides a lower-bound for $\mathrm{ht}(\mathrm{T}_n)$ since it ensures that asymptotically there always exist vertices at height $\lfloor \gamma e^{(z_+ - \epsilon)} \log n \rfloor$, for any small $\epsilon > 0$. The convergence of the height (1.10) can then be obtained by proving a corresponding upper-bound, which can be done using quite rough estimates.

This result includes the well-known asymptotics $\mathrm{ht}(\mathrm{T}_n) \sim e \log n$ as $n \to \infty$ for the uniform random tree, proved for example in [14, 46]. Using the connection from preferential attachment trees to weighted recursive trees given by Theorem 1.1, it also includes the case of preferential attachment trees with constant fitnesses, for which similar results were proved, in [46] for the height and in [30, 29] for the asymptotic

behaviour of the profile (1.8). This type of asymptotics for the profile in the case of non-constant sequences of fitnesses is new. In a subsequent work [38], Pain and the author studied the corrective terms for the asymptotic behaviour of the height (1.10) and proved that

$$\mathrm{ht}(\mathrm{T}_n) = \gamma e^{z_+} \log n - \frac{3}{2z_+} \log \log n + O_{\mathbb{P}}(1),$$

adapting techniques that were originally developed in the context of branching random walks.

Remark 1.6. Remark that (1.9) can only make sense whenever $f_\gamma(z) > 0$. Since in the case $\gamma > 1$, the function f_γ tends to $1 - \gamma < 0$ as $z \to \infty$, it is understandable that the asymptotics (1.9) cannot be valid for values of z below a certain threshold. Nevertheless, one can check that $f_\gamma(z_-) > 0$ and wonder about what part of the proof breaks down below z_-. In fact, we could (and it would be easier to) prove an analog of Theorem 1.5 for the the *weighted profile* of the tree, which corresponds to the total weight of vertices at every height instead their number, and a statement similar to (1.9) would hold for every z such that $f_\gamma(z) > 0$. What breaks down when z is below z_- is that the number of vertices at height $\lfloor \gamma e^z \log n \rfloor$ start behaving differently from their total weight, which in our proof is reflected in the fact that the Laplace transform of the profile and its "weighted" counterpart do not share the same asymptotics.

As a complement to this result, we mention that there is another case where we can compute the asymptotic height of the tree, which corresponds to sequences $\boldsymbol{w}$ that grow fast to infinity. For any sequence of weights $\boldsymbol{w}$, a quantity of interest is $\sum_{i=2}^{n} \frac{w_i}{W_i}$, which is the expected height of a vertex taken with probability proportional to its weight in T_n. When this quantity grows faster than logarithmically, we have the almost sure convergence (see Proposition 3.16 in Section 3.3)

$$\lim_{n\to\infty} \frac{\mathrm{ht}(\mathrm{T}_n)}{\sum_{i=2}^{n} \frac{w_i}{W_i}} = \lim_{n\to\infty} \frac{\mathrm{ht}(u_n)}{\sum_{i=2}^{n} \frac{w_i}{W_i}} = 1,$$

which indicates that vertices are always added close to the tip of the tree.

1.2.2 Convergence of the weight measure

We study the convergence of some natural probability measures defined on the trees $(\mathrm{T}_n)_{n\geq 1}$. Some of those results are used in the proof Theorem 1.5 and others are motivated by applications developed in the companion paper [51], in which they allow to improve a Gromov-Hausdorff to a Gromov-Hausdorff-Prokhorov scaling limit convergence for some randomly growing graphs.

Plane-tree framework. To express these results it is easier to work with plane trees. We introduce the Ulam-Harris tree $\mathbb{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$, where $\mathbb{N} := \{1, 2, \dots\}$, with the convention that $\mathbb{N}^0 = \{\emptyset\}$. Classically, a plane tree τ is defined as a non-empty subset of $\mathbb{U}$ such that

(i) if $v \in \tau$ and $v = ui$ for some $i \in \mathbb{N}$, then $u \in \tau$,

(ii) for all $u \in \tau$, there exists $\deg_\tau^+(u) \in \mathbb{N} \cup \{0\}$ such that for all $i \in \mathbb{N}$, $ui \in \tau$ iff $i \leq \deg_\tau^+(u)$.

We choose to construct our sequence $(\mathrm{T}_n)_{n\geq 1}$ of weighted recursive trees as plane trees by considering that each time a vertex is added, it becomes the right-most child of its parent. In this way the vertices $(u_1, u_2 \dots)$ of the trees $(\mathrm{T}_n)_{n\geq 1}$, listed in order of arrival, form a sequence of elements of $\mathbb{U}$. In fact, from now on, we always assume that we use

this particular embedded construction, both for the WRT and the PAT. Note that with this representation as unlabelled subsets of $\mathbb{U}$, the tree T_n itself, for any $n \geq 1$, *does not* contain information relative to the labelling (hence the weight) of its vertices, but this piece of information can be read from the sequence $(\mathrm{T}_1, \mathrm{T}_2, \ldots, \mathrm{T}_n)$.

We also denote $\partial \mathbb{U} = \mathbb{N}^{\mathbb{N}}$, which we can be interpreted as the set of infinite paths from the root to infinity, and write $\overline{\mathbb{U}} = \mathbb{U} \cup \partial \mathbb{U}$. We classically endow this set with the distance

$$
\mathrm{d}^{\exp}(u, v) = \begin{cases} 0, & \text{if } u = v, \\ \exp(- \operatorname{ht}(u \wedge v)), & \text{otherwise}, \end{cases} \tag{1.11}
$$

where $u \wedge v$ denotes the most recent common ancestor of u and v, and the height $\operatorname{ht}(u)$ of a vertex $u \in \mathbb{U}$ is defined as the only n such that $u \in \mathbb{N}^n$. Note that even when $u, v \in \partial \mathbb{U}$, their most recent common ancestor $u \wedge v$ belongs to $\mathbb{U}$, as long as $u \neq v$. Endowed with this distance, $\overline{\mathbb{U}}$ is then a complete separable metric space.

In the paper, except when proving results related to the weak convergence of measures, for which we use the topology generated by $\mathrm{d}^{\exp}$, we consider $\mathbb{U}$ as a graph and we compute distances between vertices using the corresponding graph distance, which we denote d. In particular, the height $\operatorname{ht}(u)$ of a vertex u is always its graph distance $\mathrm{d}(\emptyset, u)$ to the root $\emptyset$.

Convergence of measures. For every $n \geq 1$, we define the measure μ_n on $\mathbb{U}$, which only charges the set $\{u_1, \ldots, u_n\}$ of vertices of T_n, with for any $1 \leq k \leq n$,

$$
\mu_n(\{u_k\}) = \frac{w_k}{W_n}. \tag{1.12}
$$

We refer to μ_n as the *natural weight measure* on T_n. The following theorem classifies the possible behaviours of (μ_n) for any weight sequence, and is, in a way, an extension to the case of trees of a theorem of Pemantle [43] for time-dependent Pólya urns.

Theorem 1.7. *The sequence $(\mu_n)_{n \geq 1}$ converges almost surely weakly towards a limiting probability measure μ on $\overline{\mathbb{U}}$. There are three possible behaviours for μ:*

(i) *If $\sum_{i=1}^{\infty} w_i < \infty$, then μ is carried on $\mathbb{U}$.*

(ii) *If $\sum_{i=1}^{\infty} w_i = \infty$ and $\sum_{i=1}^{\infty} \left(\frac{w_i}{W_i} \right)^2 < \infty$, then μ is diffuse and supported on $\partial \mathbb{U}$.*

(iii) *If $\sum_{i=1}^{\infty} \left(\frac{w_i}{W_i} \right)^2 = \infty$ then μ is concentrated on one point of $\partial \mathbb{U}$.*

This convergence can be extended to other natural measures on the tree, such as the uniform measure on T_n, or some "preferential attachment measure" which charges each vertex proportionally to some affine function of its degree. This is the content of Proposition 2.4. Note that in our case of interest, when the sequence w satisfies the assumption $(\square_\gamma)$ for some $\gamma > 0$, we are in case (ii) of the theorem.

In the specific case of a WRT (resp. PAT) with a sequence of weights (resp. of fitnesses) that is constant starting from the second term, the measure μ has an explicit description: for any $u \in \mathbb{U}$, writing $T(u) := \{uv \mid v \in \overline{\mathbb{U}}\}$ for the sub-tree descending from u, the sequences

$$
\left(\frac{\mu(T(ui))}{\mu(T(u))} \right)_{i \geq 1}, \quad \text{for } u \in \mathbb{U},
$$

are independent and have an explicit GEM distribution (see [45] for a definition). Furthermore, the corresponding sequence of trees, conditionally on μ, can be described as a *split tree*. This result, along with other properties of these families of growing trees can be found in [28].

1.3 Organisation of the paper

The paper is organised as follows.

We first investigate some properties of weighted random trees $(\mathrm{T}_n)_{n\geq 1}$ with a deterministic weight sequence $\boldsymbol{w}$. In Section 2.1 we first prove Proposition 2.1 which states the scaling limit convergence of the degree sequence using elementary methods. Then in Section 2.2, we prove the weak convergence of the weight measure μ_n to some limit μ and describe three regimes for its behaviour. We also study other natural measures related to the sequence of trees (T_n) and prove that they also converge towards μ. For all these measures, our main tool consists in introducing martingales related to the mass of a subtree descending from a fixed vertex. This is the content of Theorem 1.7 and Proposition 2.4. In Section 3, we prove Theorem 1.5 about the convergence of the height and the profile of WRT. This is achieved by first proving the uniform convergence of a rescaled version of the Laplace transform of the profile on a complex domain, which is the content of Proposition 3.1. This ensures that we can use [29, Theorem 2.1] for the convergence of the profile. This convergence provides a lower-bound for the height of the tree; we then prove a matching upper-bound to obtain asymptotics for the height. We also prove Proposition 3.16, which identify the asymptotic behaviour of the height of the tree in the case where the weights increase very fast.

Then we switch to studying a sequence $(\mathrm{P}_n)_{n\geq 1}$ of preferential attachment trees with sequence of fitnesses $\mathbf{a}$. In Section 4, we present a proof of Theorem 1.1 and Corollary 1.2 using a coupling between the preferential attachment process with a sequence of Pólya urn processes and this establishes that $(\mathrm{P}_n)_{n\geq 1}$ can also be described as having distribution $\mathrm{WRT}(\mathbf{w^a})$ for a random sequence $\mathbf{w^a}$; we then prove Proposition 1.3 which relates the properties of $\mathbf{w^a}$ to that of $\mathbf{a}$. We finish the section by stating and proving Proposition 4.2 which states that the sequence $(\mathrm{M}_n^{\mathbf{a}})$ defined in (1.5) as some random multiple of $(\mathrm{W}_n^{\mathbf{a}})$ is a Markov chain. In Section 5, we identify in Proposition 5.1 the distribution of the chain $(\mathrm{M}_n^{\mathbf{a}})$ for particular sequences $\mathbf{a}$ using moment identifications. We then present two applications of this result, one concerning a model of Pólya urn with immigration and the other concerning another model of preferential attachment graphs, in Proposition 5.5.

Some technical results can be found in Appendix A.

2 Measures and degrees in weighted random trees

In this section, we work with a sequence of trees $(\mathrm{T}_n)_{n\geq 1}$ that has distribution $\mathrm{WRT}(\boldsymbol{w})$ for some deterministic sequence $\boldsymbol{w}$. We start with two statistics of the tree that are quite easy to analyze, namely the sequence of degrees of the vertices of the tree and also some natural measures defined on the tree.

2.1 Convergence of the degree sequence

We start the section by proving convergence for the sequence of degrees of the vertices in their order of creation in the WRT model. We suppose here that the sequence of weights $\boldsymbol{w}$ is such that there exists constants $C > 0$ and $0 < \gamma < 1$ for which

$$W_k \underset{k\to\infty}{\sim} C \cdot k^\gamma. \tag{2.1}$$

We write $\deg^+_{\mathrm{T}_n}(u_k)$ for the out-degree of the vertex u_k in T_n. For a fixed $k \geq 1$ remark that, as a sequence of random variables indexed by $n \geq 1$, we have the equality in distribution

$$\left(\deg^+_{\mathrm{T}_n}(u_k)\right)_{n\geq 1} \overset{(\mathrm{d})}{=} \left(\sum_{i=k}^{n-1} \mathbf{1}_{\left\{U_i \leq \frac{w_k}{W_i}\right\}}\right)_{n\geq 1}, \tag{2.2}$$

with $(U_i)_{i\geq 1}$ a sequence of independent uniform variables in $(0,1)$. With this description of the distribution of the degrees of fixed vertices, only using some law of large numbers for the convergence and Chernoff bounds for the fluctuations we obtain the following result.

Proposition 2.1. *For a sequence of weights $\boldsymbol{w}$ satisfying (2.1), the following holds.*

(i) *We have the almost sure pointwise convergence*

$$n^{-(1-\gamma)} \cdot (\deg_{\mathrm{T}_n}^+(u_1), \deg_{\mathrm{T}_n}^+(u_2), \dots) \xrightarrow[n\to\infty]{} \frac{1}{(1-\gamma)C} \cdot (w_1, w_2, \dots). \qquad (2.3)$$

(ii) *If the sequence furthermore satisfies $w_k \leq (k+1)^{\gamma-1+c'+o(1)}$ for some constant $0 \leq c' < 1-\gamma$, then there exists a deterministic function of k which goes to 0 as $k \to \infty$, also denoted $o(1)$, such that all n large enough, we have for all $k \geq 1$*

$$\deg_{\mathrm{T}_n}^+(u_k) \leq n^{1-\gamma} \cdot (k+1)^{\gamma-1+c'+o(1)}, \qquad (2.4)$$

and the convergence (2.3) holds almost surely in the space ℓ^p for all $p > \frac{1}{1-\gamma-c'}$.

Proof. To prove (i), first remark that for any $k \geq 1$ such that $w_k \neq 0$, thanks to (2.1), we have

$$\sum_{i=k}^{n-1} \frac{w_k}{W_i} \underset{n\to\infty}{\sim} w_k \cdot \frac{n^{1-\gamma}}{C(1-\gamma)}.$$

Using a version of the Borel-Cantelli lemma (see Lemma A.1 in the appendix), we get that almost surely

$$\deg_{\mathrm{T}_n}^+(u_k) = \sum_{i=k}^{n-1} \mathbf{1}_{\left\{U_i \leq \frac{w_k}{W_i}\right\}} \underset{n\to\infty}{\sim} \sum_{i=k}^{n-1} \frac{w_k}{W_i} \underset{n\to\infty}{\sim} w_k \cdot \frac{n^{1-\gamma}}{(1-\gamma)C},$$

and hence $n^{-(1-\gamma)} \cdot \deg_{\mathrm{T}_n}^+(u_k) \to \frac{w_k}{(1-\gamma)C}$. For the indices k for which $w_k = 0$, we of course have $\deg_{\mathrm{T}_n}^+(u_k) = 0$ almost surely for all $n \geq 1$, and so the convergence also holds. This finishes the proof of (i).

For the second part of the statement, we first compute

$$\mathbb{E}\left[\exp\left(\deg_{\mathrm{T}_n}^+(u_k)\right)\right] = \mathbb{E}\left[\exp\left(\sum_{i=k}^{n-1} \mathbf{1}_{\left\{U_i \leq \frac{w_k}{W_i}\right\}}\right)\right] = \prod_{i=k}^{n-1}\left(1 + (e-1)\frac{w_k}{W_i}\right)$$

$$\leq \exp\left((e-1)w_k \sum_{i=k}^{n-1} \frac{1}{W_i}\right),$$

where we have used the inequality $1 + x \leq e^x$. Now let C' be a constant such that for all $n \geq 1$, we have $\sum_{i=1}^{n-1} \frac{1}{W_i} \leq C' \cdot n^{1-\gamma}$ (such a constant exists because of the assumption (2.1)). For all $k \geq 1$, we introduce the following

$$\xi_k := \max\left(2C'(e-1)w_k, k^{\gamma-1}\log^2(k+a)\right),$$

where the real number $a > 0$ is chosen in such a way that the function $x \mapsto x^{\gamma-1}\log(x+a)$ is decreasing on $(0,\infty)$. Using the definition of C' and of ξ_k, together with Markov's inequality, we get that for any integers k and n such that $n \geq k$

$$\mathbb{P}\left(\deg_{\mathrm{T}_n}^+(u_k) \geq \xi_k \cdot n^{1-\gamma}\right) \leq \exp\left(-\xi_k \cdot n^{1-\gamma} + (e-1)w_k \sum_{i=k}^{n-1} \frac{1}{W_i}\right)$$

$$\leq \exp\left(-\frac{1}{2} \cdot \xi_k \cdot n^{1-\gamma}\right).$$

Using a union bound and the fact that $\deg^+_{T_n}(u_k) = 0$ for any $k > n$, we get that for all $n \geq 1$,

$$\mathbb{P}\left(\exists k \geq 1, \quad \deg^+_{T_n}(u_k) \geq \xi_k \cdot n^{1-\gamma}\right) \leq \sum_{k=1}^{n} \exp\left(-\frac{1}{2} \cdot \xi_k \cdot n^{1-\gamma}\right)$$

$$\leq \sum_{k=1}^{n} \exp\left(-\frac{1}{2} \cdot k^{\gamma-1} \log^2(k+a) \cdot n^{1-\gamma}\right)$$

$$\leq n \cdot \exp\left(-\frac{1}{2} \cdot \log^2(n+a)\right).$$

The last display is summable over all $n \geq 1$ and hence using the Borel-Cantelli lemma, we almost surely have for n large enough

$$\forall k \geq 1, \quad \deg^+_{T_n}(u_k) \leq n^{1-\gamma} \cdot \xi_k.$$

We can conclude by noting that under our assumptions we have $\xi_k \leq (k+1)^{\gamma-1+c'+o(1)}$. The convergence in ℓ^p for $p > \frac{1}{1-\gamma-c'}$ is just obtained by dominated convergence using the componentwise convergence (2.3) and the ℓ^p domination (2.4). □

2.2 Convergence of measures

The goal of this section is to prove Theorem 1.7, which concerns the convergence of the sequence of weight measures (μ_n) seen as measures on $\overline{\mathbb{U}}$. One of the key arguments is the fact that the weight of the subtree descending from a fixed vertex can be described using a generalised Pólya urn scheme, as studied by Pemantle [43]. We also prove Proposition 2.4, which states the weak convergence of other measures.

Time-dependent Pólya urn scheme. We start by describing an urn process, following Pemantle [43]. Let a, b be two non-negative real numbers, with $a + b > 0$, and $k \geq 1$ be an integer and $(s_n)_{n \geq k+1}$ be a sequence of non-negative real numbers. We refer to the following process as a time-dependent Pólya urn starting at time k with a red balls and b black balls and weight sequence $(s_n)_{n \geq k+1}$:

- At time k, the urn contains a red balls and b black balls[2].

- Then at every time $n \geq k + 1$, a ball is drawn at random and replaced in the urn, along with s_n additional balls of the same colour.

For any $n \geq k$ we call R_n the proportion of red balls in the urn at time n. We can easily check that $(R_n)_{n \geq k}$ is a martingale in its own filtration, with values in $[0, 1]$. As a result, it converges as $n \to \infty$ a.s. and in L^1 towards some random variable R_∞.

Characterization of the convergence of probability measures over $\overline{\mathbb{U}}$. Recall from the introduction the definition of the Ulam-Harris tree $\mathbb{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$ and its completed version $\overline{\mathbb{U}} = \mathbb{U} \cup \partial\mathbb{U}$, which is endowed with the distance $\mathrm{d}^{\exp}$ defined in (1.11). Recall that $(\overline{\mathbb{U}}, \mathrm{d}^{\exp})$ is a separable and complete metric space.

For any $u \in \mathbb{U}$, we write $T(u) := \{uv \mid v \in \overline{\mathbb{U}}\}$ the subtree descending from u. In $\overline{\mathbb{U}}$ there is an easy characterisation of the weak convergence of sequences of probability measures defined on the Borel-σ-field associated to $\mathrm{d}^{\exp}$, which a direct consequence of the Portmanteau theorem (see e.g. [6, Theorem 2.1]):

---

[2]Those numbers of balls are not required to be integers.

Lemma 2.2. *Let $(\pi_n)_{n \geq 1}$ be a sequence of Borel probability measures on $\overline{\mathbb{U}}$. Then $(\pi_n)_{n \geq 1}$ converges weakly to a probability measure π if and only if for any $u \in \mathbb{U}$,*

$$\pi_n(\{u\}) \to \pi(\{u\}) \quad and \quad \pi_n(T(u)) \to \pi(T(u)) \quad as \quad n \to \infty.$$

We provide a proof of this lemma for completeness.

Proof. We can check that the sets of the form $\{u\}$ for $u \in \mathbb{U}$, or $T(u)$ for $u \in \mathbb{U}$, are clopen (closed and open) for the topology generated by $\mathrm{d}^{\exp}$, so by the Portmanteau theorem this already proves the "only if" part of the lemma. Now reciprocally, suppose that the condition on $(\pi_n)_{n \geq 1}$ is satisfied. We can check that every open set $\mathcal{O}$ can be written as a countable disjoint union of these clopen sets (see for example [49, Lemma 1.2] for a similar statement for the topology of $\partial \mathbb{U}$), which we write $\mathcal{O} = \bigsqcup_{k \geq 1} \mathcal{O}_k$. Then, using Fatou's lemma and the σ-additivity of measures we get

$$\liminf_{n \to \infty} \pi_n(\mathcal{O}) = \liminf_{n \to \infty} \sum_{k=1}^{\infty} \pi_n(\mathcal{O}_k) \geq \sum_{k=1}^{\infty} \liminf_{n \to \infty} \pi_n(\mathcal{O}_k) = \sum_{k=1}^{\infty} \pi(\mathcal{O}_k) = \pi(\mathcal{O})$$

We conclude using the Portmanteau theorem again. $\qquad \square$

2.2.1 Proof of Theorem 1.7.

We apply the criterion given in Lemma 2.2 to our sequence $(\mu_n)_{n \geq 1}$, which, we recall, is defined in such a way that for all $n \geq 1$, the measure μ_n charges only the vertices $\{u_1, u_2, \ldots, u_n\}$ of the tree $\mathtt{T}_n$, and such that for any $1 \leq k \leq n$,

$$\mu_n(\{u_k\}) = \frac{w_k}{W_n}. \tag{2.5}$$

Proof of Theorem 1.7. We can already see that if $(W_n)_{n \geq 1}$ is bounded and hence converges to some W_∞ we have $\mu_n(\{u_k\}) \to \frac{w_k}{W_\infty}$ as $n \to \infty$. In this case it is easy to verify that (μ_n) weakly converges to the measure μ which is such that $\mu(\{u_k\}) = \frac{w_k}{W_\infty}$ for all $k \geq 1$ and $\mu(\overline{\mathbb{U}} \setminus \{u_1, u_2, \ldots\}) = 0$. In this case $\mu(\mathbb{U}) = 1$ and so μ is carried on $\mathbb{U}$.

From now on, we assume that $W_n \to \infty$ as $n \to \infty$ and show that in this case, μ_n converges weakly to some limit μ that is carried on $\partial \mathbb{U}$. In this case we have $\mu_n(\{u_k\}) = \frac{w_k}{W_n} \to 0$ as $n \to \infty$. We denote for every integers $n, k \geq 1$,

$$M_n^{(k)} := \mu_n(T(u_k)),$$

the proportion of the total mass above vertex u_k at time n. Remark that this quantity evolves as the proportion of red balls in a time-dependent Pólya urn scheme with weights $(w_i)_{i \geq k+1}$, starting at time k with W_{k-1} black balls and w_k red balls. Hence for all $k \geq 1$, the sequence $(M_n^{(k)})_{n \geq k}$ almost surely converges to a limit $M_\infty^{(k)}$. Also, for any $u \in \mathbb{U}$ such that $u \notin \{u_1, u_2, \ldots\}$, the sequence $(\mu_n(T(u)))_{n \geq 1}$ (and also $(\mu_n(\{u\}))_{n \geq 1}$) is identically equal to zero. Hence we have convergence of $(\mu_n(\{u\}))_{n \geq 1}$ and $(\mu_n(T(u)))_{n \geq 1}$ for all $u \in \mathbb{U}$.

The last step in order to prove the weak convergence of $(\mu_n)_{n \geq 1}$ is to prove that the quantities that we obtain in the limit indeed define a probability measure on $\overline{\mathbb{U}}$. If for all $u \in \mathbb{U}$ we have

$$\lim_{n \to \infty} \mu_n(T(u)) = \sum_{i=1}^{\infty} \lim_{n \to \infty} \mu_n(T(ui)), \tag{2.6}$$

then it entails that $\mu_n \underset{n \to \infty}{\to} \mu$, where μ is the unique probability measure on $\overline{\mathbb{U}}$ such that for all $u \in \mathbb{U}$,

$$\mu(\{u\}) = 0 \quad and \quad \mu(T(u)) = \lim_{n \to \infty} \mu_n(T(u)).$$

The existence of such a measure μ is ensured by the Kolmogorov extension theorem on the product space $\partial \mathbb{U} = \mathbb{N}^{\mathbb{N}}$.

For any $u \notin \{u_1, u_2, \dots\}$, the equality (2.6) is immediate, so we focus on proving it for all u_k for $k \geq 1$. For any $n, k, i \geq 1$, let

$$M_n^{(k,i)} := \mu_n(\{u_k\}) + \sum_{j=i+1}^{\infty} \mu_n(T(u_k j)) = \mu_n\left(T(u_k)\right) - \sum_{j=1}^{i} \mu_n\left(T(u_k j)\right).$$

Using the right-hand-side of the last display, and the fact that we already proved the a.s. convergence of the sequences $(\mu_n(T(u)))_{n \geq 1}$ for all $u \in \mathbb{U}$, we know that for any k, i, the quantity $M_n^{(k,i)}$ almost surely converges as $n \to \infty$ to some limit $M_\infty^{(k,i)}$. Proving (2.6) then reduces to proving that for any $k \geq 1$, we almost surely have $M_\infty^{(k,i)} \underset{i \to \infty}{\to} 0$. By construction, the sequence $(M_\infty^{(k,i)})_{i \geq 1}$ is non-negative and non-increasing, hence it converges almost surely, so it suffices to prove that its almost sure limit is 0.

We define $\tau^{(k,i)} := \inf\{n \geq 1 \mid u_n = u_k i\}$, the time when the vertex u_k receives its i-th child in the growth procedure. Conditionally on the event $\{\tau^{(k,i)} = t\}$, the process $(M_n^{(k,i)})_{n \geq t}$ evolves as the proportion of red balls in a time-dependent Pólya urn scheme, starting at time t with w_k red balls (that correspond to the weight of u_k) and $(W_t - w_k)$ blacks balls (that correspond to the total weight of other vertices in the tree), and weights $(w_m)_{m \geq t+1}$. Hence we have

$$\mathbb{E}\left[M_\infty^{(k,i)} \,\Big|\, \tau^{(k,i)} = t\right] = \mathbb{E}\left[M_t^{(k,i)} \,\Big|\, \tau^{(k,i)} = t\right] = \frac{w_k}{W_t}.$$

On the event $\{\tau^{(k,i)} = \infty\}$, we have $M_n^{(k,i)} = 0$ for $n < k$ and $M_n^{(k,i)} = \mu_n(\{u_k\})$ for $n \geq k$, which decreases almost surely to 0, so $M_\infty^{(k,i)} = 0$ a.s. on that event.

Using the crude bound $\tau^{(k,i)} \geq i$, which entails that $W_{\tau^{(k,i)}} \geq W_i$ almost surely, we get

$$\mathbb{E}\left[M_\infty^{(k,i)}\right] = \mathbb{E}\left[M_\infty^{(k,i)} \mathbf{1}_{\left\{\tau^{(k,i)} < \infty\right\}}\right] = \mathbb{E}\left[M_{\tau^{(k,i)}}^{(k,i)} \mathbf{1}_{\left\{\tau^{(k,i)} < \infty\right\}}\right] = \mathbb{E}\left[\frac{w_k}{W_{\tau^{(k,i)}}} \mathbf{1}_{\left\{\tau^{(k,i)} < \infty\right\}}\right]$$
$$\leq \frac{w_k}{W_i} \underset{i \to \infty}{\to} 0,$$

hence $M_\infty^{(k,i)} \underset{i \to \infty}{\to} 0$ in L^1, so its almost sure limit is also 0. In the end, by Lemma 2.2, the sequence of measures (μ_n) almost surely converges weakly to a limit μ, and this measure only charges the set $\partial \mathbb{U}$.

We just finished proving that, for any sequence of weight $\boldsymbol{w}$, the sequence $(\mu_n)_{n \geq 1}$ almost surely converges weakly to a probability measure μ. Furthermore, we proved that μ is carried on $\mathbb{U}$ when (W_n) is bounded and carried on $\partial \mathbb{U}$ when $W_n \to \infty$. The proof is then finished by applying the lemma stated below. $\qquad \square$

Lemma 2.3. *Suppose that $\sum_{n=1}^{\infty} w_n = \infty$ so that μ is carried on $\partial \mathbb{U}$. Then either $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2 < \infty$ and then μ is almost surely diffuse, or $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2 = \infty$ and then μ is carried on one point of $\partial \mathbb{U}$.*

Proof. For any $k \geq 1$ the process $(\mu_n(T(u_k))_{n \geq k}$ evolves as the proportion of red balls in a time-dependent Pólya urn started at time k with w_k red balls, W_{k-1} black balls and a weight sequence $(w_n)_{n \geq k+1}$. By the work of Pemantle in [44], if we assume $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2 = \infty$ then the limiting proportion $\mu(T(u_k))$ almost surely belongs to the set $\{0, 1\}$. This translates into the fact that $\mu(T(u)) \in \{0, 1\}$ almost surely for any $u \in \mathbb{U}$, which entails that μ is almost surely carried on one leaf of $\partial \mathbb{U}$.

On the contrary, we now assume that $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n} \right)^2 < \infty$ and prove that this entails that the limiting measure μ is diffuse almost surely. Consider the function $(\cdot \wedge \cdot) : \overline{\mathbb{U}} \times \overline{\mathbb{U}} \to \overline{\mathbb{U}}$ which associates to each couple (u, v) their most recent common ancestor $u \wedge v$ in the completed tree $\overline{\mathbb{U}}$. This function is continuous with respect to the distance $\mathrm{d}^{\exp}$. Then, since $\mu_n \to \mu$ almost surely weakly, we also get the following almost sure weak convergence of the push-forward of the product measure $\mu_n \otimes \mu_n$ on $\mathbb{U} \times \mathbb{U}$ by the function $(\cdot \wedge \cdot)$:

$$(\cdot \wedge \cdot)_* (\mu_n \otimes \mu_n) \to (\cdot \wedge \cdot)_* (\mu \otimes \mu). \tag{2.7}$$

We fix $n \geq 1$ and conditionally on $(\mathtt{T}_1, \mathtt{T}_2, \ldots, \mathtt{T}_n)$, let D_n and D'_n be two independent vertices taken under μ_n. Then, an argument taken from the proof of [12, Lemma 3.8] in a slightly different setting ensures that

$$\mathbb{P}\left(D_n \wedge D'_n = u_k \right) = \left(\frac{w_k}{W_k} \right)^2 \cdot \prod_{i=k+1}^{n} \left(1 - \left(\frac{w_i}{W_i} \right)^2 \right)$$

$$\xrightarrow[k \to \infty]{} p_k := \left(\frac{w_k}{W_k} \right)^2 \cdot \prod_{i=k+1}^{\infty} \left(1 - \left(\frac{w_i}{W_i} \right)^2 \right).$$

The argument goes as follows:

- with probability $\left(\frac{w_n}{W_n} \right)^2$, we have $D_n = D'_n = D_n \wedge D'_n = u_n$,

- with probability $1 - \left(\frac{w_n}{W_n} \right)^2$, it is not the case, and we can check that conditionally on this event, the vertices $[D_n]_{n-1}$ and $[D'_n]_{n-1}$ defined as the most recent ancestor in $\mathtt{T}_{n-1}$ of respectively D_n and D'_n, are independent and taken under the measure μ_{n-1}, and that $D_n \wedge D'_n = [D_n]_{n-1} \wedge [D'_n]_{n-1}$.

It suffices to then apply this in cascade to get the last display.

Note that thanks to the summability condition, the infinite product $\prod_{i=2}^{\infty} \left(1 - \left(\frac{w_i}{W_i} \right)^2 \right)$ is non-zero, and this suffices to ensure that the obtained sequence $(p_k)_{k \geq 1}$ is a probability distribution. Thanks to the weak convergence (2.7), it corresponds to the (annealed) distribution $p_k = \mathbb{P}\left(D_\infty \wedge D'_\infty = u_k \right)$, where D_∞ and D'_∞ are two independent points taken under the measure μ, conditionally on $(\mathtt{T}_n)_{n \geq 1}$. Now we can write

$$\mathbb{P}\left(\mathrm{d}^{\exp}(D_\infty, D'_\infty) \leq e^{-k} \right) = \mathbb{P}\left(\mathrm{ht}(D_\infty \wedge D'_\infty) \geq k \right) \leq \sum_{i=k+1}^{\infty} p_i,$$

where the inequality is due to the fact that the vertices $u_1, u_2, \ldots, u_k$ have a height smaller than k. Hence $\mathbb{P}\left(\mathrm{d}^{\exp}(D_\infty, D'_\infty) = 0 \right) \leq \lim_{k \to \infty} \mathbb{P}\left(\mathrm{d}^{\exp}(D_\infty, D'_\infty) \leq e^{-k} \right) = 0$. So, almost surely, two points taken independently under μ are different, and this ensures that μ is diffuse. $\qquad \square$

2.2.2 Convergence of other sequences of measures.

In the last part of this section, we study two other sequences of measures (η_n) and (ν_n) carried on the Ulam tree $\mathbb{U}$. For every $n \geq 2$, these measures only charge the vertices $\{u_1, u_2, \ldots, u_n\}$ in such a way that for any $1 \leq k \leq n$,

$$\eta_n(\{u_k\}) = \frac{b_k + \deg_{\mathtt{T}_n}^+(u_k)}{B_n + n - 1} \quad \text{and} \quad \nu_n(\{u_k\}) = \frac{1}{n},$$

where $(b_n)_{n\geq 1}$ is a sequence of real numbers such that $b_1 > -1$ and $b_n \geq 0$ for all $n \geq 2$. We write $B_n := \sum_{k=1}^{n} b_k$. We suppose that $B_n = O(n)$ and that there exists $\epsilon > 0$ such that $b_n = O(n^{1-\epsilon})$. The assumptions on the sequence $(b_n)_{n\geq 1}$ are chosen such that they are satisfied by a sequence $(a_n)_{n\geq 1}$ of fitnesses that satisfies (H_c) for some $c > 0$.

Proposition 2.4. *Under the assumptions $\sum_{n=1}^{\infty} w_n = \infty$ and $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2 < \infty$, the sequences $(\eta_n)_{n\geq 1}$ and $(\nu_n)_{n\geq 1}$ converge almost surely weakly towards the limiting measure μ on $\partial \mathbb{U}$ that is defined in Theorem 1.7.*

The convergence of the uniform measure $(\nu_n)_{n\geq 1}$ is used in the proof of Theorem 1.5, and the case of $(\eta_n)_{n\geq 1}$ is mostly motivated by the applications carried out in [51]. The rest of this section is devoted to proving Proposition 2.4. We treat the two sequences of measures separately. We use the notation $\mathcal{F}_n = \sigma(\mathtt{T}_1, \mathtt{T}_2, \ldots, \mathtt{T}_n)$ for all $n \geq 1$.

The degree measure. Consider the sequence $(\eta_n)_{n\geq 1}$ on $\overline{\mathbb{U}}$. Since the sequence $(W_n)_{n\geq 1}$ tends to infinity, we have $\eta_n(\{u\}) \to 0$ for every $u \in \mathbb{U}$. Indeed, using the equality in distribution (2.2) and Lemma A.1 in the appendix, it is easy to see that either $\sum_{i=1}^{\infty} W_i^{-1} < \infty$ in which case the degrees $\deg_{\mathtt{T}_n}^{+}(u_k)$ are eventually constant as $n \to \infty$; or $\sum_{i=1}^{\infty} W_i^{-1} = \infty$, in which case we have the almost sure asymptotic behaviour $\deg_{\mathtt{T}_n}^{+}(u_k) \sim w_k \cdot \sum_{i=k}^{n} W_i^{-1}$. Trivially in the first case and by Cesàro summation in the second, for all $k \geq 1$, we have $n^{-1} \deg_{\mathtt{T}_n}^{+}(u_k) \to 0$ almost surely as $n \to \infty$.

For all $k \geq n$, we keep the notation $M_n^{(k)} = \mu_n(T(u_k))$ introduced in the proof of Theorem 1.7 and let

$$N_n^{(k)} := \eta_n(T(u_k)).$$

We can check that

$$N_{n+1}^{(k)} = \frac{1}{B_{n+1} + n} \cdot \left((B_n + n - 1) \cdot N_n^{(k)} + (b_{n+1} + 1) \cdot \mathbf{1}_{\{u_{n+1} \in T(u_k)\}} \right).$$

Now, using that $\mathbb{P}\left(u_{n+1} \in T(u_k) \mid \mathcal{F}_n\right) = M_n^{(k)}$ and that $\mathbb{E}\left[M_{n+1}^{(k)} \,\middle|\, \mathcal{F}_n\right] = M_n^{(k)}$, we get

$$\mathbb{E}\left[N_{n+1}^{(k)} - M_{n+1}^{(k)} \,\middle|\, \mathcal{F}_n\right] = \frac{B_n + n - 1}{B_{n+1} + n} \cdot N_n^{(k)} + \frac{b_{n+1} + 1}{B_{n+1} + n} \cdot M_n^{(k)} - M_n^{(k)}$$

$$= \frac{B_n + n - 1}{B_{n+1} + n} \cdot \left(N_n^{(k)} - M_n^{(k)}\right).$$

Hence, if we denote $X_n^{(k)} := (B_n + n - 1) \cdot \left(N_n^{(k)} - M_n^{(k)}\right)$, then the last computation shows that $\left(X_n^{(k)}\right)_{n\geq k}$ is a martingale for the filtration $(\mathcal{F}_n)_{n\geq 1}$. More precisely we can write

$$X_{n+1}^{(k)} - X_n^{(k)} = \underbrace{\left((1 + b_{n+1}) - \frac{w_{n+1}}{W_{n+1}} (B_{n+1} + n) \right)}_{c_n} \cdot \left(\mathbf{1}_{\{u_{n+1} \in T(u_k)\}} - M_n^{(k)}\right),$$

hence we have

$$\mathbb{E}\left[X_{n+1}^{(k)} - X_n^{(k)} \,\middle|\, \mathcal{F}_n\right] = 0 \quad \text{and} \quad \mathbb{E}\left[\left(X_{n+1}^{(k)} - X_n^{(k)}\right)^2\right] \leq c_n^2.$$

Then, using [18, Chapter VII.9, Theorem 3], we get that if

$$\sum_{n=k}^{\infty} n^{-2} c_n^2 < \infty \tag{2.8}$$

then $\frac{X_n^{(k)}}{n} \to 0$ a.s. as $n \to \infty$, which would prove that $N_n^{(k)} \longrightarrow M_\infty^{(k)}$ as $n \to \infty$. In our case, we can verify that (2.8) holds. Indeed, we have

$$
\begin{aligned}
n^{-2} c_n^2 &= n^{-2} \left((1 + b_{n+1}) - \frac{w_{n+1}}{W_{n+1}} \left(B_{n+1} + n \right) \right)^2 \\
&\leq n^{-2} \cdot 3 \left(1 + b_{n+1}^2 + \left(\frac{w_{n+1}}{W_{n+1}} \left(B_{n+1} + n \right) \right)^2 \right) \\
&\leq 3n^{-2} + 3b_{n+1}^2 n^{-2} + \mathrm{cst} \cdot \left(\frac{w_{n+1}}{W_{n+1}} \right)^2 ,
\end{aligned}
$$

and thanks to our assumptions, each term in the last display is that of a summable series. In particular, for the second term, this can be proved using the fact that $B_n = O(n)$ and $b_{n+1} = O\left(n^{1-\epsilon}\right)$, through a computation similar to that of (A.4) in the appendix. In the end, using Lemma 2.2, we have the almost sure convergence

$$
\eta_n \to \mu \qquad \text{weakly.}
$$

The uniform measure on the vertices of $\mathtt{T}_n$**.** Consider the sequence (ν_n) on $\overline{\mathbb{U}}$. Fix $k \geq 1$. For any $n \geq k$ we can write $\nu_n(T(u_k)) = \frac{1}{n} \sum_{i=k}^n \mathbf{1}_{\{u_i \in T(u_k)\}}$. For any $i \geq k+1$, we have $\mathsf{p}_i := \mathbb{P}\left(u_i \in T(u_k) \mid \mathcal{F}_{i-1} \right) = \mu_{i-1}(T(u_k))$, which tends a.s. to some limit $\mu(T(u_k))$ as $i \to \infty$. Using Lemma A.1 in the appendix, we have

$$
\frac{\sum_{i=k+1}^n \mathbf{1}_{\{u_i \in T(u_k)\}}}{\sum_{i=k+1}^n \mathsf{p}_i} \underset{n \to \infty}{\longrightarrow} 1 \quad \text{a.s. on the event} \quad \left\{ \sum_{i=k+1}^\infty \mathsf{p}_i = \infty \right\},
$$

and also

$$
\sum_{i=k+1}^n \mathbf{1}_{\{u_i \in T(u_k)\}} \quad \text{converges a.s. on the event} \quad \left\{ \sum_{i=k+1}^\infty \mathsf{p}_i < \infty \right\}.
$$

In both cases we get $\nu_n(T(u_k)) \underset{n \to \infty}{\to} \lim_{i \to \infty} \mathsf{p}_i = \mu(T(u_k))$ almost surely. We also have for any $k \geq 1$,

$$
\nu_n(\{u_k\}) = \frac{1}{n} \underset{n \to \infty}{\to} 0 \qquad \text{and} \qquad \forall u \notin \{u_1, u_2, \dots\}, \forall n \geq 1, \ \nu_n(\{u\}) = \nu_n(T(u)) = 0,
$$

so we can conclude using Lemma 2.2 that almost surely $\nu_n \underset{n \to \infty}{\to} \mu$ weakly.

3 Height and profile of WRT

The main goal of this section is to prove Theorem 1.5 which gives asymptotics for the profile and height of the tree. Recall that we denote

$$
\mathbb{L}_n(k) := \# \left\{ 1 \leq i \leq n \mid \mathrm{ht}(u_i) = k \right\},
$$

the number of vertices at height k in the tree $\mathtt{T}_n$. In order to get information on the sequence of functions $(k \mapsto \mathbb{L}_n(k))_{n \geq 1}$ we study their Laplace transform

$$
z \mapsto \sum_{k=0}^\infty \mathbb{L}_n(k) e^{kz} = \sum_{i=1}^n e^{z \, \mathrm{ht}(u_i)} = n \cdot \int_{\mathbb{U}} e^{z \, \mathrm{ht}(u)} \mathrm{d}\nu_n(u), \tag{3.1}
$$

where the last expression is given using an integral against the probability measure ν_n defined in Section 2.2 as the uniform measure on the vertices of $\mathtt{T}_n$. The key result in

our approach is to prove the convergence of this sequence of analytic functions when appropriately rescaled, uniformly in z on an open neighbourhood of 0 in the complex plane. It then allows us to use [29, Theorem 2.1] and hence derive a convergence result for the profile. We actually start in Section 3.1 by studying the asymptotic behaviour of the similarly defined sequence of functions

$$z \mapsto \int_{\mathbb{U}} e^{z\,\mathrm{ht}(u)}\mathrm{d}\mu_n(u) = \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}, \tag{3.2}$$

where we integrate with respect to the weight measure μ_n instead of the uniform measure ν_n as before. This one is easier to study because for every fixed $z \in \mathbb{C}$, it defines a martingale as n grows, up to some deterministic scaling. Then in Section 3.2, we make use of this first result and show that up to some deterministic multiplicative constant, the two sequences of integrals appearing in (3.1) and (3.2) are almost surely equivalent when n tends to infinity.

We fix the value of $\gamma > 0$ for this whole section. Throughout this section, we always work under the assumption that $(\square_\gamma)$ holds for the sequence w. For some results, we additionally assume that their exists $p \in (1\,,2]$ such that the stronger condition $(\square_\gamma^p)$ holds, i.e.

$$W_n \underset{n\to\infty}{\bowtie} \mathrm{cst}\cdot n^\gamma \qquad \text{and} \qquad \sum_{i=n}^{2n} w_n^p \leq n^{1+(\gamma-1)p+o(1)}.$$

We let $\phi : z \mapsto \gamma(e^z - 1)$ be a function of a complex parameter z and let $z \mapsto N_n(z)$ be the following rescaled version of the Laplace transform of the profile

$$N_n(z) := n^{-(1+\phi(z))} \sum_{k=0}^{\infty} \mathbb{L}_n(k) e^{zk}. \tag{3.3}$$

The proposition below ensures that the sequence $(z \mapsto N_n(z))_{n\geq 1}$ converges uniformly on all compact subsets of some complex *domain* $\mathscr{D}$, i.e. an open, connected subset of $\mathbb{C}$, to some limiting function $z \mapsto N_\infty(z)$ which does not vanish anywhere on the set $\mathscr{D} \cap \mathbb{R}$, along with some more technical statements.

Proposition 3.1. *Suppose that the weight sequence w satisfies $(\square_\gamma^p)$ for some $\gamma > 0$ and some $p \in (1\,,2]$. Then there exists a domain $\mathscr{D} \subset \mathbb{C}$ such that $\mathscr{D} \cap \mathbb{R} = (z_-\,,z_+)$ where $z_- < 0$ and $z_+ > 0$ are defined as in (1.7), such that the following properties are satisfied:*

(i) *With probability 1, the sequence of random analytic functions $(z \mapsto N_n(z))_{n\geq 1}$ converges uniformly on all compact subsets of $\mathscr{D}$, as $n \to \infty$, to some random analytic function $z \mapsto N_\infty(z)$ which satisfies $\mathbb{P}\left(N_\infty(z) \neq 0 \text{ for all } z \in (z_-\,,z_+)\right) = 1$.*

(ii) *For every compact subset $K \subset \mathscr{D}$ and $r \in \mathbb{N}$, we can find an a.s. finite random variable $C_{K,r}$ such that for all $n \in \mathbb{N}$,*

$$\sup_{z\in K} |N_n(z) - N_\infty(z)| < C_{K,r}(\log n)^{-r}.$$

(iii) *For every compact subset $K \subset (z_-\,,z_+)$, every $0 < a < \pi$ and $r \in \mathbb{N}$,*

$$\sup_{z\in K} \left[e^{-(1+\phi(z))\log n} \int_a^\pi \left| \sum_{k=0}^{\infty} \mathbb{L}_n(k) e^{z+iu} \right| \mathrm{d}u \right] = o\big((\log n)^{-r}\big) \quad \textit{a.s. as } n \to \infty.$$

Under the results of Proposition 3.1 we can apply [29, Theorem 2.1] whose conclusions for the sequence $(k \mapsto \mathbb{L}(k))_{n \geq 1}$ are the following. For any $k \geq 0$, $n \geq 1$ and $z \in (z_-, z_+)$, we denote

$$x_n(k; z) = \frac{k - \gamma e^z \log n}{\sqrt{\gamma e^z \log n}}.$$

Then, for every integer $r \geq 0$ and every compact subset $K \subset (z_-, z_+)$, we have the convergence

$$(\log n)^{\frac{r+1}{2}} \cdot \sup_{k \in \mathbb{N}} \sup_{z \in K} \left| e^{zk - (1 + \phi(z)) \log n} \mathbb{L}_n(k) - \frac{N_\infty(z) e^{-\frac{1}{2} x_n(k;z)^2}}{\sqrt{2\pi \log n}} \sum_{j=0}^{r} \frac{G_j(x_n(k;z), z)}{(\log n)^{j/2}} \right| \xrightarrow[n \to \infty]{\text{a.s.}} 0,$$

where for all $j \geq 0$, the (random) functions $G_j(x, z)$ are polynomials of degree at most 3 in x and are entirely determined from ϕ and N_∞, with $G_0 = 1$, see [29, Equation (16)] for their complete definition. The asymptotics (1.8) and (1.9) stated in Theorem 1.5 follow from the last display. Indeed, (1.8) is obtained by letting $r = 0$ and $z = 0$ and using the fact that $N_\infty(0) = 1$ almost surely. For (1.9), we let $r = 0$, and use $k = \lfloor \gamma e^z \log n \rfloor$.

In Section 3.3, we complete the proof of Theorem 1.5 by computing the asymptotic behaviour of the height of the tree. Since the convergence of the profile already ensures that there almost surely are vertices at height $\gamma e^{(z_+ - \epsilon)} \log n$ for $\epsilon > 0$ small enough and all n large enough, it suffices to prove a corresponding upper-bound in order to finish proving the convergence (1.10) in Theorem 1.5.

3.1 Study of the Laplace transform of the weighted profile

The goal of this section is to study the sequence $\left(z \mapsto \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z \operatorname{ht}(u_i)} \right)_{n \geq 1}$. We give here a brief summary of the work presented in the section. The results to which we refer in the following paragraph are the ones that are later used in Section 3.2 to ultimately prove Proposition 3.1.

First in Section 3.1.1 we write $\sum_{i=1}^{n} \frac{w_i}{W_n} e^{z \operatorname{ht}(u_i)} = C_n(z) \cdot M_n(z)$ for some deterministic $C_n(z)$ and a martingale $M_n(z)$. We analyze the behaviour of $C_n(z)$ in Lemma 3.3. Then in Section 3.1.2 we study the asymptotic behaviour of $M_n(z)$ as $n \to \infty$ and we prove Proposition 3.6 which states that for z in open set $\mathscr{V} \subset \mathbb{C}$, we almost surely have the convergence $M_n(z) \to M_\infty(z)$ as $n \to \infty$, and this convergence is uniform in z on every compact set and happens at a polynomial speed in n. Lemma 3.5 identifies the intersection of $\mathscr{V}$ and the real line. The main ingredient for the proof of Proposition 3.6 is Lemma 3.7, which concerns the moments of $|M_n(z)|$ and $|M_{2n}(z) - M_n(z)|$, and this result is also used in the next section. Section 3.1.3 then aims at showing that $z \mapsto M_\infty(z)$ has no real zero, which is the content of Lemma 3.9.

3.1.1 Definition of $M_n(z)$ and $C_n(z)$

The following lemma is the starting point of our analysis. We use the notation $\mathcal{F}_n = \sigma(\mathrm{T}_1, \mathrm{T}_2, \ldots, \mathrm{T}_n)$.

Lemma 3.2. *For all $z \in \mathbb{C}$ and all $n \geq 1$, we have*

$$\mathbb{E}\left[\sum_{i=1}^{n+1} \frac{w_i}{W_{n+1}} e^{z \operatorname{ht}(u_i)} \,\middle|\, \mathcal{F}_n \right] = \left(1 + (e^z - 1) \frac{w_{n+1}}{W_{n+1}} \right) \cdot \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z \operatorname{ht}(u_i)}.$$

Proof. Recall that conditionally on $\mathcal{F}_n$, the $(n+1)$-st vertex u_{n+1} of T_{n+1} is a child of the

vertex $u_{K_{n+1}}$, where $\mathbb{P}\left(K_{n+1} = k \mid \mathcal{F}_n\right) = \frac{w_k}{W_n}$. We compute

$$\sum_{i=1}^{n+1} \frac{w_i}{W_{n+1}} e^{z\,\mathrm{ht}(u_i)} = \frac{W_n}{W_{n+1}} \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)} + \frac{w_{n+1}}{W_{n+1}} \cdot e^z \cdot e^{z\,\mathrm{ht}(u_{K_{n+1}})}.$$

Taking conditional expectation with respect to $\mathcal{F}_n$ yields:

$$\mathbb{E}\left[\sum_{i=1}^{n+1} \frac{w_i}{W_{n+1}} e^{z\,\mathrm{ht}(u_i)} \,\middle|\, \mathcal{F}_n\right] = \frac{W_n}{W_{n+1}} \cdot \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)} + \frac{w_{n+1}}{W_{n+1}} \cdot e^z \cdot \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}$$

$$= \left(1 + (e^z - 1)\frac{w_{n+1}}{W_{n+1}}\right) \cdot \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}.$$

This concludes the proof. $\qquad\square$

Let J be an integer that we are going to fix later on. The last result ensures that if $z \in \mathbb{C}$ is such that $\forall i \geq J,\ 1 + (e^z - 1)\frac{w_i}{W_i} \neq 0$, then we can define for all $n \geq J$

$$C_n(z) := \prod_{i=J}^{n} \left(1 + (e^z - 1)\frac{w_i}{W_i}\right) \qquad \text{and} \qquad M_n(z) := \frac{1}{C_n(z)} \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)},$$

and the sequence $(M_n(z))_{n \geq J}$ is a martingale. If J is fixed, then there exist parameters z with $\mathrm{Im}(z) = \pi \mod 2\pi$ for which the sequence $(C_n(z))_{n \geq J}$ takes the value 0. Under the assumption ($\square_\gamma$) on the sequence $\boldsymbol{w}$, we know that $\frac{w_n}{W_n} \to 0$ as $n \to \infty$. If we restrict ourselves to values of z that belong to a set of the form $\{\xi \in \mathbb{C} \mid \mathrm{Re}(\xi) < x\}$ for some $x > 0$, then

$$\left|1 + (e^z - 1)\frac{w_n}{W_n}\right| \geq 1 - |e^z - 1| \cdot \frac{w_n}{W_n} \geq 1 - (e^x + 1) \cdot \frac{w_n}{W_n} \xrightarrow[n \to \infty]{} 1 > 0,$$

hence it suffices to take J large enough in order for the sequence $(C_n(z))_{n \geq J}$ to only take non-zero values for all $z \in \{\xi \in \mathbb{C} \mid \mathrm{Re}(\xi) < x\}$ and all $n \geq J$. In what follows we work on the set

$$\mathscr{E} = \{z \in \mathbb{C} \mid \mathrm{Re}\, z < z_+\},$$

where z_+ is as defined in Proposition 3.1. For technical reasons, we also sometimes consider the larger set

$$\mathscr{E}' = \{z \in \mathbb{C} \mid \mathrm{Re}\, z < 2z_+\}.$$

Using the preceding discussion, we fix $J \geq 1$ such that the sequence $z \mapsto (C_n(z))_{n \geq J}$ does not have any zero on $\mathscr{E}'$, so that $z \mapsto (M_n(z))_{n \geq J}$ is well-defined for all $z \in \mathscr{E}'$.

We introduce the following notation. Let $F(z,n)$ and $G(z,n)$ be two functions of a complex parameter z and an integer $n \in \mathbb{N}$. For any subset of the complex plane $S \subset \mathbb{C}$ (we don't require it to be open nor connected) we write

$$F(n,z) = O_S(G(n,z)) \qquad (\text{resp.} \qquad F(n,z) = o_S(G(n,z))) \tag{3.4}$$

to express the fact that $F(n,z)$ is a big (resp. small) o of $G(n,z)$ as $n \to \infty$, uniformly on every compact $K \subset S$. Note that later in the paper, we also use this notation for random functions of z and n when such a comparison holds almost surely.

Now, we derive some information on the asymptotic behaviour of $C_n(z)$.

Lemma 3.3. *Suppose that w satisfies ($\square_\gamma$). Then there exists $\epsilon > 0$ and an analytic function $z \mapsto c(z)$ on $\mathscr{E}'$ such that*

$$C_n(z) = \exp\big(\phi(z)\log n + c(z) + O_{\mathscr{E}}\big(n^{-\epsilon}\big)\big).$$

Remark that the lemma implies that for any $z \in \mathscr{E}'$, we have

$$C_n(z) \sim e^{c(z)} \cdot n^{\phi(z)}$$

as $n \to \infty$. It is also immediate that $\mathbb{E}\left[\sum_{k=1}^n \frac{w_k}{W_n} e^{z\,\mathrm{ht}(u_k)}\right] = \mathbb{E}\left[M_J(z)\right] \cdot C_n(z)$ satisfies the same asymptotics up to a constant, as soon as z is such that $\mathbb{E}\left[M_J(z)\right] \neq 0$.

Before proving the lemma, we state the following result which follows from elementary calculus. Its proof can be found in the appendix.

Lemma 3.4. *Suppose that w satisfies ($\square_\gamma$). Then there exists $\epsilon > 0$ such that*

$$\sum_{i=n}^{+\infty} \left(\frac{w_i}{W_i}\right)^2 = O\big(n^{-\epsilon}\big) \quad \text{and also} \quad \sum_{i=1}^n \frac{w_i}{W_i} = \gamma \log n + \mathrm{cst} + O\big(n^{-\epsilon}\big).$$

Proof of Lemma 3.3. We write Log for the principal value of the complex logarithm. For $z \in \mathbb{C}$ such that $|z| < 1$ we have $\mathrm{Log}(1 + z) = \sum_{i=1}^{\infty} \frac{(-1)^{n-1}}{n} z^n$. If for every $i \geq J$ and $z \in \mathscr{E}'$, we let

$$h(i, z) = \mathrm{Log}\left(1 + (e^z - 1)\frac{w_i}{W_i}\right) - (e^z - 1)\frac{w_i}{W_i},$$

which is well-defined thanks to our choice of J, then $|h(i, z)| = O_{\mathscr{E}'}\left(\left(\frac{w_i}{W_i}\right)^2\right)$ is summable in i and the rest of the series is

$$\left|\sum_{i=n}^{\infty} h(i, z)\right| \leq \sum_{i=n}^{\infty} |h(i, z)| = O_{\mathscr{E}'}\left(\sum_{i=n}^{\infty} \left(\frac{w_i}{W_i}\right)^2\right) = O_{\mathscr{E}'}\big(n^{-\epsilon}\big), \tag{3.5}$$

for some $\epsilon > 0$, thanks to Lemma 3.4. Then we write

$$C_n(z) = \prod_{i=J}^n \left(1 + (e^z - 1)\frac{w_i}{W_i}\right) = \exp\left((e^z - 1)\sum_{i=J}^n \frac{w_i}{W_i} + \sum_{i=J}^n h(i, z)\right)$$

which yields using (3.5) and Lemma 3.4

$$C_n(z) = \exp\left((e^z - 1)(\gamma \log n + \mathrm{cst} + O_{\mathscr{E}'}\big(n^{-\epsilon}\big)) + \sum_{i=J}^{\infty} h(i, z) - \sum_{i=n+1}^{\infty} h(i, z)\right)$$

$$= \exp\left(\phi(z)\log n + \underbrace{(e^z - 1) \cdot \mathrm{cst} + \sum_{i=J}^{\infty} h(i, z)}_{c(z)} + O_{\mathscr{E}'}\big(n^{-\epsilon}\big)\right),$$

and $c(z)$ is an analytic function of z, which finishes the proof. $\square$

3.1.2 Asymptotic properties of the martingale $(M_n(z))_{n \geq 1}$.

Note that when the parameter z is a positive real number, the sequence $(M_n(z))_{n \geq 1}$ is a positive martingale and so it converges almost surely to some limit. In fact, we
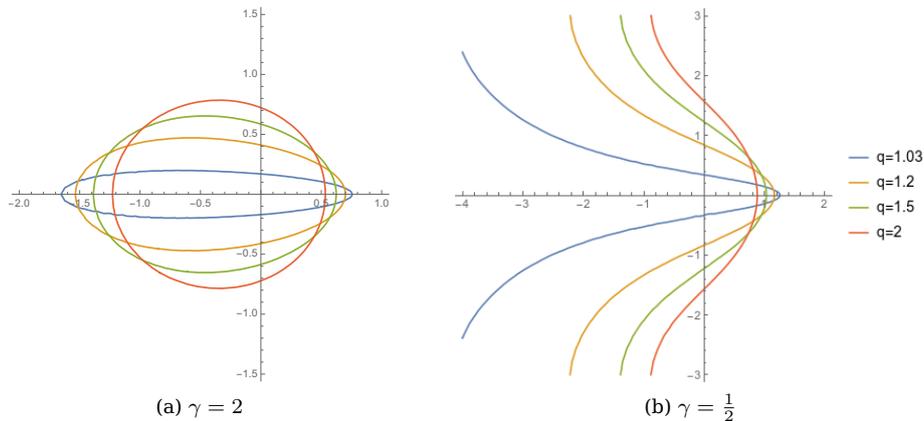
(a) $\gamma = 2$ (b) $\gamma = \frac{1}{2}$

Figure 1: The boundary of the connected component of $\mathscr{V}_q$ that contains 0, for some values of $q \in (1, 2]$, plotted for $\gamma = 2$ and $\gamma = \frac{1}{2}$.

want to prove that these martingales converge almost surely (and also in L^1) for the largest possible range of parameters z. **For the rest of Section 3.1 and also in the subsequent Section 3.2, we assume that the weight sequence w satisfies** $(\square_\gamma^p)$ **for some fixed parameters $\gamma > 0$ and $p \in (1, 2]$.**

We align our notation with the one used in [10, Theorem 2.2] which states something similar to our forthcoming Proposition 3.6 for another model, the binary search tree. For any $z \in \mathscr{E}$ and $q \in (1, p]$, we let

$$g(z, q) := \phi(q \operatorname{Re} z) - q \operatorname{Re}(\phi(z)) - q + 1 = \gamma(e^{q \operatorname{Re} z} - 1 - q \operatorname{Re}(e^z) + q) - q + 1.$$

For any $q \in (1, p]$, let $\mathscr{V}_q = \{z \in \mathscr{E} \mid g(z, q) < 0\}$, and denote

$$\mathscr{V} = \bigcup_{1 < q \leq p} \mathscr{V}_q. \tag{3.6}$$

As we prove below in Proposition 3.6, this set $\mathscr{V}$ is our set of interest, on which we can prove that the martingales $(M_n(z))_{n \geq J}$ converge nicely. Before stating anything about convergence of martingales, we state and prove an easy lemma about the intersection of $\mathscr{V}$ and the real line.

Lemma 3.5. *The set $\mathscr{V}$ is open and contains the open interval of real numbers $I_\gamma :=$ $\{x \in \mathbb{R} \mid \gamma(xe^x - e^x + 1) - 1 < 0\}$ which contains 0.*

Proof. Of course $\mathscr{V}$ is open as a union of open sets. For any real x we have $g(x, 1) = 0$. So, if $\frac{\partial g}{\partial q}(x, 1) < 0$ then there exists $q > 1$ for which $g(x, q) < 0$. Since $\frac{\partial g}{\partial q}(x, 1) = \gamma(xe^x - e^x + 1) - 1$, the set $\mathscr{V}$ contains the interval I_γ in the statement of the lemma. Since $\frac{\partial g}{\partial q}(0, 1) = -1 < 0$, we have $0 \in I_\gamma$. $\qquad \square$

Note that the interval I_γ contains the interval (z_-, z_+) defined in (1.7) and this will be particularly important later on for the proof of Proposition 3.1. We now turn to the next proposition, which justifies our definition of $\mathscr{V}$.

Proposition 3.6. *The sequence of functions $(z \mapsto M_n(z))_{n \geq J}$ converges uniformly almost surely and in L^1 towards an analytic function $z \mapsto M_\infty(z)$ on every compact subset of $\mathscr{V}$. Furthermore, for any compact subset $K \subset \mathscr{V}$, there exists a real $\epsilon(K) > 0$ such that almost surely*

$$|M_n(z) - M_\infty(z)| = O_K\left(n^{-\epsilon(K)}\right).$$

The proof of the proposition follows from the next lemma, together with Lemma A.3, stated in the appendix.

Lemma 3.7. *For any $q \in (1, p]$ and $z \in \mathscr{E}$ we have*

$$\mathbb{E}\left[|M_n(z)|^q\right] = O_{\mathscr{E}}\left(n^{0 \vee g(z,q) + o_{\mathscr{E}}(1)}\right). \tag{3.7}$$

and also

$$\mathbb{E}\left[|M_{2n}(z) - M_n(z)|^q\right] = O_{\mathscr{E}}\left(n^{(1-q) \vee g(z,q) + o_{\mathscr{E}}(1)}\right). \tag{3.8}$$

We first prove Proposition 3.6, assuming the lemma and then proceed to prove the lemma.

Proof of Proposition 3.6. Let $K \subset \mathscr{V}$ be a compact set. Without loss of generality, we can assume that $K \subset \mathscr{V}_q$ for some $q \in (1, p]$. Indeed, if that is not the case, by compactness we can write $K \subset \bigcup_{i=1}^{k} \mathscr{V}_{q_i}$, with k finite and $q_1, q_2, \dots q_k \in (1, p]$, and separately consider the compact sets

$$K_j = K \setminus \bigcup_{\substack{i=1 \\ i \neq j}}^{k} \mathscr{V}_{q_i}$$

which are such that $K_j \subset \mathscr{V}_{q_j}$, for $j = 1, 2, \dots k$.

The convergence result is then an application of Lemma A.3, on the set $\mathscr{V}_q$ with $\alpha(z) = 0$ and, say $\delta(z) = -\frac{1}{2}g(z, q) > 0$. The limiting function is analytic as a uniform limit of analytic functions. $\square$

Proof of Lemma 3.7. For any $q \in (1, p]$ and $n \geq J$, we write

$$M_{n+1}(z) - M_n(z) = M_n(z)\left(\frac{C_n(z)}{C_{n+1}(z)} - 1\right) + \frac{1}{C_{n+1}(z)} \cdot \frac{w_{n+1}}{W_{n+1}} \cdot e^{z \operatorname{ht}(u_{n+1})}.$$

Taking the q-th power of the modulus on both sides and using the inequality $|a + b|^q \leq 2^q \cdot (|a|^q + |b|^q)$, we get

$$\mathbb{E}\left[|M_{n+1}(z) - M_n(z)|^q\right]$$
$$\leq \mathbb{E}\left[|M_n(z)|^q\right] \cdot 2^q \left|\frac{C_n(z)}{C_{n+1}(z)} - 1\right|^q + 2^q \frac{1}{|C_{n+1}(z)|^q}\left(\frac{w_{n+1}}{W_{n+1}}\right)^q \cdot \mathbb{E}\left[|e^z|^{q \operatorname{ht}(u_{n+1})}\right]. \tag{3.9}$$

Using Lemma A.2 in the appendix, we have for any $n \geq J$,

$$\mathbb{E}\left[|M_{n+1}(z)|^q\right] \leq \mathbb{E}\left[|M_n(z)|^q\right] + 2^q \cdot \mathbb{E}\left[|M_{n+1}(z) - M_n(z)|^q\right].$$

Using the last display and equation (3.9), we get a recurrence inequality of the form

$$\mathbb{E}\left[|M_{n+1}(z)|^q\right] \leq (1 + a_n(z)) \cdot \mathbb{E}\left[|M_n(z)|^q\right] + b_n(z), \tag{3.10}$$

where

$$a_n(z) = 2^{2q}\left|\frac{C_n(z)}{C_{n+1}(z)} - 1\right|^q \quad \text{and} \quad b_n(z) = 2^{2q}\frac{1}{|C_{n+1}(z)|^q}\left(\frac{w_{n+1}}{W_{n+1}}\right)^q \cdot \mathbb{E}\left[|e^z|^{q \operatorname{ht}(u_{n+1})}\right].$$

Applying (3.10) in cascade we get

$$\mathbb{E}\left[|M_n(z)|^q\right] \leq \prod_{i=J}^{n-1}(1 + a_i(z)) \cdot \left(\mathbb{E}\left[|M_J(z)|^q\right] + \sum_{i=J}^{n-1} b_i(z)\right). \tag{3.11}$$

Now notice that from our assumption on the sequence $(w_n)_{n \geq 1}$ we have

$$a_n(z) = 2^{2q} \left| \frac{C_n(z)}{C_{n+1}(z)} - 1 \right|^q = 2^{2q} \left| \frac{1}{1 + (e^z - 1)\frac{w_{n+1}}{W_{n+1}}} - 1 \right|^q = O_{\mathscr{E}}\left(\left(\frac{w_{n+1}}{W_{n+1}} \right)^q \right). \quad (3.12)$$

On the other hand, since $z \in \mathscr{E}$ then $q \operatorname{Re} z \in \mathscr{E}'$, so we can use Lemma 3.3 to get

$$
\begin{aligned}
b_n(z) &= \text{cst} \cdot \left(\frac{w_{n+1}}{W_{n+1}} \right)^q \cdot |C_{n+1}(z)|^{-q} \cdot e^{q \operatorname{Re} z} \cdot \mathbb{E}\left[\sum_{k=1}^n \frac{w_k}{W_n} e^{(q \operatorname{Re} z)\operatorname{ht}(u_k)} \right] \\
&= \left(\frac{w_{n+1}}{W_{n+1}} \right)^q \cdot O_{\mathscr{E}}\left(n^{-q \operatorname{Re}(\phi(z))} \right) \cdot O_{\mathscr{E}}\left(n^{\phi(q \operatorname{Re} z)} \right) \\
&= \left(\frac{w_{n+1}}{W_{n+1}} \right)^q \cdot O_{\mathscr{E}}\left(n^{g(z,q)-1+q} \right). \quad (3.13)
\end{aligned}
$$

We conclude using the following lemma which is an application of Hölder's inequality using the assumption $(\square_\gamma^p)$.

Lemma 3.8. *For any $q \in (1, p]$ we have $\displaystyle\sum_{i=n}^{2n} \left(\frac{w_i}{W_i} \right)^q \leq n^{1-q+o(1)}$.*

Together with (3.12), this proves that $(a_n(z))_{n \geq 1}$ is summable and so $\prod_{i=J}^\infty (1 + a_i(z)) = O_{\mathscr{E}}(1)$. Also

$$\sum_{i=n}^{2n} b_i(z) = O_{\mathscr{E}}\left(n^{g(z,q)+o_{\mathscr{E}}(1)} \right),$$

and so $\sum_{i=J}^n b_i(z) = O_{\mathscr{E}}\left(n^{0 \vee g(z,q)+o_{\mathscr{E}}(1)} \right)$. Replacing this in (3.11) finishes to prove (3.7). In order to prove (3.8), we use Lemma A.2 again and write

$$
\begin{aligned}
\mathbb{E}\left[|M_{2n}(z) - M_n(z)|^q \right] &\leq 2^q \cdot \sum_{i=n}^{2n-1} \mathbb{E}\left[|M_{i+1}(z) - M_i(z)|^q \right] \\
&\underset{(3.7),\,(3.9)}{\leq} \sum_{i=n}^{2n-1} \left(a_i(z) \cdot O_{\mathscr{E}}\left(n^{0 \vee g(z,q)+o_{\mathscr{E}}(1)} \right) + b_i(z) \right) \\
&\underset{(3.12),\,(3.13)}{\leq} \sum_{i=n}^{2n-1} \left(\frac{w_{i+1}}{W_{i+1}} \right)^q \left(O_{\mathscr{E}}\left(n^{0 \vee g(z,q)+o_{\mathscr{E}}(1)} \right) + O_{\mathscr{E}}\left(n^{g(z,q)-1+q} \right) \right).
\end{aligned}
$$

Using Lemma 3.8 we get $\mathbb{E}\left[|M_{2n}(z) - M_n(z)|^q \right] = O_{\mathscr{E}}\left(n^{(1-q) \vee g(z,q)+o_{\mathscr{E}}(1)} \right)$ which finishes the proof of the lemma. $\qquad\square$

3.1.3 The limit $M_\infty(z)$ does not have real zeros

Now that we have proved that their almost surely exists a limiting function $z \mapsto M_\infty(z)$ defined on the set $\mathscr{V}$, we are interested in the possible location of the zeros of this random function. In fact, the function $z \mapsto M_\infty(z)$ is related to the function $z \mapsto N_\infty(z)$ of Proposition 3.1, for which we aim to prove that it has almost surely no zero on some real interval (z_-, z_+) which contains 0. We state a similar result for $z \mapsto M_\infty(z)$. Recall the definition of the interval I_γ in Lemma 3.5.

Lemma 3.9. *The function $M_\infty(z)$ has almost surely no zero on I_γ.*

The rest of Section 3.1.3 is devoted to proving Lemma 3.9. We start by stating and proving the following weaker statement.

Lemma 3.10. *For all $z \in I_\gamma$, we have almost surely $M_\infty(z) > 0$. As a consequence, the number of zeros of the map $(z \mapsto M_\infty(z))$ on the interval I_γ is almost surely at most countable.*

Recall from (1.1) the definition of the sequence of independent random variables $(K_2, K_3, \dots)$ that is used in the construction of the trees $(\mathsf{T}_n)_{n \geq 1}$. The main argument in the following proof consists in checking that an event of interest is measurable with respect to the tail σ-algebra for the sequence $(K_2, K_3, \dots)$ and applying Kolmogorov's $0 - 1$ law.

Proof of Lemma 3.10. Fix $N \geq J$ and $z \in I_\gamma$ and for all $n \geq N$, let

$$M_n^{(N)}(z) = \frac{1}{C_n(z)} \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\, \mathrm{d}(u_i, \mathsf{T}_N)},$$

where d denotes the graph distance in $\mathbb{U}$, and the distance between a vertex and a subset of vertices is defined the obvious way. The idea behind the introduction of $(M_n^{(N)}(z))_{n \geq N}$ is that, up to a positive multiplicative constant (i.e. a deterministic constant that depends on N and z but not on n), it has the same distribution as the sequence $(M_n(z))_{n \geq J}$ associated to the growth of the tree that one informally obtains by contracting all the vertices $\{u_1, u_2, \dots, u_N\}$ into one. The growth of such a tree can be described as that of a weighted recursive tree with weight sequence $(W_N, w_{N+1}, w_{N+2}, \dots)$, which shares the same asymptotic property ($\square_\gamma^p$) as the original sequence $(w_n)_{n \geq 1}$.

We claim the following:

(i) Due to the above remarks, $(M_n^{(N)}(z))_{n \geq N}$ is a positive martingale that satisfies the same assumptions as $(M_n(z))_{n \geq N}$ so it converges a.s. and in L^1 towards a non-negative limit, $M_\infty^{(N)}(z)$, thanks to Proposition 3.6, which we can apply here because $z \in I_\gamma \subset \mathscr{V}$.

(ii) We have $(1 \wedge e^z)^N M_n^{(N)}(z) \leq M_n(z) \leq (1 \vee e^z)^N M_n^{(N)}(z)$.

(iii) The sequence $(M_n^{(N)}(z))_{n \geq N}$, hence also its limit $M_\infty^{(N)}(z)$, is independent of the N first steps of the construction, and is hence a measurable function of the sequence $(K_n)_{n \geq N+1}$.

Using all these observations we deduce that for any $N \geq J$ we have the equality of events $\{M_\infty(z) > 0\} = \{M_\infty^{(N)}(z) > 0\}$. This proves that $\{M_\infty(z) > 0\}$ is measurable with respect to the tail σ-algebra generated by the sequence $(K_n)_{n \geq 2}$, which is a sequence of jointly independent random variables. Kolmogorov's 0-1 law then ensures that this event has probability 0 or 1. By L^1 convergence we have $\mathbb{E}\left[M_\infty(z)\right] = \mathbb{E}\left[M_J(z)\right] > 0$ and this proves that $\mathbb{P}\left(M_\infty(z) > 0\right) = 1$. It follows immediately that the limit $z \mapsto M_\infty(z)$ can only have finitely many zeros in any compact subset of I_γ almost surely, because otherwise, by analyticity of $z \mapsto M_\infty(z)$ on the connected component of $\mathscr{V}$ that contains I_γ, the function would be identically 0 on I_γ with positive probability. This ensures that the total number of zeros in I_γ is at most countable and finishes the proof. $\square$

Now, in order to prove Lemma 3.9, we use an argument of self-similarity: essentially, if we take two vertices u_i and u_j in the tree, then conditionally on the sequences of vertices that are respectively added above u_i or above u_j, the subtrees above u_i and u_j evolve as two *independent* weighted recursive trees. Using Proposition 3.6 and Lemma 3.10, the normalized Laplace transform of the weighted profile of each of those two subtrees should almost surely converge to some random analytic function on $\mathscr{V}$, which is non-negative on I_γ and has at most countably many zeros on this interval. Since

the two are independent, their zeros should not overlap and hence the sum of their contribution should result in a function that is positive on I_γ.

Proof of Lemma 3.9. Let us formalise this line of reasoning. First, recall from Theorem 1.7 the definition of the measure μ on $\partial \mathbb{U}$ and the fact that under our assumption ($\square_\gamma$) for the weight sequence $\boldsymbol{w}$, it is almost surely diffuse. This allows us to define

$$I^{(1)} := \inf \{ i \geq 1 \mid \mu(T(u_i)) \in (0,1) \}$$

and

$$I^{(2)} := \inf \{ i \geq I_1 \mid u_i \notin T(u_{I_1}) \text{ and } \mu(T(u_i)) \in (0,1) \} ,$$

and those two random variables are almost surely finite.

Consider the sequences $\left(\mathbf{1}_{\left\{ u_n \in T(u_{I^{(j)}}) \right\}} \right)_{n \geq 1}$ for $j \in \{1,2\}$, which record the times when a vertex is added to $T(u_{I^{(1)}})$ or $T(u_{I^{(2)}})$. We work conditionally on those two sequences for the rest of the proof. For $j \in \{1,2\}$, we define

$$\forall n \geq 1, \quad N_n^{(j)} := \sum_{i=1}^{n} \mathbf{1}_{\left\{ u_i \in T(u_{I^{(j)}}) \right\}} \qquad \text{and} \qquad \forall k \geq 1, \quad \tau_k^{(j)} := \inf \left\{ n \geq 1 \;\middle|\; N_n^{(j)} \geq k \right\},$$

which record respectively the number of vertices among $\{u_1, u_2, \dots, u_n\}$ that are in $T(u_{I^{(j)}})$ and conversely, the k-th time where a vertex is added to $T(u_{I^{(j)}})$ in the construction of $(\mathtt{T}_n)_{n \geq 1}$. We let $w_k^{(j)} := w_{\tau^{(j)}(k)}$ and $W_k^{(j)} := \sum_{i=1}^{k} w_i^{(j)}$, and also $u_k^{(j)} := u_{\tau_k^{(j)}}$. We also define for $k \geq 1$

$$\mathtt{T}_k^{(j)} := \left\{ u \in \mathbb{U} \;\middle|\; u_{I^{(j)}} u \in \mathtt{T}_{\tau_k^{(j)}} \right\},$$

the subtree hanging above $u_{I^{(j)}}$ at the time where it contains exactly k vertices (translated to the origin in so as to be considered as a plane tree).

Let us state the following intermediate result, which we prove at the end of the section. Note that the random sequences $(N_n^{(j)})_{n \geq 1}$, $(\tau_k^{(j)})_{k \geq 1}$ and $(w_k^{(j)})_{k \geq 1}$ for $j \in \{1,2\}$ can be read from the two sequences $\left(\mathbf{1}_{\left\{ u_n \in T(u_{I^{(j)}}) \right\}} \right)_{n \geq 1}$ for $j \in \{1,2\}$.

Lemma 3.11. *The following holds.*

(i) *For $j \in \{1,2\}$, we almost surely have $N_n^{(j)} \underset{n \to \infty}{\sim} \mu(T(u_{I^{(j)}})) \cdot n$.*

(ii) *For $j \in \{1,2\}$, the sequence $(w_k^{(j)})_{k \geq 1}$ satisfies ($\square_\gamma^p$) almost surely.*

(iii) *Conditionally on the two sequences $\left(\mathbf{1}_{\left\{ u_n \in T(u_{I^{(1)}}) \right\}} \right)_{n \geq 1}$ and $\left(\mathbf{1}_{\left\{ u_n \in T(u_{I^{(2)}}) \right\}} \right)_{n \geq 1}$, the sequences of trees $(\mathtt{T}_k^{(1)})_{k \geq 1}$ and $(\mathtt{T}_k^{(2)})_{k \geq 1}$ are independent and have respective distributions $\mathrm{WRT}((w_k^{(1)})_{k \geq 1})$ and $\mathrm{WRT}((w_k^{(2)})_{k \geq 1})$.*

Recall the discussion before Lemma 3.3. For $j \in \{1,2\}$, let $J^{(j)} \geq 1$ be the smallest integer such that for all $k \geq J^{(j)}$ and for all $z \in \mathscr{E}'$ we have $1 + (e^z - 1)\frac{w_k^{(j)}}{W_k^{(j)}} \neq 0$. Then we can define for $k \geq J^{(j)}$,

$$M_k^{(j)}(z) := \frac{1}{C_k^{(j)}(z)} \sum_{i=1}^{k} \frac{w_i^{(j)}}{W_k^{(j)}} e^{z \, \mathrm{d}(u_{I^{(j)}}, u_i^{(j)})} \quad \text{with} \quad C_k^{(j)}(z) := \prod_{i=J^{(j)}}^{k} \left(1 + (e^z - 1)\frac{w_i^{(j)}}{W_i^{(j)}} \right).$$

These processes are the martingales associated to the weighted profile of the trees $(\mathtt{T}_k^{(j)})_{k \geq 1}$ for $j \in \{1,2\}$. Thanks to Lemma 3.11(iii) those trees have respective distribution

WRT($(w_k^{(j)})_{k\geq 1}$), for $j \in \{1, 2\}$ and thanks to Lemma 3.11(ii), those weight sequences satisfy ($\square_\gamma^p$) almost surely. This allows us to apply Proposition 3.6, which entails that for $j \in \{1, 2\}$, the sequence of functions $(z \mapsto M_k^{(j)}(z))_{k\geq J^{(j)}}$ converges almost surely to an analytic limit $z \mapsto M_\infty^{(j)}(z)$ on the set $\mathscr{V}$. Now we can write, for n sufficiently large

$$
\begin{aligned}
M_n(z) &= \frac{1}{C_n(z)} \sum_{i=1}^n \frac{w_i}{W_n} e^{z \, \mathrm{ht}(u_i)} \\
&\geq \frac{C_{N_n^{(1)}}^{(1)}(z) \cdot W_{N_n^{(1)}}^{(1)}}{C_n(z) \cdot W_n} \cdot e^{z \, \mathrm{ht}(u_{I^{(1)}})} \cdot M_{N_n^{(1)}}^{(1)}(z) + \frac{C_{N_n^{(2)}}^{(2)}(z) \cdot W_{N_n^{(2)}}^{(2)}}{C_n(z) \cdot W_n} \cdot e^{z \, \mathrm{ht}(u_{I^{(2)}})} \cdot M_{N_n^{(2)}}^{(2)}(z).
\end{aligned}
\tag{3.14}
$$

Using Lemma 3.3, we have almost surely for $j \in \{1, 2\}$,

$$
C_k^{(j)}(z) = \exp\Big(\phi(z) \log k + c^{(j)}(z) + o_{\mathscr{E}}(1) \Big) \quad \text{as} \quad k \to \infty.
$$

Using the asymptotics $N_n^{(j)} \underset{n\to\infty}{=} \mu(T(u_{I^{(j)}})) \cdot n \cdot (1 + o(1))$ from Lemma 3.11(i) we get

$$
C_{N_n^{(j)}}^{(j)}(z) = \exp\Big(\phi(z) \log n + \phi(z) \log(\mu(T(u_{I^{(j)}}))) + c^{(j)}(z) + o_{\mathscr{E}}(1) \Big).
$$

From the a.s. convergence of the sequence of measures $(\mu_n)_{n\geq 1}$, see Theorem 1.7, we also get

$$
\frac{W_{N_n^{(j)}}^{(j)}}{W_n} = \mu_n(T(u_{I^{(j)}})) \underset{n\to\infty}{\to} \mu(T(u_{I^{(j)}})) > 0,
$$

which entails that for $j \in \{1, 2\}$, uniformly on all compact subsets of $\mathscr{E}$, we have the a.s. convergence

$$
\frac{W_{N_n^{(j)}}^{(j)}}{W_n} \cdot \frac{C_{N_n^{(j)}}^{(j)}(z)}{C_n(z)} \underset{n\to\infty}{\to} A_j(z) := \mu(T(u_{I^{(j)}})) \cdot \exp\Big(\phi(z) \log(\mu(T(u_{I^{(j)}}))) + c^{(j)}(z) - c(z) \Big),
$$

where the limiting function $z \mapsto A_j(z)$ is analytic and only takes positive values on $\mathscr{E} \cap \mathbb{R}$. Then, for any $z \in \mathscr{V} \cap \mathbb{R}$, taking the limit $n \to \infty$ in (3.14) yields

$$
M_\infty(z) \geq e^{z \, \mathrm{ht}(u_{I^{(1)}})} \cdot A_1(z) \cdot M_\infty^{(1)}(z) + e^{z \, \mathrm{ht}(u_{I^{(2)}})} \cdot A_2(z) \cdot M_\infty^{(2)}(z).
$$

Now, thanks to Lemma 3.10, the function $z \mapsto M_\infty^{(1)}(z)$ can only have at most countably many zeros on $I_\gamma \subset \mathscr{V} \cap \mathbb{R}$ and for any $z \in I_\gamma$, we have $M_\infty^{(2)}(z) > 0$ almost surely. Then, if we condition on the location of the zeros $z_1, z_2 \ldots$ of $M_\infty^{(1)}$ on I_γ, since $M_\infty^{(2)}$ is independent of $z_1, z_2 \ldots$, we have $M_\infty^{(2)}(z_i) > 0$ for all $i \geq 1$ almost surely. Hence M_∞ has almost surely no zeros on I_γ. $\qquad\square$

We now prove Lemma 3.11, which we used in the preceding proof.

Proof of Lemma 3.11. Point (i) follows just from Theorem 1.7 and Proposition 2.4 and the fact that for $j \in \{1, 2\}$ we have $N_n^{(j)} = n\nu_n(T(u_{I^{(j)}}))$.

We now prove (ii). In order to do that, we first prove that for $j \in \{1, 2\}$, we have

$$
\mu_n(T(u_{I^{(j)}})) \underset{n\to\infty}{\bowtie} \mu(T(u_{I^{(j)}})) \qquad \text{and} \qquad \tau_n^{(j)} \underset{n\to\infty}{\bowtie} \mu(T(u_{I^{(j)}}))^{-1} \cdot n.
\tag{3.15}
$$

We conclude from here: using the fact that w satisfies ($\square_\gamma$), we get

$$
W_n^{(j)} = W_{\tau_n^{(j)}} \cdot \mu_{\tau_n^{(j)}}(T(u_{I^{(j)}})) \underset{n\to\infty}{\bowtie} \mathrm{cst} \cdot (\tau_n^{(j)})^\gamma \cdot \mu(T(u_{I^{(j)}})) \underset{n\to\infty}{\bowtie} \mathrm{cst} \cdot n^\gamma,
$$

with a positive constant. We also have

$$\sum_{k=n}^{2n}(w_k^{(j)})^p = \sum_{k=n}^{2n}(w_{\tau_k^{(j)}})^p \leq \sum_{i=\tau_n^{(j)}}^{\tau_{2n}^{(j)}} w_i^p.$$

Because of (3.15), we have the following almost sure convergence $\frac{\tau_{2n}^{(j)}}{\tau_n^{(j)}} \to 2$ as $n \to \infty$, hence almost surely for n large enough we have $\tau_{2n}^{(j)} \leq 4\tau_n^{(j)}$, so

$$\sum_{i=\tau_n^{(j)}}^{\tau_{2n}^{(j)}} w_i^p \leq \sum_{i=\tau_n^{(j)}}^{4\tau_n^{(j)}} w_i^p \leq \sum_{i=\tau_n^{(j)}}^{2\tau_n^{(j)}} w_i^p + \sum_{i=2\tau_n^{(j)}}^{4\tau_n^{(j)}} w_i^p \leq (\tau_n^{(j)})^{1+(\gamma-1)p+o(1)} + (2\tau_n^{(j)})^{1+(\gamma-1)p+o(1)}$$

$$\leq n^{1+(\gamma-1)p+o(1)},$$

where in the two last inequalities we used the fact that w satisfies ($\square_\gamma^p$) and the almost sure linear growth of $\tau_n^{(j)}$ ensured by (3.15).

So it remains only to prove (3.15). Recall the proof of Theorem 1.7. For all $k \geq 1$ the process $(\mu_n(T(u_k)))_{n \geq k}$ is a martingale and almost surely we have

$$|\mu_{n+1}(T(u_k)) - \mu_n(T(u_k))| = \frac{w_{n+1}}{W_{n+1}} \cdot \big|\mathbf{1}_{\{u_{n+1} \in T(u_k)\}} - \mu_n(T(u_k))\big| \leq \frac{w_{n+1}}{W_{n+1}}.$$

Using successively Lemma A.2 and then Lemma 3.8, where the latter applies because w satisfies ($\square_\gamma^p$), we get

$$\mathbb{E}\left[|\mu_{2n}(T(u_k)) - \mu_n(T(u_k))|^p\right] \leq 2^p \cdot \sum_{i=n+1}^{2n}\left(\frac{w_i}{W_i}\right)^p = O\left(n^{1-p+o(1)}\right).$$

Using then Lemma A.3 with $q = p$ and $\alpha = 0$ and $\delta = (p-1)/2$, we get that, for some $\epsilon > 0$, $|\mu_n(T(u_k)) - \mu(T(u_k))| = O(n^{-\epsilon})$ almost surely. Since this is true almost surely for all $k \geq 1$, we use it with $k \in \{I^{(1)}, I^{(2)}\}$. As by definition for $j \in \{1,2\}$ we have $\mu(T(u_{I^{(j)}})) > 0$, we conclude that $\mu_n(T(u_{I^{(j)}})) \underset{n \to \infty}{\bowtie} \mu(T(u_{I^{(j)}}))$.

Then, for any $k \geq 1$, consider the process $\left(n\nu_n(T(u_k)) - \sum_{i=k+1}^n \mu_i(T(u_k))\right)_{n \geq k}$. It is easy to verify that this process is a martingale in its own filtration and that its increments are bounded by 1. Using again Lemma A.3 with $q = 2$ and $\alpha = 1$ and $\delta = 1$, we get that $n^{-1}\big|n\nu_n(T(u_k)) - \sum_{i=k+1}^n \mu_i(T(u_k))\big| = O(n^{-\epsilon})$ for some $\epsilon > 0$. Using again that for $j \in \{1,2\}$ the value $\mu(T(u_{I^{(j)}}))$ is almost surely positive, we can write $N_n^{(j)} = n\nu_n(T(u_{I^{(j)}})) \underset{n \to \infty}{\bowtie} \mu(T(u_{I^{(j)}})) \cdot n$. Using the definition of $\tau_n^{(j)}$, we can check that this entails that $\tau_n^{(j)} \underset{n \to \infty}{\bowtie} \mu(T(u_{I^{(j)}}))^{-1} \cdot n$ almost surely. This concludes the proof of (3.15) and so, (ii) is proved.

We now prove (iii). For any $k \geq 1$, we consider the sequence $\left(\mathbf{1}_{\{u_n \in T(u_k)\}}\right)_{n \geq 1}$ that encodes the labels of the vertices above u_k. Note that the limiting mass $\mu(T(u_k))$ can be computed from that sequence. Now, we sequentially reveal the sequences $\left(\mathbf{1}_{\{u_n \in T(u_1)\}}\right)_{n \geq 1}, \left(\mathbf{1}_{\{u_n \in T(u_2)\}}\right)_{n \geq 1}, \ldots, \left(\mathbf{1}_{\{u_n \in T(u_k)\}}\right)_{n \geq 1}, \ldots$ until we get to a k for which $\mu(T(u_k)) \in (0, 1)$. By definition, the first index k for which it happens is $k = I^{(1)}$.

Then we continue revealing the sequences $\left(\mathbf{1}_{\{u_n \in T(u_k)\}}\right)_{n \geq 1}$ for $k > I^{(1)}$ but only for the k's such that $u_k \notin T(u_{I^{(1)}})$ until we get to a k for which $\mu(T(u_k)) \in (0, 1)$. By definition, this second index k is $k = I^{(2)}$. Remark, and this is the key in this argument, that after determining $I^{(1)}$ and $I^{(2)}$ in this way, the only information that we have about $T(u_{I^{(1)}})$ and $T(u_{I^{(2)}})$ is the list of labels of the vertices that belong each of them (and the position of $u_{I^{(1)}}$ and $u_{I^{(2)}}$).

Now, conditionally on all this information, it is straightforward to see from the attachment dynamics that for any $j \in \{1, 2\}$, when the $(i + 1)$-st vertex attaches above $u_{I(j)}$ at time $\tau_{i+1}^{(j)}$, the label $K_{\tau_{i+1}^{(j)}}$ of the vertex to which it attaches is chosen among $\tau_1^{(j)}, \tau_2^{(j)}, \ldots \tau_i^{(j)}$ with probability proportional to their respective weight $w_{\tau_1^{(j)}}, w_{\tau_2^{(j)}}, \ldots w_{\tau_i^{(j)}}$, independently for different choices of $i \geq 1$ and $j \in \{1, 2\}$. This ensures that those trees evolve as independent weighted recursive trees and finishes the proof of (iii) and hence that of the lemma. $\qquad\square$

3.2 From the weighted to the unweighted sum.

Now we want to transfer these results of convergence to the Laplace transform of the "true" profile. Recall from (3.3) the definition of the sequence of functions $(z \mapsto N_n(z))_{n \geq 1}$. We still assume until the end of Section 3.2 that w satisfies $(\square_\gamma^p)$ for some $\gamma > 0$ and $p \in (1, 2]$.

We introduce the following quantity, for $n \geq J$,

$$
\begin{aligned}
X_n(z) &:= n^{1+\phi(z)} \cdot N_n(z) - e^z \sum_{k=J}^{n-1} C_k(z) M_k(z) \\
&= \sum_{i=1}^{n} e^{z \operatorname{ht}(u_i)} - e^z \sum_{k=J}^{n-1} \left(\sum_{i=1}^{k} \frac{w_i}{W_k} e^{z \operatorname{ht}(u_i)} \right)
\end{aligned}
$$

The goal of this subsection is to show that the quantity $X_n(z)$ is negligible as $n \to \infty$ compared to any of the two terms in the difference, for z contained in some subset of the complex plane. This way we transfer the asymptotics that we have proved for $M_n(z)$ and $C_n(z)$ in the last section to asymptotics for $N_n(z)$, which is the quantity that we want to study in the end. We start by stating and proving a lemma.

Lemma 3.12. *The process* $(X_n(z))_{n \geq J}$ *is a martingale with respect to the filtration* $(\mathcal{F}_n)$. *Furthermore, for all* $q \in (1, p]$,

$$
\mathbb{E}\left[|X_{2n}(z) - X_n(z)|^q\right] = O_{\mathscr{E}}\left(n^{1+(q\operatorname{Re}(\phi(z)) \vee \phi(q\operatorname{Re} z))+o_{\mathscr{E}}(1)}\right).
$$

Proof. This process is of course $(\mathcal{F}_n)$-adapted and integrable. For the martingale property we compute

$$
\begin{aligned}
\mathbb{E}\left[X_{n+1}(z) \mid \mathcal{F}_n\right] &= \mathbb{E}\left[X_n(z) - e^z C_n(z) M_n(z) + e^{z \operatorname{ht}(u_{n+1})} \;\Big|\; \mathcal{F}_n\right] \\
&= X_n(z) - e^z C_n(z) M_n(z) + e^z \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z \operatorname{ht}(u_i)} = X_n(z).
\end{aligned}
$$

For $z \in \mathscr{E}$ and $q \in (1, p]$, we make the following computation, using Lemma 3.3 and Lemma 3.7,

$$
\begin{aligned}
\mathbb{E}\left[|X_{n+1}(z) - X_n(z)|^q\right] &= \mathbb{E}\left[\left|-e^z C_n(z) M_n(z) + e^{z \operatorname{ht}(u_{n+1})}\right|^q\right] \\
&\leq 2^q \cdot \left(e^{q\operatorname{Re} z}|C_n(z)|^q \mathbb{E}\left[|M_n(z)|^q\right] + e^{q\operatorname{Re} z} \mathbb{E}\left[\sum_{i=1}^{n} \frac{w_i}{W_n} e^{\operatorname{ht}(u_i) q \operatorname{Re} z}\right] \right) \\
&= O_{\mathscr{E}}\left(n^{q\operatorname{Re}\phi(z)+0\vee g(z,q)+o_{\mathscr{E}}(1)}\right) + O_{\mathscr{E}}\left(n^{\phi(q\operatorname{Re} z)}\right) \\
&= O_{\mathscr{E}}\left(n^{q\operatorname{Re}\phi(z)\vee(q\operatorname{Re}\phi(z)+g(z,q))\vee\phi(q\operatorname{Re} z)+o_{\mathscr{E}}(1)}\right)
\end{aligned}
$$

and the last exponent reduces to $q \operatorname{Re} \phi(z) \vee \phi(q \operatorname{Re} z)$ because $(q \operatorname{Re} \phi(z) + g(z, q)) = \phi(q \operatorname{Re} z) + 1 - q < \phi(q \operatorname{Re} z)$. Hence, using Lemma A.2, we get

$$\mathbb{E}\left[|X_{2n}(z) - X_n(z)|^q\right] \leq 2^q \sum_{i=n}^{2n} \mathbb{E}\left[|X_{i+1}(z) - X_i(z)|^q\right] = O_{\mathscr{E}}\left(n^{1 + (q \operatorname{Re}(\phi(z)) \vee \phi(q \operatorname{Re} z)) + o_{\mathscr{E}}(1)}\right),$$

which finishes the proof of the lemma. $\qquad\square$

Recall the definition of z_+ and z_- in (1.7). We define the domain $\mathscr{D}$ to which we refer in the statement of Proposition 3.1 as the connected component of

$$\mathscr{V} \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$$

that contains 0, where $\mathscr{V}$ is defined in (3.6). In this way, $\mathscr{D}$ is a domain of $\mathbb{C}$ and $\mathscr{D} \cap \mathbb{R} = (z_-, z_+)$. Indeed, first, $\mathscr{D}$ is open and connected by definition. Then recall from Lemma 3.5 that $\mathscr{V} \cap \mathbb{R}$ contains $I_\gamma = \{x \in \mathbb{R} \mid 1 + \gamma(e^x - 1 - xe^x) > 0\}$ an open interval which contains 0 and has z_+ as its right endpoint. Now just check that $\{z \in \mathbb{R} \mid 1 + \operatorname{Re}(\phi(z)) > 0\} = (z_-, \infty)$ and that $z_- \in I_\gamma$.

For technical reasons, we also introduce the following subset of $\mathbb{C}$, here identified as $\mathbb{R} \times \mathbb{R}$,

$$\mathscr{D}' := (z_-, z_+) \times (0, 2\pi),$$

on which the process $(z \mapsto M_n(z))_{n \geq J}$, and hence also $(z \mapsto X_n(z))_{n \geq J}$, are well-defined. We further decompose $\mathscr{D}'$ into a union of open sets

$$\mathscr{D}' = \bigcup_{1 < q \leq p} \mathscr{D}'_q \quad \text{where} \quad \mathscr{D}'_q = \{z \in \mathscr{D}' \mid g(\operatorname{Re} z, q) < 0\}.$$

Lemma 3.13. *The following holds.*

(i) *For all compact $K \subset \mathscr{D}$ there exists $\epsilon(K) > 0$ such that almost surely*

$$n^{-(1 + \operatorname{Re} \phi(z))} \cdot |X_n(z)| = O_K\left(n^{-\epsilon(K)}\right).$$

(ii) *For all compact $K \subset \mathscr{D}'$, there exists $\epsilon(K) > 0$ such that*

$$n^{-(1 + \phi(\operatorname{Re} z))} \cdot \left|\sum_{i=J}^{n-1} C_i(z) M_i(z)\right| = O_K\left(n^{-\epsilon(K)}\right)$$

(iii) *For all compact $K \subset \mathscr{D}'$, there exists $\epsilon(K) > 0$ such that almost surely*

$$n^{-(1 + \phi(\operatorname{Re} z))} \cdot |X_n(z)| = O_K\left(n^{-\epsilon(K)}\right).$$

Proof. We start by proving (i). For that, we first prove that the statement holds for all compact $K \subset \mathscr{V} \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$ hence that will also be the case for $K \subset \mathscr{D}$. First, we use the same compactness argument as in the proof of Proposition 3.6 to ensure that it suffices to prove the statement for compact sets K such that $K \subset \mathscr{V}_q \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$ for some $q \in (1, p]$. Then, for such a compact set K, the result is ensured by an application of Lemma A.3 on the open set $\mathscr{V}_q \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$ with $\alpha(z) = 1 + \operatorname{Re}(\phi(z)) > 0$ and $\delta(z) = \min(q - 1, -g(z, q)) > 0$, which we can apply thanks to Lemma 3.12.

We now prove (ii). For any $q \in (1, p]$, thanks to Lemma 3.7, on the open set $\mathscr{D}'_q \subset \mathscr{E}$ we have $\mathbb{E}\left[|M_{2n}(z) - M_n(z)|^q\right] = O_{\mathscr{D}'_q}\left(n^{(1-q)\vee g(z,q)+o_{\mathscr{D}'_q}(1)}\right)$ and

$$g(z,q) = q\underbrace{(\phi(\operatorname{Re} z) - \operatorname{Re} \phi(z))}_{>0} + \underbrace{g(\operatorname{Re} z, q)}_{<0}.$$

Applying Lemma A.3 for the sequence of functions $(z \mapsto M_n(z))_{n \geq J}$ on any compact subset $K \subset \mathscr{D}'_q$ with $\alpha(z) = \phi(\operatorname{Re} z) - \operatorname{Re} \phi(z) > 0$ and $\delta(z) = \min(-1 + q + q(\phi(\operatorname{Re} z) - \operatorname{Re}(\phi(z))), -g(\operatorname{Re} z, q)) > 0$ yields:

$$n^{-\phi(\operatorname{Re} z)+\operatorname{Re}\phi(z)} \cdot M_n(z) = O_K\left(n^{-\epsilon(K)}\right).$$

Using the estimates of Lemma 3.3, we have $|C_n(z)| = O_K\left(n^{\operatorname{Re}\phi(z)}\right)$, and so $|C_i(z)M_i(z)| = O_K\left(i^{\phi(\operatorname{Re} z)-\epsilon(K)}\right)$. Hence $\sum_{i=J}^{n-1} |C_i(z)M_i(z)| = O_K\left(n^{0\vee(1+\phi(\operatorname{Re} z)-\epsilon(K))}\right)$ which finishes the proof of (ii).

Last, in order to prove (iii), we use Lemma A.3 on $\mathscr{D}'_q$ for the sequence $(z \mapsto X_n(z))_{n \geq J}$ with $\alpha(z) = 1 + \phi(\operatorname{Re} z) > 0$ and $\delta(z) = \min(-1+q+q(\phi(\operatorname{Re} z)-\operatorname{Re}(\phi(z))), -g(\operatorname{Re} z, q))$. $\quad\square$

In order to conclude, we also need the following lemma, which is a direct consequence of Lemma 3.3.

Lemma 3.14. *For any compact $K \subset \mathscr{E} \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$, there exists $\epsilon(K)$ such that*

$$\left| n^{-(1+\phi(z))} \cdot \sum_{i=J}^{n-1} C_i(z) - \frac{e^{c(z)}}{1+\phi(z)} \right| = O_K\left(n^{-\epsilon(K)}\right)$$

Proof. On any compact $K \subset \mathscr{E} \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$, using Lemma 3.3 we write

$$C_n(z) = e^{c(z)} \cdot n^{\phi(z)} \cdot (1 + O_K\left(n^{-\epsilon}\right)),$$

so that

$$\sum_{i=J}^{n-1} C_i(z) = e^{c(z)} \cdot \sum_{i=J}^{n-1} i^{\phi(z)} + e^{c(z)} \cdot \sum_{i=J}^{n-1} i^{\phi(z)} \cdot O_K\left(i^{-\epsilon}\right)$$

$$= \frac{e^{c(z)} n^{1+\phi(z)}}{1+\phi(z)} \cdot (1 + O_K\left(n^{-1}\right)) + O_K\left(n^{1+\phi(z)-\epsilon(K)}\right),$$

where in the second line, we use the fact that $\inf_{z \in K}(1 + \operatorname{Re}\phi(z)) > 0$, and we define $\epsilon(K) := \epsilon \wedge 1 \wedge \inf_{z \in K}(1 + \operatorname{Re}\phi(z))$. This proves the lemma. $\quad\square$

We can now prove Proposition 3.1.

Proof of Proposition 3.1. We start by proving simultaneously that for any $z \in \mathscr{D}$, we almost surely have

$$N_\infty(z) := \lim_{n \to \infty} N_n(z) = \frac{e^{z+c(z)}}{1+\phi(z)} \cdot M_\infty(z), \tag{3.16}$$

and also that both points (i) and (ii) of the proposition hold. Let $K \subset \mathscr{D}$ be a compact set. For $z \in K$, we write

$$\left| N_n(z) - \frac{e^{z+c(z)}}{1+\phi(z)} M_\infty(z) \right| \leq \left| n^{-(1+\phi(z))} X_n(z) \right|$$

$$+ \left| n^{-(1+\phi(z))} e^z \sum_{i=J}^{n} C_i(z) M_i(z) - \frac{e^{z+c(z)}}{1+\phi(z)} M_\infty(z) \right|$$

The first term is $O_K\big(n^{-\epsilon(K)}\big)$ thanks to Lemma 3.13(i). For the second one, we re-write the term $M_i(z)$ appearing in the sum as $M_i(z) = M_\infty(z) + (M_i(z) - M_\infty(z))$, then re-arrange the terms and use the triangle inequality to get the following upper-bound

$$
|M_\infty(z)| \cdot |e^z| \cdot \overbrace{\left| n^{-(1+\phi(z))} \cdot \sum_{i=J}^{n} C_i(z) - \frac{e^{c(z)}}{(1+\phi(z))} \right|}^{O_K\big(n^{-\epsilon(K)}\big)}
$$

$$
+ \, n^{-(1+\mathrm{Re}\,\phi(z))} \cdot |e^z| \cdot \sum_{i=J}^{n-1} \underbrace{|C_i(z)| \cdot |M_i(z) - M_\infty(z)|}_{O_K\big(i^{\mathrm{Re}\,\phi(z)-\epsilon(K)}\big)}.
$$

In the above display, we used Lemma 3.14 for the first term and then Lemma 3.3 together with Proposition 3.6 for the second term. Note that here we slightly abuse notation and use the same symbol $\epsilon(K)$ for different positive quantities. In the end, the whole expression is $O_K\big(n^{-\epsilon(K)}\big)$, for yet another value $\epsilon(K) > 0$. From (3.16), it is clear that the limiting function $z \mapsto N_\infty(z)$ is analytic and has almost surely no zero on (z_-, z_+) because of Lemma 3.9. For (iii), we prove the stronger statement: for any compact subset $K \subset (z_-, z_+)$ and $0 < a < \pi$, there exists $\epsilon(K,a) > 0$ such that almost surely,

$$
\sup_{x \in K} \sup_{a \leq \eta \leq \pi} n^{-(1+\phi(x))} \left| \sum_{i=1}^{n} e^{(x+i\eta)\,\mathrm{ht}(u_i)} \right| = O\Big(n^{-\epsilon(K,a)}\Big).
$$

For this, we write

$$
n^{-(1+\phi(x))} \left| \sum_{i=1}^{n} e^{(x+i\eta)\,\mathrm{ht}(u_i)} \right| \leq n^{-(1+\phi(x))} |X_n(x+i\eta)|
$$

$$
+ \, n^{-(1+\phi(x))} \left| \sum_{i=J}^{n-1} C_i(x+i\eta) M_i(x+i\eta) \right|.
$$

We apply points (ii) and (iii) of Lemma 3.13 to the compact $K \times [a, \pi]$ and get the desired bound. $\qquad\square$

3.3 Height of the tree

In this section, we study the behaviour of the height $\mathrm{ht}(\mathsf{T}_n)$ of the tree T_n, which is defined as the maximal height of the vertices of T_n, i.e.

$$
\mathrm{ht}(\mathsf{T}_n) = \max_{1 \leq k \leq n} \mathrm{ht}(u_k).
$$

We start by showing that under the assumption ($\square_\gamma^p$) we have the convergence (1.10). Then, for the sake of completeness, we also study the simpler case where $\log n = o\big(\sum_{i=1}^{n} \frac{w_i}{W_i}\big)$.

One key argument in our proofs is the following equality for the *annealed* moment generating function of the height of u_k, for any fixed $k \geq 2$, which can be seen as a corollary of Lemma 3.2

$$
\mathbb{E}\left[e^{z\,\mathrm{ht}(u_k)} \right] = e^z \cdot \prod_{j=2}^{k-1} \left(1 + (e^z - 1)\frac{w_j}{W_j} \right). \tag{3.17}
$$

Some elementary computations using the Chernoff bound and the last display yield the following lemma.

Lemma 3.15. *Suppose that the sequence of weights $\boldsymbol{w}$ satisfies*

$$\limsup_{n\to\infty} \frac{1}{\log n} \sum_{i=2}^{n} \frac{w_i}{W_i} \leq u \in (0\,,\infty).$$

Then almost surely we have

$$\limsup_{n\to\infty} \frac{\mathrm{ht}(\mathrm{T}_n)}{\log n} \leq u e^{z_+(u)},$$

where $z_+(u)$ is the unique positive root of $u(z e^z - e^z + 1) - 1 = 0$.

Proof. Using the expression (3.17) for the moment generating function of $\mathrm{ht}(u_n)$ we get, for any $z > 0$ and $n \geq 2$

$$\mathbb{E}\left[e^{z\,\mathrm{ht}(u_n)}\right] = e^z \cdot \prod_{j=2}^{n-1}\left(1 + (e^z - 1)\frac{w_j}{W_j}\right) \leq \exp\left(1 + (e^z - 1)\sum_{j=2}^{n-1}\frac{w_j}{W_j}\right)$$

$$\leq \exp\left((\log n)\cdot(u(e^z - 1) + o(1))\right),$$

where we use the inequality $(1 + x) \leq e^x$ and the assumption on $\boldsymbol{w}$. Then, for any $z > 0$ and $n \geq 1$, we use Markov's inequality to get

$$\mathbb{P}\left(\mathrm{ht}(u_n) \geq u e^z \log n\right) \leq e^{-u z e^z \log n}\mathbb{E}\left[e^{z\,\mathrm{ht}(u_n)}\right] \leq \exp\left(-u \log n(z e^z - e^z + 1 + o(1))\right).$$

If we take $z > 0$ such that $u(z e^z - e^z + 1) > 1$ then the right-hand-side is summable and hence using the Borel-Cantelli lemma shows that for all n large enough, we have $\mathrm{ht}(u_n) \leq u e^z \log n$. Letting $z \searrow z^+(u)$, we get the result. $\qquad\square$

We now prove the last claim (1.10) of Theorem 1.5. Here we suppose that the weight sequence $\boldsymbol{w}$ satisfies ($\square_\gamma^p$) for some $\gamma > 0$ and some $p \in (1\,,2]$.

Proof of (1.10). Recall the asymptotics (1.9) in Theorem 1.5. It ensures that there almost surely exist vertices at height $\lfloor \gamma e^z \log n \rfloor$, for any fixed $z \in (z_-\,, z_+)$ and n large enough. Hence the height of the tree T_n satisfies

$$\liminf_{n\to\infty} \frac{\mathrm{ht}(\mathrm{T}_n)}{\log n} \geq \gamma e^{z_+}.$$

For the limsup, we use Lemma 3.15 with $u = \gamma$ (this is justified by Lemma 3.4), which yields $\limsup_{n\to\infty} \frac{\mathrm{ht}(\mathrm{T}_n)}{\log n} \leq \gamma e^{z_+}$. $\qquad\square$

To finish the section, we state a proposition.

Proposition 3.16. *Let $f(n) := \sum_{i=2}^{n-1} \frac{w_i}{W_i}$. If $\log n = o(f(n))$ then we have the almost sure convergences*

$$\lim_{n\to\infty} \frac{\mathrm{ht}(\mathrm{T}_n)}{f(n)} = \lim_{n\to\infty} \frac{\mathrm{ht}(u_n)}{f(n)} = 1.$$

Proof. As we can check from its moment generating function (3.17), for $n \geq 2$ the random variable $\mathrm{ht}(u_n) - 1$ is a sum of independent Bernoulli random variables, with expectation $f(n)$. Using standard bounds for sums of Bernoulli random variables yields

$$\mathbb{P}\left(|\mathrm{ht}(u_n) - 1 - f(n)| \geq \epsilon f(n)\right) \leq 2\exp\left(-\epsilon^2 f(n)/3\right),$$

which is summable in n for any $\epsilon > 0$. The result of the proposition is then obtained using the Borel-Cantelli lemma. $\qquad\square$

4 Preferential attachment trees are weighted recursive trees

In this section, we study preferential attachment trees with fitnesses **a** as defined in the introduction. First, in Section 4.1, we prove Theorem 1.1 which allows us to see them as weighted random trees $\mathrm{WRT}(\mathbf{w^a})$ for some random weight sequence $\mathbf{w^a}$. Then in Section 4.2 we prove Proposition 1.3 which relates the a.s. asymptotic behaviour of $\mathbf{w^a}$ to the behaviour of **a**. Finally, in Section 4.3 we prove Proposition 4.2, which ensures that the sequence $\mathbf{m^a}$ obtained as the scaling limit of the degrees can be expressed as the increments of a Markov chain.

4.1 Coupling with a sequence of Pólya urns: proof of Theorem 1.1

Here we fix an arbitrary sequence **a** such that $a_1 > -1$ and $\forall n \geq 2,\ a_n \geq 0$. We recall the notation, for $n \geq 0$,

$$A_n := \sum_{i=1}^{n} a_i,$$

with the convention that $A_0 = 0$. We consider a sequence of trees $(\mathrm{P}_n)_{n \geq 1}$ evolving according to the distribution $\mathrm{PAT}(\mathbf{a})$ and we want to prove Theorem 1.1, namely that there exists a random sequence of weights $\mathbf{w^a}$ for which the sequence evolves as a $\mathrm{WRT}(\mathbf{w^a})$. The proof uses a decomposition of this process into an infinite number of Pólya urns. This is very close to what is used in the proofs of [3, Theorem 2.1] or [7, Section 1.2] in similar settings. The novelty of our approach is to express this result using weighted random trees, since it allows us to apply all the results developed in the preceding section.

Pólya urns. For us, a Pólya urn process $(\mathsf{Urn}(n))_{n \geq 0} = (X(n), \mathrm{Total}(n))_{n \geq 0}$ is a Markov chain on $E := \{(x, z) \in [0, \infty) \times (0, \infty) \mid x \leq z\}$ with transition probabilities given by the matrix P where for all $(x, z) \in E$,

$$P\left((x, z), (x+1, z+1)\right) = \frac{x}{z} \qquad \text{and} \qquad P\left((x, z), (x, z+1)\right) = \frac{z-x}{z}.$$

The quantities $X(n)$ and $\mathrm{Total}(n)$ represent respectively the number of red balls and the total number of balls at time n in a urn containing red and blacks balls where at each step, a colour is chosen at random proportionally to the current proportion in the urn and a ball of the drawn colour is added to the urn. Starting at time 0 from the state $(a, a + b)$, i.e. with a red balls and b black balls, it is well-known that the sequence $(\Delta X(n))_{n \geq 1} = (X(n) - X(n-1))_{n \geq 1}$ of random variables is *exchangeable*, and an application of de Finetti's representation theorem ensures that it has the same distribution as i.i.d. samples of Bernoulli random variables with a *random* parameter β, which has distribution $\mathrm{Beta}(a, b)$, where we use the convention that $\mathrm{Beta}(a, b) = \delta_1$ if $b = 0$.

Note that the process $(\mathsf{Urn}(n))_{n \geq 0}$ is entirely determined from $(\Delta X(n))_{n \geq 1}$ and that the random variable β is a measurable function of the sequence $(\mathsf{Urn}(n))_{n \geq 0}$ because it can almost surely be obtained as $\beta = \lim_{n \to \infty} \frac{X(n)}{n}$.

Nested structure of urns in the tree. For all $k \geq 1$ we define the following process in $n \geq k$

$$W_k(n) := A_k + \sum_{i=1}^{k} \deg_{\mathrm{P}_n}^{+}(u_i),$$

the "total activity" of the vertices $\{u_1, u_2, \ldots, u_k\}$, for which we remark that for any $k \geq 1$ we have

$$W_k(k) = A_k + k - 1 \qquad \text{and} \qquad W_k(k+1) = A_k + k.$$

Imagine that P_n is already constructed and that we add a new vertex u_{n+1} to the tree. We choose its parent in a downward sequential way:

- we first determine whether the parent is u_n, this happens with probability

$$\frac{a_n + \deg_{\mathsf{P}_n}^+(u_n)}{W_n(n)} = 1 - \frac{W_{n-1}(n)}{W_n(n)},$$

- then with the complementary probability $\frac{W_{n-1}(n)}{W_n(n)}$ it is not, so conditionally on this we determine whether it is u_{n-1}, this happens with (conditional) probability

$$\frac{a_{n-1} + \deg_{\mathsf{P}_n}^+(u_{n-1})}{W_{n-1}(n)} = 1 - \frac{W_{n-2}(n)}{W_{n-1}(n)}.$$

- then with the complementary probability $\frac{W_{n-2}(n)}{W_{n-1}(n)}$ it is not, etc... We continue this process until we stop at some u_i.

Now we fix $k \geq 1$ and introduce the following time-change: for all $N \geq 0$, we let

$$\theta_k(N) := \inf\left\{ n \geq k+1 \mid W_{k+1}(n) = A_{k+1} + k + N \right\},$$

be the N-th time that a vertex is attached on one of the vertices $\{u_1, \ldots, u_{k+1}\}$ after time $k+1$, where by definition, we have $\theta_k(0) = k+1$. Remark that it can be the case that $\theta_k(N)$ is not defined for large N, if there is only a finite number of vertices attaching to $\{u_1, \ldots, u_{k+1}\}$. We ignore this possibility for the moment, and only consider sequences $\mathbf{a}$ for which $A_n = O(n)$, for which this almost surely does not happen. In this case for all $N \geq 0$ we set

$$\mathsf{Urn}_k(N) := (W_k(\theta_k(N)), W_{k+1}(\theta_k(N))) = (W_k(\theta_k(N)), A_{k+1} + k + N).$$

Now, the three following facts are the key observations in order to prove Theorem 1.1:

(i) for all $k \geq 1$, the process $\mathsf{Urn}_k = (\mathsf{Urn}_k(N))_{N \geq 0}$ has the distribution of a Pólya urn starting from the state $(A_k + k, A_{k+1} + k)$,

(ii) those process are jointly independent for $k \geq 1$,

(iii) the whole sequence $(\mathsf{P}_n)_{n \geq 1}$ is a function of the trajectories of the processes $(\mathsf{Urn}_k, \ k \geq 1)$.

Point (i) already follows from the discussion above. A moment of thought shows that (ii) holds as well: of course the processes $(W_k(n), W_{k+1}(n))_{n \geq k+1}$ for different k are not independent at all but the point is that they only interact through the time-changes $(\theta_k(\cdot), k \geq 1)$. Last, for (iii), note that we can reconstruct the tree P_n at time n from the random variables $(W_i(k))_{1 \leq i, k \leq n}$ and that these random variables can be entirely determined using the system of equations below

$$\begin{cases} W_k(n) = \mathsf{Urn}_k(W_{k+1}(n) - (A_{k+1} + k)), & \text{for } 1 \leq k \leq n-1, \\ W_n(n) = A_n + n - 1. \end{cases}$$

Reversing the construction and using the exchangeability. We now reverse the construction and start with an independent family $(\mathsf{Urn}_k,\ k \geq 1)$ of processes that have, for each $k \geq 1$, the distribution of a Pólya urn starting from the state $(A_k + k, A_{k+1} + k)$, so that they have the joint same distribution as the ones described in (i) and (ii). From what we did above, the sequence $(\mathsf{P}_n)_{n \geq 1}$ that they determine through (iii) has distribution $\mathrm{PAT}(\mathbf{a})$. A moment of thought shows that this argument actually still holds for a completely arbitrary sequence of fitnesses $\mathbf{a}$.

Now, using de Finetti's theorem, each of the processes Urn_k can be produced by sampling $\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1})$ and adding a red ball at each step independently with probability β_k and a black ball with probability $1 - \beta_k$. This is of course done independently for different $k \geq 1$.

In terms of our downward sequential procedure defined above for finding the parent of each newcomer, it amounts to saying that each time that we have to choose between attaching to u_{k+1} or attaching to a vertex among $\{u_1, \ldots, u_k\}$, the former is chosen with probability $1 - \beta_k$ and the latter with probability β_k. Let us verify that the law of $(\mathsf{P}_n)_{n \geq 1}$ conditionally on the sequence $(\beta_k)_{k \geq 1}$ can indeed be expressed as $\mathrm{WRT}(\mathbf{w^a})$ where the random sequence of weights $\mathbf{w^a} = (\mathsf{w}_n^{\mathbf{a}})_{n \geq 1}$ is defined from the sequence $(\beta_k)_{k \geq 1}$ as

$$\forall n \geq 1, \qquad \mathsf{W}_n^{\mathbf{a}} = \prod_{i=1}^{n-1} \beta_i^{-1} \qquad \text{and} \qquad \mathsf{w}_n^{\mathbf{a}} = \mathsf{W}_n^{\mathbf{a}} - \mathsf{W}_{n-1}^{\mathbf{a}},$$

with the convention that $\mathsf{W}_1^{\mathbf{a}} = 1$ and $\mathsf{W}_0^{\mathbf{a}} = 0$. We reason conditionally on the sequence $(\beta_k)_{k \geq 1}$ (or equivalently the sequence $(\mathsf{w}_n^{\mathbf{a}})_{n \geq 1}$). When determining the parent of u_{n+1}, whose label we denote J_{n+1} as in (1.2), we successively try to attach to $u_n, u_{n-1}, \ldots$ until we stop at $u_{J_{n+1}}$. Using the independence, we get that for every $k \in \{1, 2, \ldots, n\}$,

$$\mathbb{P}\left(J_{n+1} = k \mid \mathsf{P}_1, \ldots, \mathsf{P}_n, (\beta_i)_{i \geq 1}\right) = \beta_{n-1}\beta_{n-2}\ldots\beta_k(1 - \beta_{k-1}) = \frac{\mathsf{W}_k^{\mathbf{a}} - \mathsf{W}_{k-1}^{\mathbf{a}}}{\mathsf{W}_n^{\mathbf{a}}} = \frac{\mathsf{w}_k^{\mathbf{a}}}{\mathsf{W}_n^{\mathbf{a}}}.$$

This proves Theorem 1.1. Let us explain how Corollary 1.2 follows from the proof that we developed here. From the discussion in the previous paragraph, in the case of a sequence $\mathbf{a}$ for which $A_n = O(n)$, each of the processes $(\mathsf{Urn}_k(N))_{N \geq 0}$ for $k \geq 1$ is a measurable function of $(\mathsf{P}_n)_{n \geq 1}$, and hence the associated β_k also is. In the end, the sequence $(\mathsf{w}_n^{\mathbf{a}})_{n \geq 1}$ is a measurable function of $(\mathsf{P}_n)_{n \geq 1}$ and it is easy to check that it corresponds to the one described in the statement of Corollary 1.2.

4.2 Proof of Proposition 1.3

Recall the definition of $(\mathsf{W}_n^{\mathbf{a}})_{n \geq 1}$ as the random sequence of cumulated weights defined Theorem 1.1, whose distribution depends on a sequence $\mathbf{a}$ of fitnesses, and is expressed using a sequence of independent Beta-distributed random variables $(\beta_k)_{k \geq 1}$, where

$$\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1}).$$

We now prove Proposition 1.3, which relates the growth of $(\mathsf{W}_n^{\mathbf{a}})_{n \geq 1}$ to that of $(A_n)_{n \geq 1}$.

Proof of Proposition 1.3. As in [21, Proof of Lemma 1.1], we introduce

$$X_n := \prod_{i=1}^{n-1} \frac{\beta_i}{\mathbb{E}[\beta_i]}. \tag{4.1}$$

It is easy to see that X_n is a positive martingale, hence it almost surely converges to a limit X_∞ as $n \to \infty$. Now, using the fact that the $(\beta_n)_{n \geq 1}$ are independent and that for

any integer $q \geq 0$, the q-th moment of a random variable with $\text{Beta}(a, b)$ distribution is given by

$$\frac{\Gamma(a+q)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+q)} = \prod_{k=0}^{q-1} \frac{a+k}{a+b+k}, \tag{4.2}$$

we can compute

$$\begin{aligned}
\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i^p\right] &= \prod_{i=1}^{n-1} \left(\prod_{k=0}^{p-1} \frac{i+A_i+k}{i+A_{i+1}+k}\right) \\
&= \prod_{k=0}^{p-1} \left(\frac{1+A_1+k}{n+A_n+k-1} \prod_{i=2}^{n-1} \frac{i+A_i+k}{i+A_i+k-1}\right) \\
&= \left(\prod_{k=0}^{p-1} \frac{1+A_1+k}{n+A_n+k-1}\right) \cdot \prod_{k=0}^{p-1}\prod_{i=2}^{n-1} \left(1 + \frac{1}{i+A_i+k-1}\right). \tag{4.3}
\end{aligned}$$

Now since by assumption $\mathbf{a}$ satisfies (H_c), there exists $\epsilon > 0$ such that $A_n = c \cdot n + O(n^{1-\epsilon})$ and without loss of generality we can assume that $\epsilon < 1$. For all $k \in [\![0, p-1]\!]$ we can write

$$n + A_n + k - 1 \underset{n\to\infty}{=} (c+1)n + O(n^{1-\epsilon}) \quad \text{and so} \quad \frac{1}{n+A_n+k-1} \underset{n\to\infty}{=} \frac{1}{(c+1)n} + O(n^{-1-\epsilon}).$$

Hence

$$\begin{aligned}
\prod_{k=0}^{p-1}\prod_{i=2}^{n-1}\left(1 + \frac{1}{i+A_i+k-1}\right) &= \prod_{k=0}^{p-1}\prod_{i=2}^{n-1}\left(1 + \frac{1}{(c+1)i} + O(i^{-1-\epsilon})\right) \\
&= \exp\left(\sum_{k=0}^{p-1}\sum_{i=2}^{n}\left(\frac{1}{(c+1)i} + O(i^{-1-\epsilon})\right)\right) \\
&= \exp\left(\frac{p}{c+1}\log n + \text{cst} + O(n^{-\epsilon})\right) \\
&= \text{cst} \cdot n^{\frac{p}{c+1}}\left(1 + O(n^{-\epsilon})\right). \tag{4.4}
\end{aligned}$$

In the end, since $\left(\prod_{k=0}^{p-1}\frac{1+A_1+k}{n+A_n+k-1}\right) = \text{cst} \cdot \prod_{k=0}^{p-1}\frac{1}{(c+1)n+O(n^{1-\epsilon})} = \text{cst} \cdot n^{-p} \cdot (1 + O(n^{-\epsilon}))$, we get, using (4.3) and (4.4),

$$\prod_{i=1}^{n-1}\mathbb{E}\left[\beta_i^p\right] = C_p \cdot n^{-p+p/(c+1)} \cdot (1 + O(n^{-\epsilon})) \tag{4.5}$$

where C_p is a positive constant which depends on p and the sequence $\mathbf{a}$. This entails that, under our assumptions, for any $p \geq 1$, we have

$$\mathbb{E}\left[X_n^p\right] = \frac{\prod_{i=1}^{n-1}\mathbb{E}\left[\beta_i^p\right]}{\prod_{i=1}^{n-1}\mathbb{E}\left[\beta_i\right]^p} \underset{n\to\infty}{\to} \frac{C_p}{C_1^p},$$

which shows that this martingale is bounded in L^p for all $p \geq 1$ and hence it is uniformly integrable. Consequently, it converges a.s. and in L^p to a limit random variable X_∞, with moments determined by

$$\forall p \geq 1, \qquad \mathbb{E}\left[X_\infty^p\right] = \frac{C_p}{C_1^p}. \tag{4.6}$$

Furthermore, we have

$$\mathbb{E}\left[(X_{n+1} - X_n)^2\right] = \mathbb{E}\left[X_n^2\left(\frac{\beta_n}{\mathbb{E}\left[\beta_n\right]} - 1\right)^2\right] \leq \mathbb{E}\left[X_n^2\right] \cdot \frac{\text{Var}\left(\beta_n\right)}{\mathbb{E}\left[\beta_n\right]^2}. \tag{4.7}$$

Since $\beta_n \sim \mathrm{Beta}(n + A_n, a_{n+1})$, we get

$$\mathbb{E}\left[\beta_n\right] = \frac{n + A_n}{n + A_{n+1}} \to 1 \qquad \text{and} \qquad \mathrm{Var}\left(\beta_n\right) = \frac{a_{n+1}(n + A_n)}{(n + A_{n+1})^2(n + A_{n+1} + 1)} = O\left(\frac{a_{n+1}}{n^2}\right). \tag{4.8}$$

Using (4.7), (4.8), Lemma A.2 and then summing over $n \le k \le 2n - 1$ and using the fact that $\mathbf{a}$ satisfies (H_c) we get that

$$\mathbb{E}\left[(X_{2n} - X_n)^2\right] = O\left(\frac{\sum_{k=n+1}^{2n} a_k}{n^2}\right) = O\left(\frac{A_{2n}}{n^2}\right) = O\left(n^{-1}\right).$$

Using Lemma A.3, we get that almost surely, for any $\epsilon' < \frac{1}{2}$,

$$|X_n - X_\infty| = O\left(n^{-\epsilon'}\right).$$

Since $\beta_i > 0$ almost surely for every $i \ge 1$, the event $\{X_\infty = 0\}$ is a tail event for the filtration generated by the β_i's and has hence probability 0 or 1, thanks to Kolmogorov's 0-1 law. In the end, it has probability 0 because $\mathbb{E}\left[X_\infty\right] = 1$. We deduce that

$$\begin{aligned}
(\mathsf{W}_n^{\mathbf{a}})^{-1} = \prod_{i=1}^{n-1} \beta_i = X_n \cdot \prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i\right] \\
= X_\infty \cdot \left(1 + O\left(n^{-\epsilon'}\right)\right) \cdot C_1 \cdot n^{-1 + \frac{1}{c+1}} \cdot \left(1 + O(n^{-\epsilon})\right) \\
= C_1 \cdot X_\infty \cdot n^{-1 + \frac{1}{c+1}} \cdot \left(1 + O\left(n^{-(\epsilon \wedge \epsilon')}\right)\right).
\end{aligned} \tag{4.9}$$

Hence, we have,

$$\mathsf{W}_n^{\mathbf{a}} \underset{n \to \infty}{\bowtie} Z \cdot n^{\frac{c}{(c+1)}} \qquad \text{with} \qquad Z := \frac{1}{X_\infty \cdot C_1}. \tag{4.10}$$

Whenever $a_n \le n^{c' + o(1)}$ as $n \to \infty$, we can show the following (we postpone the proof to the end of the section)

Lemma 4.1. *For any $\delta > 0$ small enough, we have*

$$\mathbb{P}\left(1 - \beta_k > k^{-1 + c' + \delta}\right) \le \exp\left(-(k+1)^{c' + \delta + o(1)}\right).$$

Since the last quantity is summable in k we can use the Borel-Cantelli lemma (and a sequence of δ going to 0) to show that almost surely $1 - \beta_k \le k^{-1 + c' + o^\omega(1)}$ as $k \to \infty$, where the term $o^\omega(1)$ denotes a random function of k that almost surely tends to 0 as $k \to \infty$. Combining this with (4.10), we finish proving the proposition by writing

$$\mathsf{w}_k^{\mathbf{a}} = \mathsf{W}_k^{\mathbf{a}} - \mathsf{W}_{k-1}^{\mathbf{a}} = \mathsf{W}_k^{\mathbf{a}} \cdot (1 - \beta_{k-1}) \le (k+1)^{c' - 1/(c+1) + o^\omega(1)}. \qquad \square$$

We finish by giving a proof of Lemma 4.1.

Proof of Lemma 4.1. Let $x \ge 0$ and $y > 1$ and let X be a random variable with distribution $\mathrm{Beta}(x + 1, y)$ and Y with distribution $\mathrm{Beta}(x, 1)$, independent of X. By standard results on Beta distributions, the product $Z = X \cdot Y$ has distribution $\mathrm{Beta}(x, y)$.

Then for any $z \in [0, 1]$ we have, using the explicit expression of the density of X,

$$\begin{aligned}
\mathbb{P}\left(Z > z\right) \le \mathbb{P}\left(X > z\right) &= \frac{\Gamma\left(x + 1 + y\right)}{\Gamma\left(x + 1\right)\Gamma\left(y\right)} \int_z^1 u^x(1 - u)^{y-1}\,\mathrm{d}u \\
&\le \frac{\Gamma\left(x + 1 + y\right)}{\Gamma\left(x + 1\right)\Gamma\left(y\right)} \cdot \exp\left(-(y-1)z\right) \int_z^1 u^x\,\mathrm{d}u \\
&\le \frac{\Gamma\left(x + 1 + y\right)}{\Gamma\left(x + 2\right)\Gamma\left(y\right)} \cdot \exp\left(-(y-1)z\right),
\end{aligned}$$

and the last display in increasing in x. We later use this inequality for well-chosen sequences (x_n), (y_n) and (z_n) taking place of the values of x, y, z. First, remark that for any two non-negative sequences (x_n) and (y_n) with (y_n) going to infinity and $x_n = o(y_n)$, we have the following estimate using Stirling's approximation:

$$\log \left(\frac{\Gamma(x_n + 1 + y_n)}{\Gamma(x_n + 2)\,\Gamma(y_n)} \right) \underset{n \to \infty}{=} (x_n + 1)\log(y_n) \cdot (1 + o(1)).$$

Now we apply the above computations for every $n \geq 1$ with $z_n := n^{-1+c'+\delta}$ to the random variables $(1 - \beta_n)$ which have distribution $\mathrm{Beta}(x_n, y_n)$, with $x_n := a_{n+1}$ and $y_n := A_n + n$. In particular, in this context we have $x_n = a_n \leq (n+1)^{c'+o(1)}$ and $y_n = A_n + n = (n+1)^{1+o(1)}$, so that the all of the above applies and

$$\begin{aligned}
\log \mathbb{P}\left((1 - \beta_n) > n^{-1+c'+\delta}\right) &\leq (x_n + 1)\log(y_n)(1 + o(1)) - (y_n - 1)z_n \\
&\leq (n+1)^{c'+o(1)} - (n+1)^{c'+\delta+o(1)} \\
&\leq -(n+1)^{c'+\delta+o(1)},
\end{aligned}$$

which is what we wanted. $\qquad\square$

4.3 The distribution of the limiting sequence

We stay in the same setting as Section 4.2. Suppose that we are working with a sequence of fitnesses $\mathbf{a}$ that satisfies (H_c) for some $c > 0$. The sequence $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ is defined in (1.5) as some random multiple of the sequence $(\mathsf{W}_n^{\mathbf{a}})_{n \geq 1}$, whose construction is described in Theorem 1.1 from a sequence $(\beta_n)_{n \geq 1}$ of independent random variables with $\beta_n \sim \mathrm{Beta}(A_n + n, a_{n+1})$, so that for all $n \geq 1$,

$$\mathsf{M}_n^{\mathbf{a}} = \frac{c+1}{Z} \cdot \prod_{k=1}^{n-1} \beta_k^{-1},$$

where the random variable Z is the one that appears in (4.10), and depends on the whole sequence $(\beta_n)_{n \geq 1}$.

Proposition 4.2. *For any sequence $\mathbf{a}$ that satisfies the condition (H_c), the sequence $(\mathsf{M}_k^{\mathbf{a}})_{k \geq 1}$ is a (possibly time-inhomogeneous) Markov chain such that for all $k \geq 1$, $\mathsf{M}_{k+1}^{\mathbf{a}}$ is independent of $\beta_1, \beta_2, \ldots, \beta_k$. The fact that for all $k \geq 1$ we have $\mathsf{M}_k^{\mathbf{a}} = \beta_k \cdot \mathsf{M}_{k+1}^{\mathbf{a}}$ with $\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1})$ independent of $\mathsf{M}_{k+1}^{\mathbf{a}}$ characterises the backward transitions of the chain.*

Proof. We follow the same steps as [21, Lemma 1.1]. Recall the definition of the random variable X_∞ as the limit of the sequence $(X_n)_{n \geq 1}$ defined in (4.1), the definition (4.5) of the constant C_1 and their relation to the random variable Z. We have

$$\mathsf{M}_1^{\mathbf{a}} = \frac{c+1}{Z} = (C_1 \cdot (c+1) \cdot X_\infty) \quad \text{and for } k \geq 2, \quad \mathsf{M}_k^{\mathbf{a}} = \mathsf{M}_1^{\mathbf{a}} \cdot \prod_{i=1}^{k-1} \beta_i^{-1}. \tag{4.11}$$

It then follows that we can write, for $k \geq 1$,

$$\mathsf{M}_{k+1}^{\mathbf{a}} = C_1 \cdot (c+1) \cdot X_\infty \cdot \prod_{i=1}^{k} \beta_i^{-1} = C_1 \cdot (c+1) \cdot \lim_{n \to \infty} \frac{\prod_{i=k+1}^{n-1} \beta_i}{\prod_{i=1}^{n-1} \mathbb{E}[\beta_i]},$$

which ensures that $\mathsf{M}_{k+1}^{\mathbf{a}}$ is independent of $\beta_1, \beta_2, ..., \beta_k$. The limit in the last equality exists almost surely thanks to the results of the preceding section.

Now we prove the Markov property of the chain. Let $k \geq 1$. Because of the definition of the chain as a product, the distribution of M_{k+1}^{a} conditional on the past trajectory $M_1^{a}, M_2^{a}, \ldots, M_k^{a}$ is the same as the distribution of M_{k+1}^{a} conditional on $M_k^{a}, \beta_1, \ldots, \beta_{k-1}$. Since $M_{k+1}^{a} = \beta_k^{-1} \cdot M_k^{a}$ and that β_k and M_k^{a} are both independent of $\beta_1, \ldots, \beta_{k-1}$, this conditional distribution corresponds to the one of M_{k+1}^{a} conditional on the present state of the chain M_k^{a}. $\qquad\square$

Computing the moments. In some cases where the sequence a is sufficiently regular, we can compute explicitly every moment of the random variable M_k^{a} for every $k \geq 1$. Indeed, using (4.6) and (4.11) and the independence, we get

$$
\begin{aligned}
\mathbb{E}\left[(M_k^{a})^p\right] &= \mathbb{E}\left[\left(C_1 \cdot (c+1) \cdot \lim_{n\to\infty} \frac{\prod_{i=k}^{n-1}\beta_i}{\prod_{i=1}^{n-1}\mathbb{E}\left[\beta_i\right]}\right)^p\right] \\
&= C_1^p \cdot (c+1)^p \cdot \lim_{n\to\infty} \frac{\prod_{i=k}^{n-1}\mathbb{E}\left[\beta_i^p\right]}{\left(\prod_{i=1}^{n-1}\mathbb{E}\left[\beta_i\right]\right)^p} \\
&= \frac{(c+1)^p \cdot C_p}{\prod_{i=1}^{k-1}\mathbb{E}\left[\beta_i^p\right]}.
\end{aligned}
\tag{4.12}
$$

In general, if the collection $(\mu_p)_{p\geq 1}$ of p-th moments of some random variable satisfies the so-called Carleman's condition: $\sum_{p=1}^{\infty} \mu_{2p}^{-1/(2p)} = \infty$, then its distribution is uniquely determined from those moments.

5 Examples and applications

In this section, we compute the explicit distribution of (M_n^{a}) for some particular sequences a. We then describe some applications of our results to a model of Pólya urn with immigration and then to a model of preferential attachment graphs.

5.1 The limit chain for particular sequences a

As stated in the preceding section, we can compute the distribution of M_k^{a} for some fixed k by the expression of its moments (4.12), provided that they satisfy Carleman's condition. Knowing these distributions and the backward transitions given in Proposition 4.2 then characterizes the law of the whole process. For two particular examples, this law has an explicit expression.

Proposition 5.1. *In the two following cases, the distribution of the chain (M_n^{a}) is explicit.*

(i) *If a is of the form $a = (a, b, b, b, \ldots)$ with $a > -1$ and $b > 0$, then the sequence $(M_n^{a})_{n\geq 1}$ is a Mittag-Leffler Markov chain $\mathrm{MLMC}\left(\frac{1}{b+1}, \frac{a}{b+1}\right)$.*

(ii) *If a is of the form $a = (a, b_1, b_2, \ldots, b_\ell, b_1, b_2, \ldots b_\ell, b_1, \ldots)$, periodic of period ℓ starting from the second term and such that $a > -1$ and $b_1, b_2, \ldots, b_\ell$ are non-negative integers with at least one being non-zero, then, letting $S = b_1 + b_2 + \cdots + b_\ell$, the sequence $\frac{\ell^{\frac{-\ell}{S+\ell}}}{S+\ell} \cdot (M_n^{a})_{n\geq 1}$ has the distribution of an Intertwined Product of Generalised Gamma Processes with parameters $(a, b_1, b_2, \ldots, b_\ell)$, which we denote $\mathrm{IPGGP}(a, b_1, b_2, \ldots, b_\ell)$.*

Note that the two cases (i) and (ii) are not mutually exclusive. We prove the two points of this proposition in separate subsections. The proper definitions of the distributions to which we refer in the statement are given along the proof.

5.1.1 Mittag-Leffler Markov chains

Let us study the case where the underlying preferential attachment tree has a sequence of fitnesses $\mathbf{a}$ that is of the form $(a, b, b, b, \dots)$. We start by recalling the definitions of Mittag-Leffler distributions and Mittag-Leffler Markov chains and introduced in [21], and also studied in [26].

Mittag-Leffler distributions. Let $0 < \alpha < 1$ and $\theta > -\alpha$. The generalized Mittag-Leffler $\mathrm{ML}(\alpha, \theta)$ distribution has pth moment

$$\frac{\Gamma(\theta)\Gamma(\theta/\alpha + p)}{\Gamma(\theta/\alpha)\Gamma(\theta + p\alpha)} = \frac{\Gamma(\theta + 1)\Gamma(\theta/\alpha + p + 1)}{\Gamma(\theta/\alpha + 1)\Gamma(\theta + p\alpha + 1)} \tag{5.1}$$

and the collection of p-th moments for $p \in \mathbb{N}$ uniquely characterizes this distribution thanks to Carleman's criterion.

Mittag-Leffler Markov Chains. For any $0 < \alpha < 1$ and $\theta > -\alpha$, we introduce the (a priori) inhomogenous Markov chain $(\mathsf{M}_n^{\alpha,\theta})_{n \geq 1}$, the distribution of which we call the Mittag-Leffler Markov chain of parameters (α, θ), or $\mathrm{MLMC}(\alpha, \theta)$. This type of Markov chain was already defined in [21], for some parameters α and θ. It is a Markov chain such that for any $n \geq 1$,

$$\mathsf{M}_n^{\alpha,\theta} \sim \mathrm{ML}\left(\alpha, \theta + n - 1\right),$$

and the transition probabilities are characterised by the following equality in law:

$$\left(\mathsf{M}_n^{\alpha,\theta}, \mathsf{M}_{n+1}^{\alpha,\theta}\right) = \left(B_n \cdot \mathsf{M}_{n+1}^{\alpha,\theta}, \mathsf{M}_{n+1}^{\alpha,\theta}\right),$$

with $B_n \sim \mathrm{Beta}\left(\frac{\theta + n - 1}{\alpha} + 1, \frac{1}{\alpha} - 1\right)$, independent of $\mathsf{M}_{n+1}^{\alpha,\theta}$. These chains are constructed (for some values of θ depending on α) in [21]. Our proof of Proposition 5.1(i) ensures that these chains exist for any choice of parameters $0 < \alpha < 1$ and $\theta > -\alpha$. Let us mention that the proof of [21, Lemma 1.1] is still valid for the whole range of parameters $0 < \alpha < 1$ and $\theta > -\alpha$, which proves that these Markov chains are in fact time-homogeneous. We provide, in a later paragraph, another proof of this time-homogeneity using a more conceptual argument that relies on preferential attachment trees.

The limiting Markov chain is a Mittag-Leffler. Recall the definition of the sequence $(\beta_k)_{k \geq 1}$ and their respective distributions $\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1})$. From our assumption that $\mathbf{a} = a, b, b, b \dots$ we have for all $k \geq 1$,

$$(A_k + k, a_{k+1}) = (1 + a + (k - 1)b, b).$$

Proof of Proposition 5.1 (i). For $p \geq 1$, we can make the following computation, using (4.2), one change of indices and several times the property of the Gamma function that for any $z > 0$ we have $\Gamma(z + 1) = z\Gamma(z)$:

$$\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i^p\right] = \prod_{i=1}^{n-1} \frac{\Gamma\left(1 + a + p + (b+1)(i-1)\right)\Gamma\left(a + (b+1)i\right)}{\Gamma\left(1 + a + (b+1)(i-1)\right)\Gamma\left(a + (b+1)i + p\right)}$$

$$= \frac{\Gamma\left(1 + a + p\right)}{\Gamma\left(1 + a\right)} \cdot \frac{\Gamma\left(a + (b+1)(n-1)\right)}{\Gamma\left(a + (b+1)(n-1) + p\right)} \cdot \frac{\Gamma\left(\frac{a+p}{b+1} + n - 1\right)}{\Gamma\left(\frac{a}{b+1} + n - 1\right)} \cdot \frac{\Gamma\left(1 + \frac{a}{b+1}\right)}{\Gamma\left(1 + \frac{a+p}{b+1}\right)}.$$

$$\tag{5.2}$$

Using Stirling formula, we can then compute the numbers C_p introduced in (4.5),

$$C_p = (b+1)^{-p} \cdot \frac{\Gamma\left(1+a+p\right)\Gamma\left(1+\frac{a}{b+1}\right)}{\Gamma\left(1+a\right)\Gamma\left(1+\frac{a+p}{b+1}\right)}. \tag{5.3}$$

Using (4.12), the moments of M_k are given, for any $p \in \mathbb{N}$ by the formula:

$$\mathbb{E}\left[(\mathsf{M}_k^{\mathbf{a}})^p\right] = \frac{(b+1)^p \cdot C_p}{\prod_{i=1}^{k-1}\mathbb{E}\left[\beta_i^p\right]} \underset{(5.3),\,(5.2)}{=} \frac{\Gamma\left(\frac{a}{b+1}+k-1\right)\Gamma\left(a+(b+1)(k-1)+p\right)}{\Gamma\left(a+(b+1)(k-1)\right)\Gamma\left(\frac{a+p}{b+1}+k-1\right)}$$

These moments identify using (5.1) the distribution of $\mathsf{M}_k^{\mathbf{a}}$ for all $k \geq 1$ as

$$\mathsf{M}_k^{\mathbf{a}} \sim \mathrm{ML}\left(\frac{1}{b+1}, \frac{a}{b+1}+k-1\right).$$

From this, and the form of the backward transitions, we can identify $(\mathsf{M}_k^{\mathbf{a}})_{k\geq 1}$ as having a distribution $\mathrm{MLMC}\left(\frac{1}{b+1}, \frac{a}{b+1}\right)$. □

Time-homogeneity of MLMC. We keep the notation from the previous paragraph, working with a sequence $\mathbf{a} = a, b, b, b \ldots$, and show the time-homogeneity of the corresponding Mittag-Leffler Markov chain $(\mathsf{M}_k^{\mathbf{a}})_{k\geq 1} \sim \mathrm{MLMC}\left(\frac{1}{b+1}, \frac{a}{b+1}\right)$ using its connection with preferential attachment trees.

For any $x > -1$, consider the sequence $\boldsymbol{x} = x, b, b, b \ldots$ and $(\mathsf{P}_n^x)_{n\geq 1} \sim \mathrm{PAT}(\boldsymbol{x})$ in such a way that, using Theorem 1.4,

$$\mathsf{M}_1^{\boldsymbol{x}} = \lim_{n\to\infty} n^{-1/(b+1)} \cdot \deg_{\mathsf{P}_n^x}^+(u_1) \quad \text{and} \quad \mathsf{M}_2^{\boldsymbol{x}} = \lim_{n\to\infty} n^{-1/(b+1)} \cdot (\deg_{\mathsf{P}_n^x}^+(u_1) + \deg_{\mathsf{P}_n^x}^+(u_2)).$$

By choosing x appropriately, we can make $(\mathsf{M}_1^{\boldsymbol{x}}, \mathsf{M}_2^{\boldsymbol{x}})$ have the distribution of any of the couples $(\mathsf{M}_k^{\mathbf{a}}, \mathsf{M}_{k+1}^{\mathbf{a}})$ for $k \geq 1$. Thus, in order to prove the time-homogeneity of the transitions, it suffices to prove that the conditional distribution of $\mathsf{M}_2^{\boldsymbol{x}}$ with respect to $\mathsf{M}_1^{\boldsymbol{x}}$ does not depend on x.

Recall from Section 1.2.2 in the introduction that we see $(\mathsf{P}_n^x)_{n\geq 1}$ as an increasing sequence of plane trees, defined as subsets of $\mathbb{U}$. Also recall that for any $u \in \mathbb{U}$, we denote $T(u)$ the subtree descending from u in $\mathbb{U}$. At every time $n \geq 1$, we can consider the sequence $(\#(\mathsf{P}_n^x \cap T(1)), \#(\mathsf{P}_n^x \cap T(2)), \ldots)$, which counts the number of vertices in the subtrees descending from the children of $u_1 = \emptyset$ in P_n^x, in order of creation (completed by an sequence of zeros). We can check that this sequence evolves as n grows with the same distribution as the number of customers seating at different tables in a Chinese Restaurant Process with seating plan $(\frac{1}{b+1}, \frac{x}{b+1})$, see [45, Section 3.2] for a definition.

Then, conditionally on the evolution of this sequence, every time that a vertex is added to one of those subtrees, it is attached to any vertex already present in the subtree with probability proportional to its out-degree plus b (and in particular this does not depend on the value of x).

Thanks to [45, Corollary 3.9], two Chinese Restaurant Processes with respective seating plan $(\frac{1}{b+1}, \frac{x}{b+1})$ and $(\frac{1}{b+1}, \frac{x'}{b+1})$ with $x, x' > -1$ have a density with respect to each other and this density is a function of the scaling limit of the number of tables created in the process, which corresponds in our case to M_1^x.

These observations allow us to conclude that the distribution of $(\mathsf{P}_n^x)_{n\geq 1}$ for any $x > -1$ has a positive density with respect to $(\mathsf{P}_n^0)_{n\geq 1}$, and this density is a function of M_1^x. From here, it is clear that conditionally on $\mathsf{M}_1^{\boldsymbol{x}}$, the distribution of the quantity $(\mathsf{M}_2^{\boldsymbol{x}} - \mathsf{M}_1^{\boldsymbol{x}}) = \lim_{n\to\infty} \deg_{\mathsf{P}_n^x}^+(u_2)$ does not depend on x, which concludes the argument.

5.1.2 Products of generalised Gamma random variables.

The following paragraphs aim at proving Proposition 5.1(ii). In the first and second paragraphs, we define the families of distributions of GGP and IPGGP-processes. Some special cases of these processes already appeared in [42, 40]. In the third one we prove that the distribution of $(\mathsf{M}_k^{\mathbf{a}})_{k \geq 1}$ belongs to this family whenever the sequence $\mathbf{a}$ is of the form assumed in Proposition 5.1(ii).

Construction of a $\mathrm{GPP}(z,r)$**-process.** For $z, r > 0$ real numbers, let $(Z_i)_{i \geq 1}$ be a family of independent variables with the following distribution:

$$Z_1 \sim \mathrm{Gamma}\left(\frac{z}{r}\right) \quad \text{and for } i \geq 2, \quad Z_i \sim \mathrm{Gamma}(1) = \mathrm{Exp}(1),$$

where, for any $k > 0$, the distribution $\mathrm{Gamma}(u)$ has density $x \mapsto \frac{x^{u-1}e^{-x}}{\Gamma(u)}\mathbf{1}_{\{x>0\}}$ with respect to the Lebesgue measure. Then for all $k \geq 1$ we define G_k as,

$$\mathsf{G}_k := \left(\sum_{i=1}^{k} Z_i\right)^{\frac{1}{r}}.$$

We say that the process $(\mathsf{G}_k)_{k \geq 1}$ has the distribution of a *Generalised Gamma process* with parameters (z,r) which we denote $\mathrm{GPP}(z,r)$.

Note that, using standard distributional equalities with Gamma and Beta distributions, for every $k \geq 1$, we have $(\mathsf{G}_k)^r \sim \mathrm{Gamma}\left(k - 1 + \frac{z}{r}\right)$ and

$$V_k := \left(\frac{\mathsf{G}_k}{\mathsf{G}_{k+1}}\right)^r \sim \mathrm{Beta}\left(k - 1 + \frac{z}{r}, 1\right), \quad \text{so that} \quad V_k^{1/r} = \frac{\mathsf{G}_k}{\mathsf{G}_{k+1}} \sim \mathrm{Beta}\left(r(k-1) + z, 1\right), \tag{5.4}$$

and $V_k^{1/r}$ is independent of G_{k+1}. In fact, we can further show that $V_1, V_2, \ldots, V_k, \mathsf{G}_{k+1}$ are jointly independent with the corresponding distribution and that this characterizes the finite dimensional marginals of this process.

Remark 5.2. For $z = r$, the process $(\mathsf{G}_k)_{k \geq 1}$ has exactly the distribution of the points of a Poisson process on $(0, \infty)$ with intensity $r \cdot t^{r-1}\mathrm{d}t$, listed in increasing order.

Intertwined Products of GGP**-processes.** Let a be a real number with $a > -1$ and $b_1, b_2, \ldots, b_\ell$ be non-negative integers where at least one is non-zero. We let $B_r := \sum_{s=1}^{r} b_s$ for all $0 \leq r \leq \ell$, with the convention that $B_0 := 0$. We also let $S = B_\ell$. Then we define the set

$$\mathcal{S} := \{1, 2, \ldots, S + \ell - 1\} \setminus \{B_r + r \mid 1 \leq r \leq \ell - 1\}$$

$$= \bigcup_{r=1}^{\ell} \{r + B_{r-1}, r + B_{r-1} + 1, \ldots, r + B_{r-1} + b_r - 1\}$$

Start with independent GPP processes $\left\{\mathsf{G}^{(q)} \mid q \in \mathcal{S}\right\}$ indexed by $\mathcal{S}$ such that for all $q \in \mathcal{S}$,

$$\mathsf{G}^{(q)} \sim \mathrm{GPP}(a + q, \ell + S).$$

Now $\mathsf{G} = (\mathsf{G}_k)_{k \geq 1}$ is defined in such a way that for all $n \geq 1$ and $1 \leq r \leq \ell$ we have

$$\mathsf{G}_{\ell \cdot (n-1)+r} = \left(\prod_{\substack{q \in \mathcal{S} \\ 1 \leq q \leq r-1+B_{r-1}}} \mathsf{G}_{n+1}^{(q)}\right) \cdot \left(\prod_{\substack{q \in \mathcal{S} \\ r+B_{r-1} \leq q \leq S+\ell-1}} \mathsf{G}_n^{(q)}\right). \tag{5.5}$$

The process $(\mathsf{G}_k)_{k\geq 1}$ defined above is said to have distribution of an *Intertwined Product of Generalized Gamma Processes* with parameters $(a, b_1, b_2, \ldots, b_\ell)$, which we denote $\mathrm{IPGGP}(a, b_1, b_2, \ldots, b_\ell)$. Its finite dimensional marginals can be obtained in the same way as it was done in the preceding paragraph for Generalized Gamma processes.

Identification of the limiting chain. Fix $\ell \geq 1$ and $b_1, b_2, \ldots, b_\ell \geq 0$ some integers (where at least one is non-zero) and suppose that the sequence $\mathbf{a}$ has the following form,

$$\mathbf{a} = (a, b_1, b_2, \ldots, b_\ell, b_1, b_2, \ldots b_\ell, b_1, \ldots)$$

meaning that the sequence is periodic with period ℓ starting from the second term, with $a > -1$.

For any $j \geq 0$ and $1 \leq r \leq \ell$ we have

$$\beta_{j\ell+r} \sim \mathrm{Beta}\left(a + B_{r-1} + r + j(\ell + S), b_r\right),$$

for the $(\beta_k)_{k\geq 1}$ as defined in Theorem 1.1. We fix $1 \leq r \leq \ell$. For any $j \geq 1, p \geq 1$, we use the moments (4.2) of a Beta random variable and a telescoping argument to write

$$
\begin{aligned}
\mathbb{E}\left[(\beta_{j\ell+r})^p\right] &= \prod_{q=0}^{p-1} \frac{a + B_{r-1} + r + j(\ell + S) + q}{a + B_{r-1} + r + j(\ell + S) + q + b_r} \\
&= \prod_{q=0}^{b_r-1} \frac{a + B_{r-1} + r + j(\ell + S) + q}{a + B_{r-1} + r + j(\ell + S) + q + p}.
\end{aligned}
$$

Using the last display, we get that for any $n \geq 1$,

$$
\prod_{j=0}^{n} \mathbb{E}\left[(\beta_{j\ell+r})^p\right] = \prod_{q=0}^{b_r-1} \frac{\Gamma\left(\frac{a+B_{r-1}+r+q}{\ell+S} + n + 1\right)\Gamma\left(\frac{a+B_{r-1}+r+q+p}{\ell+S}\right)}{\Gamma\left(\frac{a+B_{r-1}+r+q}{\ell+S}\right)\Gamma\left(\frac{a+B_{r-1}+r+q+p}{\ell+S} + n + 1\right)}.
$$

Using Stirling's approximation we get that for a fixed $1 \leq r \leq \ell$,

$$
\prod_{j=0}^{n} \mathbb{E}\left[(\beta_{j\ell+r})^p\right] \underset{n\to\infty}{\sim} n^{-\frac{p\cdot b_r}{\ell+S}} \cdot \prod_{q=0}^{b_r-1} \frac{\Gamma\left(\frac{a+B_{r-1}+r+q+p}{\ell+S}\right)}{\Gamma\left(\frac{a+B_{r-1}+r+q}{\ell+S}\right)}.
$$

Hence, recalling the definition of C_p in (4.5) and accounting for all the possible values of r we get

$$
C_p = \ell^{\frac{pS}{\ell+S}} \cdot \prod_{r=1}^{\ell} \prod_{q=0}^{b_r-1} \frac{\Gamma\left(\frac{a+B_{r-1}+r+q+p}{\ell+S}\right)}{\Gamma\left(\frac{a+B_{r-1}+r+q}{\ell+S}\right)} = \ell^{\frac{pS}{\ell+S}} \cdot \prod_{i\in\mathcal{S}} \frac{\Gamma\left(\frac{a+i+p}{\ell+S}\right)}{\Gamma\left(\frac{a+i}{\ell+S}\right)}.
$$

Then using (4.12) with $c = S/\ell$,

$$
\begin{aligned}
&\mathbb{E}\left[(\mathsf{M}^{\mathbf{a}}_{\ell\cdot(n-1)+r})^p\right] \\
&= \frac{(c+1)^p \cdot C_p}{\prod_{i=1}^{\ell\cdot(n-1)+r-1} \mathbb{E}[\beta_i^p]} \\
&= \left(\frac{S+\ell}{\ell}\right)^p \cdot \ell^{\frac{pS}{S+\ell}} \cdot \prod_{t=1}^{r-1} \prod_{j=0}^{b_t-1} \frac{\Gamma\left(n + \frac{a+B_{t-1}+t+j+p}{\ell+S}\right)}{\Gamma\left(n + \frac{a+B_{t-1}+t+j}{\ell+S}\right)} \cdot \prod_{t=r}^{\ell} \prod_{j=0}^{b_t-1} \frac{\Gamma\left(n-1 + \frac{a+B_{t-1}+t+j+p}{\ell+S}\right)}{\Gamma\left(n-1 + \frac{a+B_{t-1}+t+j}{\ell+S}\right)} \\
&= \left((S+\ell)\cdot(\ell^{\frac{-\ell}{S+\ell}})\right)^p \cdot \prod_{\substack{q\in\mathcal{S} \\ 1\leq q\leq r-1+B_{r-1}}} \frac{\Gamma\left(n + \frac{a+q+p}{\ell+S}\right)}{\Gamma\left(n + \frac{a+q}{\ell+S}\right)} \cdot \prod_{\substack{q\in\mathcal{S} \\ r+B_{r-1}\leq q\leq S+\ell-1}} \frac{\Gamma\left(n-1 + \frac{a+q+p}{\ell+S}\right)}{\Gamma\left(n-1 + \frac{a+q}{\ell+S}\right)}.
\end{aligned}
$$

Using the last display and the fact that random variable with distribution $\mathrm{Gamma}(u)$ has p-th moment equal to $\frac{\Gamma(u+p)}{\Gamma(u)}$, we can identify the distribution of the one-dimensional marginals $\ell^{\frac{\ell}{S+\ell}} \cdot \mathsf{M}_k^{\mathbf{a}}$ for any $k \geq 1$ with the ones of the process described in (5.5). The identification of the distribution of the process $\ell^{\frac{\ell}{S+\ell}} \cdot (\mathsf{M}_k^{\mathbf{a}})_{k \geq 1}$ as $\mathrm{IPGGP}(a, b_1, b_2, \ldots, b_\ell)$ is then obtained by comparing their finite dimensional distribution which are characterized by Proposition 4.2 and, respectively, (5.5) together with the discussion below (5.4).

Sparse sequences. Let us treat a particular example of parameters $a, b_1, b_2, \ldots b_\ell$ for which the distribution $\mathrm{IPGGP}(a, b_1, b_2, \ldots, b_\ell)$ has a simpler description than the general case. Suppose that only one of the parameters $b_1, b_2, \ldots, b_\ell$ is non-zero, say b_ℓ for example. Keeping the notation introduced above, the corresponding set $\mathcal{S}$ contains only b_ℓ elements $\mathcal{S} = \{\ell, \ell+1, \ldots, \ell + b_\ell - 1\}$. Following the definition (5.5), the process $(\mathsf{G}_k)_{k \geq 1}$ with distribution $\mathrm{IPGGP}(a, 0, 0, \ldots, 0, b_\ell)$ is constant on every interval $[\![(n-1)\ell + 1, n\ell]\!]$ for any integer $n \geq 1$ and, the process $(\mathsf{G}_{(n-1)\ell+1})_{n \geq 1}$ is just given by a product of b_ℓ independent GGP-processes

$$\mathsf{G}_{(k-1)\ell+1} = \prod_{\ell \leq q \leq \ell + b_\ell - 1} \mathsf{G}_k^{(q)},$$

where for all $\ell \leq q \leq \ell + b_\ell - 1$, the process $(\mathsf{G}_k^{(q)})_{k \geq 1}$ has distribution $\mathrm{GPP}(a + q, \ell + b_\ell)$.

In the particular case where $a = 1$ and $(b_1, b_2, \ldots, b_{\ell-1}, b_\ell) = (0, 0, \ldots, 0, 1)$, the picture is even simpler because the last display becomes a product over only one term. We can check using Remark 5.2 that the process $(\mathsf{G}_{(k-1)\ell+1})_{k \geq 1}$ has then exactly the distribution of the points of a Poisson process on $(0, \infty)$ with intensity $(\ell+1)t^\ell \mathrm{d}t$, listed in increasing order, which was already noted in [40, Remark 2].

5.2 Application to Pólya urns with immigration

Define the following generalisation of Pólya's urn, which depends on a sequence of numbers $(a_n)_{n \geq 1}$: start at time 1 with an urn containing a_1 red balls. At every time $n \geq 2$, we sample a ball uniformly at random from the urn, return it to the urn with 1 additional ball of the same colour, plus an immigration of a_n additional black balls. The outcome of the first step being deterministic, it is equivalent to consider that we start at time 2 with $a_1 + 1$ red balls and a_2 white balls in the urn, so that we allow ourselves to consider any (possibly negative) value $a_1 > -1$. This model was studied in the sequence of paper [41, 42, 40] in specific cases of periodic immigration and also studied in [1] with a larger class of periodic immigration.

We denote R_n the number of red balls in the urn at time n and state a scaling limit result for R_n when $n \to \infty$. We also identify the speed of convergence and the Gaussian fluctuations around the limit, provided that the immigration is sufficiently regular.

Recall from the introduction the assumption (H_c) defined for a real number $c > 0$. We introduce the following more precise assumption of the same type, for any $c > 0$ and $\delta > 0$.

$$A_n = c \cdot n \cdot \left(1 + O\left(n^{-\delta}\right)\right). \tag{H_c^δ}$$

Remark that for any $0 < \delta < \frac{1}{2}$, this assumption is satisfied for periodic sequences $\mathbf{a}$, and almost surely satisfied by sequences of i.i.d. non-negative random variables with a second moment.

Proposition 5.3. *Assume that the sequence $\mathbf{a} = (a_n)_{n \geq 1}$ satisfies (H_c) for some $c > 0$. Then for $D_n := n^{-\frac{1}{c+1}} \cdot R_n$ the following holds:*

(i) *We have the following almost sure convergence,*

$$D_n \underset{n \to \infty}{\longrightarrow} D_\infty,$$

where D_∞ has the same law as $\mathsf{M}_1^{\mathbf{a}}$, defined in (1.5).

(ii) *If we additionally assume that $\mathbf{a}$ satisfies (H_c^δ) for some $\delta > \frac{1}{2(c+1)}$ then we have*

$$n^{\frac{1}{2(c+1)}} \cdot \frac{D_\infty - D_n}{\sqrt{D_n}} \underset{n \to \infty}{\overset{(d)}{\longrightarrow}} \mathcal{N}(0,1).$$

Remark 5.4. If the sequence $\mathbf{a}$ has one of the particular forms treated in Proposition 5.1 of the previous section, we can identify the distribution of the limiting random variable as being Mittag-Leffler or a product of independent generalized Gamma random variables. This gives us an alternative proof for the similar statement [1, Theorem 3.8].

Proof. Let $(\mathsf{P}_n)_{n \geq 1}$ be a sequence of trees with distribution $\mathrm{PAT}(\mathbf{a})$ and consider the sequence $(a_1 + \deg_{\mathsf{P}_n}^+(u_1))_{n \geq 1}$. It is easy to see that this sequence has exactly the same distribution as the sequence $(R_n)_{n \geq 1}$ of numbers of red balls in a Pólya urn with immigration sequence $\mathbf{a}$, hence it suffices to study the former quantity.

Since the sequence $\mathbf{a}$ satisfies our assumption (H_c) for some $c > 0$ then using (1.6) we can write the following almost sure convergence

$$n^{-\frac{1}{c+1}} \cdot \deg_{\mathsf{P}_n}^+(u_1) \underset{n \to \infty}{\to} \mathsf{M}_1^{\mathbf{a}},$$

where the sequence $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ is defined in (1.5), so this proves (i).

We turn to the proof of (ii). We prove this convergence in two steps, by first proving some corresponding result for the degree of the first vertex in a WRT, and then using Theorem 1.1 and Proposition 1.3 to transfer the result to the corresponding PAT distribution.

Let $(\mathsf{T}_n)_{n \geq 1}$ be a sequence of trees with distribution $\mathrm{WRT}(\boldsymbol{w})$ with a sequence $\boldsymbol{w}$ satisfying the following assumption

$$W_n \underset{n \to \infty}{=} \frac{n^\gamma}{1-\gamma} \cdot \left(1 + o\left(n^{-\frac{1-\gamma}{2}}\right)\right), \tag{5.6}$$

for some $\gamma \in (0,1)$. In this context, recalling (2.2), the degree of the first vertex can be written as

$$\deg_{\mathsf{T}_n}^+(u_1) = \sum_{i=2}^n \mathbf{1}_{\left\{U_i \leq \frac{w_1}{W_n}\right\}}$$

$$= w_1 \cdot n^{1-\gamma} + w_1 \cdot \left(\sum_{i=2}^n \frac{1}{W_n} - n^{1-\gamma}\right) + \sum_{i=2}^n \left(\mathbf{1}_{\left\{U_i \leq \frac{w_1}{W_n}\right\}} - \frac{w_1}{W_n}\right).$$

Now, using our assumption on the sequence $(W_n)_{n \geq 1}$ we get $\frac{1}{W_n} = \frac{1-\gamma}{n^\gamma} + o\left(n^{-\frac{\gamma}{2}-\frac{1}{2}}\right)$, so that

$$\left(\sum_{i=2}^n \frac{1}{W_n} - n^{1-\gamma}\right) = o\left(n^{\frac{1-\gamma}{2}}\right).$$

Rearranging the terms, we get

$$n^{\frac{1-\gamma}{2}} \cdot \left(n^{-(1-\gamma)} \cdot \deg_{\mathsf{T}_n}^+(u_1) - w_1\right) = n^{-\frac{1-\gamma}{2}} \cdot \sum_{i=2}^n \left(\mathbf{1}_{\left\{U_i \leq \frac{w_1}{W_n}\right\}} - \frac{w_1}{W_n}\right) + o(1),$$

and using the Lindeberg-Feller theorem (see [16, Theorem 3.4.5] for example), we get that the expression on the RHS of the last display converges in distribution when $n \to \infty$ to a Gaussian distribution $\mathcal{N}(0, w_1)$. Recalling that $n^{-(1-\gamma)} \cdot \deg_{\mathrm{T}_n}^+(u_1) \to w_1$ a.s. as $n \to \infty$, we can also write using Slutsky's lemma

$$n^{\frac{1-\gamma}{2}} \cdot \frac{(n^{-(1-\gamma)} \cdot \deg_{\mathrm{T}_n}^+(u_1) - w_1)}{\sqrt{n^{-(1-\gamma)} \cdot \deg_{\mathrm{T}_n}^+(u_1)}} \xrightarrow[n \to \infty]{(\mathrm{d})} \mathcal{N}(0, 1). \tag{5.7}$$

Now we transfer this result to the case of preferential attachment trees. For this, thanks to Theorem 1.1 and Proposition 1.3, it suffices to prove that a satisfies the condition (H_c^δ) with $\delta > \frac{1}{2(c+1)}$ then the corresponding sequence $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ defined (1.5) almost surely satisfies (5.6) for $\gamma = \frac{c}{c+1}$. From Proposition 1.3 and the definition of $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ as a scaled version of $(\mathsf{W}_n^{\mathbf{a}})_{n \geq 1}$, we already know that we have $\mathsf{M}_n^{\mathbf{a}} \underset{n \to \infty}{=} \frac{1}{1-\gamma} \cdot n^\gamma \cdot (1 + O(n^{-\epsilon}))$ almost surely, for $\gamma = \frac{c}{c+1}$ and *some* $\epsilon > 0$. In our case, this is not enough, because we need something quantitative on ϵ. Going along the proof of Proposition 1.3 again, we get from (4.9) that we can in fact write

$$\mathsf{M}_n^{\mathbf{a}} \underset{n \to \infty}{=} \frac{1}{1 - \gamma} \cdot n^\gamma \cdot \left(1 + O\left(n^{-\zeta}\right)\right)$$

for any $\zeta < \delta \wedge \frac{1}{2}$, so that (5.6) is almost surely satisfied by $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ if

$$\delta > \frac{1 - \gamma}{2} = \frac{1}{2(c+1)},$$

which is what we assumed. Now, thanks to Theorem 1.1, conditionally on the sequence $(\mathsf{M}_n^{\mathbf{a}})_{n \geq 1}$ the distribution of $(\mathsf{P}_n)_{n \geq 1}$ is $\mathrm{WRT}((\mathsf{m}_n^{\mathbf{a}})_{n \geq 1})$. Applying (5.7) in this case finishes to prove (ii). □

5.3 Applications to some other models of preferential attachment

We present here another model of preferential attachment which appears in the literature, for example in [40]. This model does not produce a tree as ours does, but we can couple them in such a way that some of their features coincide. We only focus on one particular model of graph here but the method presented here can adapt to other similar models.

A model of (m, α)-preferential attachment Let S be a non-empty graph, with vertex-set $\{v_1^{(1)}, \ldots, v_1^{(k)}\}$ which have degrees $(d_1, \ldots d_k)$, and $m \geq 2$ an integer and $\alpha > -m$ a real number such that $\alpha + d_i > 0$ for all $1 \leq i \leq k$. The model is then defined as follows: we let $\mathsf{G}_1 = \mathsf{S}$. Then, at any time $n \geq 1$, the graph G_{n+1} is constructed from the graph G_n by:

- adding a new vertex labelled v_{n+1} with m outgoing edges,

- choosing sequentially to which other vertex each of these edges are pointed, each vertex being chosen with probability proportional to α plus its degree (the degree of the vertices are updated after each edge-creation).

The degree of a vertex in a graph refers in this section to the number of edges incident to it. Here the growth procedure in fact produces *multigraphs*, in which it is possible for two vertices to be connected to each other by more than one edge. In this case, all those edges contribute in the count of their degree.

We can couple this model to a preferential attachment tree with sequence of fitnesses **a** defined as:

$$\mathbf{a} = (w(\mathtt{S}), \underbrace{0, 0, \ldots, 0}_{m-1}, m + \alpha, \underbrace{0, 0, \ldots, 0}_{m-1}, m + \alpha, 0, 0 \ldots),$$

where $w(\mathtt{S}) := d_1 + d_2 + \cdots + d_k + k\alpha$.

Indeed, we first construct $(\mathtt{P}_n)$ with distribution $\mathrm{PAT}(\mathbf{a})$. Then, for any $n \geq 1$, consider the tree $\mathtt{P}_{1+m(n-1)}$ and for all $2 \leq i \leq n$, merge together each vertex with fitness $m + \alpha$ together with the $m - 1$ vertices with fitness 0 that arrived immediately before it. If $\mathtt{G}_1$ only contains one vertex, it is immediate that the obtained sequence of graphs has exactly the same distribution as $(\mathtt{G}_n)_{n\geq 1}$. For general seed graphs $\mathtt{S}$, we can still use the same construction and the obtained sequence of graphs has the same evolution as some sequence $(\widetilde{\mathtt{G}}_n)_{n\geq 1}$ which would be obtained from $(\mathtt{G}_n)_{n\geq 1}$ by merging all the vertices $\{v_1^{(1)}, \ldots, v_1^{(k)}\}$ into a unique vertex v_1.

Note that a similar construction would also be possible if the degrees of the vertices $v_2, v_3, \ldots$ were given by a sequence of integers $(m_2, m_3, \ldots)$ instead of all being equal to some constant value m. This is for example the case in the model studied in [13], where the degrees are random.

We have the following convergence for degrees of vertices in the graph, as $n \to \infty$. The distribution $\mathrm{Dir}(x_1, x_2, \ldots, x_k)$ appearing in the statement is a Dirichlet distribution with parameters $x_1, x_2, \ldots, x_k$ on the k-dimensional simplex.

Proposition 5.5. *The following convergence holds almost surely in any ℓ^p with $p > 2 + \frac{\alpha}{m}$:*

$$n^{-\frac{1}{2+\alpha/m}} (\deg_{\mathtt{G}_n}(v_1^{(1)}), \deg_{\mathtt{G}_n}(v_1^{(2)}), \ldots, \deg_{\mathtt{G}_n}(v_1^{(k)}), \deg_{\mathtt{G}_n}(v_2), \deg_{\mathtt{G}_n}(v_3), \ldots)$$
$$\xrightarrow[n\to\infty]{} (\mathsf{N}_1 \cdot B^{(1)}, \mathsf{N}_1 \cdot B^{(2)}, \ldots \mathsf{N}_1 \cdot B^{(k)}, \mathsf{N}_2 - \mathsf{N}_1, \mathsf{N}_3 - \mathsf{N}_2, \ldots),$$

where

$$(B^{(1)}, B^{(2)}, \ldots B^{(k)}) \sim \mathrm{Dir}(d_1 + \alpha, d_2 + \alpha, \ldots, d_k + \alpha),$$

and the process $(\mathsf{N}_n)_{n\geq 1}$ is independent of $(B^{(1)}, B^{(2)}, \ldots, B^{(k)})$.

Furthermore, whenever $\alpha \in \mathbb{Z}$ with $\alpha > -m$ or $m = 1$ then the distribution of $(\mathsf{N}_n)_{n\geq 1}$ is explicit and given by:

- *if $\alpha \in \mathbb{Z}$ with $\alpha > -m$, then*

$$\frac{m^{\frac{-2m}{2m+\alpha}}}{2m+\alpha} \cdot (\mathsf{N}_n)_{n\geq 1} \sim \mathrm{IPGGP}(w(\mathtt{S}), \underbrace{0, 0, \ldots, 0}_{m-1}, m + \alpha)$$

- *if $m = 1$, then*

$$(\mathsf{N}_n)_{n\geq 1} \sim \mathrm{MLMC}\left(\frac{1}{2+\alpha}, \frac{w(\mathtt{S})}{2+\alpha}\right),$$

where those distributions are the ones that appear in Proposition 5.1.

This result strengthens the one of [40, Theorem 1, Theorem 2 and Proposition 1] which corresponds (up to some definition convention) to the case $\alpha = 1 - m$. We emphasize that the convergence here is almost sure in an ℓ^p space.

Proof of Proposition 5.5. Using the coupling argument, we know that we can construct jointly the sequence of graphs $(\mathtt{G}_n)_{n\geq 1}$ and a sequence of trees $(\mathtt{P}_n)_{n\geq 1} \sim \mathrm{PAT}(\mathbf{a})$ with fitness sequence

$$\mathbf{a} = (w(S), \underbrace{0, 0, \ldots, 0}_{m-1}, m + \alpha, \underbrace{0, 0, \ldots, 0}_{m-1}, m + \alpha, \ldots),$$

in such a way that for every $n \geq 1$, the sequence

$$\left(\sum_{j=1}^{k} \left(\deg_{\mathsf{G}_n}(v_1^{(j)}) - d_j \right), (\deg_{\mathsf{G}_n}(v_2) - m), \ldots, (\deg_{\mathsf{G}_n}(v_n) - m), 0, 0 \ldots \right)$$

coincides with

$$(\deg^+_{\mathsf{P}_{1+(n-1)m}}(u_1), \deg^+_{\mathsf{P}_{1+(n-1)m}}(u_{1+m}), \deg^+_{\mathsf{P}_{1+(n-1)m}}(u_{1+2m}), \ldots).$$

Using this connection, together with Theorem 1.1, Proposition 1.3 and Proposition 2.1 we get

$$n^{-\frac{1}{2+\alpha/m}} (\deg_{\mathsf{G}_n}(v_1) + \deg_{\mathsf{G}_n}(v_2) + \cdots + \deg_{\mathsf{G}_n}(v_k), \deg_{\mathsf{G}_n}(u_2), \deg_{\mathsf{G}_n}(u_3), \ldots)$$
$$\xrightarrow[n \to \infty]{} (\mathsf{N}_1, \mathsf{N}_2 - \mathsf{N}_1, \mathsf{N}_3 - \mathsf{N}_2, \ldots),$$

almost surely in ℓ^p for all $p > 2 + \frac{\alpha}{m}$, for some random sequence $(\mathsf{N}_n)_{n \geq 1}$. Note that the time-change between $(\mathsf{G}_n)_{n \geq 1}$ and $(\mathsf{P}_n)_{n \geq 1}$ is responsible for an extra factor in the scaling, so that the sequence $(\mathsf{N}_n)_{n \geq 1}$ has the distribution of $m^{\frac{m}{2m+\alpha}} \cdot (\mathsf{M}_n^{\mathsf{a}})_{n \geq 1}$. In the case $\alpha \in \mathbb{Z}$ or $m = 1$, Proposition 5.1 identifies the distribution of the limiting sequence.

Last, the convergence of $\frac{1}{\sum_{j=1}^{k} \deg_{\mathsf{G}_n}(v_1^{(j)})} \cdot (\deg_{\mathsf{G}_n}(v_1^{(1)}), \deg_{\mathsf{G}_n}(v_1^{(2)}), \ldots, \deg_{\mathsf{G}_n}(v_1^{(k)}))$ just follows from the classical result of convergence for the proportion of balls in a Pólya urn. $\qquad \square$

A Technical proofs and results

This appendix contains some technical results that are used throughout this paper. Let start by stating a useful conditional version of the Borel-Cantelli lemma.

Lemma A.1. *Let $(\mathcal{F}_n)_{n \geq 1}$ be a filtration and let $(B_n)_{n \geq 1}$ be a sequence of events adapted to this filtration. For all $n \geq 1$, let $\mathsf{p}_n := \mathbb{P}(B_n \mid \mathcal{F}_{n-1})$. We have*

$$\frac{\sum_{i=1}^{n} \mathbf{1}_{B_i}}{\sum_{i=1}^{n} \mathsf{p}_i} \xrightarrow[n \to \infty]{} 1 \qquad \textit{a.s. on the event} \quad \left\{ \sum_{i=1}^{\infty} \mathsf{p}_i = \infty \right\}$$

and also

$$\sum_{i=1}^{n} \mathbf{1}_{B_i} \quad \textit{converges a.s. on the event} \quad \left\{ \sum_{i=1}^{\infty} \mathsf{p}_i < \infty \right\}.$$

Proof. The first convergence is the content of [16, Theorem 5.4.11] and the second one is an application of [16, Theorem 5.4.9] to the martingale $(\sum_{i=1}^{n}(\mathbf{1}_{B_i} - \mathsf{p}_i))_{n \geq 1}$. $\qquad \square$

The following lemma is a rewriting of [5, Lemma 1]. We provide the proof for completeness.

Lemma A.2. *Let $(M_n)_{n \geq 1}$ be a complex-valued martingale with finite q-th moment for some $q \in [1, 2]$. Then for every $n \geq 1$ we have*

$$\mathbb{E}\left[|M_{n+1}|^q\right] \leq \mathbb{E}\left[|M_n|^q\right] + 2^q \cdot \mathbb{E}\left[|M_{n+1} - M_n|^q\right].$$

Proof. Let $X_{n+1} := M_{n+1} - M_n$ and let X'_{n+1} be a random variable such that conditionally on $(M_1, \ldots, M_n)$ the random variable X'_{n+1} is independent of, and has the same

distribution as X_{n+1}. Then

$$
\begin{aligned}
\mathbb{E}\left[|M_{n+1}|^q\right] &= \mathbb{E}\left[\left|\mathbb{E}\left[M_{n+1} - X'_{n+1} \mid M_1, \ldots M_{n+1}\right]\right|^q\right] \\
&\leq \mathbb{E}\left[\left|M_{n+1} - X'_{n+1}\right|^q\right] \\
&= \mathbb{E}\left[\left|M_n + X_{n+1} - X'_{n+1}\right|^q\right] \\
&\leq \mathbb{E}\left[|M_n|^q\right] + \mathbb{E}\left[\left|X_{n+1} - X'_{n+1}\right|^q\right] \\
&\leq \mathbb{E}\left[|M_n|^q\right] + 2^q \cdot \mathbb{E}\left[|X_{n+1}|^q\right],
\end{aligned}
$$

where the first equality comes from the fact that $\mathbb{E}\left[X'_{n+1} \mid M_1, \ldots M_{n+1}\right] = 0$. The first inequality is the one of Jensen for conditional expectation, applied to the convex function $z \mapsto |z|^q$. The second inequality is due to Clarkson, see [55, Lemma 1], and can be applied because the distribution of $X_{n+1} - X'_{n+1}$ conditional on M_n is symmetric and $1 \leq q \leq 2$. The last inequality comes from the triangle inequality for the L^q-norm. $\qquad\square$

Let us state another result about martingales, which we use numerous times throughout the paper. Recall our uniform big-O and small-o notation, introduced in (3.4).

Lemma A.3. *Suppose that $(z \mapsto Z_n(z))_{n\geq 1}$ is a sequence of analytic functions on some open set $\mathscr{O} \subset \mathbb{C}$, adapted to some filtration $(\mathcal{G}_n)$. Suppose that for every $z \in \mathscr{O}$, the sequence $(Z_n(z))_{n\geq 1}$ is a martingale with respect to the filtration $(\mathcal{G}_n)$. If there exists a parameters $q > 1$ and continuous functions $\alpha : \mathscr{O} \to \mathbb{R}$ and $\delta : \mathscr{O} \to (0\,,\infty)$ such that for all $n \geq 1$ we have*

$$
\mathbb{E}\left[|Z_{2n}(z) - Z_n(z)|^q\right] = O_{\mathscr{O}}\left(n^{\alpha(z)q - \delta(z) + o_{\mathscr{O}}(1)}\right),
$$

then for any compact subset $K \subset \mathscr{O}$, there exists $\epsilon(K) > 0$ such that

(i) *if $\alpha > 0$ on $\mathscr{O}$ we have $n^{-\alpha(z)} \cdot |Z_n(z) - Z_1(z)| = O_K\left(n^{-\epsilon(K)}\right)$ almost surely and also in expectation,*

(ii) *if $\alpha \leq 0$ on $\mathscr{O}$, the almost sure limit $Z_\infty(z)$ exists for $z \in \mathscr{O}$ and we have $n^{-\alpha(z)} \cdot |Z_n(z) - Z_\infty(z)| = O_K\left(n^{-\epsilon(K)}\right)$ almost surely and also in expectation.*

Proof of Lemma A.3. First, without loss of generality, we can consider that the term $o_{\mathscr{O}}(1)$ is identically equal to 0, otherwise we just replace the function $z \mapsto \delta(z)$ by $z \mapsto \frac{1}{2} \cdot \delta(z)$. Second, by compactness, it is sufficient to prove the result for a small disk around each $x \in K$. Since $\mathscr{O}$ is an open set, let $\rho > 0$ be such that $D(x, 2\rho) \subset \mathscr{O}$, where $D(x, 2\rho)$ is the closed disk in the complex plane with centre x and radius 2ρ. We denote

$$
\underline{\alpha} = \inf_{D(x,2\rho)} \alpha, \quad \overline{\alpha} = \sup_{D(x,2\rho)} \alpha, \quad \underline{\delta} = \inf_{D(x,2\rho)} \delta,
$$

and choose ρ small enough so that $\underline{\alpha} - \overline{\alpha} + \frac{1}{q}\underline{\delta} > 0$. Then if we let $\xi : [0\,, 2\pi] \to \mathbb{C}$ such that $\xi(t) = x + 2\rho e^{it}$, we have for any n and m, using the Cauchy formula

$$
\sup_{z \in D(x,\rho)} |Z_n(z) - Z_m(z)| \leq \pi^{-1} \int_0^{2\pi} |Z_n(\xi(t)) - Z_m(\xi(t))| \mathrm{d}t.
$$

Now,

$$
\begin{aligned}
\sup_{2^s \leq n \leq 2^{s+1}} \sup_{z \in D(x,\rho)} |Z_n(z) - Z_{2^s}(z)| &\leq \pi^{-1} \sup_{2^s \leq n \leq 2^{s+1}} \int_0^{2\pi} |Z_n(\xi(t)) - Z_{2^s}(\xi(t))| \mathrm{d}t \\
&\leq \pi^{-1} \int_0^{2\pi} \sup_{2^s \leq n \leq 2^{s+1}} |Z_n(\xi(t)) - Z_{2^s}(\xi(t))| \mathrm{d}t. \quad \text{(A.1)}
\end{aligned}
$$

Using sequentially Jensen's inequality, Doob's maximal inequality in L^q and the assumption of the lemma gives us, for $z \in \mathrm{D}(x, 2\rho)$,

$$
\begin{aligned}
\mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} |Z_n(z) - Z_{2^s}(z)|\right] &\leq \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} |Z_n(z) - Z_{2^s}(z)|^q\right]^{\frac{1}{q}} \\
&\leq \frac{q}{q-1} \cdot \mathbb{E}\left[|Z_{2^{s+1}}(z) - Z_{2^s}(z)|^q\right]^{\frac{1}{q}} \\
&\underset{s\to\infty}{=} O_{\mathrm{D}(x,2\rho)}\left(2^{\left(\alpha(z) - \frac{1}{q} \cdot \delta(z)\right)s}\right).
\end{aligned}
\tag{A.2}
$$

So using (A.1), Fubini's theorem and (A.2), we get

$$
\begin{aligned}
\mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_{2^s}(z)|\right] &\leq \pi^{-1} \int_0^{2\pi} \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} |Z_n(\xi(t)) - Z_{2^s}(\xi(t))|\right] \mathrm{d}t \\
&\underset{s\to\infty}{=} O\left(2^{\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)s}\right).
\end{aligned}
\tag{A.3}
$$

Now we treat the two cases $\alpha > 0$ and $\alpha \leq 0$ separately. Remark that the quantity $\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)$ is negative when $\alpha \leq 0$, but can be of any sign in the case $\alpha > 0$.

    • For $\alpha > 0$ and $r \geq 0$, we let

$$
A_r := 2^{-\underline{\alpha} r} \cdot \sum_{s=0}^{r} \sup_{2^s \leq n \leq 2^{s+1}} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_{2^s}(z)|.
$$

Using (A.3), we have

$$
\mathbb{E}\left[A_r\right] \leq \mathrm{cst} \cdot 2^{-\underline{\alpha} r} \cdot \sum_{s=0}^{r} 2^{\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)s} \leq \mathrm{cst} \cdot 2^{-\left(\underline{\alpha} - 0 \vee \left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)\right)r}.
$$

Thanks to our assumptions, the number $\beta := \underline{\alpha} - 0 \vee \left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)$ is positive. Using Markov's inequality and the last display yields

$$
\mathbb{P}\left(A_r > 2^{-\frac{\beta}{2}r}\right) \leq 2^{\frac{\beta}{2}r} \cdot \mathbb{E}\left[A_r\right] \leq \mathrm{cst} \cdot 2^{-\frac{\beta}{2}r},
$$

which is summable, so the Borel-Cantelli lemma ensures that $A_r = O\left(2^{-\frac{\beta}{2}r}\right)$ almost surely as $r \to \infty$. Now for any $n \geq 1$, there is a unique integer r_n such that $2^{r_n} \leq n < 2^{r_n+1}$, namely $r_n = \lfloor \log_2(n) \rfloor$, and we write

$$
n^{-\alpha(z)} \cdot |Z_n(z) - Z_1(z)| \leq n^{-\alpha(z)} \sup_{1 \leq k \leq n} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_1(z)| \leq A_{r_n},
$$

which proves point (i), because almost surely $A_{r_n} = O\left(n^{-\frac{\beta}{2}}\right)$ as $n \to \infty$.

    • For $\alpha \leq 0$, the reasoning is similar so we use the same notation for slightly different quantities. For any integer $r \geq 0$ we let

$$
A_r := 2^{-\underline{\alpha}(r+1)} \cdot \sum_{s=r}^{\infty} \sup_{2^s \leq k \leq 2^{s+1}} \sup_{z \in \mathrm{D}(x,\rho)} |Z_k(z) - Z_{2^s}(z)|
$$

Then, thanks to (A.3), we have

$$
\mathbb{E}\left[A_r\right] \leq \mathrm{cst} \cdot 2^{-\underline{\alpha}(r+1)} \cdot \sum_{s=r}^{\infty} 2^{\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)s} \leq \mathrm{cst} \cdot 2^{-\left(\underline{\alpha} - \overline{\alpha} + \frac{1}{q} \cdot \underline{\delta}\right)r},
$$

and thanks to our assumption the number $\beta := \underline{\alpha} - \overline{\alpha} + \frac{1}{q} \cdot \underline{\delta}$ is positive. Using the same arguments as in the case $\alpha > 0$ we have $A_r = O\left(2^{-\frac{\beta}{2}r}\right)$ almost surely as $r \to \infty$ and taking $r_n := \lfloor \log_2(n) \rfloor$ yields

$$n^{-\alpha(z)} \sup_{k \geq n} \sup_{z \in \mathrm{D}(x,\rho)} |Z_k(z) - Z_n(z)| \leq A_{r_n}.$$

Again we almost surely have $A_{r_n} = O\left(n^{-\frac{\beta}{2}}\right)$ as $n \to \infty$. This ensures that the sequence of functions $(z \mapsto Z_n(z))_{n \geq 1}$ is almost surely a Cauchy sequence for the uniform convergence on the disc $\mathrm{D}(x,\rho)$ (so that its limit $z \mapsto Z_\infty(z)$ is well-defined on the disk) and that (ii) is satisfied. $\qquad\square$

Finally, we give a proof of Lemma 3.4.

Proof of Lemma 3.4. From the assumption, we know that there exists $\epsilon > 0$ such that $W_n = \mathrm{cst} \cdot n^\gamma + O(n^{\gamma-\epsilon})$ as $n \to \infty$. Without loss of generality, we can assume that $\epsilon < 1$. Then it is immediate that $w_n = W_{n+1} - W_n = O(n^{\gamma-\epsilon})$. Then

$$\sum_{i=n}^{2n} \left(\frac{w_i}{W_i}\right)^2 \leq \frac{1}{W_n^2} \cdot \max_{n \leq i \leq 2n} w_i \cdot \sum_{i=n}^{2n} w_i \leq \frac{W_{2n}}{W_n^2} \cdot \max_{n \leq i \leq 2n} w_i = O\left(n^{-\epsilon}\right), \qquad \text{(A.4)}$$

and the first point follows by summing over intervals of the type $[\![n2^k, n2^{k+1}]\!]$.

Now write

$$\frac{W_1}{W_n} = \prod_{i=2}^{n} \frac{W_{i-1}}{W_i} = \prod_{i=2}^{n} \left(1 - \frac{w_i}{W_i}\right) = \exp\left(\sum_{i=2}^{n} \log\left(1 - \frac{w_i}{W_i}\right)\right).$$

Since $\frac{w_i}{W_i} \to 0$ as $i \to \infty$, we get

$$\log\left(1 - \frac{w_i}{W_i}\right) = -\frac{w_i}{W_i} + O\left(\left(\frac{w_i}{W_i}\right)^2\right)$$

Putting everything together, we get

$$\begin{aligned}
\sum_{i=2}^{n} \frac{w_i}{W_i} &= -\sum_{i=2}^{n} \log\left(1 - \frac{w_i}{W_i}\right) + \sum_{i=2}^{n} O\left(\left(\frac{w_i}{W_i}\right)^2\right) \\
&= \log W_n - \log W_1 + \sum_{i=2}^{\infty} O\left(\left(\frac{w_i}{W_i}\right)^2\right) - O\left(\sum_{i=n+1}^{\infty}\left(\frac{w_i}{W_i}\right)^2\right) \\
&= \log W_n + \mathrm{cst} + O\left(n^{-\epsilon}\right).
\end{aligned}$$

Last, just remark that $\log W_n = \log(\mathrm{cst} \cdot n^\gamma \cdot (1 + O(n^{-\epsilon}))) = \gamma \log n + \mathrm{cst} + O(n^{-\epsilon})$, which finishes the proof. $\qquad\square$

References

[1] Cyril Banderier, Philippe Marchal, and Michael Wallner, *Periodic Pólya urns, the density method and asymptotics of Young tableaux*, Annals of Probability **48** (2020), no. 4, 1921–1965. MR-4124529

[2] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512. MR-2091634

[3] Noam Berger, Christian Borgs, Jennifer T. Chayes, and Amin Saberi, *Asymptotic behavior and distributional limits of preferential attachment graphs*, The Annals of Probability **42** (2014), no. 1, 1–40. MR-3161480

[4] Shankar Bhamidi, *Universal techniques to analyze preferential attachment trees: Global and local analysis*, Preprint (2007).

[5] John D. Biggins, *Uniform convergence of martingales in the branching random walk*, The Annals of Probability **20** (1992), no. 1, 137–151. MR-1143415

[6] Patrick Billingsley, *Convergence of probability measures*, 2nd ed., Wiley series in probability and statistics. Probability and statistics section, Wiley, New York, 1999. MR-1700749

[7] Benjamin Bloem-Reddy and Peter Orbanz, *Preferential attachment and vertex arrival times*, arXiv:1710.02159 (2017). MR-3653681

[8] Konstantin A. Borovkov and Vladimir A. Vatutin, *On the asymptotic behaviour of random recursive trees in random environments*, Advances in Applied Probability **38** (2006), no. 4, 1047–1070. MR-2285693

[9] Brigitte Chauvin, Michael Drmota, and Jean Jabbour-Hattab, *The profile of binary search trees*, Annals of Applied Probability (2001), 1042–1062. MR-1878289

[10] Brigitte Chauvin, Thierry Klein, Jean-François Marckert, and Alain Rouault, *Martingales and profile of binary search trees*, Electronic Journal of Probability **10** (2005), 420–435. MR-2147314

[11] Bo Chen, Daniel Ford, and Matthias Winkel, *A new family of Markov branching trees: the alpha-gamma model*, Electronic Journal of Probability **14** (2009), 400–430. MR-2480547

[12] Nicolas Curien and Bénédicte Haas, *Random trees constructed by aggregation*, Ann. Inst. Fourier **67** (2017), no. 5, 1963–2001. MR-3732681

[13] Maria Deijfen, Henri van den Esker, Remco van der Hofstad, and Gerard Hooghiemstra, *A preferential attachment model with random initial degrees*, Arkiv för Matematik **47** (2009), no. 1, 41–72. MR-2480915

[14] Luc Devroye, *Branching processes in the analysis of the heights of trees*, Acta Informatica **24** (1987), no. 3, 277–298. MR-0894557

[15] Michael Drmota, *Random trees: an interplay between combinatorics and probability*, Springer Science & Business Media, 2009. MR-2484382

[16] Rick Durrett, *Probability: Theory and examples*, 4 ed., Cambridge University Press, 2010. MR-2722836

[17] Güler Ergün and Geoff J Rodgers, *Growing random networks with fitness*, Physica A: Statistical Mechanics and its Applications **303** (2002), no. 1-2, 261–272.

[18] William Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971. MR-0270403

[19] Daniel J. Ford, *Probabilities on cladograms: introduction to the alpha model*, arXiv:0511246 (2005). MR-2708802

[20] Nikolaos Fountoulakis, Tejas Iyer, Cécile Mailler, and Henning Sulzbach, *Dynamical models for random simplicial complexes*, arXiv:1910.12715 (2019).

[21] Christina Goldschmidt and Bénédicte Haas, *A line-breaking construction of the stable trees*, Electron. J. Probab. **20** (2015), 1–24. MR-3317158

[22] Bénédicte Haas and Robin Stephenson, *Scaling limits of k-ary growing trees*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **51** (2015), no. 4, 1314–1341. MR-3414449

[23] Bénédicte Haas and Robin Stephenson, *Scaling limits of multi-type Markov Branching trees*, arXiv:1912.07296 (2019). MR-3835788

[24] Ella Hiesmayr and Ümit Işlak, *Asymptotic results on hoppe trees and their variations*, Journal of Applied Probability **57** (2020), no. 2, 441–457. MR-4125458

[25] Tejas Iyer, *Degree distributions in recursive trees with fitnesses*, arXiv:2005.02197 (2020).

[26] Lancelot F James, *Generalized Mittag-Leffler distributions arising as limits in preferential attachment models*, arXiv:1509.07150 (2015).

[27] Svante Janson, *Limit theorems for triangular urn schemes*, Probab. Theory Related Fields **134** (2006), no. 3, 417–452. MR-2226887

[28] Svante Janson, *Random recursive trees and preferential attachment trees are random split trees*, Combinatorics, Probability and Computing **28** (2019), no. 1, 81–99. MR-3917907

[29] Zakhar Kabluchko, Alexander Marynych, and Henning Sulzbach, *General Edgeworth expansions with applications to profiles of random trees*, The Annals of Applied Probability **27** (2017), no. 6, 3478–3524. MR-3737930

[30] Zsolt Katona, *Width of a scale-free tree*, Journal of Applied Probability **42** (2005), no. 3, 839–850. MR-2157524

[31] Tiffany Y. Y. Lo, *Weak local limit of preferential attachment random trees with additive fitness*, arXiv:2103.00900 (2021).

[32] Bas Lodewijks and Marcel Ortgiese, *A phase transition for preferential attachment models with additive fitness*, Electronic Journal of Probability **25** (2020), 1 – 54. MR-4193887

[33] Bas Lodewijks and Marcel Ortgiese, *The maximal degree in random recursive graphs with random weights*, arXiv:2007.05438 (2020).

[34] Cécile Mailler and Gerónimo Uribe Bravo, *Random walks with preferential relocations and fading memory: a study through random recursive trees*, J. Stat. Mech.: Theory Exp. **2019** (2019), no. 9, 093206. MR-4021476

[35] Philippe Marchal, *A note on the fragmentation of a stable tree*, Fifth Colloquium on Mathematics and Computer Science, Discrete Math. Theor. Comput. Sci. Proc., AI, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008, pp. 489–499. MR-2508809

[36] Tamás Móri, *The maximum degree of the Barabási–Albert random tree*, Combinatorics, Probability and Computing **14** (2005), no. 3, 339–348. MR-2138118

[37] Mark EJ Newman, *The structure and function of complex networks*, SIAM review **45** (2003), no. 2, 167–256. MR-2010377

[38] Michel Pain and Delphin Sénizergues, *Correction terms for the height of weighted recursive trees*, arXiv:2101.01156 (2021).

[39] Erol Peköz, Adrian Röllin, and Nathan Ross, *Pólya urns with immigration at random times*, Bernoulli **25** (2019), no. 1, 189–220. MR-3892317

[40] Erol Peköz, Adrian Röllin, and Nathan Ross, *Joint degree distributions of preferential attachment random graphs*, Advances in Applied Probability **49** (2017), no. 2, 368–387. MR-3668381

[41] Erol A. Peköz, Adrian Röllin, and Nathan Ross, *Degree asymptotics with rates for preferential attachment random graphs*, The Annals of Applied Probability **23** (2013), no. 3, 1188–1218. MR-3076682

[42] Erol A. Peköz, Adrian Röllin, and Nathan Ross, *Generalized gamma approximation with rates for urns, walks and trees*, The Annals of Probability **44** (2016), no. 3, 1776–1816. MR-3502594

[43] Robin Pemantle, *A time-dependent version of Pólya's urn*, Journal of Theoretical Probability **3** (1990), no. 4, 627–637. MR-1067672

[44] Robin Pemantle, *A survey of random processes with reinforcement*, Probability Surveys **4** (2007), 1–79. MR-2282181

[45] Jim Pitman, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006, Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. MR-2245368

[46] Boris Pittel, *Note on the heights of random recursive trees and random m-ary search trees*, Random Structures & Algorithms **5** (1994), no. 2, 337–347. MR-1262983

[47] Jean-Luc Rémy, *Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire*, RAIRO Inform. Théor. **19** (1985), 179–195. MR-0803997

[48] Nathan Ross and Yuting Wen, *Scaling limits for some random trees constructed inhomogeneously*, Electron. J. Probab. **23** (2018), 1–35. MR-3761565

[49] Pierre Rousselin, *Marches aléatoires sur les arbres aléatoires*, Ph.D. thesis, 2018.

[50] Eva-Maria Schopp et al., *A functional limit theorem for the profile of b-ary trees*, The Annals of Applied Probability **20** (2010), no. 3, 907–950. MR-2680553

[51] Delphin Sénizergues, *Growing random graphs with a preferential attachment structure*, arXiv:2002.09999 (2020).

[52] Robert T. Smythe and Hosam M. Mahmoud, *A survey of recursive trees*, Theory of Probability and Mathematical Statistics **51** (1995), 1–27. MR-1445048

[53] Henning Sulzbach, *A functional limit law for the profile of plane-oriented recursive trees*, Fifth Colloquium on Mathematics and Computer Science (Uwe Roesler, ed.), DMTCS Proceedings, vol. DMTCS Proceedings vol. AI, Fifth Colloquium on Mathematics and Computer Science, Discrete Mathematics and Theoretical Computer Science, 2008, pp. 339–350. MR-2508798

[54] Remco van der Hofstad, *Random graphs and complex networks. Vol. 1*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2017. MR-3617364

[55] Bengt von Bahr and Carl-Gustav Esseen, *Inequalities for the rth absolute moment of a sum of random variables,* $1 \leq r \leq 2$, Ann. Math. Statist **36** (1965), 299–303. MR-0170407