Electronic Journal of Probability

Electron. J. Probab. **19** (2014), no. 56, 1–29.
ISSN: 1083-6489 DOI: 10.1214/EJP.v19-3164

Local limits of conditioned Galton-Watson trees: the condensation case

Romain Abraham* Jean-François Delmas†

Abstract

We provide a complete picture of the local convergence of critical or sub-critical Galton-Watson trees conditioned on having a large number of individuals with out-degree in a given set. The generic case, where the limit is a random tree with an infinite spine has been treated in a previous paper. We focus here on the non-generic case, where the local limit is a random tree with a node with infinite out-degree. This case corresponds to the so-called condensation phenomenon.

Keywords: Galton-Watson ; random tree ; local-limit ; non-extinction, branching process.
AMS MSC 2010: 60J80 ; 60B10.
Submitted to EJP on November 26, 2013, final version accepted on June 20, 2014.
Supersedes HAL:hal-00909604.

1 Introduction

Conditioning critical or sub-critical Galton-Watson (GW) trees comes from the seminal work of Kesten, [12]. Let $p = (p(n), n \in \mathbb{N})$ be an offspring distribution such that:

$$p(0) > 0, \; p(0) + p(1) < 1. \tag{1.1}$$

Let $\mu(p) = \sum_{n=0}^{+\infty} np(n)$ be its mean. If $\mu(p) < 1$ (resp. $\mu(p) = 1$, $\mu(p) > 1$), we say that the offspring distribution and the associated GW tree are sub-critical (resp. critical, super-critical). In the critical and sub-critical cases, the tree is a.s. finite, but Kesten considered in [12] the local limit of a sub-critical or critical tree conditioned to have height greater than n. When n goes to infinity, this conditioned tree converges in distribution to the so-called size-biased GW tree. This random tree has an infinite spine on which are grafted a random number of independent GW trees with the same offspring distribution p. This limit tree can be seen as the GW tree conditioned on non-extinction.

Since then, other conditionings have been considered for critical GW trees: large total progeny see Kennedy [11] and Geiger and Kaufmann [6], large number of leaves

*Laboratoire MAPMO, CNRS, UMR 7349, Fédération Denis Poisson, FR 2964, Université d'Orléans, Orléans, FRANCE. E-mail: romain.abraham@univ-orleans.fr
†Université Paris-Est, École des Ponts, CERMICS, 6-8 av. Blaise Pascal, Champs-sur-Marne, Marne-la-Vallée, FRANCE. E-mail: delmas@cermics.enpc.fr

see Curien and Kortchemski [4]. We are interested in this paper in the conditioning on having a large number of individuals with out-degree belonging to a given set $\mathcal{A}$ which has been introduced in Rizzolo [18] and also appears in Kortchemski [13]. In Abraham and Delmas [1], it is proven that, in the critical case, the local limit of a GW tree conditioned to have a large number of individuals with out degree in a given set $\mathcal{A}$ is still the GW tree conditioned on non-extinction.

However, the results are different in the sub-critical case. Let $\mathcal{A} \subset \mathbb{N}$. We set

$$p(\mathcal{A}) = \sum_{k\in\mathcal{A}} p(k).$$

We first define for an offspring distribution p that satisfies (1.1) and a set $\mathcal{A}$ such that $p(\mathcal{A}) > 0$ a modified offspring distribution $p_{\mathcal{A},\theta}$ by:

$$\forall k \geq 0, \quad p_{\mathcal{A},\theta}(k) = \begin{cases} c_{\mathcal{A}}(\theta)\theta^k p(k) & \text{if } k \in \mathcal{A}, \\ \theta^{k-1} p(k) & \text{if } k \in \mathcal{A}^c, \end{cases} \tag{1.2}$$

where the normalizing constant $c_{\mathcal{A}}(\theta)$ is given by:

$$c_{\mathcal{A}}(\theta) = \frac{\theta - \mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]}{\theta\mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]}, \tag{1.3}$$

where X is a random variable distributed according to p. Let $I_{\mathcal{A}}$ be the set of positive θ for which $p_{\mathcal{A},\theta}$ is a probability distribution. If p is sub-critical, according to Lemma 5.2, either there exists (a unique) $\theta^c_{\mathcal{A}} \in I_{\mathcal{A}}$ such that $p_{\mathcal{A},\theta^c_{\mathcal{A}}}$ is critical or $\theta^*_{\mathcal{A}} := \max I_{\mathcal{A}} \in I_{\mathcal{A}}$ and $p_{\mathcal{A},\theta^*_{\mathcal{A}}}$ is sub-critical. We shall say, see Definition 5.3, that p is **generic** for the set $\mathcal{A}$ in the former case and that p is **non-generic** for the set $\mathcal{A}$ in the latter case. See Lemma 5.4 and Remark 5.5 on the non-generic property.

For a tree $\mathbf{t}$, let $\mathcal{L}_{\mathcal{A}}(\mathbf{t})$ be the set of nodes of $\mathbf{t}$ whose number of offspring belongs to $\mathcal{A}$ and $L_{\mathcal{A}}(\mathbf{t})$ be its cardinal (see definition in Section 6). It is proven in [1] that, for every $\theta \in I_{\mathcal{A}}$, if τ is a GW tree with offspring distribution p and $\tau_{\mathcal{A},\theta}$ is a GW tree with offspring distribution $p_{\mathcal{A},\theta}$, then the conditional distributions of τ given $\{L_{\mathcal{A}}(\tau) = n\}$ and that of $\tau_{\mathcal{A},\theta}$ given $\{L_{\mathcal{A}}(\tau_{\mathcal{A},\theta}) = n\}$ are the same. Therefore, if p is generic for the set $\mathcal{A}$, that is there exists a $\theta^c_{\mathcal{A}} \in I_{\mathcal{A}}$ such that $p_{\mathcal{A},\theta^c_{\mathcal{A}}}$ is critical, then the GW tree τ conditioned on $L_{\mathcal{A}}(\tau)$ being large converges to the size-biased GW tree associated with $p_{\mathcal{A},\theta^c_{\mathcal{A}}}$.

When the sub-critical offspring distribution is non-generic for $\mathbb{N}$, a condensation phenomenon has been observed when conditioning with respect to the total population size, see Jonnsson and Stefansson [8] and Janson [7]: the limiting tree is no more the size-biased tree but a tree that contains a single node with infinitely many offspring. The goal of this paper is to give a short proof of this result and to show that such a condensation also appears when p is non-generic for $\mathcal{A}$ and conditioning by $L_{\mathcal{A}}(\tau)$ being large. This and [1] give a complete description of the limit in distribution of a critical or sub-critical GW tree τ conditioned on $\{L_{\mathcal{A}}(\tau) = n\}$ as n goes to infinity.

We summarize this complete description following Janson ([7], Section 5). Let p be an offspring distribution that satisfies (1.1) which is critical or sub-critical (that is $\mu(p) \leq 1$). Let $\tau^*(p)$ denote the random tree which is defined by:

i) There are two types of nodes: *normal* and *special*.

ii) The root is special.

iii) Normal nodes have offspring distribution p.

iv) Special nodes have offspring distribution the biased distribution $\tilde{p}$ on $\mathbb{N} \cup \{+\infty\}$ defined by:

$$\tilde{p}(k) = \begin{cases} k\,p(k) & \text{if } k \in \mathbb{N}, \\ 1-\mu & \text{if } k = +\infty. \end{cases}$$

v) The offsprings of all the nodes are independent of each others.

vi) All the children of a normal node are normal.

vii) When a special node gets a finite number of children, one of them is selected uniformly at random and is special while the others are normal.

viii) When a special node gets an infinite number of children, all of them are normal.

Notice that:

- If p is critical, then a.s. $\tau^*(p)$ has one infinite spine and all its nodes have finite degrees. This is the size-biased GW tree considered in [12].
- If $\mu(p) < 1$ then a.s. $\tau^*(p)$ has exactly one node of infinite degree and no infinite spine. This tree has been considered in [8, 7].

Definition 1.1. *Let $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. We define $p^*_{\mathcal{A}}$ as:*

- ***critical case*** *($\mu(p) = 1$):*

$$p^*_{\mathcal{A}} = p.$$

- ***sub-critical and generic for*** *$\mathcal{A}$ ($\mu(p) < 1$ and there exists (a unique) $\theta^c_{\mathcal{A}} \in I_{\mathcal{A}}$ such that $\mu(p_{\mathcal{A},\theta^c_{\mathcal{A}}}) = 1$):*

$$p^*_{\mathcal{A}} = p_{\mathcal{A},\theta^c_{\mathcal{A}}}.$$

- ***sub-critical and non-generic for*** *$\mathcal{A}$ ($\mu(p) < 1$ and $\mu(p_{\mathcal{A},\theta^*_{\mathcal{A}}}) < 1$):*

$$p^*_{\mathcal{A}} = p_{\mathcal{A},\theta^*_{\mathcal{A}}}, \quad \textit{with } \theta^*_{\mathcal{A}} = \max I_{\mathcal{A}}. \tag{1.4}$$

Remark 1.2. The uniqueness of $\theta^c_{\mathcal{A}}$ is not immediate and follows from Lemma 5.2.

We state our main result (the convergence of random discrete trees is precisely defined in Section 2 and GW trees are presented in Section 3).

Theorem 1.3. *Let τ be a GW tree with offspring distribution p which satisfies (1.1) and $\mu(p) \leq 1$. Let $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. We have the following convergence in distribution:*

$$\text{dist}\left(\tau \,\middle|\, L_{\mathcal{A}}(\tau) = n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p^*_{\mathcal{A}})\right), \tag{1.5}$$

where the limit is understood along the infinite sub-sequence $\{n \in \mathbb{N}^*; \mathbb{P}(L_{\mathcal{A}}(\tau) = n) > 0\}$, as well as:

$$\text{dist}\left(\tau \,\middle|\, L_{\mathcal{A}}(\tau) \geq n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p^*_{\mathcal{A}})\right). \tag{1.6}$$

The theorem has already been proven in the critical case and the sub-critical generic case in [1]. We concentrate here on the case of the sub-critical non-generic case. The non-generic case for $\mathcal{A} = \mathbb{N}$, $0 \in \mathcal{A}$, $0 \notin \mathcal{A}$ are respectively proven in Sections 4, 6 and 7. Let us add that a sub-critical offspring distribution p is either generic for all $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$, or non-generic at least for $\{0\}$ and maybe for some other sets and generic for other sets $\mathcal{A}$ such that $p(\mathcal{A}) > 0$, see Lemma 5.4. It is not possible for a sub-critical offspring distribution p to be non-generic for all $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$ unless the radius of convergence of its generating function is 1, see Remark 5.5. By considering the last example of Remark 5.5, we exhibit a distribution p which is non-generic for $\{0\}$ but generic for $\mathbb{N}$. Thus the associated GW tree conditioned on having n vertices converges in distribution (as n goes to infinity) to a tree with an infinite spine whereas the same tree conditioned on having n leaves converges in distribution to a tree with a node of infinite degree.

Let us add that we only focus here on local limits. Global limits of conditioned Galton-Watson trees have also been studied in the generic case first by Aldous [2] (conditioning on the total population size) and then by Rizzolo [18] for the conditioning studied here. Global limits on non-generic trees have also been studied by Kortchemski [14] when conditioned to the total population size. Let us add that global limits require regularity for the offspring distribution (such as belonging to the domain of attraction of stable laws) whereas we only require here a first moment to obtain a local limit.

In Section 2, we recall the setting of the discrete trees (which is close to [1], but has to include discrete trees with nodes of infinite degree). We also give in Lemma 2.2, in the same spirit of Lemma 2.1 in [1], a convergence determining class which is the key result to prove the convergence in the non-generic case. Section 3 is devoted to some remarks on GW trees. We study in detail the distribution $p_{\mathcal{A},\theta}$ defined by (1.2) in Section 5. The proof of Theorem 1.3 is given in the following three sections. More precisely, the case $\mathcal{A} = \mathbb{N}$ is presented in Section 4. This provides a short and self-contained proof of the results from [8, 7]. The case $0 \in \mathcal{A}$ can be handled in the same spirit, see Section 6, using that the set $\mathcal{L}_{\mathcal{A}}(\tau)$ can be encoded into a GW tree $\tau^{\mathcal{A}}$, see Mimami [15] or [18]. Notice that if $0 \notin \mathcal{A}$, then $\mathcal{L}_{\mathcal{A}}(\tau)$, when non empty, can also be encoded into a GW tree $\tau^{\mathcal{A}}$, see [18]. However, we didn't use this result, but rather use in Section 7 a more technical version of the previous proofs to treat the case $0 \notin \mathcal{A}$. We prove in the appendix, Section 8, consequences of the strong ratio limit property we used in the previous sections.

2 The set of discrete trees

We recall Neveu's formalism [17] for ordered rooted trees. We set

$$\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^*)^n$$

the set of finite sequences of positive integers with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. For $n \geq 0$ and $u = (u_1, \ldots, u_n) \in \mathcal{U}$, let $|u| = n$ be the length of u and:

$$|u|_\infty = \max(|u|, u_1, u_2, \ldots, u_{|u|})$$

with the convention $|\emptyset| = |\emptyset|_\infty = 0$. We will call $|u|_\infty$ the norm of u although it is not a norm since $\mathcal{U}$ is not even a vector space. If u and v are two sequences of $\mathcal{U}$, we denote by uv the concatenation of the two sequences, with the convention that $uv = u$ if $v = \emptyset$ and $uv = v$ if $u = \emptyset$. The set of ancestors of $\emptyset$ is $A_\emptyset = \{\emptyset\}$ and of $u \neq \emptyset$ is:

$$A_u = \{v \in \mathcal{U}; \text{there exists } w \in \mathcal{U},\, w \neq \emptyset, \text{ such that } u = vw\}. \tag{2.1}$$

The most recent common ancestor of a subset $\mathbf{s}$ of $\mathcal{U}$, denoted by MRCA($\mathbf{s}$), is the unique element v of $\bigcap_{u\in\mathbf{s}} A_u$ with maximal length $|v|$. For $u, v \in \mathcal{U}$, we denote by $u < v$ the lexicographic order on $\mathcal{U}$ i.e. $u < v$ if either $u \in A_v$ or, if we set $w = \text{MRCA}(\{u, v\})$, then $u = wiu'$ and $v = wjv'$ for some $i, j \in \mathbb{N}^*$ with $i < j$.

A tree $\mathbf{t}$ is a subset of $\mathcal{U}$ that satisfies:

- $\emptyset \in \mathbf{t}$,
- If $u \in \mathbf{t}$, then $A_u \subset \mathbf{t}$.

- For every $u \in \mathbf{t}$, there exists $k_u(\mathbf{t}) \in \mathbb{N} \cup \{+\infty\}$ such that, for every positive integer i, $ui \in \mathbf{t}$ iff $1 \le i \le k_u(\mathbf{t})$.

The integer $k_u(\mathbf{t})$ represents the number of offsprings of the vertex $u \in \mathbf{t}$. (Notice that $k_u(\mathbf{t})$ has to be finite in [1], whereas $k_u(\mathbf{t})$ might take the value $+\infty$ here.) The vertex $u \in \mathbf{t}$ is called a leaf if $k_u(\mathbf{t}) = 0$ and it is said infinite if $k_u(\mathbf{t}) = +\infty$. By convention, we shall set $k_u(\mathbf{t}) = -1$ if $u \notin \mathbf{t}$. The vertex $\emptyset$ is called the root of $\mathbf{t}$. We set:

$$|\mathbf{t}| = \mathrm{Card}\ (\mathbf{t}).$$

Let $\mathbf{t}$ be a tree. The set of its leaves is $\mathcal{L}_0(\mathbf{t}) = \{u \in \mathbf{t}; k_u(\mathbf{t}) = 0\}$. Its height and its "norm" are resp. defined by

$$H(\mathbf{t}) = \sup\{|u|,\ u \in \mathbf{t}\} \quad \text{and} \quad H_\infty(\mathbf{t}) = \sup\{|u|_\infty,\ u \in \mathbf{t}\} = \max(H(\mathbf{t}), \sup\{k_u(\mathbf{t}),\ u \in \mathbf{t}\});$$

they can be infinite. For $u \in \mathbf{t}$, we define the sub-tree $\mathcal{S}_u(\mathbf{t})$ of $\mathbf{t}$ "above" u as:

$$\mathcal{S}_u(\mathbf{t}) = \{v \in \mathcal{U},\ uv \in \mathbf{t}\}.$$

For $u \in \mathbf{t} \setminus \mathcal{L}_0(\mathbf{t})$, we also define the forest $\mathcal{F}_u(\mathbf{t})$ "above" u as the following sequence of trees:

$$\mathcal{F}_u(\mathbf{t}) = (\mathcal{S}_{ui}(\mathbf{t});\ i \in \mathbb{N}^*, i \le k_u(\mathbf{t})).$$

For $u \in \mathbf{t} \setminus \{\emptyset\}$, we also define the sub-tree $\mathcal{S}^u(\mathbf{t})$ of $\mathbf{t}$ "below" u as:

$$\mathcal{S}^u(\mathbf{t}) = \{v \in \mathbf{t}; u \notin A_v\}.$$

Notice that $u \in \mathcal{S}^u(\mathbf{t})$.

For $v = (v_k, k \in \mathbb{N}^*) \in (\mathbb{N}^*)^{\mathbb{N}^*}$, we set $\bar{v}_n = (v_1, \dots, v_n)$ for $n \in \mathbb{N}$, with the convention that $\bar{v}_0 = \emptyset$ and $\bar{\mathbf{v}} = \{\bar{v}_n, n \in \mathbb{N}\}$ defines an infinite spine or branch. We denote by $\mathbb{T}_\infty$ the set of trees. We denote by $\mathbb{T}_0$ the subset of finite trees,

$$\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T}_\infty;\ |\mathbf{t}| < +\infty\},$$

by $\mathbb{T}_\infty^{(h)}$ the subset of trees with norm less than h,

$$\mathbb{T}_\infty^{(h)} = \{\mathbf{t} \in \mathbb{T}_\infty;\ H_\infty(\mathbf{t}) \le h\},$$

by $\mathbb{T}_0^*$ the subset of trees with no infinite branch,

$$\mathbb{T}_0^* = \{\mathbf{t} \in \mathbb{T}_\infty; \forall v \in (\mathbb{N}^*)^{\mathbb{N}}, \bar{\mathbf{v}} \not\subset \mathbf{t}\},$$

and by $\mathbb{T}_2$ the subset of trees with no infinite branch and with exactly one infinite vertex,

$$\mathbb{T}_2 = \{\mathbf{t} \in \mathbb{T}_\infty; \mathrm{Card}\ \{u \in \mathbf{t}; k_u(\mathbf{t}) = +\infty\} = 1\} \cap \mathbb{T}_0^*.$$

Notice that $\mathbb{T}_0$ is countable and $\mathbb{T}_2$ is uncountable.

For $h \in \mathbb{N}$, the restriction function $r_{h,\infty}$ from $\mathbb{T}_\infty$ to $\mathbb{T}_\infty$ is defined by:

$$r_{h,\infty}(\mathbf{t}) = \{u \in \mathbf{t},\ |u|_\infty \le h\}.$$

Sets $r_{h,\infty}(\mathbf{t})$ are called "left-balls" in [8]. We endow the set $\mathbb{T}_\infty$ with the ultra-metric distance

$$d_\infty(\mathbf{t}, \mathbf{t}') = 2^{-\max\{h \in \mathbb{N},\ r_{h,\infty}(\mathbf{t}) = r_{h,\infty}(\mathbf{t}')\}}.$$

A sequence $(\mathbf{t}_n, n \in \mathbb{N})$ of trees converges to a tree $\mathbf{t}$ with respect to the distance d_∞ if and only if, for every $h \in \mathbb{N}$,

$$r_{h,\infty}(\mathbf{t}_n) = r_{h,\infty}(\mathbf{t}) \qquad \text{for } n \text{ large enough},$$

that is for all $u \in \mathcal{U}$, $\lim_{n\to+\infty} k_u(\mathbf{t}_n) = k_u(\mathbf{t}) \in \mathbb{N} \cup \{-1,+\infty\}$. The Borel σ-field associated with the distance d_∞ is the smallest σ-field containing the singletons for which the restrictions functions $(r_{h,\infty}, h \in \mathbb{N})$ are measurable. With this distance, the restriction functions are contractant. Since $\mathbb{T}_0$ is dense in $\mathbb{T}_\infty$ and $(\mathbb{T}_\infty, d_\infty)$ is complete and compact, we get that $(\mathbb{T}_\infty, d_\infty)$ is a compact Polish metric space.

Remark 2.1. In [1], we considered

$$\mathbb{T} = \{\mathbf{t} \in \mathbb{T}_\infty;\, k_u(\mathbf{t}) < +\infty\, \forall u \in \mathbf{t}\}$$

the subset of trees with no infinite vertex. On $\mathbb{T}$, we defined the distance:

$$d(\mathbf{t},\mathbf{t}') = 2^{-\max\{h\in\mathbb{N},\ r_h(\mathbf{t})=r_h(\mathbf{t}')\}},$$

with $r_h(\mathbf{t}) = \{u \in \mathbf{t},\ |u| \le h\}$. Notice that $(\mathbb{T}, d)$ is Polish but not compact and that $\mathbb{T}$ is not closed in $(\mathbb{T}_\infty, d_\infty)$. If a sequence $(\mathbf{t}_n, n \in \mathbb{N}^*)$ converges in $(\mathbb{T}, d)$ then it converges in $(\mathbb{T}_\infty, d_\infty)$. And if a sequence $(\mathbf{t}_n, n \in \mathbb{N}^*)$ of elements of $\mathbb{T}$ converges in $(\mathbb{T}_\infty, d_\infty)$ to a limit in $\mathbb{T}$ then it converges to the same limit in $(\mathbb{T}, d)$.

Consider the closed ball $B_\infty(\mathbf{t}, 2^{-h}) = \{\mathbf{t}' \in \mathbb{T}_\infty; d_\infty(\mathbf{t},\mathbf{t}') \le 2^{-h}\}$ for some $\mathbf{t} \in \mathbb{T}_\infty$ and $h \in \mathbb{N}$ and notice that:

$$B_\infty(\mathbf{t}, 2^{-h}) = r_{h,\infty}^{-1}(\{r_{h,\infty}(\mathbf{t})\}).$$

Since the distance is ultra-metric, the closed balls are open and the open balls are closed, and the intersection of two balls is either empty or one of them. We deduce that the family $((r_{h,\infty}^{-1}(\{\mathbf{t}\}), \mathbf{t} \in \mathbb{T}_\infty^{(h)}), h \in \mathbb{N})$ is a π-system, and Theorem 2.3 in [3] implies that this family is convergence determining for the convergence in distribution. Let $(T_n, n \in \mathbb{N}^*)$ and T be $\mathbb{T}_\infty$-valued random variables. We denote by $\mathrm{dist}\ (T)$ the distribution of the random variable T (which is uniquely determined by the sequence of distributions of $r_{h,\infty}(T)$ for every $h \ge 0$), and we denote:

$$\mathrm{dist}\ (T_n) \underset{n\to+\infty}{\longrightarrow} \mathrm{dist}\ (T)$$

for the convergence in distribution of the sequence $(T_n, n \in \mathbb{N}^*)$ to T. Notice that this convergence in distribution is equivalent to the finite dimensional convergence in distribution of $(k_u(T_n), u \in \mathcal{U})$ to $(k_u(T), u \in \mathcal{U})$ as n goes to infinity.

We deduce from the portmanteau theorem that the sequence $(T_n, n \in \mathbb{N}^*)$ converges in distribution to T if and only if for all $h \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}_\infty^{(h)}$:

$$\lim_{n\to+\infty} \mathbb{P}(r_{h,\infty}(T_n) = \mathbf{t}) = \mathbb{P}(r_{h,\infty}(T) = \mathbf{t}).$$

As we shall only consider $\mathbb{T}_0$-valued random variables that converge in distribution to a $\mathbb{T}_2$-valued random variable, we give an other characterization of convergence in distribution that holds for this restriction. To present this result, we introduce some notations. If $v = (v_1, \ldots, v_n) \in \mathcal{U}$, with $n > 0$, and $k \in \mathbb{N}$, we define the shift of v by k as $\theta(v,k) = (v_1 + k, v_2, \ldots, v_n)$. If $\mathbf{t} \in \mathbb{T}_0$, $\mathbf{s} \in \mathbb{T}_\infty$ and $x \in \mathbf{t}$ we denote by:

$$\mathbf{t} \circledast (\mathbf{s}, x) = \mathbf{t} \cup \{x\theta(v, k_x(\mathbf{t})),\ v \in \mathbf{s} \setminus \{\emptyset\}\}$$

the tree obtained by grafting the tree $\mathbf{s}$ at x on "the right" of the tree $\mathbf{t}$, with the convention that $\mathbf{t} \circledast (\mathbf{s}, x) = \mathbf{t}$ if $\mathbf{s} = \{\emptyset\}$ is the tree reduced to its root. Notice that if x is a leaf of $\mathbf{t}$ and $\mathbf{s} \in \mathbb{T}$, then this definition coincides with the one given in [1].

For every $\mathbf{t} \in \mathbb{T}_0$ and every $x \in \mathbf{t}$, we consider the set of trees obtained by grafting a tree at x on "the right" of $\mathbf{t}$:

$$\mathbb{T}(\mathbf{t}, x) = \{\mathbf{t} \circledast (\mathbf{s}, x),\ \mathbf{s} \in \mathbb{T}_\infty\}$$

as well as for $k \in \mathbb{N}$:

$$\mathbb{T}(\mathbf{t}, x, k) = \{\mathbf{s} \in \mathbb{T}(\mathbf{t}, x);\, k_x(\mathbf{s}) = k\} \quad \text{and} \quad \mathbb{T}_+(\mathbf{t}, x, k) = \{\mathbf{s} \in \mathbb{T}(\mathbf{t}, x);\, k_x(\mathbf{s}) \geq k\}$$

the subsets of $\mathbb{T}(\mathbf{t}, x)$ such that the number of offspring of x are resp. k and k or more. It is easy to see that $\mathbb{T}_+(\mathbf{t}, x, k)$ is closed. It is also open, as for all $\mathbf{s} \in \mathbb{T}_+(\mathbf{t}, x, k)$ we have that $B_\infty(\mathbf{s}, 2^{-\max(k, H_\infty(\mathbf{t}))-1}) \subset \mathbb{T}_+(\mathbf{t}, x, k)$.

Moreover, notice that the set $\mathbb{T}_2$ is a Borel subset of the set $\mathbb{T}$. The next lemma gives another criterion for the convergence in distribution in $\mathbb{T}_0 \cup \mathbb{T}_2$. Its proof is very similar to the proof of Lemma 2.1 in [1].

Lemma 2.2. *Let $(T_n, n \in \mathbb{N}^*)$ and T be $\mathbb{T}_\infty$-valued random variables which belong a.s. to $\mathbb{T}_0 \cup \mathbb{T}_2$. The sequence $(T_n, n \in \mathbb{N}^*)$ converges in distribution to T if and only if for every $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$ and $k \in \mathbb{N}$, we have:*

$$\lim_{n\to+\infty} \mathbb{P}(T_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = \mathbb{P}(T \in \mathbb{T}_+(\mathbf{t}, x, k)) \quad \textit{and} \quad \lim_{n\to+\infty} \mathbb{P}(T_n = \mathbf{t}) = \mathbb{P}(T = \mathbf{t}). \tag{2.2}$$

Remark 2.3. Let

$$\mathbb{T}_1 = \{\mathbf{t} \in \mathbb{T}; \exists!\, v \in (\mathbb{N}^*)^{\mathbb{N}^*} \text{ s.t. } \bar{\mathbf{v}} \subset \mathbf{t}\},$$

be the subset of trees with only one infinite spine (or branch). We give in [1] a characterization of the convergence in $\mathbb{T}_0 \cup \mathbb{T}_1$ as follows. Let $(T_n, n \in \mathbb{N}^*)$ and T be $\mathbb{T}$-valued random variables which belong a.s. to $\mathbb{T}_0 \cup \mathbb{T}_1$. The sequence $(T_n, n \in \mathbb{N}^*)$ converges in distribution to T if and only if (2.2) holds for every $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathcal{L}_0(\mathbf{t})$ and $k = 0$. In a sense, the convergence in $\mathbb{T}_0 \cup \mathbb{T}_1$ is thus easier to check.

Proof. The subclass $\mathcal{F} = \{\mathbb{T}_+(\mathbf{t}, x, k) \bigcap (\mathbb{T}_0 \bigcup \mathbb{T}_2)\,,\ \mathbf{t} \in \mathbb{T}_0,\ x \in \mathbf{t}, k \in \mathbb{N}\} \cup \{\{\mathbf{t}\},\ \mathbf{t} \in \mathbb{T}_0\}$ of Borel sets on $\mathbb{T}_0 \bigcup \mathbb{T}_2$ forms a π-system since we have

$$\mathbb{T}_+(\mathbf{t}_1, x_1, k_1) \cap \mathbb{T}_+(\mathbf{t}_2, x_2, k_2) = \begin{cases} \mathbb{T}_+(\mathbf{t}_1, x_1, k_1) & \text{if } \mathbf{t}_1 \in \mathbb{T}(\mathbf{t}_2, x_2) \text{ and } x_2 \in A_{x_1}, \\ \mathbb{T}_+(\mathbf{t}_1, x_1, k_1 \vee k_2) & \text{if } \mathbf{t}_1 \in \mathbb{T}(\mathbf{t}_2, x_2) \text{ and } x_1 = x_2, \\ \{\mathbf{t}_1\} & \text{if } \mathbf{t}_1 \in \mathbb{T}(\mathbf{t}_2, x_2) \text{ and } x_2 \notin A_{x_1} \cup \{x_1\}, \\ \emptyset & \text{in the other (non-symmetric) cases.} \end{cases}$$

For every $h \in \mathbb{N}$ and every $\mathbf{t} \in \mathbb{T}_\infty^{(h)}$, we have that $\mathbf{t}'$ belongs to $r_{h,\infty}^{-1}(\{\mathbf{t}\}) \bigcap \mathbb{T}_2$ if and only if $\mathbf{t}'$ belongs to some $\mathbb{T}_+(\mathbf{s}, x, k) \bigcap \mathbb{T}_2$ with $x \in \mathbf{t}$ such that $|x|_\infty = h$ and $\mathbf{s}$ belongs to $r_{h,\infty}^{-1}(\{\mathbf{t}\}) \bigcap \mathbb{T}_0$ with $x \in \mathbf{s}$. Since $\mathbb{T}_0$ is countable, we deduce that $\mathcal{F}$ generates the Borel σ-field on $\mathbb{T}_0 \cup \mathbb{T}_2$. In particular $\mathcal{F}$ is a separating class in $\mathbb{T}_0 \cup \mathbb{T}_2$. Since $A \in \mathcal{F}$ is closed and open as well, according to Theorem 2.3 of [3], to prove that the family $\mathcal{F}$ is a convergence determining class, it is enough to check that, for all $\mathbf{t} \in \mathbb{T}_0 \cup \mathbb{T}_2$ and $h \in \mathbb{N}$, there exists $A \in \mathcal{F}$ such that:

$$\mathbf{t} \in A \subset B_\infty(\mathbf{t}, 2^{-h}). \tag{2.3}$$

If $\mathbf{t} \in \mathbb{T}_0$, this is clear as $\{\mathbf{t}\} = B_\infty(\mathbf{t}, 2^{-h})$ for all $h > H_\infty(\mathbf{t})$. If $\mathbf{t} \in \mathbb{T}_2$, for all $\mathbf{s} \in \mathbb{T}_0$ and $x \in \mathbf{s}$ such that $\mathbf{t} \in \mathbb{T}_+(\mathbf{s}, x, k)$, with $k = k_x(\mathbf{s})$, we have $\mathbf{t} \in \mathbb{T}_+(\mathbf{s}, x, k) \subset B_\infty(\mathbf{t}, 2^{-|x|_\infty})$. Since we can find such a $\mathbf{s}$ and x such that $|x|_\infty$ is arbitrary large, we deduce that (2.3) is satisfied. This proves that the family $\mathcal{F}$ is a convergence determining class in $\mathbb{T}_0 \cup \mathbb{T}_2$. Since, for $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$ and $k \in \mathbb{N}$, the sets $\mathbb{T}_+(\mathbf{t}, x, k)$ and $\{\mathbf{t}\}$ are open and closed, we deduce from the portmanteau theorem that if $(T_n, n \in \mathbb{N}^*)$ converges in distribution to T, then (2.2) holds for every $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$ and $k \in \mathbb{N}$. $\square$

3 GW trees

3.1 Definition

Let $p = (p(n), n \in \mathbb{N})$ be a probability distribution on the set of the non-negative integers. We assume that p satisfies (1.1). Let $g(z) = \sum_{k\in\mathbb{N}} p(k)\, z^k$ be the generating function of p. We denote by $\rho(p)$ its radius of convergence and we will write ρ for $\rho(p)$ when it is clear from the context. We say that p is aperiodic if $\{k; p(k) > 0\} \subset d\mathbb{N}$ implies $d = 1$.

A $\mathbb{T}$-valued random variable τ is a Galton-Watson (GW) tree with offspring distribution p if the distribution of $k_\emptyset(\tau)$ is p and for $n \in \mathbb{N}^*$, conditionally on $\{k_\emptyset(\tau) = n\}$, the sub-trees $(\mathcal{S}_1(\tau), \mathcal{S}_2(\tau), \dots, \mathcal{S}_n(\tau))$ are independent and distributed as the original tree τ. Equivalently, for every $h \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_\infty^{(h)}$, we have:

$$\mathbb{P}(r_{h,\infty}(\tau) = \mathbf{t}) = \prod_{u\in r_{h-1,\infty}(\mathbf{t})} p(k_u(\mathbf{t})).$$

In particular, the restriction of the distribution of τ on the set $\mathbb{T}_0$ is given by:

$$\forall \mathbf{t} \in \mathbb{T}_0, \quad \mathbb{P}(\tau = \mathbf{t}) = \prod_{u\in\mathbf{t}} p(k_u(\mathbf{t})). \tag{3.1}$$

The GW tree is called critical (resp. sub-critical, super-critical) if $\mu(p) = 1$ (resp. $\mu(p) < 1$, $\mu(p) > 1$). In the critical and sub-critical case, we have that a.s. τ belongs to $\mathbb{T}_0$.

Let $\mathbb{P}_k$ be the distribution of the forest $\tau^{(k)} = (\tau_1, \dots, \tau_k)$ of i.i.d. GW trees with offspring distribution p. We set:

$$|\tau^{(k)}| = \sum_{j=1}^k |\tau_j|.$$

When there is no confusion, we shall write τ for $\tau^{(k)}$.

3.2 Condensation tree

Assume that p satisfies (1.1) with $\mu(p) < 1$. Recall the definition of the tree $\tau^*(p)$ in the introduction. Remark that, as $\mu(p) < 1$, the tree $\tau^*(p)$ belongs a.s. to $\mathbb{T}_2$.

For $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$, we set:

$$D(\mathbf{t}, x) = \frac{\mathbb{P}(\tau = \mathcal{S}^x(\mathbf{t}))}{p(0)} \mathbb{P}_{k_x(\mathbf{t})}(\tau = \mathcal{F}_x(\mathbf{t})).$$

For $z \in \mathbb{R}$, we set $z_+ = \max(z, 0)$. Let X be a random variable with distribution p. The following lemma is elementary.

Lemma 3.1. *Assume that p satisfies (1.1) and $\mu(p) < 1$. The distribution of $\tau^*(p)$ is also characterized by: a.s. $\tau^*(p) \in \mathbb{T}_2$ and for $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$, $k \in \mathbb{N}$,*

$$\mathbb{P}(\tau^*(p) \in \mathbb{T}_+(\mathbf{t}, x, k)) = D(\mathbf{t}, x) \left(1 - \mu(p) + \mathbb{E}\left[(X - k_x(\mathbf{t}))_+ \mathbf{1}_{\{X\geq k\}}\right]\right). \tag{3.2}$$

In particular, we have that if $x \in \mathcal{L}_0(\mathbf{t})$:

$$\mathbb{P}(\tau^*(p) \in \mathbb{T}(\mathbf{t}, x),\, k_x(\tau^*(p)) = +\infty) = (1 - \mu(p)) \frac{\mathbb{P}(\tau = \mathbf{t})}{p(0)}$$

and

$$\mathbb{P}(\tau^*(p) \in \mathbb{T}(\mathbf{t}, x)) = \frac{\mathbb{P}(\tau = \mathbf{t})}{p(0)}\cdot \tag{3.3}$$

Remark 3.2. Assume that p satisfies (1.1) and $\mu(p) = 1$. Then, according to [1], the distribution of $\tau^*(p)$ is characterized by $\tau^*(p) \in \mathbb{T}_1$ a.s. and (3.3) for $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathcal{L}_0(\mathbf{t})$.

Remark 3.3. Let $\tau^S(p)$ denote the limit (in distribution) of a critical or sub-critical GW tree τ conditionally on $\{H(\tau) = n\}$ or $\{H(\tau) \geq n\}$ as n goes to infinity (see [12, 1] for the existence of this limit). The distribution of $\tau^S(p)$ is characterized by the properties i) to vii) with $\tilde{p}$ in iv) replaced by the size-biased distribution p°:

$$p^\circ(k) = \frac{k\, p(k)}{\mu} \text{ for } k \in \mathbb{N}.$$

Remark that, when p is critical, the definitions of $\tau^*(p)$ and $\tau^{\mathrm{S}}(p)$ coincide. We have that a.s. $\tau^{\mathrm{S}}(p)$ belongs to $\mathbb{T}_1$. Following [1], we notice that the distribution of $\tau^{\mathrm{S}}(p)$ is characterized by: a.s. $\tau^{\mathrm{S}}(p) \in \mathbb{T}_1$ and for all $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathcal{L}_0(\mathbf{t})$,

$$\mathbb{P}(\tau^{\mathrm{S}}(p) \in \mathbb{T}(\mathbf{t}, x)) = \frac{\mathbb{P}(\tau = t)}{\mu(p)^{|x|} p(0)}. \tag{3.4}$$

4 Conditioning on the total population size ($\mathcal{A} = \mathbb{N}$)

We prove Theorem 1.3 for the special case $\mathcal{A} = \mathbb{N}$ and concentrate on the case p non-generic for $\mathbb{N}$. The results of this section appear already in [7] see also [8]. We provide here an elementary proof relying on the strong ratio limit property of random walks on the integers.

Recall the definition of $p_{\mathbb{N},\theta}$ in (1.2) and that $I_{\mathbb{N}}$ is the set of positive θ for which $p_{\mathbb{N},\theta}$ is a probability distribution. It is easy to check that $\mu(p_{\mathbb{N},\theta})$ is increasing in θ. Following [7], we define a non-generic distribution for $\mathbb{N}$ as follows.

Definition 4.1. *Let p be a distribution on $\mathbb{N}$ satisfying (1.1). We say p is non-generic for $\mathbb{N}$ if $\lim_{\theta \uparrow \rho(p)} \mu(p_{\mathbb{N},\theta}) < 1$.*

4.1 The non-generic case with $\rho(p) = 1$

Notice that if p satisfies (1.1) and if $\rho(p) = 1$ and $\mu(p) < 1$, then p is in particular non-generic for $\mathbb{N}$.

Theorem 4.2. *Assume that p satisfies (1.1), $\rho(p) = 1$ and $\mu(p) < 1$. We have that:*

$$\mathrm{dist}\,(\tau \bigm| |\tau| = n) \underset{n\to+\infty}{\longrightarrow} \mathrm{dist}\,(\tau^*(p)), \tag{4.1}$$

where the limit is understood along the infinite sub-sequence $\{n \in \mathbb{N}^*; \mathbb{P}(|\tau| = n) > 0\}$, and:

$$\mathrm{dist}\,(\tau \bigm| |\tau| \geq n) \underset{n\to+\infty}{\longrightarrow} \mathrm{dist}\,(\tau^*(p)). \tag{4.2}$$

Proof. For simplicity, we shall assume that p is aperiodic, that is $\mathbb{P}(|\tau| = n) > 0$ for all n large enough. The adaptation to the periodic case is left to the reader.

Recall that $\rho(p) = 1$. Let $k \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$, $\ell = k_x(\mathbf{t})$ and $m = |\mathbf{t}|$. We have:

$$\mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k), |\tau| = n) = D(\mathbf{t}, x) \sum_{j \geq \max(\ell+1, k)} p(j) \mathbb{P}_{j-\ell}(|\tau| = n - m).$$

Let $(X_n, n \in \mathbb{N}^*)$ be a sequence of independent random variables taking values in $\mathbb{N}$ with distribution p and set $S_n = \sum_{k=1}^n X_k$. Let us recall Dwass formula (see [5]): for every $k \in \mathbb{N}^*$ and every $n \geq k$, we have

$$\mathbb{P}_k(|\tau| = n) = \frac{k}{n} \mathbb{P}(S_n = n - k). \tag{4.3}$$

Let τ_n be distributed as τ conditionally on $\{|\tau| = n\}$. Using Dwass formula (4.3), we have

$$\begin{aligned}\mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) &= \frac{\mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k),\ |\tau| = n)}{\mathbb{P}(|\tau| = n)} \\ &= D(\mathbf{t}, x) \sum_{j \geq \max(\ell+1,k)} p(j) \frac{\mathbb{P}_{j-\ell}(|\tau| = n - m)}{\mathbb{P}(|\tau| = n)} \\ &= D(\mathbf{t}, x) \sum_{j \geq \max(\ell+1,k)} p(j) n \frac{j - \ell}{n - m} \frac{\mathbb{P}(S_{n-m} = n - m - j + \ell)}{\mathbb{P}(S_n = n - 1)} \cdot \end{aligned}$$

We then set

$$\delta_n^0(k, \ell) = \frac{1}{\mathbb{P}(S_n = n)} \sum_{j \geq k} p(j)\ \mathbb{P}(S_n = n + \ell - j) \tag{4.4}$$

and

$$\delta_n^1(k, \ell) = \frac{1}{\mathbb{P}(S_n = n)} \sum_{j \geq k} jp(j)\ \mathbb{P}(S_n = n + \ell - j). \tag{4.5}$$

We get:

$$\begin{aligned}\mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = D(\mathbf{t}, x) \frac{n}{n - m} \frac{\mathbb{P}(S_{n-m} = n - m)}{\mathbb{P}(S_n = n - 1)} \\ \Big(\delta^1_{n-m}(\max(\ell + 1, k), \ell) - \ell \delta^0_{n-m}(\max(\ell + 1, k), \ell)\Big).\end{aligned}$$

Then use the strong ratio limit property (8.2) as well as its consequences (8.3) and (8.4), to get that:

$$\lim_{n \to +\infty} \mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = D(\mathbf{t}, x) \left(1 - \mu(p) + \sum_{j \geq \max(\ell+1,k)} (j - \ell) p(j)\right). \tag{4.6}$$

Thanks to (3.2), we get:

$$\lim_{n \to +\infty} \mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = \mathbb{P}(\tau^*(p) \in \mathbb{T}_+(\mathbf{t}, x, k)).$$

Then use Lemma 2.2 to get (4.1). Since $\mathrm{dist}\ (\tau\ \big|\ |\tau| \geq n)$ is a mixture of $\mathrm{dist}\ (\tau\ \big|\ |\tau| = k)$ for $k \geq n$, we deduce that (4.2) holds. □

Remark 4.3. Assume $\mu(p) = 1$. Then the proof of (4.6) still holds and we get in particular that for all $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$:

$$\lim_{n \to +\infty} \mathbb{P}(\tau_n \in \mathbb{T}(\mathbf{t}, x)) = \frac{\mathbb{P}(\tau = \mathbf{t})}{p(0)} \cdot$$

Then the application $\mathbb{T}(\mathbf{t}, x) \mapsto \mathbb{P}(\tau = \mathbf{t})/p(0)$ can be extended into a probability distribution on $\mathbb{T}_1$ which is given by the distribution of $\tau^*(p)$ (also equal to the distribution of τ^{S} defined in Remark 3.3). Then use Remark 2.3 to get that $\mathrm{dist}\ (\tau | \, |\tau| = n)$ converges to $\mathrm{dist}\ (\tau^*(p))$.

4.2 The non-generic case with $\rho(p) > 1$

We consider the case $\rho(p) > 1$. The offspring distribution $p_{\mathbb{N},\theta}$ of (1.2) has generating function:

$$g_\theta(z) = \frac{g(\theta z)}{g(\theta)} \cdot$$

Notice that if p is non-generic for $\mathbb{N}$, we have $I_{\mathbb{N}} = (0, \rho(p)]$ and $p^*_{\mathbb{N}}$ defined by (1.4) is $p^*_{\mathbb{N}} = p_{\mathbb{N},\rho(p)}$. According to [11] (see also Proposition 5.5 in [1] for a more general setting), if $\tau_{\mathbb{N},\theta}$ denotes a GW tree with offspring distribution $p_{\mathbb{N},\theta}$, then the distribution of $\tau_{\mathbb{N},\theta}$ conditionally on $|\tau_{\mathbb{N},\theta}|$ does not depend on $\theta \in I_{\mathbb{N}}$.

Corollary 4.4. *Assume that p satisfies (1.1) and is non-generic for $\mathbb{N}$. We have that:*

$$\text{dist}\,\left(\tau \;\middle|\; |\tau| = n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p^*_{\mathbb{N}})),$$

where the limit is understood along the infinite sub-sequence $\{n \in \mathbb{N}^*;\, \mathbb{P}(|\tau| = n) > 0\}$, and:

$$\text{dist}\,\left(\tau \;\middle|\; |\tau| \geq n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p^*_{\mathbb{N}})).$$

Proof. Let us first remark that, by construction $\rho(p^*_{\mathbb{N}}) = 1$ so that we can apply Theorem 4.2 to the tree $\tau_{\mathbb{N},\rho(p)}$. The first convergence is then a direct consequence of (4.1) and the fact that τ conditionally on $\{|\tau| = n\}$ is distributed as $\tau_{\mathbb{N},\rho(p)}$ conditionally on $\{|\tau_{\mathbb{N},\rho(p)}| = n\}$. The proof of the second convergence is similar to the proof of (4.2). $\square$

4.3 The general case

Theorem 4.2 and Corollary 4.4 (which concern the non-generic case) combined with Proposition 4.6 and Corollary 5.9 from [1] (which concern the generic case) complete the proof of Theorem 1.3 for the case $\mathcal{A} = \mathbb{N}$. This gives a complete description of the asymptotic distribution of critical and sub-critical GW trees conditioned to have a large total population size.

5 Generic and non-generic distributions

Let p be a distribution on $\mathbb{N}$ satisfying (1.1) and let X be a random variable with distribution p. Recall that $\rho(p)$ denotes the radius of convergence of the generating function g of p. Let $\mathcal{A} \subset \mathbb{N}$ be such that $p(\mathcal{A}) > 0$. We consider the modified distribution $p_{\mathcal{A},\theta}$ on $\mathbb{N}$ given by (1.2) and let $I_{\mathcal{A}}$ be the set of positive θ for which $p_{\mathcal{A},\theta}$ is a probability distribution. We have $\theta \in I_{\mathcal{A}}$ if and only if $\theta > 0$ and:

$$\left(\mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right] < +\infty \quad \text{and} \quad \mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] < \theta\right) \quad \text{or} \quad \mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] = \theta. \tag{5.1}$$

In (5.1) the former case corresponds to $c_{\mathcal{A}}(\theta) > 0$ and the latter to $c_{\mathcal{A}}(\theta) = 0$.

Notice $I_{\mathcal{A}}$ is an interval of $(0, +\infty)$ which contains 1. We have $\inf I_{\mathcal{A}} = 0$ if $0 \in \mathcal{A}$ and $1 > \inf I_{\mathcal{A}} \geq p(0)$ if $0 \notin \mathcal{A}$. Let:

$$\theta^*_{\mathcal{A}} = \sup I_{\mathcal{A}} \in [1, \rho(p)]. \tag{5.2}$$

We deduce from the definition of $p_{\mathcal{A},\theta}$ the following rule of composition, for $\theta \in I_{\mathcal{A}}$ and $\theta q \in I_{\mathcal{A}}$:

$$p_{\mathcal{A},\theta q} = \left(p_{\mathcal{A},\theta}\right)_{\mathcal{A},q}. \tag{5.3}$$

The generating function, $g_{\mathcal{A},\theta}$, of $p_{\mathcal{A},\theta}$ is given by:

$$g_{\mathcal{A},\theta}(z) = \mathbb{E}\left[(z\theta)^X \left(\frac{1}{\theta}\mathbf{1}_{\{X\in\mathcal{A}^c\}} + c_{\mathcal{A}}(\theta)\mathbf{1}_{\{X\in\mathcal{A}\}}\right)\right].$$

And we have:

$$\mu(p_{\mathcal{A},\theta}) = \mathbb{E}\left[X\theta^{X-1}\mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] + c_{\mathcal{A}}(\theta)\mathbb{E}\left[X\theta^X\mathbf{1}_{\{X\in\mathcal{A}\}}\right]. \tag{5.4}$$

Let:

$$\theta^c_{\mathcal{A}} = \inf\{\theta \in I_{\mathcal{A}};\, \mu(p_{\mathcal{A},\theta}) = 1\}, \tag{5.5}$$

with the convention that $\inf \emptyset = +\infty$. Notice that the function $\theta \mapsto \mu(p_{\mathcal{A},\theta})$ is continuous over $I_\mathcal{A}$ and $\mathcal{C}^1$ on $\overset{o}{I}_\mathcal{A}$ the interior of $I_\mathcal{A}$.

Lemma 5.1. *Let p be a distribution on $\mathbb{N}$ satisfying (1.1) and $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. For $\theta \in \overset{o}{I}_\mathcal{A}$, we have:*

$$\frac{d}{d\theta}\mu(p_{\mathcal{A},\theta}) > 0 \quad \text{if} \quad \mu(p_{\mathcal{A},\theta}) \leq 1.$$

If $0 \in \mathcal{A}$, then the function $\theta \mapsto \mu(p_{\mathcal{A},\theta})$ is increasing over $I_\mathcal{A}$.

Proof. Since p satisfies (1.1), it is easy to check that $p_{\mathcal{A},\theta}$ satisfies (1.1) for all $\theta \in I_\mathcal{A}$ such that $\theta < \theta^*_\mathcal{A}$. For the first part of the Lemma, thanks to the composition rule, it is enough to prove that $\frac{d}{d\theta}\mu(p_{\mathcal{A},\theta}) > 0$ at $\theta = 1$ if $\mu(p) \leq 1$, with p satisfying (1.1) and $\rho(p) > 1$.

Let $\theta \in I_\mathcal{A}$. We have:

$$\mu(p_{\mathcal{A},\theta}) - \mathbb{E}[X] = \frac{h_\mathcal{A}(\theta)}{\theta \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right]},$$

with

$$\begin{aligned} h_\mathcal{A}(\theta) = \mathbb{E}\left[X\theta^X \mathbf{1}_{\{X \in \mathcal{A}^c\}}\right] \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right] + \theta \mathbb{E}\left[X\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right] \\ - \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}^c\}}\right] \mathbb{E}\left[X\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right] - \theta \mathbb{E}[X] \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right]. \end{aligned}$$

Of course we have $h_\mathcal{A}(1) = 0$. The function $h_\mathcal{A}$ is of class $\mathcal{C}^\infty$ on $[0, \rho(p))$. Notice that

$$\left(\frac{d}{d\theta}\mu(p_{\mathcal{A},\theta})\right)_{|\theta=1} = \frac{h'_\mathcal{A}(1)}{p(\mathcal{A})}.$$

Moreover,

$$h'_\mathcal{A}(1) = \mathbb{E}\left[(X-1)(Xp(\mathcal{A}) - \mathbb{E}\left[X\mathbf{1}_{\{X \in \mathcal{A}\}}\right])\right] = p(\mathcal{A})\mathbb{E}\left[X(X-1)\right] + (1 - \mathbb{E}[X])\mathbb{E}\left[X\mathbf{1}_{\{X \in \mathcal{A}\}}\right].$$

In particular, we deduce from this last expression that $h'_\mathcal{A}(1) > 0$ if $\mathbb{E}[X] \leq 1$. This ends the proof of the first part of the lemma.

Let us assume that $0 \in \mathcal{A}$. We set $\Gamma(\mathcal{A}) = \mathbb{E}\left[X\mathbf{1}_{\{X \in \mathcal{A}\}}\right]/p(\mathcal{A})$. If $\mathbb{E}[X] = \mu(p) > 1$, notice that $h'_\mathcal{A}(1)/p(\mathcal{A})$ is minimal when $\Gamma(\mathcal{A})$ is maximal. We have for $k \notin \mathcal{A}$:

$$\Gamma(\mathcal{A} \cup \{k\}) - \Gamma(\mathcal{A}) = \frac{p(k)}{p(\mathcal{A} \cup \{k\})}\left(k - \Gamma(\mathcal{A})\right).$$

By induction, this implies that $\Gamma(\mathcal{A} \cup \{j; j \geq k\}) > \Gamma(\mathcal{A})$ as soon as $k > \Gamma(\mathcal{A})$ as well as $\Gamma\left(\mathcal{A} \cap (\{j; j \geq k\} \cup \{0\})\right) < \Gamma(\mathcal{A})$ as soon as $k < \Gamma(\mathcal{A})$. Therefore, we have that $\Gamma(\mathcal{A})$ is maximal (for all subsets $\mathcal{A}$ of $\mathbb{N}$ containing 0) for $\mathcal{A}$ of the form $\mathcal{A}_n = \{0\} \cup \{k; k \geq n\}$.

By considering $h'_{\mathcal{A}_n}(1) - h'_{\mathcal{A}_{n-1}}(1)$, it is then easy to check that the function $n \mapsto h'_{\mathcal{A}_n}(1)$ is first non-decreasing and then non-increasing. Since $h'_{\mathcal{A}_0}(1)$ and $h'_{\mathcal{A}_\infty}(1)$ are positive, we get that $h'_{\mathcal{A}_n}(1)$ is positive for all $n \in \mathbb{N}$ and thus $h'_\mathcal{A}(1)$ is positive. This ends the proof of the second part of the lemma. $\square$

Let us consider the equation on $I_\mathcal{A}$:

$$\mu(p_{\mathcal{A},\theta}) = 1. \tag{5.6}$$

Lemma 5.2. *Let p be a distribution on $\mathbb{N}$ satisfying (1.1) and $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. Equation (5.6) has at most one solution in $I_\mathcal{A}$. If there is no solution to Equation (5.6), then we have $\mu(p) < 1$, $\theta^*_\mathcal{A}$ belongs to $I_\mathcal{A}$ and $\mu(p_{\mathcal{A},\theta^*_\mathcal{A}}) < 1$.*

The (unique) solution of (5.6) in $I_\mathcal{A}$, if it exists, is denoted $\theta^c_\mathcal{A}$. Notice that $p_{\mathcal{A},\theta^c_\mathcal{A}}$ is critical.

Proof. Lemma 5.1 directly implies that Equation (5.6) has at most one solution.

If $0 \in \mathcal{A}$, then we have $\inf_{I_\mathcal{A}} \mu(p_{\mathcal{A},\theta}) = p(1)\mathbf{1}_{\{1\in\mathcal{A}^c\}} < 1$. If $0 \notin \mathcal{A}$, then set $q = \min I_\mathcal{A} \in (0,1)$. Notice that $c_\mathcal{A}(q) = 0$ and $\mathbb{E}\left[q^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] = q$. Use that the function $\theta \mapsto \mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]$ is convex and less than the identity map on $(q,1]$ to deduce that $\mathbb{E}\left[Xq^{X-1}\mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]$ is strictly less than 1. Then use (5.4) to deduce that:

$$\lim_{\theta\downarrow q} \mu(p_{\mathcal{A},\theta}) = \mathbb{E}\left[Xq^{X-1}\mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] < 1.$$

In conclusion, we deduce that $\inf_{I_\mathcal{A}} \mu(p_{\mathcal{A},\theta}) < 1$. Hence, if $\mu(p) \geq 1$ then Equation (5.6) has at least one solution.

From what precedes, if there is no solution to Equation (5.6), this implies that $\mu(p) < 1$ and thus:

$$\mu(p_{\mathcal{A},\theta}) < 1 \quad \text{for all } \theta \in I_\mathcal{A}. \tag{5.7}$$

We only need to consider the case $\theta^*_\mathcal{A} > 1$. Since $\theta^*_\mathcal{A} \leq \rho(p)$, we have $\rho(p) > 1$. Since $\mu(p) < 1$, the interval $J = \{\theta; g(\theta) < \theta\}$ is non-empty and $\inf J = 1$. On $J \cap I_\mathcal{A}$, we deduce from (1.3) that $\theta c_\mathcal{A}(\theta) > 1$ and then from (5.4) that $\mu(p_{\mathcal{A},\theta}) > g'(\theta)$ and thus $g'(\theta) < 1$. Notice that this implies that $I_\mathcal{A} \bigcap (1,+\infty)$ is a subset of $\bar{J}$ the closure of J. The properties on g imply that $\bar{J} = \{\theta; g(\theta) \leq \theta\}$. This clearly implies that (5.1) holds for $\theta^*_\mathcal{A}$ that is $\theta^*_\mathcal{A} \in I_\mathcal{A}$. Then conclude using (5.7). □

We give the definition of a non-generic distribution which generalizes Definition 4.1.

Definition 5.3. *Let p be a distribution on $\mathbb{N}$ satisfying (1.1) and $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. If Equation (5.6) has a (unique) solution in $I_\mathcal{A}$, then p is called generic for $\mathcal{A}$. If Equation (5.6) has no solution, then p is called non-generic for $\mathcal{A}$.*

In the next lemma, we write ρ for $\rho(p)$.

Lemma 5.4. *Let p be a distribution on $\mathbb{N}$ satisfying (1.1) such that $\mu(p) < 1$.*

- *If $\rho = +\infty$ or ($\rho < +\infty$ and $g'(\rho) \geq 1$), then p is generic for any $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$.*
- *If $\rho = 1$ (and thus $g'(1) < 1$), then p is non-generic for all $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$.*
- *If $1 < \rho < +\infty$ and $g'(\rho) < 1$ (and thus $g(\rho) < \rho$), then p is non-generic for $\{0\}$ and p is generic for $\{k\}$ for all k large enough and such that $p(k) > 0$. Furthermore p is non-generic for $\mathcal{A} \subset \mathbb{N}$ (with $p(\mathcal{A}) > 0$) if and only if:*

  $$\mathbb{E}[Y|Y \in \mathcal{A}] < \frac{\rho - \rho g'(\rho)}{\rho - g(\rho)},$$

  with Y distributed as $p_{\mathbb{N},\rho}$, that is $\mathbb{E}[f(Y)] = \mathbb{E}[f(X)\rho^X]/g(\rho)$ for every non-negative measurable function f. We also have $\theta^*_\mathcal{A} = \rho$.

Remark 5.5. We give some consequences and remarks related to the previous Lemma.

1. If p is generic for $\{0\}$ then it is generic for all $\mathcal{A} \subset \mathbb{N}$ with $p(\mathcal{A}) > 0$.
2. If $\mathcal{A}$ and $\mathcal{B}$ are disjoint subsets of $\mathbb{N}$ such that $p(\mathcal{A}) > 0$ and $p(\mathcal{B}) > 0$, then if p is non-generic for $\mathcal{A}$ and for $\mathcal{B}$ then it is non-generic for $\mathcal{A} \bigcup \mathcal{B}$.
3. If $\mathcal{A}$ and $\mathcal{B}$ are disjoint subsets of $\mathbb{N}$ such that $p(\mathcal{A}) > 0$ and $p(\mathcal{B}) > 0$, then if p is generic for $\mathcal{A}$ and for $\mathcal{B}$ then it is generic for $\mathcal{A} \bigcup \mathcal{B}$.
4. Assume $\rho(p) > 1$ and $\mathcal{A} \subset \mathcal{B}$ with $p(\mathcal{B}) > p(\mathcal{A}) > 0$.

- Then p non-generic for $\mathcal{A}$ does not imply in general that p is non-generic for $\mathcal{B}$. (See case (6) below with $\mathcal{A} = \{0\}$ and $\mathcal{B} = \mathbb{N}$.)
- Then p non-generic for $\mathcal{B}$ does not imply in general that p is non-generic for $\mathcal{A}$. (Let p satisfying (1.1) be such that $\rho(p) > 1$ and p non-generic for $\mathcal{B} = \mathbb{N}$. Then, according to Lemma 5.4, there exists k large enough such that $p(k) > 0$ and p is generic for $\mathcal{A} = \{k\}$.)

5. Let Y be as in the last statement of Lemma 5.4. According to the second part of the proof of Lemma 5.1, we get that there exists $n_0 \in \mathbb{N}^*$ such that:

$$\sup_{\mathcal{A} \ni 0} \mathbb{E}[Y | Y \in \mathcal{A}] = \mathbb{E}[Y | Y \in \mathcal{A}_{n_0}],$$

with $\mathcal{A}_n = \{0\} \cup \{k; k \geq n\}$. In particular, if p is non-generic for $\mathcal{A}_{n_0}$ then it is non-generic for all $\mathcal{A}$ containing 0.

6. Let G be a generating function with radius of convergence $\rho_G = 1$. Let $c \in (0, 1)$. Let p be the distribution with generating function:

$$g(z) = \frac{G(cz)}{G(c)}.$$

The radius of convergence of g is thus $\rho = 1/c$ and we have:

$$g_{\mathbb{N},\rho}(z) = G(z) \quad \text{and} \quad g_{\{0\},\rho}(z) = \frac{cG(z)}{G(c)} + 1 - \frac{c}{G(c)}.$$

Therefore, we have:

$$g'_{\mathbb{N},\rho}(1) = G'(1) \quad \text{and} \quad g'_{\{0\},\rho}(1) = \frac{cG'(1)}{G(c)}.$$

If $G'(1) = 1$, then we have $G(c) > c$. This implies $g'_{\{0\},\rho}(1) < g'_{\mathbb{N},\rho}(1) = 1$. Thus p is generic for $\mathbb{N}$ but non generic for $\{0\}$.

Proof. For $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$ and $\theta \in I_{\mathcal{A}}$, notice that:

$$\mu(p_{\mathcal{A},\theta}) - 1 = G_{\mathcal{A}}(\theta) \frac{\theta - g(\theta)}{\theta} - (1 - g'(\theta)) \quad \text{with} \quad G_{\mathcal{A}}(\theta) = \frac{\mathbb{E}\left[X\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right]}{\mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right]}. \tag{5.8}$$

If $\rho = +\infty$ or $\rho < +\infty$ and $g'(\rho) \geq 1$, then there exists $q > 1$ finite such that $g'(q) = 1$ which implies that q satisfies (5.1). We also have $g(q) < q$. This implies, thanks to (5.8), that $\mu(p_{\mathcal{A},q}) > 1$. Therefore, p is generic for $\mathcal{A}$.

If $\rho < +\infty$ and $g'(\rho) < 1$, then we have $g(\rho) < \rho$ and ρ satisfies (5.1). This implies that $\theta^*_{\mathcal{A}} = \rho \in I_{\mathcal{A}}$. According to Lemma 5.2, p is non-generic for $\mathcal{A}$ if and only if $\mu(p_{\mathcal{A},\rho}) < 1$ that is, using (5.8):

$$G_{\mathcal{A}}(\rho) < \frac{\rho - \rho g'(\rho)}{\rho - g(\rho)}.$$

We have $G_{\{0\}}(\rho) = 0$ and thus p is non-generic for $\{0\}$. For k such that $p(k) > 0$, we have $G_{\{k\}}(\rho) = k$ and thus p is generic for k large enough such that $p(k) > 0$. To conclude, notice that $G_{\mathcal{A}}(\rho) = \mathbb{E}[Y | Y \in \mathcal{A}]$. □

6 Vertices with a given number of children I: case $0 \in \mathcal{A}$

Assume $0 \in \mathcal{A} \subset \mathbb{N}$ and $\mathcal{A} \neq \mathbb{N}$. Assume that p satisfies (1.1), $\mu(p) < 1$. We prove Theorem 1.3 for p non-generic for $\mathcal{A}$.

In what follows, we denote by X a random variable distributed according to p. We consider only $\mathbb{P}(X \in \mathcal{A}) < 1$, as the case $\mathbb{P}(X \in \mathcal{A}) = 1$ corresponds to $\mathcal{A} = \mathbb{N}$ of Section 4. For $\mathbf{t} \in \mathbb{T}_0$, we set $\mathcal{L}_\mathcal{A}(\mathbf{t}) = \{u \in \mathbf{t}, k_u(\mathbf{t}) \in \mathcal{A}\}$ the set of nodes whose number of children belongs to $\mathcal{A}$ and define $L_\mathcal{A}(\mathbf{t}) = \mathrm{Card}\,(\mathcal{L}_\mathcal{A}(\mathbf{t}))$.

For a tree $\mathbf{t} \in \mathbb{T}_0$, following [15, 18], we can map the set $\mathcal{L}_\mathcal{A}(\mathbf{t})$ onto a tree $\mathbf{t}^\mathcal{A}$. We first define a map ϕ from $\mathcal{L}_\mathcal{A}(\mathbf{t})$ on $\mathcal{U}$ and a sequence $(\mathbf{t}_k)_{1\leq k\leq n}$ of trees (where $n = L_\mathcal{A}(\mathbf{t})$) as follows. Recall that we denote by $<$ the lexicographic order on $\mathcal{U}$. Let $u^1 < \cdots < u^n$ be the ordered elements of $\mathcal{L}_\mathcal{A}(\mathbf{t})$.

- $\phi(u^1) = \emptyset$, $\mathbf{t}_1 = \{\emptyset\}$.
- For $1 < k \leq n$, set $w^k = MRCA(\{u^{k-1}, u^k\})$ the most recent common ancestor of u^{k-1} and u^k and recall that $S_{w^k}(\mathbf{t})$ denotes the tree above w^k. We set $\mathrm{s} = \{w^k u, u \in S_{w^k}$ the sub-tree above w^k and $v = \min(\mathcal{L}_\mathcal{A}(\mathbf{s}))$. Then, we set

  $$\phi(u^k) = \phi(v)(k_{\phi(v)}(\mathbf{t}_{k-1}) + 1)$$

  the concatenation of the node $\phi(v)$ with the integer $k_{\phi(v)}(\mathbf{t}_{k-1}) + 1$, and

  $$\mathbf{t}_k = \mathbf{t}_{k-1} \cup \{\phi(u^k)\}.$$

  In other words, $\phi(u^k)$ is a child of $\phi(v)$ in $\mathbf{t}_k$ and we add it "on the right" of the other children (if any) of $\phi(v)$ in the previous tree $\mathbf{t}_{k-1}$ to get $\mathbf{t}_k$.

It is clear by construction that $\mathbf{t}_k$ is a tree for every $k \leq n$. We set $\mathbf{t}^\mathcal{A} = \mathbf{t}_n$. Then ϕ is a one-to-one map from $\mathcal{L}_\mathcal{A}(\mathbf{t})$ onto $\mathbf{t}^\mathcal{A}$. The construction of the tree $\mathbf{t}^\mathcal{A}$ is illustrated on Figure 1. Notice that $L_\mathcal{A}(\mathbf{t})$ is just the total progeny of $\mathbf{t}^\mathcal{A}$.

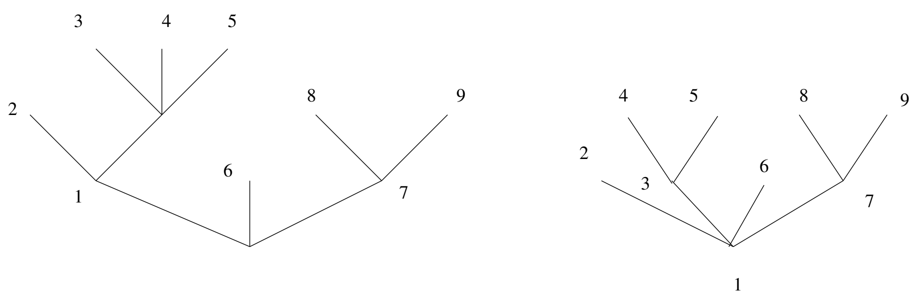

Figure 1: left: a tree $\mathbf{t}$, right: the tree $\mathbf{t}^\mathcal{A}$ for $\mathcal{A} = \{0, 2\}$

If τ is a GW tree with offspring distribution p, the tree $\tau^\mathcal{A}$ associated with $\mathcal{L}_\mathcal{A}(\tau)$ is then, according to [18] Theorem 6 (for the particular case $0 \in \mathcal{A}$), a GW tree whose offspring distribution $p^\mathcal{A}$ is defined as follows. Let N, Y'' and $(Y'_k, k \in \mathbb{N})$ be independent random variables such that N is geometric with parameter $p(\mathcal{A})$, Y'' is distributed as X conditionally on $\{X \in \mathcal{A}\}$ and $(Y'_k, k \in \mathbb{N})$ are independent random variables distributed as $X - 1$ conditionally on $\{X \in \mathcal{A}^c\}$. We set:

$$X^\mathcal{A} = \sum_{k=1}^{N-1} Y'_k + Y'', \tag{6.1}$$

with the convention that $\sum_\emptyset = 0$. Then $p^\mathcal{A}$ is the distribution of $X^\mathcal{A}$. Let $g^\mathcal{A}$ denote its generating function:

$$g^\mathcal{A}(z) = \frac{z\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]}{z - \mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]}. \tag{6.2}$$

An elementary computation gives:

$$\mu(p^{\mathcal{A}}) = 1 - \frac{1-\mu(p)}{p(\mathcal{A})} \quad \text{and} \quad g^{\mathcal{A}}(\theta) = \frac{1}{c_{\mathcal{A}}(\theta)} \cdot \tag{6.3}$$

We recover that if τ is critical ($\mu(p) = 1$) then $\tau^{\mathcal{A}}$ is critical as $\mu(p^{\mathcal{A}}) = 1$, see also [18] Lemma 6. Notice in particular that for all $k \in \mathcal{A}$:

$$p^{\mathcal{A}}(k) = \mathbb{P}(X^{\mathcal{A}} = k) \geq \mathbb{P}(N = 1, Y'' = k) = p(k), \tag{6.4}$$

and for $k \in \mathcal{A}^c$:

$$p^{\mathcal{A}}(k-1) = \mathbb{P}(X^{\mathcal{A}} = k-1) \geq \mathbb{P}(N = 2, Y_1' = k-1, Y'' = 0) = p(0)p(k). \tag{6.5}$$

Lemma 6.1. *Assume that p satisfies (1.1), $\mu(p) < 1$. Then $p^{\mathcal{A}}$ satisfies (1.1), $\mu(p^{\mathcal{A}}) < 1$ and $\rho(p^{\mathcal{A}}) = \rho(p)$ if $\rho(p) = 1$ or if $\rho(p) > 1$ and $g'(\rho(p)) < 1$.*

Proof. First, (6.4) implies that $p^{\mathcal{A}}(0) \geq p(0)$. Direct computation from (6.2) implies:

$$p^{\mathcal{A}}(0)+p^{\mathcal{A}}(1) = \frac{1}{(1-p(1)\mathbf{1}_{\{1\in\mathcal{A}^c\}})^2}\left(p(0)(1-p(1)\mathbf{1}_{\{1\in\mathcal{A}^c\}}) + p(1)\mathbf{1}_{\{1\in\mathcal{A}\}} + p(0)p(2)\mathbf{1}_{\{2\in\mathcal{A}^c\}}\right).$$

If $1 \in \mathcal{A}$, using (1.1) we get:

$$p^{\mathcal{A}}(0) + p^{\mathcal{A}}(1) \leq p(0) + p(1) + p(0)p(2) < 1.$$

If $1 \in \mathcal{A}^c$, using (1.1) we get:

$$\begin{aligned} p^{\mathcal{A}}(0) + p^{\mathcal{A}}(1) &\leq \frac{1}{(1-p(1))^2}\Big(p(0)(1-p(1)) + p(0)(1-p(0)-p(1))\Big) \\ &< \frac{1}{(1-p(1))^2}\Big(p(0)(1-p(1)) + (1-p(1))(1-p(0)-p(1)\Big) \\ &= 1. \end{aligned}$$

We deduce that $p^{\mathcal{A}}$ satisfies (1.1).

Equation (6.3) with $\mu(p) < 1$ implies directly $\mu(p^{\mathcal{A}}) < 1$.

Let $\rho_{\mathcal{A}}$ be the radius of convergence of the series given by $\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]$ and $\rho_{\mathcal{A}^c}$ be the radius of convergence of the series given by $\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]$. We get that $\min(\rho_{\mathcal{A}}, \rho_{\mathcal{A}^c}) = \rho(p)$. Using (6.1), we get:

$$g^{\mathcal{A}}(z) \geq \mathbb{P}(N=2)\mathbb{E}\left[z^{Y_1'}\right]\mathbb{E}\left[z^{Y''}\right] = \frac{1}{z}\,\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right].$$

This implies that $\rho(p^{\mathcal{A}}) \leq \rho(p)$. In particular, we get that $\rho(p^{\mathcal{A}}) = 1$ if $\rho(p) = 1$. If $\rho(p) > 1$ and $g'(\rho(p)) < 1$, then we have $g(\rho(p)) < \rho(p)$ and (6.2) is well defined for $z = \rho(p)$. This implies $\rho(p^{\mathcal{A}}) \geq \rho(p)$ and thus $\rho(p^{\mathcal{A}}) = \rho(p)$. □

6.1 The case $\rho(p) = 1$

We state now the main result of this section.

Theorem 6.2. *Assume that p satisfies (1.1), $\mu(p) < 1$ and $\rho(p) = 1$. We have that:*

$$\text{dist}\left(\tau \,\middle|\, L_{\mathcal{A}}(\tau) = n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p)\right), \tag{6.6}$$

where the limit is understood along the infinite sub-sequence $\{n \in \mathbb{N}^*; \mathbb{P}(L_{\mathcal{A}}(\tau) = n) > 0\}$, as well as

$$\text{dist}\left(\tau \,\middle|\, L_{\mathcal{A}}(\tau) \geq n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p)\right). \tag{6.7}$$

Proof. For simplicity, we shall assume that $p^{\mathcal{A}}$ is aperiodic. The adaptation to the periodic case is left to the reader. We define for $j \in \mathbb{N}$ and $n \geq 2$:

$$n_j = n - \mathbf{1}_{\{j \in \mathcal{A}\}}. \tag{6.8}$$

Let $k \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$, $\ell = k_x(\mathbf{t})$ and $m = |\mathbf{t}^{\mathcal{A}}| - \mathbf{1}_{\{x \in \mathcal{L}_{\mathcal{A}}(\mathbf{t})\}}$. We have:

$$\mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k), L_{\mathcal{A}}(\tau) = n) = D(\mathbf{t}, x) \sum_{j \geq \max(\ell+1,k)} p(j) \mathbb{P}_{j-\ell}(|\tau^{\mathcal{A}}| = n_j - m).$$

Let $(X_n, n \in \mathbb{N}^*)$ be independent random variables taking values in $\mathbb{N}$ with distribution $p^{\mathcal{A}}$ and set $S_n = \sum_{k=1}^n X_k$. According to Dwass formula (4.3), we have:

$$\mathbb{P}_{j-\ell}(\,|\tau^{\mathcal{A}}| = n_j - m) = \frac{j-\ell}{n_j - m} \mathbb{P}(S_{n_j - m} = n_j - m - j + \ell).$$

Let τ_n be distributed as τ conditionally on $\{L_{\mathcal{A}}(\tau) = n\}$. Then we have, using the definition of $\delta^{1,\mathcal{A}}_{n-m}$ and $\delta^{0,\mathcal{A}}_{n-m}$ in (8.6) and (8.7):

$$\begin{aligned}\mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) &= D(\mathbf{t}, x) \sum_{j \geq \max(\ell+1,k)} p(j) n \, \frac{j-\ell}{n_j - m} \\ &\qquad\qquad \frac{\mathbb{P}(S_{n_j - m} = n_j - m - j + \ell)}{\mathbb{P}(S_n = n-1)} \\ &= D(\mathbf{t}, x) \frac{n}{n-m} \frac{\mathbb{P}(S_{n-m} = n-m)}{\mathbb{P}(S_n = n-1)} \\ &\qquad \left(\delta^{1,\mathcal{A}}_{n-m}(\max(\ell+1,k), \ell) - \ell \delta^{0,\mathcal{A}}_{n-m}(\max(\ell+1,k), \ell) \right).\end{aligned}$$

Then use the generalizations of the strong ratio limit properties (8.2), (8.9) and (8.10) to get that:

$$\lim_{n \to +\infty} \mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = D(\mathbf{t}, x) \left(1 - \mu(p) + \sum_{j \geq \max(\ell,k)} (j-\ell) p(j) \right).$$

Thanks to (3.2), we get:

$$\lim_{n \to +\infty} \mathbb{P}(\tau_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = \mathbb{P}(\tau^*(p) \in \mathbb{T}_+(\mathbf{t}, x, k)).$$

Then use Lemma 2.2 to get (6.6). Since $\mathrm{dist}\;(\tau \mid L_{\mathcal{A}}(\tau) \geq n)$ is a mixture of $\mathrm{dist}\;(\tau \mid L_{\mathcal{A}}(\tau) = k)$ for $k \geq n$, we deduce that (6.7) holds. □

6.2 The case $\rho(p) > 1$

We consider the case p non-generic for $\mathcal{A}$ with $\rho(p) > 1$. In particular, we have $g'(\rho(p)) < 1$ and $g(\rho(p)) < \rho(p)$ thanks to Lemma 5.4. Recall the offspring distribution $p_{\mathcal{A},\theta}$ defined by (1.2). Notice that the normalizing constant $c_{\mathcal{A}}(\theta)$ is given by:

$$c_{\mathcal{A}}(\theta) = \frac{\theta - \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}^c\}}\right]}{\theta \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right]} = \frac{1}{g^{\mathcal{A}}(\theta)} \cdot \tag{6.9}$$

Notice that $p_{\mathcal{A},1} = p$. Since $\rho(p)$ is also the radius of convergence of $g^{\mathcal{A}}$, see Lemma 6.1, we deduce that $p_{\mathcal{A},\theta}$ is well defined for $\theta \in [0, \rho(p)]$ and $\theta^*_{\mathcal{A}} = \rho(p)$. Let $g_{\mathcal{A},\theta}$ be the generating function of $p_{\mathcal{A},\theta}$.

According to [11] if $\mathcal{A} = \{0\}$ and Proposition 5.5 in [1] for the general setting, if $\tau_{\mathcal{A},\theta}$ denotes a GW tree with offspring distribution $p_{\mathcal{A},\theta}$, then the distribution of $\tau_{\mathcal{A},\theta}$ conditionally on $L_{\mathcal{A}}(\tau_{\mathcal{A},\theta})$ does not depend on $\theta \in [0, \rho(p)]$.

Remark 6.3. It is easy to check that:

$$\left(g_{\mathcal{A},\theta}\right)^{\mathcal{A}}(z) = \frac{g^{\mathcal{A}}(\theta z)}{g^{\mathcal{A}}(\theta)} = \left(g^{\mathcal{A}}\right)_{\mathbb{N},\theta}(z). \tag{6.10}$$

The distribution of $\tau_{\mathcal{A},\theta}$ is the distribution of τ "shifted" by θ such that the conditional distribution given the number of vertices having a number of children in $\mathcal{A}$ is the same. Then, according to (6.10), the tree $\left(\tau_{\mathcal{A},\theta}\right)^{\mathcal{A}}$ of vertices having a number of children in $\mathcal{A}$ associated with $\tau_{\mathcal{A},\theta}$ is distributed as the distribution of $\tau^{\mathcal{A}}$ "shifted" by θ such that the conditional distribution given the total number of vertices is the same.

The proof of the following corollary is similar to the one of Corollary 4.4.

Corollary 6.4. *Assume that p satisfies (1.1) and is non-generic for $\mathcal{A}$. Let $p^*_{\mathcal{A}} = p_{\mathcal{A},\rho(p)}$. We have that:*

$$\mathrm{dist}\left(\tau \;\middle|\; L_{\mathcal{A}}(\tau) = n\right) \underset{n\to+\infty}{\longrightarrow} \mathrm{dist}\left(\tau^*(p^*_{\mathcal{A}})\right),$$

where the limit is understood along the infinite sub-sequence $\{n \in \mathbb{N}^*;\, \mathbb{P}(L_{\mathcal{A}}(\tau) = n) > 0\}$, as well as

$$\mathrm{dist}\left(\tau \;\middle|\; L_{\mathcal{A}}(\tau) \geq n\right) \underset{n\to+\infty}{\longrightarrow} \mathrm{dist}\left(\tau^*(p^*_{\mathcal{A}})\right).$$

This result with Proposition 4.6 and Corollary 5.7 in [1] ends the proof of Theorem 1.3 for the case $0 \in \mathcal{A}$, and gives a complete description of the asymptotic distribution of critical and sub-critical GW trees conditioned to have a large number vertices with given number of children.

7 Vertices with a given number of children II: case $0 \notin \mathcal{A}$

Let $\mathcal{A} \subset \mathbb{N}$. We assume in this section that $0 \notin \mathcal{A}$ and $p(\mathcal{A}) > 0$. We prove Theorem 1.3 for p non-generic for $\mathcal{A}$. Notice we follow the spirit of the case $0 \in \mathcal{A}$.

7.1 Setting and notations

Although the construction of the previous section also holds in that case with a different offspring distribution, we failed to get analogues to formulas (6.4) and (6.5) which are crucially used in the proof of Lemma 8.4 . Therefore, we prefer to map $\mathcal{L}_{\mathcal{A}}(\tau)$ onto a forest $\mathcal{F}_{\mathcal{A}}(\tau)$ of independent GW trees. Let us describe this map.

Let $\mathbf{t} \in \mathbb{T}_0$. We define a map $\tilde{\phi}$ from $\mathcal{L}_{\mathcal{A}}(\mathbf{t})$ into the set $\bigcup_{n\geq 1} \mathbb{T}_0^n$ of forests of finite trees as follows.

First, for $u \in \mathbf{t}$ we define $S^{\mathcal{A}}_u(\mathbf{t})$ the sub-tree rooted at u with no progeny in $\mathcal{A}$ by

$$S^{\mathcal{A}}_u(\mathbf{t}) = \{w \in uS_u(\mathbf{t}),\ A_w \cap A_u^c \cap \mathcal{L}_{\mathcal{A}}(\mathbf{t}) = \emptyset\}.$$

For $u \in \mathbf{t}$, we define $C^{\mathcal{A}}_u(\mathbf{t})$ as the leaves of $S^{\mathcal{A}}_u(\mathbf{t})$ that belong to $\mathcal{A}$.

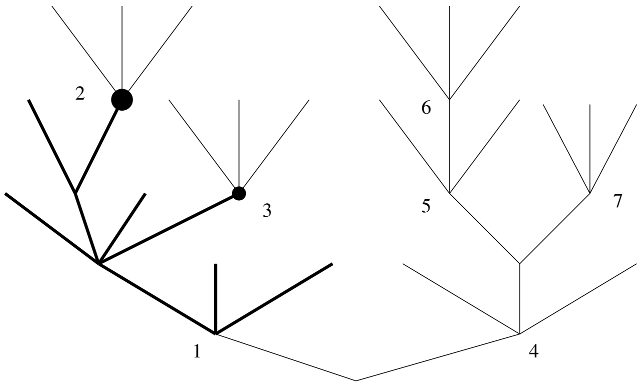

Figure 2: The sub-tree $S_1^{\mathcal{A}}(\mathbf{t})$ in bold for $\mathcal{A} = \{3\}$, and the elements of $C_1^{\mathcal{A}}(\mathbf{t})$.

We set

$$\tilde{S}_\emptyset^{\mathcal{A}}(\mathbf{t}) = \begin{cases} S_\emptyset^{\mathcal{A}}(\mathbf{t}) & \text{if } \emptyset \notin \mathcal{L}_{\mathcal{A}}(\mathbf{t}) \\ \{\emptyset\} & \text{if } \emptyset \in \mathcal{L}_{\mathcal{A}}(\mathbf{t}) \end{cases}$$

and we set $\tilde{C}_\emptyset^{\mathcal{A}}(\mathbf{t})$ the set of leaves of $\tilde{S}_\emptyset^{\mathcal{A}}(\mathbf{t})$ that belong to $\mathcal{L}_{\mathcal{A}}(\mathbf{t})$.

Let $\tilde{N}_\emptyset(\mathbf{t}) = \mathrm{Card}\ (\tilde{C}_\emptyset^{\mathcal{A}}(\mathbf{t}))$. Then the range of $\tilde{\phi}$ belongs to $\mathbb{T}_0^{\tilde{N}_\emptyset(\mathbf{t})}$. Moreover if $u_1 < u_2 < \cdots < u_{\tilde{N}_\emptyset(\mathbf{t})}$ are the elements of $\tilde{C}_\emptyset^{\mathcal{A}}(\mathbf{t})$ ranked in lexicographic order, we set for every $1 \le i \le \tilde{N}_\emptyset(\mathbf{t})$

$$\tilde{\phi}(u_i) = \emptyset^{(i)}$$

where $\emptyset^{(i)}$ denotes the root of the i-th tree in $\mathbb{T}_0^{\tilde{N}_\emptyset(\mathbf{t})}$.

We then construct $\tilde{\phi}$ recursively: if $u \in \mathcal{L}_{\mathcal{A}}(\mathbf{t})$ and $\tilde{\phi}(u) = v^{(i)}$ (which is an element of the i-th tree), then we denote by $u_1 < \cdots < u_k$ the elements of $C_u^{\mathcal{A}}(\mathbf{t})$ ranked in lexicographic order and we set for $1 \le j \le k$

$$\tilde{\phi}(u_j) = vj^{(i)}.$$

Finally, we set $\mathcal{F}_{\mathcal{A}}(\mathbf{t}) = \tilde{\phi}(\mathbf{t})$.

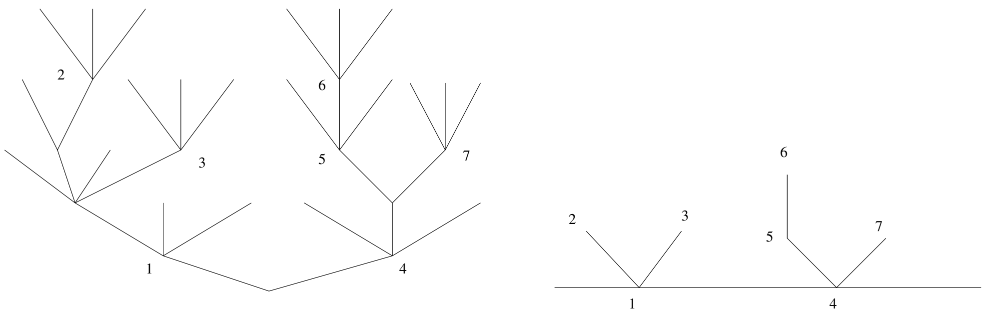

Figure 3: A tree $\mathbf{t}$ and the forest $\mathcal{F}_{\mathcal{A}}(\mathbf{t})$ for $\mathcal{A} = \{3\}$.

Let τ be a Galton-Watson tree with offspring distribution p. Let us describe the distribution of $\mathcal{F}_{\mathcal{A}}(\tau)$.

We define the offspring distribution $\tilde{p}$ by

$$\begin{cases}\tilde{p}(k) = p(k)\mathbf{1}_{\{k\in\mathcal{A}^c\}} & \text{for } k \geq 1,\\ \tilde{p}(0) = p(0) + p(\mathcal{A}). \end{cases}$$

Then $\tilde{S}_{\emptyset}^{\mathcal{A}}(\tau)$ is distributed as a (sub-critical) GW tree with offspring distribution $\tilde{p}$. In particular, if we denote by L the number of leaves of $\tilde{S}_{\emptyset}^{\mathcal{A}}(\tau)$, then we have

$$\mathbb{E}[L] = \frac{p(0) + p(\mathcal{A})}{1 - \mathbb{E}[X\mathbf{1}_{\{X\in\mathcal{A}^c\}}]}$$

where X is a random variable distributed according to p. Moreover, conditionally given L, the random variable $N := N_{\emptyset}(\tau)$ has a binomial distribution with parameter $(L, p(\mathcal{A})/(p(0) + p(\mathcal{A})))$.

In that case, let $\hat{X}^{\mathcal{A}}$ be the random variable

$$\hat{X}^{\mathcal{A}} = \sum_{k=1}^{Z'} N_k \tag{7.1}$$

where Z' is distributed as X conditionally given $\{X \in \mathcal{A}\}$ and $(N_k, k \in \mathbb{N})$ is a sequence of independent random variables, independent of Z', and distributed as N. We denote by $\hat{p}^{\mathcal{A}}$ the law of $\hat{X}^{\mathcal{A}}$. Then the forest $\mathcal{F}_{\mathcal{A}}(\tau)$ is distributed as N independent GW trees with offspring distribution $\hat{p}^{\mathcal{A}}$.

Lemma 7.1. *Assume that p statisfies (1.1), $\mu(p) < 1$. Then $\hat{p}^{\mathcal{A}}$ satisfies (1.1), $\mu(\hat{p}^{\mathcal{A}}) < 1$ and $\rho(\hat{p}^{\mathcal{A}}) = 1$ if $\rho(p) = 1$.*

Proof. Assume that p statisfies (1.1), $\mu(p) < 1$. Elementary computation gives:

$$\mathbb{E}\left[\hat{X}^{\mathcal{A}}\right] = \frac{\mathbb{E}\left[X\mathbf{1}_{\{X\in\mathcal{A}\}}\right]}{1 - \mathbb{E}\left[X\mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]}.$$

Since $\mu(p) < 1$, we get $\mu(\hat{p}^{\mathcal{A}}) < 1$.

Let $k \in \mathcal{A}$ such that $p(k) > 0$. We have:

$$\hat{p}^{\mathcal{A}}(0) \geq \frac{p(k)}{p(\mathcal{A})}\mathbb{P}(N = 0)^k > 0.$$

If $p(\mathcal{A} \cap \{1\}^c) > 0$, then choose $k > 1$, $k \in \mathcal{A}$ such that $p(k) > 0$. we have:

$$\hat{p}^{\mathcal{A}}(k) \geq \frac{p(k)}{p(\mathcal{A})}\mathbb{P}(N = 1)^k > 0.$$

If $p(\mathcal{A} \cap \{1\}^c) = 0$, then $\hat{X}^{\mathcal{A}}$ is distributed as N. Furthermore $\mathbb{P}(L > 1) > 1$ as p satisfies (1.1). Therefore, we have $\mathbb{P}(\hat{X}^{\mathcal{A}} = 2) = \mathbb{P}(N = 2) > 0$. We deduce that $\hat{p}^{\mathcal{A}}$ satisfies (1.1).

Let $\hat{g}^{\mathcal{A}}$ be the generating function of $\hat{X}^{\mathcal{A}}$. Set $\varphi(z) = (p(0) + p(\mathcal{A})z)/(p(0) + p(\mathcal{A}))$. For $z > 1$, we have $\varphi(z) > 1$ and:

$$\begin{aligned}\hat{g}^{\mathcal{A}}(z) = \mathbb{E}\left[\mathbb{E}\left[\varphi(z)^L\right]^{Z'}\right] &= \frac{1}{p(\mathcal{A})}\,\mathbb{E}\left[\mathbb{E}\left[\varphi(z)^L\right]^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]\\ &\geq \frac{1}{p(\mathcal{A})}\,\mathbb{E}\left[\mathbb{E}\left[\varphi(z)^X\mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right],\end{aligned}$$

where we used that L is stochastically larger than $X\mathbf{1}_{\{X\in\mathcal{A}^c\}}$ in the inequality. Since $\rho(p) = 1$, we deduce that $\hat{g}^{\mathcal{A}}(z) = +\infty$. This implies $\rho(\hat{p}^{\mathcal{A}}) = 1$. $\square$

7.2 Main result

We recall that $L_{\mathcal{A}}(\tau)$ is aperiodic since $0 \notin \mathcal{A}$, see the proof of Theorem 5.1 in [1].

Theorem 7.2. *Assume that p satisfies (1.1) and $\mu(p) < 1$ and $\rho(p) = 1$. We have that:*

$$\text{dist}\left(\tau \;\middle|\; L_{\mathcal{A}}(\tau) = n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p)\right), \tag{7.2}$$

as well as

$$\text{dist}\left(\tau \;\middle|\; L_{\mathcal{A}}(\tau) \geq n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p)\right). \tag{7.3}$$

Proof. It is enough to prove that for all $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$ and $k \in \mathbb{N}$:

$$\lim_{n\to+\infty} \mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k) \;\big|\; L_{\mathcal{A}}(\tau) = n) = \mathbb{P}\left(\tau^*(p) \in \mathbb{T}_+(\mathbf{t}, x, k)\right). \tag{7.4}$$

Set $M_0 = 0$ and $M_n = \sum_{k=1}^n N_k$ for $n \in \mathbb{N}^*$. Let $m = L_{\mathcal{A}}(\mathbf{t}) - \mathbf{1}_{\{k_x(\mathbf{t})\in\mathcal{A}\}}$ and $\ell = k_x(\mathbf{t})$. Recall (6.8). We have

$$\begin{aligned}
\mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k),\ & L_{\mathcal{A}}(\tau) = n) \\
&= D(\mathbf{t}, x) \sum_{j\geq \max(\ell+1,k)} p(j)\mathbb{P}_{j-\ell}(L_{\mathcal{A}}(\tau) = n_j - m) \\
&= D(\mathbf{t}, x) \sum_{j\geq \max(\ell+1,k)} p(j)\frac{j-\ell}{n_j - m}\mathbb{E}\left[N\mathbf{1}_{\{S_{n_j-m}+M_{j-1-\ell}+N=n_j-m\}}\right],
\end{aligned}$$

where we used Dwass formula (4.3) for the last equality where $S_n = \sum_{k=1}^n X_k$ with $(X_k, k \in \mathbb{N}^*)$ independent random variables distributed as $\hat{X}^{\mathcal{A}}$, see also (8.16). In particular, we have, using the definition of $B_{n-m,\ell}$ in (8.17),

$$\mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k) |\, L_{\mathcal{A}}(\tau) = n) = D(\mathbf{t}, x)\left(B_{n-m,\ell} - \sum_{j=\ell+1}^{k-1} p(j)(j-\ell)\, a_{n-m,j}\right), \tag{7.5}$$

with:

$$a_{n,j} = \frac{n}{n_j}\,\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n_j}+M_{j-1-\ell}+N=n_j\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}.$$

Notice that Lemma 8.6 implies that $\lim_{n\to+\infty} a_{n,j} = 1$. Then use Lemma 8.9 to get:

$$\begin{aligned}
\lim_{n\to+\infty} \mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t}, x, k) |\, L_{\mathcal{A}}(\tau) = n) &= D(\mathbf{t}, x)\left(1 - \ell + \mathbb{E}\left[(X-\ell)_+\mathbf{1}_{\{X\geq k\}}\right]\right) \\
&= \mathbb{P}(\tau^*(p) \in \mathbb{T}_+(\mathbf{t}, x, k)).
\end{aligned}$$

This ends the proof. ◻

Using Lemma 7.1, we easily get the following Corollary.

Corollary 7.3. *Assume that p satisfies (1.1), is non-generic for $\mathcal{A}$. Let $p^*_{\mathcal{A}} = p_{\mathcal{A},\rho(p)}$. We have that:*

$$\text{dist}\left(\tau \;\middle|\; L_{\mathcal{A}}(\tau) = n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p^*_{\mathcal{A}})\right),$$

as well as

$$\text{dist}\left(\tau \;\middle|\; L_{\mathcal{A}}(\tau) \geq n\right) \underset{n\to+\infty}{\longrightarrow} \text{dist}\left(\tau^*(p^*_{\mathcal{A}})\right).$$

This result with Proposition 4.6 and Corollary 5.7 in [1] for the generic case ends the proof of Theorem 1.3 for $0 \notin \mathcal{A}$ and gives a complete description of the asymptotic distribution of critical and sub-critical GW trees conditioned to have a large population.

8 Appendix

8.1 Strong ratio limit property

Let $(X_n, n \in \mathbb{N})$ be independent random variables taking values in $\mathbb{N}$ with distribution $p = (p(k), k \in \mathbb{N})$. We assume that:

$$\mu(p) = 1 \text{ or } (\mu(p) < 1 \text{ and, for all } \theta > 0,\ \mathbb{E}\left[e^{\theta X_1}\right] = +\infty). \tag{8.1}$$

Remark 8.1. Notice that if $\rho(p) = 1$, then we have $\mathbb{E}\left[e^{\theta X_1}\right] = +\infty$ for every positive θ.

Let $S_n = \sum_{k=1}^n X_k$. We assume that p is aperiodic (that is $\mathbb{P}(S_n = n) > 0$ for all n large enough). According to [10] or [16], we have the following strong ratio limit property for all $m, k \in \mathbb{Z}$:

$$\lim_{n\to+\infty} \frac{\mathbb{P}(S_{n-m} = n-k)}{\mathbb{P}(S_n = n)} = 1. \tag{8.2}$$

We deduce the following corollary. Recall the definition of δ_n^0 and δ_n^1 of (4.4) and (4.5).

Corollary 8.2. *Assume that p satisfies (8.1) and is aperiodic. For all $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, we have:*

$$\lim_{n\to+\infty} \delta_n^0(k,\ell) = \sum_{j\geq k} p(j). \tag{8.3}$$

and

$$\lim_{n\to+\infty} \delta_n^1(k,\ell) = 1 - \mu(p) + \sum_{j\geq k} jp(j). \tag{8.4}$$

Proof. Since $\mathbb{P}(S_{n+1} = n+\ell) = \sum_{j\in\mathbb{N}} p(j)\ \mathbb{P}(S_n = n+\ell-j)$, we have:

$$\delta_n^0(k,\ell) = \frac{\mathbb{P}(S_{n+1} = n+\ell)}{\mathbb{P}(S_n = n)} - \sum_{j<k} p(j)\ \frac{\mathbb{P}(S_n = n+\ell-j)}{\mathbb{P}(S_n = n)}.$$

Then use (8.2) to get (8.3).

Notice that, by exchangeability:

$$\sum_{j\in\mathbb{N}} jp(j)\ \mathbb{P}(S_n = n+\ell-j) = \mathbb{E}\left[X_1 \mathbf{1}_{\{S_{n+1}=n+\ell\}}\right] = \frac{n+\ell}{n+1}\mathbb{P}(S_{n+1} = n+\ell).$$

Thus we have:

$$\delta_n^1(k,\ell) = \frac{n+\ell}{n+1}\frac{\mathbb{P}(S_{n+1} = n+\ell)}{\mathbb{P}(S_n = n)} - \sum_{j<k} jp(j)\ \frac{\mathbb{P}(S_n = n+\ell-j)}{\mathbb{P}(S_n = n)}.$$

Then use (8.2) to get:

$$\lim_{n\to+\infty} \delta_n^1(k,\ell) = 1 - \sum_{j<k} jp(j).$$

Since $1 - \sum_{j<\ell} jp(j) = 1 - \mu(p) + \sum_{j\geq\ell} jp(j)$, this gives (8.4). □

We end this Section with a generalization of the strong ratio limit property.

Lemma 8.3. *Assume that p satisfies (8.1) and is aperiodic. Let N be an integer valued random variable independent of $(X_n, n \in \mathbb{N})$ such that $0 < \mathbb{E}[N] < +\infty$. Then for all $m, k \in \mathbb{Z}$, we have:*

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n-m}+N=n-k\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = 1. \tag{8.5}$$

Note that if p is periodic, then (8.5) still holds along the sub-sequence for which the denominator is positive.

Proof. We shall mimic the proof of the strong ratio limit property provided in [16]. Since p is aperiodic, the denominator of (8.5) is positive for n large enough and it is enough to prove the result for $m = 1$ and k such that $p(k) > 0$. Denote $\bar{p}_n(k) = \sum_{i=1}^n \mathbf{1}_{\{X_i=k\}}/n$. We have:

$$\mathbb{E}\left[N\bar{p}_n(k)\mathbf{1}_{\{S_n+N=n\}}\right] = \mathbb{E}\left[N\mathbf{1}_{\{X_n=k\}}\mathbf{1}_{\{S_n+N=n\}}\right] = p(k)\mathbb{E}\left[N\mathbf{1}_{\{S_{n-1}+N=n-k\}}\right].$$

Therefore, we have

$$\begin{aligned}
\left|\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n-m}+N=n-k\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} - 1\right| &= \left|\frac{\mathbb{E}\left[N\bar{p}_n(k)\mathbf{1}_{\{S_n+N=n\}}\right]}{p(k)\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} - 1\right| \\
&\leq \frac{\mathbb{E}\left[N|\bar{p}_n(k) - p(k)|\mathbf{1}_{\{S_n+N=n\}}\right]}{p(k)\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} \\
&\leq \frac{\varepsilon}{p(k)} + \frac{2\mathbb{E}\left[N\mathbf{1}_{\{|\bar{p}_n(k)-p(k)|>\varepsilon\}}\mathbf{1}_{\{S_n+N=n\}}\right]}{p(k)\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}
\end{aligned}$$

for every $\varepsilon > 0$.

The proof will be complete as soon as we prove that:

$$J_n = \frac{\mathbb{E}\left[N\mathbf{1}_{\{|\bar{p}_n(k)-p(k)|>\varepsilon\}}\mathbf{1}_{\{S_n+N=n\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}$$

converges to 0 for all $\varepsilon > 0$. Notice that:

$$J_n \leq \frac{\mathbb{E}\left[N\mathbf{1}_{\{|\bar{p}_n(k)-p(k)|>\varepsilon\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = \frac{\mathbb{P}(|\bar{p}_n(k) - p(k)| > \varepsilon)}{\mathbb{P}(S_n = n)} \frac{\mathbb{E}[N]\mathbb{P}(S_n = n)}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}.$$

According to [7], Eq. (11.15), or to [16], end of page 2954, we have $\lim_{n\to+\infty} \mathbb{P}(|\bar{p}_n(k) - p(k)| > \varepsilon)/\mathbb{P}(S_n = n) = 0$. By Fatou and using the strong ratio limit property, we have:

$$\limsup_{n\to+\infty} \frac{\mathbb{E}[N]\mathbb{P}(S_n = n)}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} \leq 1.$$

Since $\varepsilon > 0$ is arbitrary, we deduce that $\lim_{n\to+\infty} J_n = 0$. □

8.2 Generalization of the strong ratio limit property I

Assume that p satisfies (8.1) and is aperiodic. Let X be a random variable taking values in $\mathbb{N}$ with distribution p. Recall that g denotes the generating function of p.

Let $\mathcal{A} \subset \mathbb{N}$ such that $0 \in \mathcal{A}$. Let $p^{\mathcal{A}}$ be the distribution on $\mathbb{N}$ with generating function $g^{\mathcal{A}}$ given by (6.2) and $X^{\mathcal{A}}$ distributed according to $p^{\mathcal{A}}$. Recall $\mu(p^{\mathcal{A}})$ is given by (6.3). In particular $\mu(p) = 1$ (resp. $\mu(p) \leq 1$) implies $\mu(p^{\mathcal{A}}) = 1$ (resp. $\mu(p^{\mathcal{A}}) \leq 1$). And from the proof of Lemma 6.1, we get that $\mathbb{E}\left[e^{\theta X}\right] = +\infty$ for all $\theta > 0$ implies that $\mathbb{E}\left[e^{\theta X_{\mathcal{A}}}\right] = +\infty$ for all $\theta > 0$.

Let $(X_n, n \in \mathbb{N})$ be independent random variables, independent of X, taking values in $\mathbb{N}$ with distribution $p^{\mathcal{A}}$. Let $S_n = \sum_{k=1}^n X_k$. We assume that $p^{\mathcal{A}}$ is aperiodic (that is $\mathbb{P}(S_n = n) > 0$ for all n large enough). In particular the strong ratio limit property (8.2) holds as well as (8.3) and (8.4) hold with p replaced by $p^{\mathcal{A}}$.

Recall (6.8), that is $n_j = n - \mathbf{1}_{\{i\in\mathcal{A}\}}$, and let:

$$\delta_n^{0,\mathcal{A}}(k,\ell) = \frac{1}{\mathbb{P}(S_n = n)} \sum_{j\geq k} p(j)\ \frac{n}{n_j}\mathbb{P}(S_{n_j} = n_j + \ell - j) \tag{8.6}$$

and

$$\delta_n^{1,\mathcal{A}}(k,\ell) = \frac{1}{\mathbb{P}(S_n = n)} \sum_{j\geq k} jp(j)\ \frac{n}{n_j} \mathbb{P}(S_{n_j} = n_j + \ell - j). \tag{8.7}$$

We stress that in (4.4) and (4.5), $(S_n, n \in \mathbb{N})$ is a random walk with increments distributed according to p; whereas in (8.6) and (8.7), $(S_n, n \in \mathbb{N})$ is a random walk with increments distributed according to $p^{\mathcal{A}}$.

Lemma 8.4. *Assume that p satisfies (8.1) and $p^{\mathcal{A}}$ is aperiodic. For all $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, we have:*

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right]}{\mathbb{P}(S_n = n)} = 1, \tag{8.8}$$

$$\lim_{n\to+\infty} \delta_n^{0,\mathcal{A}}(k,\ell) = \sum_{j\geq k} p(j) \tag{8.9}$$

and

$$\lim_{n\to+\infty} \delta_n^{1,\mathcal{A}}(k,\ell) = 1 - \mu(p) + \sum_{j\geq k} jp(j). \tag{8.10}$$

Proof. We define:

$$a_n(j) = p(j) \frac{\mathbb{P}(S_{n_j} = n_j + \ell - j)}{\mathbb{P}(S_n = n)} \frac{n}{n_j}$$

as well as

$$b_n(j) = p^{\mathcal{A}}(j) \frac{\mathbb{P}(S_{n-1} = n + \ell - j - 1)}{\mathbb{P}(S_n = n)} + \frac{p^{\mathcal{A}}(j-1)}{p(0)} \frac{\mathbb{P}(S_n = n + \ell - j)}{\mathbb{P}(S_n = n)},$$

with the convention that $p^{\mathcal{A}}(-1) = 0$. Notice that, as $\mu(p) \leq 1$, we have $p(0) > 0$.

Thanks to Lemma 6.1, $p^{\mathcal{A}}$ satisfies (8.1). Using the strong ratio limit property (that is (8.2) with $p^{\mathcal{A}}$ instead of p), we have $\lim_{n\to+\infty} a_n(j) = p(j)$ and $\lim_{n\to+\infty} b_n(j) = p^{\mathcal{A}}(j) + p^{\mathcal{A}}(j-1)/p(0)$. We have:

$$\sum_{j\in\mathbb{N}} b_n(j) = \frac{\mathbb{P}(S_n = n + \ell - 1)}{\mathbb{P}(S_n = n)} + \frac{1}{p(0)} \frac{\mathbb{P}(S_{n+1} = n + \ell + 1)}{\mathbb{P}(S_n = n)}.$$

We deduce from the strong ratio limit property (that is (8.2) with $p^{\mathcal{A}}$ instead of p) that:

$$\lim_{n\to+\infty} \sum_{j\in\mathbb{N}} b_n(j) = 1 + \frac{1}{p(0)} = \sum_{j\in\mathbb{N}} \lim_{n\to+\infty} b_n(j).$$

Then use (6.4) and (6.5) to get that $a_n(j) \leq 2b_n(j)$ for $n \geq 2$. Then use the dominated convergence theorem (see Theorem 1.21 in [9]) to get that:

$$\lim_{n\to+\infty} \sum_{j\in\mathbb{N}} a_n(j) = \sum_{j\in\mathbb{N}} \lim_{n\to+\infty} a_n(j) = 1.$$

Notice that $\sum_{j\in\mathbb{N}} a_n(j) = \mathbb{E}\left[\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right] / \mathbb{P}(S_n = n)$ to deduce that (8.8) holds. Since $\delta_n^{0,\mathcal{A}}(k,\ell) = \sum_{j\geq k} a_n(j)$, the proof of (8.9) is then similar to the proof of (8.3).

Set $c_n(\ell) = \delta_n^{1,\mathcal{A}}(0,\ell)$ that is:

$$c_n(\ell) = \frac{\mathbb{E}\left[\frac{n}{n_X} X\mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right]}{\mathbb{P}(S_n = n)}.$$

According to Lemma 8.5 below, (8.2) and (8.8), we have that $\lim_{n\to+\infty} c_n(\ell) = 1$ for all $\ell \in \mathbb{Z}$. Then arguing as in the proof of (8.4), we easily get (8.10). $\square$

Lemma 8.5. *For all $\ell \in \mathbb{Z}$, $n \geq 2$, we have:*

$$\mathbb{E}\left[\frac{n}{n_X} X \mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right] = \ell \mathbb{E}\left[\frac{n}{n_X} \mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right] - (\ell-1)\mathbb{P}(S_n = n+\ell-1). \tag{8.11}$$

Proof. We first prove (8.11) for $\ell \leq 0$. Let $k \geq 1$. By decomposing according to the number of children of the root of the first tree in the forest, we have:

$$\mathbb{P}_k(|\tau^{\mathcal{A}}| = n) = \sum_{j \in \mathbb{N}} p(j) \mathbb{P}_{j+k-1}(|\tau^{\mathcal{A}}| = n_j),$$

with the convention that $\mathbb{P}_0(\cdot) = 0$. Then using Dwass formula (4.3) in each side of this equality, we get:

$$k\mathbb{P}(S_n = n-k) = \mathbb{E}\left[\frac{n}{n_X}(X+k-1)\mathbf{1}_{\{X+S_{n_X}=n_X-k+1\}}\right].$$

Take $\ell = 1-k$ to get that (8.11) holds for $\ell \leq 0$.

Unfortunately, we didn't get a similar proof for $\ell \geq 1$ and we prove (8.11) for $\ell \geq 1$ by induction. Let $\ell \geq 0$. Assume that (8.11) holds for all $\ell' \leq \ell$ and all $n \geq 2$, and let us prove it holds for $\ell+1$ and all $n \geq 2$. We have:

$$\mathbb{E}\left[\frac{n+1}{n_X+1} X \mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right] = A_1 + \mathbb{E}\left[\frac{n_X-n}{n_X(n_X+1)} X \mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right], \tag{8.12}$$

with

$$A_1 = \mathbb{E}\left[\frac{n}{n_X} X \mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right].$$

Using (8.11), we have:

$$\begin{aligned}
A_1 &= \sum_{j \in \mathbb{N}} p^{\mathcal{A}}(j) \mathbb{E}\left[\frac{n}{n_X} X \mathbf{1}_{\{X+S_{n_X}=n_X+1+\ell-j\}}\right] \\
&= p^{\mathcal{A}}(0)\mathbb{E}\left[\frac{n}{n_X} X \mathbf{1}_{\{X+S_{n_X}=n_X+1+\ell\}}\right] \\
&\qquad + \sum_{j \in \mathbb{N}^*} p^{\mathcal{A}}(j)\left((\ell+1-j)\mathbb{E}\left[\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X}=n_X+\ell+1-j\}}\right] - (\ell-j)\mathbb{P}(S_n = n+\ell-j)\right).
\end{aligned}$$

So we have:

$$A_1 = p^{\mathcal{A}}(0)A_2 + A_3 - \mathbb{E}\left[(\ell - X_1)\mathbf{1}_{\{S_{n+1}=n+\ell\}}\right], \tag{8.13}$$

with

$$A_2 = \mathbb{E}\left[\frac{n}{n_X} X \mathbf{1}_{\{X+S_{n_X}=n_X+1+\ell\}}\right] - (\ell+1)\mathbb{E}\left[\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X}=n_X+\ell+1\}}\right] + \ell\mathbb{P}(S_n = n+\ell) \tag{8.14}$$

and

$$A_3 = \mathbb{E}\left[(\ell+1-X_1)\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right].$$

We compute the last term of (8.13). We have:

$$\mathbb{E}\left[(\ell - X_1)\mathbf{1}_{\{S_{n+1}=n+\ell\}}\right] = \mathbb{E}\left[\left(\ell - \frac{S_{n+1}}{n+1}\right)\mathbf{1}_{\{S_{n+1}=n+\ell\}}\right] = \frac{n}{n+1}(\ell-1)\mathbb{P}(S_{n+1} = n+\ell).$$

We compute A_3:

$$\begin{aligned} A_3 &= \mathbb{E}\left[\left(\ell+1-\frac{S_{n_X+1}}{n_X+1}\right)\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right]\\ &= \mathbb{E}\left[\left(\ell+1-\frac{n_X+1+\ell-X}{n_X+1}\right)\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right]\\ &= \ell\mathbb{E}\left[\frac{n}{n_X+1}\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right]+\mathbb{E}\left[\frac{n}{n_X(n_X+1)}\,X\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right]. \end{aligned}$$

Plugging the result in (8.12), we get:

$$\begin{aligned} &\mathbb{E}\left[\frac{n+1}{n_X+1}\,X\mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right]\\ &\qquad= p^{\mathcal{A}}(0)A_2+\ell\mathbb{E}\left[\frac{n}{n_X+1}\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right]\\ &\qquad\qquad+\mathbb{E}\left[\frac{1}{n_X+1}\,X\mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right]-\frac{n}{n+1}(\ell-1)\mathbb{P}(S_{n+1}=n+\ell). \end{aligned}$$

We obtain, using that $(n+1)_X = n_X+1$ and (8.11) with $n+1$ instead of n:

$$\begin{aligned} p^{\mathcal{A}}(0)A_2 &= \frac{n}{n+1}\mathbb{E}\left[\frac{n+1}{n_X+1}\,X\mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right]-\frac{\ell n}{n+1}\mathbb{E}\left[\frac{n+1}{n_X+1}\mathbf{1}_{\{X+S_{n_X+1}=n_X+\ell+1\}}\right]\\ &\quad+\frac{n}{n+1}(\ell-1)\mathbb{P}(S_{n+1}=n+\ell)\\ &= 0. \end{aligned}$$

Recall (8.14). The fact that $A_2=0$ gives exactly that (8.11) holds with ℓ replaced by $\ell+1$. This proves the induction and ends the proof of the lemma. □

8.3 Generalization of the strong ratio limit property II

We use notations from Sections 7.1 and 7.2, see in particular (7.1) and thereafter for the definitions of $\hat{X}^{\mathcal{A}}$, $\hat{p}^{\mathcal{A}}$ and N. We assume that $(N_k, k\in\mathbb{N}^*)$ are independent random variables distributed as N, $(X_k, k\in\mathbb{N}^*)$ are independent random variables distributed as $\hat{X}^{\mathcal{A}}$, and that the two sequences are independent. Let $M_0=0$, $S_0=0$ and for $n\in\mathbb{N}^*$:

$$M_n=\sum_{k=1}^n N_k, \qquad S_n=\sum_{k=1}^n X_k.$$

We have the following result.

Lemma 8.6. *Assume $\hat{p}^{\mathcal{A}}$ is aperiodic, with $\mu(\hat{p}^{\mathcal{A}})<1$ and $\rho(\hat{p}^{\mathcal{A}})=1$. Let $m\in\mathbb{N}$ and $k\in\mathbb{Z}$, we have:*

$$\lim_{n\to+\infty}\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N+M_m=n-k\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}=1.$$

Proof. Let

$$c_{n,\ell}=\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n-\ell-k\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}.$$

Denote by $q=(q(\ell),\ell\in\mathbb{N})$ the distribution of M_k and by $r=(r(\ell),\ell\in\mathbb{N})$ the distribution of S_m. We have, thanks to Lemma 8.3, that $\lim_{n\to+\infty}c_{n,\ell}=1$ and:

$$\lim_{n\to+\infty}\sum_{\ell\in\mathbb{N}}r(\ell)c_{n,\ell}=\lim_{n\to+\infty}\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n+m}+N=n-k\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}=1=\sum_{\ell\in\mathbb{N}}r(\ell)\lim_{n\to+\infty}c_{n,\ell}.$$

Let j_0 such that $\mathbb{P}(X_1 = j_0) > 0$. Notice that:

$$r(\ell) = \mathbb{P}(S_m = \ell) \geq \mathbb{P}(X_1 + \ldots + X_m = mj_0, M_m = \ell, N_{m+1} + \ldots N_{mj_0} = 0).$$

We deduce that there exists $c > 0$ such that $q(\ell) \leq Cr(\ell)$ for all $\ell \in \mathbb{N}$. By dominated convergence (see Theorem 1.21 in [9]), we deduce that $\lim_{n\to+\infty} \sum_{\ell\in\mathbb{N}} q(\ell)c_{n,\ell} = \sum_{\ell\in\mathbb{N}} q(\ell) \lim_{n\to+\infty} c_{n,\ell} = 1$. □

Let p_N be the distribution of N. We have, using the decomposition of the GW tree with respect to the descendants of $\emptyset$ in $\mathcal{F}_\mathcal{A}(\tau)$ and Dwass formula (4.3):

$$\mathbb{P}(L_\mathcal{A}(\tau) = n) = \sum_{j\in\mathbb{N}} p_N(j)\mathbb{P}_j(|\tau^\mathcal{A}| = n) = \frac{1}{n}\,\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]. \tag{8.15}$$

More generally, we have

$$\mathbb{P}_j(L_\mathcal{A}(\tau) = n) = \frac{1}{n}\,\mathbb{E}\left[M_j\mathbf{1}_{\{S_n+M_j=n\}}\right] = \frac{j}{n}\mathbb{E}\left[N\mathbf{1}_{\{S_n+M_{j-1}+N=n\}}\right], \tag{8.16}$$

with N independent of S_n and M_{j-1}.

We set for $\ell \in \mathbb{Z}$:

$$B_{n,\ell} = \sum_{j>\ell} p(j)(j-\ell)\frac{n}{n_j}\,\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n_j}+M_{j-1-\ell}+N=n_j\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}. \tag{8.17}$$

The next lemma is the analogue of Lemma 8.5 in our current setting.

Lemma 8.7. *Assume $\hat{p}^\mathcal{A}$ is aperiodic, with $\mu(\hat{p}^\mathcal{A}) < 1$ and $\rho(\hat{p}^\mathcal{A}) = 1$. For $\ell \leq 0$, we have $\lim_{n\to+\infty} B_{n,\ell} = 1 - \ell$.*

Proof. Recall that $\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right] = n\,\mathbb{P}(L_\mathcal{A}(\tau) = n)$. Let $k \geq 0$. By decomposing τ under $\mathbb{P}_{k+1}$ with respect to the number of children of the first tree in the forest, we get:

$$\begin{aligned}
\mathbb{P}_{k+1}(L_\mathcal{A}(\tau) = n) &= \sum_{j\in\mathbb{N}} p(j)\,\mathbb{P}_{k+j}(L_\mathcal{A}(\tau) = n_j)\\
&= \sum_{j\in\mathbb{N}} p(j)\,\frac{k+j}{n_j}\mathbb{E}\left[N\mathbf{1}_{\{S_{n_j}+M_{k+j-1}+N=n_j\}}\right]\\
&= B_{n,-k}\frac{1}{n}\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right].
\end{aligned}$$

Then use (8.16) and Lemma 8.6 to deduce that:

$$\lim_{n\to+\infty} \frac{n\,\mathbb{P}_{k+1}(L_\mathcal{A}(\tau) = n)}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = k+1.$$

This gives the lemma. □

In order to extend Lemma 8.7 in a weaker form for $\ell > 0$, we give a preliminary lemma. Set for $\ell \geq k$, $\ell, k \in \mathbb{Z}$:

$$C_{n,\ell}(k) = \mathbb{E}\left[\frac{n}{n_X}\,N(X-\ell)_+\mathbf{1}_{\{S_{n_X}+M_{X-k-1}+N=n_X\}}\right].$$

Notice that for $\ell \in \mathbb{Z}$:

$$C_{n,\ell}(\ell) = nB_{n,\ell}\mathbb{P}(L_\mathcal{A}(\tau) = n). \tag{8.18}$$

We define $z_+ = \max(z, 0)$.

Lemma 8.8. *Assume $\hat{p}^{\mathcal{A}}$ is aperiodic, with $\mu(\hat{p}^{\mathcal{A}}) < 1$ and $\rho(\hat{p}^{\mathcal{A}}) = 1$. We have for $k \in \mathbb{Z}$ such that $k \leq \ell$:*

$$\lim_{n\to+\infty} \frac{C_{n,\ell}(k)}{C_{n,\ell}(\ell)} = 1.$$

Proof. Notice that $nN(X-\ell)_+/n_X$ is integrable. Mimicking the proof of Lemma 8.3 and using that n_X takes only two possible values a.s., we get for $m, k \in \mathbb{Z}$:

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[\frac{n}{n_X} N(X-\ell)_+ \mathbf{1}_{\{S_{n_X-m}+M_{X-1-\ell}+N=n_X-k\}}\right]}{\mathbb{E}\left[\frac{n}{n_X} N(X-\ell)_+ \mathbf{1}_{\{S_{n_X}+M_{X-1-\ell}+N=n_X\}}\right]} = 1.$$

Then mimicking the proof of Lemma 8.6, we get for $m \in \mathbb{N}$ and $k \in \mathbb{Z}$:

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[\frac{n}{n_X} N(X-\ell)_+ \mathbf{1}_{\{S_{n_X}+M_{X-1-\ell+m}+N=n_X-k\}}\right]}{\mathbb{E}\left[\frac{n}{n_X} N(X-\ell)_+ \mathbf{1}_{\{S_{n_X}+M_{X-1-\ell}+N=n_X\}}\right]} = 1.$$

Then take $m = \ell - k \geq 0$ to get the result. ☐

Lemma 8.9. *Assume $\hat{p}^{\mathcal{A}}$ is aperiodic, with $\mu(\hat{p}^{\mathcal{A}}) < 1$ and $\rho(\hat{p}^{\mathcal{A}}) = 1$. For $\ell > 0$, we have:*

$$\lim_{n\to+\infty} B_{n,\ell} = 1 - \mu + \mathbb{E}\left[(X-\ell)_+\right].$$

Proof. Let $\ell \geq -1$. We have:

$$C_{n,\ell}(-1) = C_{n,0}(-1) - \sum_{j=0}^{\ell-1} p(j)(j-\ell)\mathbb{E}\left[\frac{n}{n_j} N\mathbf{1}_{\{S_{n_j}+M_j+N=n_j\}}\right] - \ell\mathbb{E}\left[\frac{n}{n_X} N\mathbf{1}_{\{S_{n_X}+M_X+N=n_X\}}\right], \quad (8.19)$$

with the convention that $\sum_\emptyset = 0$. Recall that $\lim_{n\to+\infty} B_{n,-1} = 2$ and $\lim_{n\to+\infty} B_{n,0} = 1$, thanks to Lemma 8.7 and thus (8.18) implies that:

$$C_{n,-1}(-1) \sim 2\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right] \text{ and } C_{n,0}(0) \sim \mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right].$$

We deduce from Lemma 8.8 that

$$\lim_{n\to+\infty} \frac{C_{n,0}(-1)}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = \lim_{n\to+\infty} \frac{C_{n,0}(-1)}{C_{n,0}(0)} = 1.$$

We deduce from (8.19) with $\ell = -1$ and Lemma 8.6 that:

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[\frac{n}{n_X} N\mathbf{1}_{\{S_{n_X}+M_X+N=n_X\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = 1. \quad (8.20)$$

Let $\ell \geq 1$. We deduce from (8.19) with $\ell \geq 1$, (8.18), (8.15), Lemma 8.6 and (8.20) that:

$$\lim_{n\to+\infty} B_{n,\ell} = 1 - \sum_{j=0}^{\ell-1} p(j)(j-\ell) - \ell = 1 - \mu + \mathbb{E}\left[(X-\ell)_+\right].$$

☐

Acknowledgments. The authors wish to thank an anonymous referee whose precise comments improved considerably the presentation of that paper

References

[1] R. ABRAHAM and J.-F. DELMAS. Local limits of conditioned galton-watson trees: the infinite spine case. *Elec. J. of Probab.*, 19:1–19, 2014.

[2] D. ALDOUS. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993. MR-1207226

[3] P. BILLINGSLEY. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication. MR-1700749

[4] N. CURIEN and I. KORTCHEMSKI. Random non-crossing plane configurations: a conditioned Galton-Watson tree approach. *Random Struct. and Alg.*, To appear, 2014.

[5] M. DWASS. The total progeny in a branching process and a related random walk. *J. Appl. Probability*, 6:682–686, 1969. MR-0253433

[6] J. GEIGER and L. KAUFMANN. The shape of large Galton-Watson trees with possibly infinite variance. *Random Struct. and Alg.*, 25(3):311–335, 2004. MR-2086163

[7] S. JANSON. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012. MR-2908619

[8] T. JONNSSON and S. STEFANSSON. Condensation in nongeneric trees. *J. Stat. Phys.*, 142:277–313, 2011. MR-2764126

[9] O. KALLENBERG. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, 1997. MR-1464694

[10] J. G. KEMENY. A probability limit theorem requiring no moments. *Proc. Amer. Math. Soc.*, 10:607–612, 1959. MR-0107292

[11] D. KENNEDY. The Galton-Watson process conditioned on the total progeny. *J. Appl. Probability*, 12(4):800–806, 1975. MR-0386042

[12] H. KESTEN. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986. MR-0871905

[13] I. KORTCHEMSKI. Invariance principles for Galton–Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.*, 122(9):3126–3172, 2012. MR-2946438

[14] I. KORTCHEMSKI. Limit theorems for conditioned non-generic galton-watson trees. *Ann. de l'Inst. Henri Poincaré*, To appear, 2014.

[15] N. MIMAMI. On the number of vertices with a given degree in a Galton-Watson tree. *Adv. in Appl. Probab.*, 37(1):229–264, 2005. MR-2135161

[16] J. NEVEU. Sur le théorème ergodique de Chung-Erdős. *C. R. Acad. Sci. Paris*, 257:2953–2955, 1963. MR-0159362

[17] J. NEVEU. Arbres et processus de Galton-Watson. *Ann. de l'Inst. Henri Poincaré*, 22:199–207, 1986. MR-0850756

[18] D. RIZZOLO. Scaling limits of Markov branching trees and Galton-Watson trees conditioned on the number of vertices with out-degree in a given set. *Ann. de l'Inst. Henri Poincaré*, To appear, 2014.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])
- Non profit, sponsored by IMS[2], BS[3], PKP[4]
- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
[2]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[3]BS: Bernoulli Society http://www.bernoulli-society.org/
[4]PK: Public Knowledge Project http://pkp.sfu.ca/
[5]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm