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

Romain Abraham, Jean-François Delmas

LOCAL LIMITS OF CONDITIONED GALTON-WATSON TREES II: THE CONDENSATION CASE

ROMAIN ABRAHAM AND JEAN-FRANÇOIS DELMAS

Abstract. We provide a complete picture of the local convergence of critical or subcritical Galton-Watson tree 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 limit is a random tree with a node with infinite out-degree. This case corresponds to the so-called condensation phenomenon.

1. Introduction

Conditioning critical or sub-critical Galton-Watson (GW) trees comes from the seminal work of Kesten, [10]. Let $p = (p(n), n \in \mathbb{N})$ be an offspring distribution such that:

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

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 [10] the 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 [9] and Geiger and Kaufmann [5], large number of leaves see Curien and Kortchemski [3]. In [1], we generalized those previous results by conditioning the GW tree to have a large number of individuals whose number of offspring belongs to a set $\mathcal{A} \subset \mathbb{N}$. Let

$$p(\mathcal{A}) = \sum_{k \in \mathcal{A}} p(k). \tag{2}$$

If $p(\mathcal{A}) > 0$, then the limiting tree is again the same size-biased tree as for Kesten [10].

However, the results are different in the subcritical case. We first define for an offspring distribution p that satisfies (1) and a set $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$ a modified offspring distribution $p_{\mathcal{A},\theta}$ by:

$$\forall k \geq 0, \quad p_{\mathcal{A},\theta}(k) = \begin{cases} c_{\mathcal{A}}(\theta)\theta^k p(k) & \text{if } k \in \mathcal{A}, \\ \theta^{k-1} p(k) & \text{if } k \in \mathcal{A}^c, \end{cases} \tag{3}$$

Date: November 26, 2013.

2010 *Mathematics Subject Classification.* 60J80, 60B10.

Key words and phrases. Galton-Watson, random tree, condensation, non-extinction, branching process.

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

$$c_{\mathcal{A}}(\theta) = \frac{\theta - \mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]}{\theta\mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]}, \tag{4}$$

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 $p_{\theta^c_{\mathcal{A}},\mathcal{A}}$ is critical, then the GW tree τ conditioned on $L_{\mathcal{A}}(\tau)$ being large converges to the size-biased 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 [7] and Janson [6]: 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 subcritical GW tree τ conditioned on $\{L_{\mathcal{A}}(\tau) = n\}$ as n goes to infinity.

We summarize this complete description as follows. Let p be an offspring distribution that satisfies (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 tree considered in [10].
- 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 [7, 6].

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

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

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

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

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

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

$$p^*_{\mathcal{A}} = p_{\mathcal{A},\theta^*_{\mathcal{A}}}, \quad \text{with } \theta^*_{\mathcal{A}} = \max I_{\mathcal{A}}. \tag{5}$$

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 1.2. *Let τ be a GW tree with offspring distribution p which satisfies (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:*

$$\text{dist } (\tau \mid L_{\mathcal{A}}(\tau) = n) \underset{n\to+\infty}{\longrightarrow} \text{dist } (\tau^*(p^*_{\mathcal{A}})), \tag{6}$$

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

$$\text{dist } (\tau \mid L_{\mathcal{A}}(\tau) \geq n) \underset{n\to+\infty}{\longrightarrow} \text{dist } (\tau^*(p^*_{\mathcal{A}})). \tag{7}$$

The theorem has already been proven in the critical case and the subcritical generic case in [1]. We concentrate here on the case of the subcritical 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 subcritical 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 eventually for 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 subcritical offspring distribution p to be non-generic for all $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$, 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 an infinite node.

In Section 2, we recall the setting of the discrete trees (which is close to [1], but has to include discrete trees with infinite nodes). 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 (3) in Section 5. The proof of Theorem 1.2 is given in the following three sections. More precisely, the case $\mathcal{A} = \mathbb{N}$ is presented in Section 4. This provides an elementary and self-contained proof of the results from [7, 6]. 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 [11] or [14]. 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 [14]. 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 [13] 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 $u = (u_1, \ldots, u_n) \in \mathcal{U}$, we set $|u| = n$ the length of u and:

$$|u|_\infty = \max(|u|, (u_i, 1 \leq i \leq |u|))$$

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 u is the set:

$$A_u = \{v \in \mathcal{U}; \text{there exists } w \in \mathcal{U},\, w \neq \emptyset, \text{ such that } u = vw\}. \tag{8}$$

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 $u \in A_v$ 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 $1 \leq i \leq k_u(\mathbf{t})$.

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:

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

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:

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

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 $\bar{v}_n = (v_1, \ldots, v_n)$ 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 a tree consisting of 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,

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

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,

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

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:

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

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 $r_h(\mathbf{t}) = \{u \in \mathbf{t},\ |u| \leq h\}$. 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 $B_\infty(\mathbf{t}, 2^{-h}) = \{\mathbf{t}' \in \mathbb{T}_\infty; d_\infty(\mathbf{t}, \mathbf{t}') \leq 2^{-h}\}$ 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 [2] 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 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 $h \geq 0$), and we denote:

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

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 $v = (v_1, \dots, v_n) \in \mathcal{U}$, with $n > 0$, and $k \in \mathbb{N}$, we define the shift of v by k as $\theta(v, k) = (v_1 + k, v_2, \dots, v_n)$. 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:*

$$\lim_{n\to+\infty} \mathbb{P}(T_n \in \mathbb{T}_+(\mathbf{t}, x, k)) = \mathbb{P}(T \in \mathbb{T}_+(\mathbf{t}, x, k)) \quad \text{and} \quad \lim_{n\to+\infty} \mathbb{P}(T_n = \mathbf{t}) = \mathbb{P}(T = \mathbf{t}). \tag{9}$$

Remark 2.3. Let

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

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 (9) 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 [2], 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:

$$\mathbf{t} \in A \subset B_\infty(\mathbf{t}, 2^{-h}). \tag{10}$$

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 (10) 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 (9) holds for every $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathbf{t}$ and $k \in \mathbb{N}$. □

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). Let $g(z) = \sum_{k \in \mathbb{N}} p(k)\, z^k$ be the generating function of p. We denote by $\rho(p)$ its convergence radius 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 $(\mathcal{S}_1(\tau), \mathcal{S}_2(\tau), \ldots, \mathcal{S}_n(\tau))$ 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:

$$\forall \mathbf{t} \in \mathbb{T}_0, \quad \mathbb{P}(\tau = \mathbf{t}) = \prod_{u \in \mathbf{t}} p(k_u(\mathbf{t})). \tag{11}$$

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 $\tau^{(k)} = (\tau_1, \ldots, \tau_k)$ 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.** We say that the offspring distribution p is non-generic if g has convergence radius 1 and $\mu(p) = g'(1) < 1$. The corresponding GW tree is also called non-generic.

Assume that p satisfies (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 $\mathbb{T}_2$ if p is non-generic.

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

$$\mathbb{P}(\tau^*(p) \in \mathbb{T}_+(\mathbf{t}, x, k)) = D(\mathbf{t}, x) \left(1 - \mu(p) + \mathbb{E}\left[(X - k_x(\mathbf{t}))_+ \mathbf{1}_{\{X \geq k\}}\right]\right). \tag{12}$$

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

$$\mathbb{P}(\tau^*(p) \in \mathbb{T}(\mathbf{t}, x)) = \frac{\mathbb{P}(\tau = \mathbf{t})}{p(0)}. \tag{13}$$

Remark 3.2. 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. 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 $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathcal{L}_0(\mathbf{t})$,

$$\mathbb{P}(\tau^{\mathrm{S}}(p) \in \mathbb{T}(\mathbf{t}, x)) = \frac{\mathbb{P}(\tau = t)}{\mu(p)^{|x|} p(0)}. \tag{14}$$

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

We prove Theorem 1.2 for $\mathcal{A} = \mathbb{N}$ and p non-generic for $\mathbb{N}$. The results of this section appear already in [6] see also [7]. It is a special case of Theorem 1.2 with $\mathcal{A} = \mathbb{N}$. We provide here an elementary proof relying on the strong ratio limit property of random walks on the integers.

4.1. **The case $\rho(p)=1$.** We first consider the case $\rho(p)=1$ and $\mu(p)<1$.

Theorem 4.1. *Assume that p satisfies (1) and is non-generic for $\mathbb{N}$. We have that:*

$$\text{dist}\,(\tau \;\big|\; |\tau|=n) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p)), \tag{15}$$

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

$$\text{dist}\,(\tau \;\big|\; |\tau|\geq n) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p)). \tag{16}$$

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 $\rho(p)=1$. 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 [4]): for every $k\in\mathbb{N}^*$ and every $n\geq k$, we have

$$\mathbb{P}_k(|\tau|=n)=\frac{k}{n}\mathbb{P}(S_n=n-k). \tag{17}$$

Let τ_n be distributed as τ conditionally on $\{|\tau|=n\}$. Using Dwass formula (17), we have

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

We then set

$$\delta_n^0(k,\ell)=\frac{1}{\mathbb{P}(S_n=n)}\sum_{j\geq k}p(j)\ \mathbb{P}(S_n=n+\ell-j) \tag{18}$$

and

$$\delta_n^1(k,\ell)=\frac{1}{\mathbb{P}(S_n=n)}\sum_{j\geq k}jp(j)\ \mathbb{P}(S_n=n+\ell-j). \tag{19}$$

We get:

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

Then use the strong ratio limit property (44) as well as its consequences (45) and (46), to get that:

$$\lim_{n\to+\infty}\mathbb{P}(\tau_n\in\mathbb{T}_+(\mathbf{t},x,k))=D(\mathbf{t},x)\left(1-\mu(p)+\sum_{j\geq\max(\ell+1,k)}(j-\ell)p(j)\right). \tag{20}$$

Thanks to (12), 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 (15). Since dist $(\tau \;\big|\; |\tau| \geq n)$ is a mixture of dist $(\tau \;\big|\; |\tau| = k)$ for $k \geq n$, we deduce that (16) holds. □

Remark 4.2. The proof of (20) also holds if $\mu = 1$. In this case we get in particular that for all $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$:

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

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 3.2). Then use Remark 2.3 to get that dist $(\tau \,|\, |\tau| = n)$ converges to dist $(\tau^*(p))$.

4.2. **The case $\rho(p) > 1$.** We consider the case $\rho(p) > 1$. The offspring distribution $p_{\mathbb{N},\theta}$ of (3) has generating function:

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

Recall $I_{\mathbb{N}}$ is the set of positive θ for which $p_{\mathbb{N},\theta}$ is a well defined probability distribution. Furthermore, according to [9] (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 $\theta \in I_{\mathbb{N}}$. It is easy to check that $\mu(p_{\mathbb{N},\theta})$ is increasing in θ. Following [6], we shall say that p is non-generic for $\mathbb{N}$ if $\lim_{\theta\uparrow\rho(p)} \mu(p_{\mathbb{N},\theta}) < 1$. In that case, we have $I_{\mathbb{N}} = (0, \rho(p)]$ and $p^*_{\mathbb{N}}$ defined by (5) is $p^*_{\mathbb{N}} = p_{\mathbb{N},\rho(p)}$.

Corollary 4.3. *Assume that p satisfies (1) and is non-generic for $\mathbb{N}$. We have that:*

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

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

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

Proof. The first convergence is a direct consequence of (15) 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 (16). □

This result with Proposition 4.6 and Corollary 5.9 in [1] ends the proof of Theorem 1.2 for the case $\mathcal{A} = \mathbb{N}$ and 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) and let X be a random variable with distribution p. Recall $\rho(p)$ denotes the convergence radius of the generating function g of p. Let $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. We consider the modified distribution $p_{\mathcal{A},\theta}$ on $\mathbb{N}$ given by (3) 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:

$$\mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right] < +\infty \quad \text{and} \quad \mathbb{E}\left[\theta^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] \leq \theta. \tag{21}$$

In particular, $I_{\mathcal{A}}$ 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:

$$\theta^*_{\mathcal{A}} = \sup I_{\mathcal{A}} \in [1, \rho(p)]. \tag{22}$$

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

$$p_{\mathcal{A},\theta q} = (p_{\mathcal{A},\theta})_{\mathcal{A},q}\,. \tag{23}$$

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

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

And we have:

$$\mu(p_{\mathcal{A},\theta}) = \mathbb{E}\left[X\theta^{X-1}\mathbf{1}_{\{X\in\mathcal{A}^c\}}\right] + c_{\mathcal{A}}(\theta)\mathbb{E}\left[X\theta^X\mathbf{1}_{\{X\in\mathcal{A}\}}\right]. \tag{24}$$

Let:

$$\theta^c_{\mathcal{A}} = \inf\{\theta \in I_{\mathcal{A}}; \mu(p_{\mathcal{A},\theta}) = 1\}, \tag{25}$$

with the convention that $\inf\emptyset = +\infty$. Notice that the function $\theta \mapsto \mu(p_{\mathcal{A},\theta})$ is continuous over $I_{\mathcal{A}}$.

Lemma 5.1. *Let p be a distribution on $\mathbb{N}$ satisfying (1) and $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. The function $\theta \mapsto \mu(p_{\mathcal{A},\theta})$ is increasing over $(0, \theta^c_{\mathcal{A}} + \varepsilon) \bigcap I_{\mathcal{A}}$ for some strictly positive ε depending on p. If $0 \in \mathcal{A}$, then the function $\theta \mapsto \mu(p_{\mathcal{A},\theta})$ is increasing over $I_{\mathcal{A}}$.*

Proof. Notice it is enough to consider $\theta < \theta^*_{\mathcal{A}}$. Since p satisfies (1), it is easy to check that $p_{\mathcal{A},\theta}$ satisfies (1) for all $\theta \in I_{\mathcal{A}}$ such that $\theta < \theta^*_{\mathcal{A}}$. Thanks to the composition rule, it is enough to prove that $\theta \mapsto \mu_{\mathcal{A},\theta}$ is increasing at $\theta = 1$ if $\mu(p) \leq 1 + \varepsilon$ for some $\varepsilon > 0$, with p satisfying (1) and $\rho(p) > 1$.

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

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

with

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

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

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

In particular, we deduce from this last expression that $h'_{\mathcal{A}}(1) > 0$ if $\mathbb{E}[X] \leq 1$. However, since $p(\mathcal{A})\mathbb{E}\left[X(X-1)\right] > 0$ as p satisfies (1), we deduce that $h'_{\mathcal{A}}(1) > 0$ as soon as $\mathbb{E}[X] < 1 + \varepsilon$ for some small positive ε. This ends the proof of the first part of the lemma.

Let us assume that $0 \in \mathcal{A}$. Thanks to the first part, if $\mathbb{E}[X] = \mu(p) > 1$, elementary computations yield that $h'_{\mathcal{A}}(1)/\mathbb{P}(\mathcal{A})$ is minimal, that is $\mathbb{E}\left[X\mathbf{1}_{\mathcal{A}}(X)\right]/\mathbb{P}(\mathcal{A})$ is maximal, (for all subsets $\mathcal{A}$ of $\mathbb{N}$ containing 0) for $\mathcal{A}$ of the form $\mathcal{A}_n = \{0\} \cup \{k; k \geq n\}$. 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. □

Let us consider the equation:

$$\mu(p_{\mathcal{A},\theta}) = 1. \tag{26}$$

Lemma 5.2. *Let p be a distribution on $\mathbb{N}$ satisfying (1) and $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. Equation (26) has at most one solution. If there is no solution to Equation (26), 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 (26), it 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 (26) has at most one solution.

If $0 \in \mathcal{A}$, then we have $\inf_{I_{\mathcal{A}}} \mu(p_{\mathcal{A},\theta}) = p(1)\mathbf{1}_{A^c}(1) < 1$. 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 (24) 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 (26) has at least one solution.

From what precedes, if there is no solution to Equation (26), this implies that $\mu(p) < 1$ and thus:

$$\mu(p_{\mathcal{A},\theta}) < 1 \quad \text{for all } \theta \in I_{\mathcal{A}}. \tag{27}$$

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 (4) that $\theta c_{\mathcal{A}}(\theta) > 1$ and then from (24) that $\mu(p_{\mathcal{A},\theta}) > g'(\theta)$ and thus $g'(\theta) < 1$. Notice 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 (21) holds for $\theta^*_{\mathcal{A}}$ that is $\theta^*_{\mathcal{A}} \in I_{\mathcal{A}}$. Then conclude using (27). □

Definition 5.3. *Let p be a distribution on $\mathbb{N}$ satisfying (1) and $\mathcal{A} \subset \mathbb{N}$ such that $p(\mathcal{A}) > 0$. If Equation (26) has a (unique) solution, then p is called generic for $\mathcal{A}$. If Equation (26) 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) 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 $g'(1) < 1$, 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) 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) 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:

$$\mu(p_{\mathcal{A},\theta}) - 1 = G_{\mathcal{A}}(\theta)\frac{\theta - g(\theta)}{\theta} - (1 - g'(\theta)) \quad \text{with} \quad G_{\mathcal{A}}(\theta) = \frac{\mathbb{E}\left[X\theta^X \mathbf{1}_A(X)\right]}{\mathbb{E}\left[\theta^X \mathbf{1}_A(X)\right]}. \tag{28}$$

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 (21). We also have $g(q) < q$. This implies, thanks to (28), 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 (21). 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 (28):

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

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

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), $\mu(p) < 1$. We prove Theorem 1.2 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 $L_{\mathcal{A}}(\mathbf{t}) = \text{Card}\,(\mathcal{L}_{\mathcal{A}}(\mathbf{t}))$.

For a tree $\mathbf{t} \in \mathbb{T}_0$, following [11, 14], 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 $\mathbf{s} = \{w^k u, u \in S_{w^k}$ the subtree 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 $\mathbf{t}^{\mathcal{A}}$.

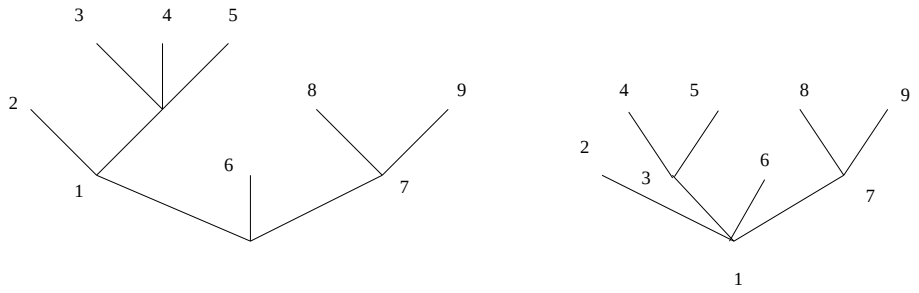

Figure 1. left: a tree $\mathbf{t}$, right: the tree $\mathbf{t}^{\mathcal{A}}$ for $\mathcal{A} = \{0, 2\}$

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 [14] 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 $\{X \notin \mathcal{A}\}$. We set:

$$X_{\mathcal{A}} = \sum_{k=1}^{N-1} Y'_k + Y'', \tag{29}$$

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:

$$g^{\mathcal{A}}(z) = \frac{z\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]}{z - \mathbb{E}\left[z^X \mathbf{1}_{\{X\notin\mathcal{A}\}}\right]}. \tag{30}$$

An elementary computation gives:

$$\mu(p^{\mathcal{A}}) = 1 - \frac{1-\mu(p)}{p(\mathcal{A})} \quad \text{and} \quad g^{\mathcal{A}}(\theta) = \frac{1}{c_{\mathcal{A}}(\theta)}\,. \tag{31}$$

We recover that if τ is critical ($\mu(p) = 1$) then $\tau^{\mathcal{A}}$ is critical as $\mu(p^{\mathcal{A}}) = 1$, see also [14] Lemma 6. Notice in particular that for all $k \in \mathcal{A}$:

$$p^{\mathcal{A}}(k) = \mathbb{P}(X_{\mathcal{A}} = k) \geq \mathbb{P}(N = 1, Y'' = k) = p(k), \tag{32}$$

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

$$p^{\mathcal{A}}(k-1) = \mathbb{P}(X_{\mathcal{A}} = k-1) \geq \mathbb{P}(N = 2, Y'_1 = k-1) = p(\mathcal{A})p(k). \tag{33}$$

Lemma 6.1. *Assume that p satisfies (1), $\mu(p) < 1$. Then $p^{\mathcal{A}}$ satisfies (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. Since (32) implies $p^{\mathcal{A}}(0) \geq p(0)$ and that $\mu(p) < 1$ with (31) implies $\mu(p^{\mathcal{A}}) < 1$, we deduce that $p^{\mathcal{A}}$ satisfies (1).

Let $\rho_{\mathcal{A}}$ be the convergence radius of the serie given by $\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}\}}\right]$ and $\rho_{\mathcal{A}^c}$ be the convergence radius of the series given by $\mathbb{E}\left[z^X \mathbf{1}_{\{X\in\mathcal{A}^c\}}\right]$. We get that $\min(\rho_{\mathcal{A}}, \rho_{\mathcal{A}^c}) = \rho(p)$. We deduce that the convergence radius of $g^{\mathcal{A}}$ is $\rho(p)$ if $\rho(p) = 1$ or if $\rho(p) > 1$ and $g'(\rho(p)) < 1$. □

6.1. **The case $\rho(p) = 1$.** We state now the main result of this section.

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

$$\text{dist}\,(\tau \;\big|\; L_{\mathcal{A}}(\tau) = n) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p)), \tag{34}$$

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

$$\text{dist}\,(\tau \;\big|\; L_{\mathcal{A}}(\tau) \geq n) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p)). \tag{35}$$

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 $n \geq 2$:

$$n_j = n - \mathbf{1}_{\mathcal{A}}(j). \tag{36}$$

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 (17), 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 (47) and (48):

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

Then use the generalizations of the strong ratio limit properties (44), (50) and (51) 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 (12), 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 (34). Since dist $(\tau \mid L_{\mathcal{A}}(\tau) \geq n)$ is a mixture of dist $(\tau \mid L_{\mathcal{A}}(\tau) = k)$ for $k \geq n$, we deduce that (35) 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 $g'(\rho) < 1$ and $g(\rho) < \rho$ thanks to Lemma 5.4. Recall the offspring distribution $p_{\mathcal{A},\theta}$ defined by (3). Notice that the normalizing constant $c_{\mathcal{A}}(\theta)$ is given by:

$$c_{\mathcal{A}}(\theta) = \frac{\theta - \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}^c\}}\right]}{\theta \mathbb{E}\left[\theta^X \mathbf{1}_{\{X \in \mathcal{A}\}}\right]} = \frac{1}{g^{\mathcal{A}}(\theta)}. \tag{37}$$

Notice that $p_{\mathcal{A},1} = p$. Since $\rho(p)$ is also the convergence radius 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 [9] 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:

$$(g_{\mathcal{A},\theta})^{\mathcal{A}}(z) = \frac{g^{\mathcal{A}}(\theta z)}{g^{\mathcal{A}}(\theta)} = \left(g^{\mathcal{A}}\right)_{\mathbb{N},\theta}(z). \tag{38}$$

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 (38), the tree $(\tau_{\mathcal{A},\theta})^{\mathcal{A}}$ 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 4.3.

Corollary 6.4. *Assume that p satisfies (1) and is non-generic for $\mathcal{A}$. Let $p^*_{\mathcal{A}} = p_{\mathcal{A},\rho(p)}$. We have that:*

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

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

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

This result with Proposition 4.6 and Corollary 5.7 in [1] ends the proof of Theorem 1.2 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 1.2 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 (32) and (33). 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 $S_u^{\mathcal{A}}(\mathbf{t})$ the subtree rooted at u with no progeny in $\mathcal{A}$ by

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

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

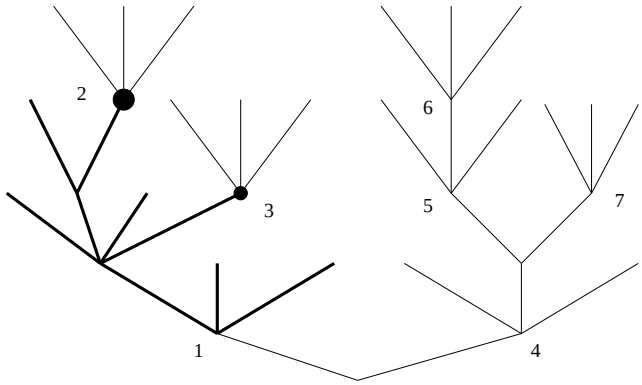

FIGURE 2. The subtree $S_1^{\mathcal{A}}(\mathbf{t})$ in bold for $\mathcal{A} = \{3\}$, and the elements of $C_1^{\mathcal{A}}(\mathbf{t})$.

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 $\tilde{N}_{\emptyset}(\mathbf{t}) = \text{Card}\ (\tilde{C}_{\emptyset}^{\mathcal{A}}(\mathbf{t}))$. 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 $1 \leq i \leq \tilde{N}_{\emptyset}(\mathbf{t})$

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

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

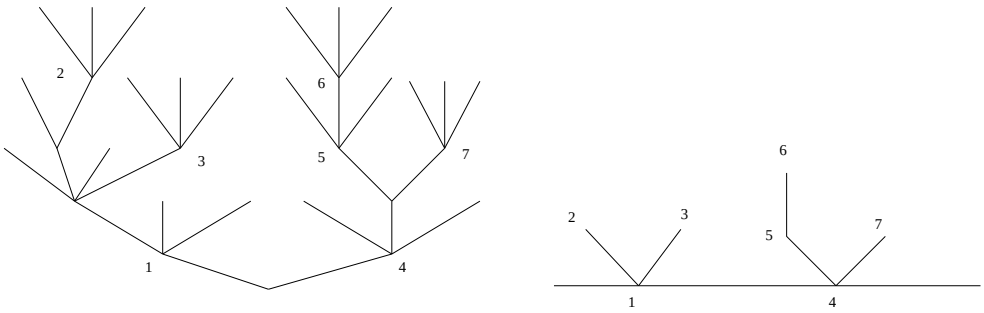

FIGURE 3. A tree $\mathbf{t}$ and the forest $\mathcal{F}_{\mathcal{A}}(\mathbf{t})$ for $\mathcal{A} = \{3\}$.

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

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

Then $\tilde{S}_\emptyset^{\mathcal{A}}(\tau)$ is distributed as a (subcritical) 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

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

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 $(L, p(\mathcal{A})/(p(0)+p(\mathcal{A})))$.

Let $X^{\mathcal{A}}$ be the random variable

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

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 $p^{\mathcal{A}}$ the law of $X^{\mathcal{A}}$. Then the forest $\mathcal{F}_{\mathcal{A}}(\tau)$ is distributed as N independent GW trees with offspring distribution $p^{\mathcal{A}}$.

7.2. **Main result.** We recall that $L_{\mathcal{A}}(\tau)$ is aperiodic since $0 \notin \mathcal{A}$, see [1].

Theorem 7.1. *Assume that p satisfies (1) and $\mu(p) < 1$ and $\rho(p) = 1$. We have that:*

$$\text{dist}\,(\tau \;\big|\; L_{\mathcal{A}}(\tau) = n) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p)), \tag{39}$$

as well as

$$\text{dist}\,(\tau \;\big|\; L_{\mathcal{A}}(\tau) \geq n) \underset{n\to+\infty}{\longrightarrow} \text{dist}\,(\tau^*(p)). \tag{40}$$

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

$$\lim_{n\to+\infty} \mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t},x,k),\ L_{\mathcal{A}}(\tau) = n) = D(\mathbf{t},x)\mathbb{P}\left(\tau^*(p) \in \mathbb{T}_+(\mathbf{t},x,k)\right). \tag{41}$$

Set $M_0 = 0$ and $M_n = \sum_{k=1}^n N_k$ for $n \in \mathbb{N}^*$. Let $m = L_{\mathcal{A}}(\mathbf{t}) - \mathbf{1}_{\mathcal{A}}(k_x(\mathbf{t}))$ and $\ell = k_x(\mathbf{t})$. Recall (36). We have

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

where we used Dwass formula (17) 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 $X^{\mathcal{A}}$, see also (58). Recall (59). In particular, we have:

$$\mathbb{P}(\tau \in \mathbb{T}_+(\mathbf{t},x,k)|\, L_{\mathcal{A}}(\tau) = n) = D(\mathbf{t},x)\left(B_{n-m,\ell} - \sum_{j=\ell+1}^{k-1} p(j)(j-\ell)\, a_{n-m,j}\right), \tag{42}$$

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. □

Corollary 7.2. *Assume that p satisfies (1), is non-generic for $\mathcal{A}$. Let $p^*_{\mathcal{A}} = p_{\mathcal{A},\rho(p)}$. We have that:*

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

as well as

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

This result with Proposition 4.6 and Corollary 5.7 in [1] for the generic case ends the proof of Theorem 1.2 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:

$$\mu(p) \leq 1 \text{ and either } \mu(p) = 1 \text{ or, for all } \theta > 0,\ \mathbb{E}\left[\mathrm{e}^{\theta X_1}\right] = +\infty. \tag{43}$$

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 [8] or [12], we have the following strong ratio limit property for all $m, k \in \mathbb{Z}$:

$$\lim_{n\to+\infty} \frac{\mathbb{P}(S_{n-m} = n-k)}{\mathbb{P}(S_n = n)} = 1. \tag{44}$$

We deduce the following corollary. Recall the definition of δ_n^0 and δ_n^1 of (18) and (19).

Corollary 8.1. *Assume that p satisfies (43) and is aperiodic. For all $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, we have:*

$$\lim_{n\to+\infty} \delta_n^0(k,\ell) = \sum_{j\geq k} p(j). \tag{45}$$

and

$$\lim_{n\to+\infty} \delta_n^1(k,\ell) = 1 - \mu(p) + \sum_{j\geq k} jp(j). \tag{46}$$

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 (44) to get (45).

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 (44) 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 (46). □

8.2. **Generalization of the strong ratio limit property I.** Assume that p satisfies (43) and is aperiodic. Let X be a random variable taking values in $\mathbb{N}$ with distribution p. Recall g denote 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 (30) and $X_{\mathcal{A}}$ distributed according to $p^{\mathcal{A}}$. Recall $\mu(p^{\mathcal{A}})$ is given by (31). 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 $\mathbb{E}\left[\mathrm{e}^{\theta X}\right] = +\infty$ for all $\theta > 0$ implies that $\mathbb{E}\left[\mathrm{e}^{\theta X_{\mathcal{A}}}\right] = +\infty$ 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 (44) holds as well as (45) and (46) hold with p replaced by $p^{\mathcal{A}}$.

Recall (36), that is $n_j = n - \mathbf{1}_{\mathcal{A}}(j)$, and let:

$$\delta_n^{0,\mathcal{A}}(k,\ell) = \frac{1}{\mathbb{P}(S_n = n)} \sum_{j \geq k} p(j)\ \frac{n}{n_j} \mathbb{P}(S_{n_j} = n_j + \ell - j) \tag{47}$$

and

$$\delta_n^{1,\mathcal{A}}(k,\ell) = \frac{1}{\mathbb{P}(S_n = n)} \sum_{j \geq k} jp(j)\ \frac{n}{n_j} \mathbb{P}(S_{n_j} = n_j + \ell - j). \tag{48}$$

We stress that in (18) and (19), $(S_n, n \in \mathbb{N})$ is a random walk with increments distributed according to p; whereas in (47) and (48), $(S_n, n \in \mathbb{N})$ is a random walk with increments distributed according to $p^{\mathcal{A}}$.

Lemma 8.2. *Assume that p satisfies (43) and is aperiodic. For all $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, we have:*

$$\lim_{n \to +\infty} \frac{\mathbb{E}\left[\frac{n}{n_X} \mathbf{1}_{\{X + S_{n_X} = n_X + \ell\}}\right]}{\mathbb{P}(S_n = n)} = 1, \tag{49}$$

$$\lim_{n \to +\infty} \delta_n^{0,\mathcal{A}}(k,\ell) = \sum_{j \geq k} p(j) \tag{50}$$

and

$$\lim_{n \to +\infty} \delta_n^{1,\mathcal{A}}(k,\ell) = 1 - \mu(p) + \sum_{j \geq k} jp(j). \tag{51}$$

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

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

with the convention that $p^{\mathcal{A}}(-1) = 0$.

Thanks to the strong ratio limit property (that is (44) with $p^{\mathcal{A}}$ instead of p), we have $\lim_{n \to +\infty} a_n(j) = p(j)$ and $\lim_{n \to +\infty} b_n(j) = p^{\mathcal{A}}(j) + p^{\mathcal{A}}(j-1)/p(\mathcal{A})$. We have:

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

We deduce from the strong ratio limit property (that is (44) with $p^{\mathcal{A}}$ instead of p) that:

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

Then use (32) and (33) to get that $a_n(j) \leq 2b_n(j)$ for $n \geq 2$ and the dominated convergence theorem 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 (49) holds. Since $\delta_n^{0,\mathcal{A}}(k,\ell) = \sum_{j\geq k} a_n(j)$, the proof of (50) is then similar to the proof of (45).

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 8.3 below, (44) and (49), we have that $\lim_{n\to+\infty} c_n(\ell) = 1$ for all $\ell\in\mathbb{Z}$. Then arguing as in the proof of (46), we easily get (51). □

Lemma 8.3. *For all $\ell\in\mathbb{Z}$, $n\geq 2$, we have:*

$$\mathbb{E}\left[\frac{n}{n_X}X\mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right] = \ell\mathbb{E}\left[\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X}=n_X+\ell\}}\right] - (\ell-1)\mathbb{P}(S_n = n+\ell-1). \tag{52}$$

Proof. We first prove (52) 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 (17) 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 (52) holds for $\ell\leq 0$.

Unfortunately, we didn't get a similar proof for $\ell\geq 1$ and we prove (52) for $\ell\geq 1$ by induction. Let $\ell\geq 0$. Assume that (52) 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:

$$\mathbb{E}\left[\frac{n+1}{n_X+1}X\mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right] = A_1 + \mathbb{E}\left[\frac{n_X-n}{n_X(n_X+1)}X\mathbf{1}_{\{X+S_{n_X+1}=n_X+1+\ell\}}\right], \tag{53}$$

with

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

Using (52), we have:

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

So we have:

$$A_1 = p^{\mathcal{A}}(0)A_2 + A_3 - \mathbb{E}\left[(\ell-X_1)\mathbf{1}_{\{S_{n+1}=n+\ell\}}\right], \tag{54}$$

with

$$A_2 = \mathbb{E}\left[\frac{n}{n_X} X \mathbf{1}_{\{X+S_{n_X}=n_X+1+\ell\}}\right] - (\ell+1)\mathbb{E}\left[\frac{n}{n_X}\mathbf{1}_{\{X+S_{n_X}=n_X+\ell+1\}}\right] + \ell\mathbb{P}(S_n = n+\ell) \tag{55}$$

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 (54). 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 (53), 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 (52) with $n+1$ instead of n:

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

Recall (55). The fact that $A_2 = 0$ gives exactly that (52) 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 Section 7.2. We have the following generalization of the strong ratio limit property.

Lemma 8.4. *Assume that $p^{\mathcal{A}}$ is aperiodic, $\mu(p^{\mathcal{A}}) < 1$, $\rho(p^{\mathcal{A}}) = 1$ and $\mathbb{E}\left[\mathrm{e}^{\theta X_{\mathcal{A}}}\right] = +\infty$ for all $\theta > 0$. Then for all $m, k \in \mathbb{Z}$, we have:*

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n-m}+N=n-k\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = 1. \tag{56}$$

Note that if $p^{\mathcal{A}}$ is periodic, then (56) still holds along the subsequence for which the denominator is positive.

Proof. We shall mimic the proof of the strong ratio limit property provided in [12]. Since $p^{\mathcal{A}}$ is aperiodic, the denominator of (56) is positive for n large enough and it is enough to prove the result for $m = 1$ and k such that $p^{\mathcal{A}}(k) > 0$. Denote $\hat{p}_n^{\mathcal{A}}(k) = \sum_{i=1}^n \mathbf{1}_{\{X_i=k\}}/n$. We have:

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

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

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

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

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

According to [12], since $p^{\mathcal{A}}$ is non-generic with $\rho(p^{\mathcal{A}}) = 1$, we have $\lim_{n\to+\infty} \mathbb{P}(|\hat{p}_n^{\mathcal{A}}(k) - p^{\mathcal{A}}(k)| > \varepsilon)/\mathbb{P}(S_n = n) = 0$. 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 $\lim_{n\to+\infty} J_n = 0$. □

Remark 8.5. Notice that, from the proof of the lemma, we see that N could be replaced by any non-negative integrable random variable independent of $(X_k, k \in \mathbb{N}^*)$.

Recall that $M_0 = 0$ and for $n \in \mathbb{N}^*$:

$$M_n = \sum_{k=1}^n N_k.$$

We assume that $(N_k, k \in \mathbb{N}^*)$ and $(X_k, k \in \mathbb{N}^*)$ are independent. We have the following result.

Lemma 8.6. *Assume $p^{\mathcal{A}}$ is aperiodic, with $\mu(p^{\mathcal{A}}) < 1$ and $\rho(p^{\mathcal{A}}) = 1$. 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 8.4, 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 $\mathbb{P}(Z_1 = j_0) > 0$. Notice that:

$$r(\ell) = \mathbb{P}(S_m = \ell) \geq \mathbb{P}(Z_1 + \ldots + Z_m = mj_0, M_m = \ell, N_{m+1} + \ldots N_{mj_0} = 0).$$

We deduce that there exists $c > 0$ such that $q(\ell) \leq Cr(\ell)$ for all $\ell \in \mathbb{N}$. By dominated convergence, 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 $\mathcal{A}$ and Dwass formula (17):

$$\mathbb{P}(L_{\mathcal{A}}(\tau)=n)=\sum_{j\in\mathbb{N}}p_N(j)\mathbb{P}_j(|\tau^{\mathcal{A}}|=n)=\frac{1}{n}\,\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]. \tag{57}$$

More generally, we have

$$\mathbb{P}_j(L_{\mathcal{A}}(\tau)=n)=\frac{1}{n}\,\mathbb{E}\left[M_j\mathbf{1}_{\{S_n+M_j=n\}}\right]=\frac{j}{n}\mathbb{E}\left[N\mathbf{1}_{\{S_n+M_{j-1}+N=n\}}\right], \tag{58}$$

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

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

$$B_{n,\ell}=\sum_{j>\ell}p(j)(j-\ell)\frac{n}{n_j}\,\frac{\mathbb{E}\left[N\mathbf{1}_{\{S_{n_j}+M_{j-1-\ell}+N=n_j\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]}. \tag{59}$$

The next lemma is the analogue of Lemma 8.3 in our current setting.

Lemma 8.7. *For $\ell\le 0$, we have $\lim_{n\to+\infty}B_{n,\ell}=1-\ell$.*

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

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

Then use (58) 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 $\ell\ge k$, $\ell,k\in\mathbb{Z}$:

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

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

$$C_{n,\ell}(\ell)=nB_{n,\ell}\mathbb{P}(L_{\mathcal{A}}(\tau)=n). \tag{60}$$

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

Lemma 8.8. *Assume $p^{\mathcal{A}}$ is aperiodic, non-generic with $\rho(p^{\mathcal{A}})=1$. We have for $k\in\mathbb{Z}$ such that $k\le\ell$:*

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

Proof. Notice that $nN(X-\ell)_+/n_X$ is integrable. Mimicking the proof of Lemma 8.4 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 $p^{\mathcal{A}}$ is aperiodic, non-generic with $\rho(p^{\mathcal{A}}) = 1$. 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:

$$C_{n,\ell}(-1) = C_{n,0}(-1) - \sum_{j=0}^{\ell-1} p(j)(j-\ell)\mathbb{E}\left[\frac{n}{n_j} N \mathbf{1}_{\{S_{n_j}+M_j+N=n_j\}}\right] - \ell\mathbb{E}\left[\frac{n}{n_X} N \mathbf{1}_{\{S_{n_X}+M_X+N=n_X\}}\right], \tag{61}$$

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 (60) implies that:

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

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 (61) with $\ell = -1$ and Lemma 8.6 that:

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[\frac{n}{n_X} N \mathbf{1}_{\{S_{n_X}+M_X+N=n_X\}}\right]}{\mathbb{E}\left[N\mathbf{1}_{\{S_n+N=n\}}\right]} = 1. \tag{62}$$

Let $\ell \geq 1$. We deduce from (61) with $\ell \geq 1$, (60), (57), Lemma 8.6 and (62) that:

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

□

References

[1] R. ABRAHAM and J. DELMAS. Local limits of conditioned galton-watson trees I: the infinite spine case. arXiv:1304.4035, 2013.

[2] 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.

[3] N. CURIEN and I. KORTCHEMSKI. Random non-crossing plane configurations: a conditioned Galton-Watson tree approach. *Random Struct. and Alg.*, To appear, 2013.

[4] M. DWASS. The total progeny in a branching process and a related random walk. *J. Appl. Probability*, 6:682–686, 1969.
[5] 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.
[6] S. JANSON. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012.
[7] T. JONNSSON and S. STEFANSSON. Condensation in nongeneric trees. *J. Stat. Phys.*, 142:277–313, 2011.
[8] J. G. KEMENY. A probability limit theorem requiring no moments. *Proc. Amer. Math. Soc.*, 10:607–612, 1959.
[9] D. KENNEDY. The Galton-Watson process conditioned on the total progeny. *J. Appl. Probability*, 12(4):800–806, 1975.
[10] H. KESTEN. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986.
[11] 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.
[12] J. NEVEU. Sur le théorème ergodique de Chung-Erdős. *C. R. Acad. Sci. Paris*, 257:2953–2955, 1963.
[13] J. NEVEU. Arbres et processus de Galton-Watson. *Ann. de l'Inst. Henri Poincaré*, 22:199–207, 1986.
[14] 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. arXiv:1105.2528, 2013.

Romain Abraham, Laboratoire MAPMO, CNRS, UMR 7349, Fédération Denis Poisson, FR 2964, Université d'Orléans, B.P. 6759, 45067 Orléans cedex 2, France.
E-mail address: romain.abraham@univ-orleans.fr

Jean-François Delmas, Université Paris-Est, CERMICS (ENPC), F-77455 Marne La Vallée, France.
E-mail address: delmas@cermics.enpc.fr