



CONDITIONING BIENAYMÉ–GALTON–WATSON TREES TO HAVE LARGE SUB-POPULATIONS

ROMAIN ABRAHAM,* *Université d'Orléans, Université de Tours, CNRS*
HONGWEI BI,** *University of International Business and Economics*
JEAN-FRANÇOIS DELMAS,*** *École des Ponts*

Abstract

We study the local limit in distribution of Bienaymé–Galton–Watson trees conditioned on having large sub-populations. Assuming a generic and aperiodic condition on the offspring distribution, we prove the existence of a limit given by a Kesten's tree associated with a certain critical offspring distribution.

Keywords: Bienaymé–Galton–Watson tree; generic probability distribution; local limit

2010 Mathematics Subject Classification: Primary 60J80
Secondary 60B10

1. Introduction

The local limit of large Bienaymé–Galton–Watson (BGW) trees has been extensively studied in recent years. The classical result of Kesten [21] describes the local limit of a critical or subcritical BGW tree conditioned on reaching at least height h : it converges locally in distribution as h tends to infinity to the so-called size-biased tree or Kesten's tree, which is a tree with an infinite spine. We refer to Section 2.5 for a precise description of Kesten's tree.

Over the years, motivated by various points of view from theoretical probability, combinatorics, biology, or physics, other conditionings have been considered, such as a large total population [13, 20], a large number of leaves [9, 18], a large number of protected nodes [1], and the existence of an individual with a large number of out-degree or children [14, 15]. Janson [16] surveyed the local limit of BGW trees when conditioned on a large total population, and Abraham and Delmas [3, 4] provided a general framework, which describes in full generality the local limit of critical or subcritical BGW trees conditioned on having a large sub-population. Note that in [3, 16, 18, 29] the local limit may exhibit a condensation phenomenon, as one node of the limiting tree has an infinite number of children. With other conditioning, such as a large population at late generation [2, 6], or with exponential weight given by the total height of the tree among trees with a given large total population size [12], the local limit is a tree with countably many infinite branches. The local limit of large multi-type Galton–Watson

Received 17 July 2025; accepted 17 July 2025.

* Postal address: Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, Tours, France. Email: romain.abraham@univ-orleans.fr

** Postal address: School of Insurance and Economics, University of International Business and Economics, Beijing, China. Email: bihw@uibe.edu.cn

*** Postal address: CERMICS, École des Ponts, Champs-sur-Marne, France. Email: jean-francois.delmas@enpc.fr

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust.

trees has also been considered in [7, 25], and also in [26, 28, 31] when conditioning on a linear combination of the sizes of the sub-populations with a given type.

One can also consider scaling limits of BGW trees (seen as metric space), the so-called Lévy continuum trees, as initiated by Aldous [8] and generalized by Duquesne and Le Gall [11]. Let us mention that there is a large recent literature on this subject. In particular, Marzouk [23] considered the scaling limit of random trees with a prescribed degree sequence, and Kargin [19] and Kortchemski and Marzouk [22] the scaling limit of BGW trees conditioned on its total population size and the number of leaves.

Motivated by these last works, we investigate the local limit of BGW trees when conditioning on its total population size and the number of leaves being large. More generally, let $\mathcal{A} = (A_i)_{i \in \llbracket 1, J \rrbracket}$, where $J \in \mathbb{N}^*$ and $\llbracket a, b \rrbracket = [a, b] \cap \mathbb{N}$ (with $\mathbb{N}^* = \{1, 2, \dots\}$ and $\mathbb{N} = \mathbb{N}^* \cup \{0\}$), be a finite collection of pairwise disjoint subsets of \mathbb{N} . Intuitively, for a BGW tree with offspring distribution p , we have the number of nodes with out-degree in A_i of order $p(A_i)$ times the total population size of the BGW tree when it is large (where $p(A) = \sum_{n \in A} p(n)$ for $A \subset \mathbb{N}$); see [17, 30] for a precise statement. In this paper we choose an arbitrary probability distribution $\alpha = (\alpha_i)_{i \in \llbracket 1, J \rrbracket}$ on $\llbracket 1, J \rrbracket$, called a *direction*, and we want to condition the BGW tree to be large with the asymptotic proportions of nodes with out-degree in A_i equal to α_i for all $i \in \llbracket 1, J \rrbracket$ up to a normalizing constant. This normalizing constant is equal to 1 if \mathcal{A} is a partition of the support of the offspring distribution. We note that, if the complement of \mathcal{A} , say A_0 , has positive probability ($p(A_0) > 0$), there is, in our setting, no condition on the nodes with out-degree in A_0 . Our main objective is to prove the local convergence of such a BGW tree conditioned on having large sub-populations, identify the local limit as a Kesten’s tree that is a random tree with an infinite spine, and describe the possible directions and conditions on the offspring distribution for such a convergence to hold.

Our present result generalizes [3, 4], which are devoted to the uni-dimensional case $J = 1$. Let us stress that other conditionings than population size are considered in [3] and that conditioning on one given sub-population to be large might lead to a local limit having an infinite spine as in [3] or to a condensation phenomenon (that is, a tree having a node with infinite out-degree) as presented in [4]. The condensation phenomenon for a nongeneric offspring distribution is an open question for $J \geq 2$.

More precisely, let $p = (p(n))_{n \in \mathbb{N}}$ be a probability distribution on \mathbb{N} ; its support is $\text{supp}(p) = \{n \in \mathbb{N} : p(n) > 0\}$. We assume that p is nontrivial in the sense that $p(0) > 0$ and $p(\{0, 1\}) < 1$. We denote by $\mu(p) \in (0, +\infty]$ its mean. We denote by \mathcal{T}_p a BGW tree with offspring distribution p . Note that we do not even assume that $\mu(p)$ is finite; however, since $p(0)$ is positive, the tree \mathcal{T}_p is finite with positive probability. For a locally finite discrete tree \mathbf{t} , we denote by $L_{\mathcal{A}}(\mathbf{t})$ the number of its nodes with out-degree (or number of children) in A . For example, $L_{\mathbb{N}}(\mathbf{t})$ represents the total number of nodes in the tree \mathbf{t} , $L_{\mathbb{N}^*}(\mathbf{t})$ its number of internal nodes, and $L_{\{0\}}(\mathbf{t})$ its number of leaves. We simply denote by $L_{\mathcal{A}}(\mathbf{t})$ the vector of the number of nodes in \mathbf{t} with out-degree in the finite collection \mathcal{A} of pairwise disjoint sets:

$$L_{\mathcal{A}}(\mathbf{t}) = (L_{A_i}(\mathbf{t}))_{i \in \llbracket 1, J \rrbracket} \in \tilde{\mathbb{N}}^J \quad \text{with } \tilde{\mathbb{N}} = \mathbb{N} \cup \{\infty\}. \tag{1}$$

For instance, if $\mathcal{A} = (\{0\}, \mathbb{N}^*)$, the vector $L_{\mathcal{A}}(\mathbf{t})$ represents the couple (number of leaves, number of internal nodes) of the tree \mathbf{t} .

In Theorem 4.1 we completely characterize the nontrivial probability distributions p' on \mathbb{N} such that $\text{supp}(p') \subset \text{supp}(p)$ and the distribution of the BGW trees \mathcal{T}_p and $\mathcal{T}_{p'}$ conditionally on $L_{\mathcal{A}}$ are the same, that is, more formally, for all $\mathbf{n} = (n_i)_{i \in \llbracket 1, J \rrbracket} \in \mathbb{N}^J$ such that

$\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}) > 0$ (and, thus, $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$), we have

$$\text{dist}(\mathcal{T}_p \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) = \text{dist}(\mathcal{T}_{p'} \mid L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}).$$

Such probability distributions are called (p, \mathcal{A}) -compatible. The (p, \mathcal{A}) -compatible probability distributions can be continuously parametrized by a parameter (θ, β) in a subset of $[0, +\infty] \times \mathbb{R}_+^J$. Recall that A_0 denotes the complement of $\bigcup_{i \in \llbracket 1, J \rrbracket} A_i$ in \mathbb{N} , which might be empty or not. When the parameter θ is positive and finite, then the (p, \mathcal{A}) -compatible probability distribution $\tilde{p}_{\theta, \beta}$ associated with the parameter (θ, β) is given by

$$\tilde{p}_{\theta, \beta}(n) = \beta_i \theta^n p(n) \quad \text{for } n \in A_i \text{ and } i \in \llbracket 0, J \rrbracket, \text{ where } \beta_0 = \theta^{-1}.$$

The fact that such exponentially tilted probability distributions are (p, \mathcal{A}) -compatible was already observed in [4] for $J = 1$ and in [31] for the multi-type BGW tree setting; see also the recent preprint [26]. In comparison with these papers, we give here an exhaustive description of the (p, \mathcal{A}) -compatible probability distributions under a very mild hypothesis on the offspring distribution p . In particular, it is possible to observe degenerate cases when the parameter θ can take the values 0 and ∞ ; see (21) and (22).

For $\mathbf{x} = (x_i)_{i \in \llbracket 1, J \rrbracket} \in \mathbb{R}^J$, we set $|\mathbf{x}| = \sum_{i \in \llbracket 1, J \rrbracket} |x_i|$ the L^1 norm of \mathbf{x} . As in [25], it is interesting to have a fixed (asymptotic) proportion of sub-populations, and thus, consider the local limit of \mathcal{T}_p conditionally on $\{L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}\}$ when $\mathbf{n}/|\mathbf{n}|$ converges to some direction $\alpha \in \mathbb{R}_+^J$ such that $|\alpha| = 1$, as $|\mathbf{n}|$ tends to infinity. Thus, it is natural to consider among the (p, \mathcal{A}) -compatible probability distributions those which are in the direction $\alpha = (\alpha_i)_{i \in \llbracket 1, J \rrbracket}$, that is,

$$\tilde{p}_{\theta, \beta}(A_i) \propto \alpha_i \quad \text{for all } i \in \llbracket 1, J \rrbracket.$$

From this, we can write α as a function of (θ, β) , (but for some pathological cases, see Remark 3.3 for details), and similarly β as a function of (θ, α) . This gives an elementary reparametrization $(p_{\theta, \alpha})$ of the family $(\tilde{p}_{\theta, \beta})$ by the parameter θ and its direction α . The family $(\tilde{p}_{\theta, \beta})$ is explicitly given in Lemma 5.1 and the possible directions in Proposition 5.1. In particular, if the entries of α are all positive then α is a possible direction. Furthermore, if α is a possible direction then the set $I_\alpha \subset [0, +\infty]$ of possible values for the parameter θ is an interval; see Section 5.2. As observed in previous works, the existence of a critical parameter θ_α such that $\mu(p_{\theta_\alpha, \alpha}) = 1$ is a key point to obtain the local limit of the conditioned BGW tree. Proposition 5.2 asserts that the mean function $\theta \mapsto \mu(p_{\theta, \alpha})$ is increasing when $\mu(p_{\theta, \alpha}) \leq 1$, and thus, there is at most one such critical parameter θ_α . Let us stress this result is not obvious as the map $\theta \mapsto \mu(p_{\theta, \alpha})$ is not monotone in general (see the example in Remark 5.2). When the critical parameter θ_α exists, the distribution p is called *generic for the direction α* , and we set

$$p_\alpha = p_{\theta_\alpha, \alpha} \quad (\text{and, thus, } \mu(p_\alpha) = 1). \tag{2}$$

We provide necessary and sufficient conditions for p to be generic in the direction α in Theorem 5.1 that are similar to those obtained in [4] when $J = 1$. We do not study further the relation between the sets $\mathcal{A} = (A_i)_{i \in \llbracket 1, J \rrbracket}$ and the fact that p is generic, and refer to [4] again to appreciate the complexity already when $J = 1$. For example, the offspring distribution p is generic for any direction α if $\{0\} \subset A_0$ (that is, the conditioning does not involve the number of leaves), $p(1) = 0$ and $0 < p(0) < 1$ (this rules out most of the pathological cases), and p has all exponential moments (this rules out all possible condensation phenomena). Let us mention that if A_0 is empty then p can not be generic in all directions, and there are even more

restrictions on the directions for which p is generic if some of the other A_i are singletons. For instance, if $p(0) + p(2) = 1$ with $0 < p(0) < 1$, the BGW tree is binary and, for $\mathcal{A} = (\{0\}, \{2\})$ almost surely (a.s.), we have $L_{\{0\}}(\mathcal{T}_p) = L_{\{2\}}(\mathcal{T}_p) + 1$; then one can check that there is only one direction for which p is generic given by $\alpha = (\frac{1}{2}, \frac{1}{2})$ and that the corresponding critical probability distribution p_α is given by $p_\alpha(0) = p_\alpha(2) = \frac{1}{2}$.

Let \mathcal{T}_p^* denote the Kesten’s tree associated with the probability distribution p when $\mu(p) = 1$, that is, the local limit in distribution of \mathcal{T}_p conditioned to have height at least h , as h tends to infinity. Before giving the main result of the paper, we recall the following hypotheses.

- (H1) p is a *nontrivial* probability distribution on \mathbb{N} (there is no moment condition). Without loss of generality we assume that the sets $\mathcal{A} = (A_i)_{i \in \llbracket 1, J \rrbracket}$ are pairwise disjoint subsets of the support of p .
- (H2) $\alpha \in \mathbb{R}_+^J$ with $|\alpha| = 1$ is a *possible direction* (this condition is always satisfied if all the entries of α are positive).
- (H3) p is *generic* in the direction α .
- (H4) p is *aperiodic* in the sense of Definition 6.1. (In particular, being aperiodic depends on the sets \mathcal{A} and the direction α ; see also Remark 6.1.)

We are now ready to state the main result of the paper. Recall that $L_{\mathcal{A}}(\mathbf{t})$, defined in (1), denotes the number of nodes in the tree \mathbf{t} with out-degree in the finite collection \mathcal{A} of pairwise disjoint sets, \mathcal{T}_p is the BGW tree with offspring distribution p , and that $\mathcal{T}_{p_\alpha}^*$ is the Kesten’s tree associated with p_α defined in (2).

Theorem 1.1. *Assume that hypotheses (H1)–(H4) hold. We have the local limit in distribution*

$$\text{dist}(\mathcal{T}_p \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) \xrightarrow{|\mathbf{n}| \rightarrow \infty} \text{dist}(\mathcal{T}_{p_\alpha}^*)$$

along any sequence (\mathbf{n}) in \mathbb{N}^J such that the sub-populations are large, that is, $\lim |\mathbf{n}| = \infty$; the conditioning is legit, that is, $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$, and the direction is strictly α , that is, $\lim_{|\mathbf{n}| \rightarrow \infty} \mathbf{n}/|\mathbf{n}| = \alpha$ and, with $\alpha = (\alpha_i)_{i \in \llbracket 1, J \rrbracket}$ and $\mathbf{n} = (n_i)_{i \in \llbracket 1, J \rrbracket}$, for all $j \in \llbracket 1, J \rrbracket$,

$$\alpha_j = 0 \implies n_j = 0. \tag{3}$$

Remark 1.1. (On the strict convergence of the sequence \mathbf{n} to the direction α .) To be complete, we also refer to Lemma 6.1 on the existence of a sequence of \mathbf{n} in \mathbb{N}^J satisfying the hypothesis above. Assume that the offspring distribution p is generic in the direction α and that α has some zero entries. We provide in Section 6.5 an example where removing condition (3) on the sequence of $\mathbf{n} = (n_i)_{i \in \llbracket 1, J \rrbracket}$ such that $\lim_{|\mathbf{n}| \rightarrow \infty} \mathbf{n}/|\mathbf{n}| = \alpha$ prevents us from obtaining the local limit of conditioned BGW trees from Theorem 1.1.

The proof of Theorem 1.1 relies on two ingredients. The first one is the use of Rizzolo’s transformation from [27] to reduce the problem to the A_0 empty case. The second is the existence of a local limit for multi-type BGW trees conditioned to the sub-populations of each type to be large (with the proportion given by the positive left eigenvector of the mean matrix) obtained in [3]. One could also use the results of Pénisson [25], but this would require stronger hypothesis; see Remark 6.4 for further comments. We note that Corollary 3.5 in

Abraham *et al.* [7] gives Theorem 1.1 in the very specific case where $\mu(p) = 1$, $p(A_0) = 0$, and the direction α is that naturally given by $p: \alpha_i = p(A_i)$.

Remark 1.2. (*Conditioning on the total population size and the number of leaves.*) Motivated by the scaling limits of BGW trees conditioned on its total population size and the number of leaves to be both large considered in [19, 22], we give as a consequence of the theorem above the corresponding local limit of such BGW trees in Remark 6.2. Conditioning on the total population size and the number of leaves amounts to considering $A_1 = \mathbb{N}^* \cap \text{supp}(p)$ and $A_2 = \{0\}$. Note the directions α can be written as $(a, 1 - a)$. As explained in Remark 6.2, we have to consider two cases.

If the support of p is reduced to two elements, say 0 and k (with $k \geq 2$ as p is nontrivial), then hypothesis (H4) is not satisfied. However, in this case the conditioning is equivalent to conditioning on the total population size and the existence of the local limit is then given by [4, 16]. Furthermore, there is only one possible direction for which p is generic; it is given by $a = 1/k$.

If the support of p is not reduced to two elements then, provided the smallest subgroup in \mathbb{Z} containing $\{x - y: x, y \in \text{supp}(p) \cap \mathbb{N}^*\}$ is \mathbb{Z} itself, hypothesis (H4) holds. In this case, hypothesis (H2) is satisfied if and only if $a \in (0, 1)$. It is easy to check that p is generic in the direction α if and only if there exists a positive finite root (which is then θ_α) to

$$g(\theta) = p(0) + a\theta g'(\theta),$$

where $g(\theta)$ is the probability generating function of p . The critical probability measure p_α is given by $p_\alpha(0) = 1 - a$ and $p_\alpha(n) = \theta_\alpha^{n-1} p(n) / g'(\theta_\alpha)$ for $n \in \mathbb{N}^*$; and we can apply Theorem 1.1.

Remark 1.3. (*On nongeneric distribution.*) If p is not generic in the possible direction α because of condition (i) in Theorem 5.1 then, as in the $J = 1$ case studied in [3], we conjecture the existence of a condensation phenomenon at the limit: the existence of a node of the local limit at finite height with an infinite degree. The first step to prove this would be considering the condensation for nongeneric multi-type BGW trees. Note that the others conditions in Theorem 5.1 might happen for probability distributions with bounded supports; see Remark 5.3(a).

The rest of the paper is structured as follows. We introduce in Section 2 the general notation and the framework of discrete trees, BGW trees, and Kesten's tree. We define the (p, \mathcal{A}) -compatible probability distributions in Section 3 and characterize all the (p, \mathcal{A}) -compatible probability distributions in Section 4 (handling the degenerate cases $\theta \in \{0, +\infty\}$ and the $0 \notin A_0$ case are delicate). We study in Section 5 the existence of the critical parameter θ_α and, thus, the probability distribution p_α that appears in the limiting Kesten's tree from Theorem 1.1. Eventually, we prove the main theorem in Section 6; see Theorem 6.1 as well as Remark 1.1.

2. Background and notation

2.1. General notation

We denote by $\mathbb{R}_+^* = (0, \infty)$ (respectively $\mathbb{N}^* = \{1, 2, \dots\}$) the set of positive real numbers (respectively integers) and by $\mathbb{R}_+ = [0, \infty)$ (respectively $\mathbb{N} = \{0, 1, \dots\}$) the set of nonnegative real numbers (respectively integers). For $i, j \in \mathbb{N}$ such that $i \leq j$, note that $\llbracket i, j \rrbracket = \mathbb{N} \cap [i, j]$. Let

$J \in \mathbb{N}^*$. For $\mathbf{x} = (x_j)_{j \in \llbracket 1, J \rrbracket} \in \mathbb{R}^J$, we set $|\mathbf{x}| = \sum_{j=1}^J |x_j|$. We consider the simplex

$$\Delta_J = \{\mathbf{x} \in \mathbb{R}_+^J : |\mathbf{x}| = 1\}.$$

We set $\mathbf{1} \in \mathbb{R}^J$ (respectively $\mathbf{0} \in \mathbb{R}^J$) the vector of \mathbb{R}^J with all its coordinates equal to 1 (respectively 0).

Let $p = (p(n))_{n \in \mathbb{N}}$ be a probability distribution on \mathbb{N} and $\text{supp}(p) = \{n \in \mathbb{N} : p(n) > 0\}$ be its support. For $A \subset \mathbb{N}$, we set

$$p(A) = \sum_{n \in A} p(n) \quad \text{and} \quad g_A(r) = \sum_{n \in A} r^n p(n) \quad \text{for } r \geq 0,$$

where the sum over an empty set is 0 by convention. In particular, we have $g_A = 0$ for any set $A \subset \mathbb{N}$ such that $p(A) = 0$. We also denote by ρ_A the radius of convergence of g_A :

$$\rho_A = \sup\{r \geq 1 : g_A(r) < +\infty\}. \tag{4}$$

For simplicity, when $A = \mathbb{N}$, we write $g(r) = g_{\mathbb{N}}(r)$ and $\rho = \rho_{\mathbb{N}}$. We write the mean of p by

$$\mu(p) = \sum_{n \in \mathbb{N}} np(n).$$

We say that a probability distribution p is *critical* (respectively *subcritical*) if $\mu(p) = 1$ (respectively $\mu(p) < 1$). The probability distribution p is *nontrivial* if

$$p(0) > 0 \quad \text{and} \quad p(0) + p(1) < 1. \tag{5}$$

2.2. The set of discrete trees

We consider ordered rooted trees in the framework of [24]. More precisely, let $\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^*)^n$ be the set of finite sequences of positive integers with the convention that $(\mathbb{N}^*)^0 = \{\emptyset\}$. Note that $H(u) = n$, the generation or the height of u if $u = (u_1, \dots, u_n) \in (\mathbb{N}^*)^n$. For $u, v \in \mathcal{U}$, denote by uv the concatenation of u and v , with the convention that $uv = u$ if $v = \emptyset$ and $uv = v$ if $u = \emptyset$. The set of ancestors of u is the set

$$\text{An}(u) = \{v \in \mathcal{U} : \text{there exists } w \in \mathcal{U} \text{ such that } u = vw\}.$$

The most recent common ancestor of $\mathbf{s} \subset \mathcal{U}$, denoted by $M(\mathbf{s})$, is the unique element u of $\bigcap_{u \in \mathbf{s}} \text{An}(u)$ with maximal height $H(u)$. Let $<$ be the usual lexicographic order on \mathcal{U} .

A tree \mathbf{t} is a subset of \mathcal{U} that satisfies $\emptyset \in \mathbf{t}$; if $u \in \mathbf{t}$ then $\text{An}(u) \subset \mathbf{t}$; for $u \in \mathbf{t}$, there exists $k_u(\mathbf{t}) \in \mathbb{N}^*$, called the out-degree of u , such that, for every $i \in \mathbb{N}^*$, $ui \in \mathbf{t}$ if and only if $i \in \llbracket 1, k_u(\mathbf{t}) \rrbracket$. The vertex \emptyset is called the root of \mathbf{t} . The vertex $u \in \mathbf{t}$ is called a leaf if $k_u(\mathbf{t}) = 0$. We set $k_u(\mathbf{t}) = -1$ if $u \notin \mathbf{t}$. Let $L_A(\mathbf{t})$ be the number of vertices of the tree \mathbf{t} whose out-degree belongs to $A \subset \mathbb{N}$:

$$L_A(\mathbf{t}) = \text{Card}(\mathcal{L}_A(\mathbf{t})) \quad \text{with} \quad \mathcal{L}_A(\mathbf{t}) = \{u \in \mathbf{t} : k_u(\mathbf{t}) \in A\}.$$

We simply write $\sharp \mathbf{t} = L_{\mathbb{N}}(\mathbf{t})$ for the cardinal of \mathbf{t} , and $L_n(\mathbf{t})$ for $L_{\{n\}}(\mathbf{t})$ when $n \in \mathbb{N}$. Note that we have

$$\sum_{u \in \mathbf{t}} k_u(\mathbf{t}) = \sharp \mathbf{t} - 1. \tag{6}$$

Then we get from (6),

$$L_0(\mathbf{t}) = 1 + \sum_{k \in \mathbb{N}^*} (k - 1)L_k(\mathbf{t}). \tag{7}$$

For $u \in \mathbf{t}$, we define the subtree above u by $\{v \in \mathcal{U} : uv \in \mathbf{t}\}$ and the fringe subtree by

$$\mathbf{s} = \{uv \in \mathcal{U} : uv \in \mathbf{t}\}. \tag{8}$$

We denote by \mathbb{T} the set of trees, \mathbb{T}_0 the subset of finite trees, and \mathbb{T}_1 the set of trees with only one infinite branch:

$$\mathbb{T}_1 = \left\{ \mathbf{t} \in \mathbb{T} \setminus \mathbb{T}_0 : \lim_{n \rightarrow \infty} H(M(\{u \in \mathbf{t} : H(u) = n\})) = \infty \right\}.$$

Let $H(\mathbf{t}) = \sup\{H(u) : u \in \mathbf{t}\}$ be the height of the tree \mathbf{t} and, for $h \in \mathbb{N}^*$, let $\mathbb{T}^{(h)} = \{\mathbf{t} \in \mathbb{T} : H(\mathbf{t}) \leq h\}$ be the set of trees with height less or equal to h .

2.3. Local convergence of trees

For $h \in \mathbb{N}$ and a tree $\mathbf{t} \in \mathbb{T}$, let $r_h(\mathbf{t}) = \{u \in \mathbf{t} : H(u) \leq h\}$ be the tree \mathbf{t} truncated at level h . Let $(T_n)_{n \in \mathbb{N}}$ and T be \mathbb{T} -valued random variables. We say that the sequence $(T_n)_{n \in \mathbb{N}}$ converges locally in distribution towards T if, for all $h \in \mathbb{N}$ and $\mathbf{t} \in \mathbb{T}^{(h)}$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(r_h(T_n) = \mathbf{t}) = \mathbb{P}(r_h(T) = \mathbf{t}), \tag{9}$$

and writing $\text{dist}(T)$ for the distribution of the random variable T , we denote it by

$$\lim_{n \rightarrow \infty} \text{dist}(T_n) = \text{dist}(T).$$

For $\mathbf{t} \in \mathbb{T}$ and $x \in \mathcal{L}_0(\mathbf{t})$, we consider a the set of trees $\mathbb{T}(\mathbf{t}, x) = \{\mathbf{t} \circledast (\tilde{\mathbf{t}}, x) : \tilde{\mathbf{t}} \in \mathbb{T}\}$, where

$$\mathbf{t} \circledast (\tilde{\mathbf{t}}, x) = \{u \in \mathbf{t}\} \cup \{xv : v \in \tilde{\mathbf{t}}\}$$

is the tree obtained by grafting $\tilde{\mathbf{t}}$ on the leaf x of \mathbf{t} . We recall from [4, Lemma 2.1], that if $(T_n)_{n \in \mathbb{N}}$ and T are $\mathbb{T}_0 \cup \mathbb{T}_1$ -valued random variables, then the sequence $(T_n)_{n \in \mathbb{N}}$ converges locally in distribution towards T if and only if, for all $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_n \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(T \in \mathbb{T}(\mathbf{t}, x)) \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{P}(T_n = \mathbf{t}) = \mathbb{P}(T = \mathbf{t}). \tag{10}$$

In practice, when dealing with conditioned BGW trees, checking conditions (10), when the local limit has an infinite spine, is usually easier than checking condition (9); see [5, Section 3.2] for more details and examples.

2.4. BGW trees

Let p be a probability distribution on \mathbb{N} . A \mathbb{T} -valued random variable τ is a BGW tree with offspring distribution p if $k_{\emptyset}(\tau)$ is distributed as p and the branching property is satisfied: for $n \in \mathbb{N}^*$, conditionally on $\{k_{\emptyset}(\tau) = n\}$, the subtrees $(S_1(\tau), \dots, S_n(\tau))$ are independent and distributed as τ . We denote by \mathcal{T}_p the BGW tree with offspring distribution p . For all finite trees $\mathbf{t} \in \mathbb{T}_0$, we have

$$\mathbb{P}(\mathcal{T}_p = \mathbf{t}) = \prod_{u \in \mathbf{t}} p(k_u(\mathbf{t})) = \prod_{n \in \mathbb{N}} p(n)^{L_n(\mathbf{t})}, \tag{11}$$

with the convention that $0^0 = 1$. When (5) holds and p is critical or subcritical, then a.s. \mathcal{T}_p is finite (that is, $\mathcal{T}_p \in \mathbb{T}_0$) and in this case (11) completely characterizes the distribution of \mathcal{T}_p .

2.5. Kesten’s tree

Let p be a critical probability distribution on \mathbb{N} (and, thus, $\mu(p) = 1$) satisfying (5). We denote by $p^* = (p^*(n))_{n \in \mathbb{N}}$, with $p^*(n) = np(n)$, the corresponding size-biased distribution. The so-called Kesten’s tree, \mathcal{T}_p^* , is a \mathbb{T}_1 -valued random tree defined as the local limit in distribution, when n tends to infinity, of a BGW tree conditioned to have a height larger than n :

$$\lim_{n \rightarrow \infty} \text{dist}(\mathcal{T}_p \mid H(\mathcal{T}_p) = n) = \text{dist}(\mathcal{T}_p^*).$$

Informally, it is the skeleton of a two-type BGW tree, where individuals are of type s (survivor) or n (normal); the root is of type s; each individual of type s has a random number of children with offspring distribution p^* , all of them of type n but for one uniformly chosen at random that is of type s; each individual of type n has a random number of children with offspring distribution p , all of them of type n. Its distribution is completely characterized by $\mathbb{P}(\mathcal{T}_p^* \in \mathbb{T}_1) = 1$ and (recall p is critical)

$$\mathbb{P}(\mathcal{T}_p^* \in \mathbb{T}(\mathbf{t}, x)) = \frac{\mathbb{P}(\mathcal{T}_p = \mathbf{t})}{p(0)} \quad \text{for all } \mathbf{t} \in \mathbb{T}_0 \text{ and } x \in \mathcal{L}_0(\mathbf{t}). \tag{12}$$

We refer to [4, Section 2.4] for more details on this equality.

Remark 2.1. (On (12).) Let p be critical. Let $r_h^*(\mathcal{T}_p^*)$ denote the Kesten’s tree \mathcal{T}_p^* whose spine is cut at level $h \geq 0$. By summing (12) over x a leaf of \mathbf{t} at height h and then summing over \mathbf{t} , we deduce that, with F a nonnegative function defined on \mathbb{T}_0 ,

$$\mathbb{E}[F(r_h^*(\mathcal{T}_p^*))] = \frac{1}{p(0)} \mathbb{E}[Z_h^0 F(\mathcal{T}_p)], \tag{13}$$

where Z_h^0 is the number of leaves of the BGW tree \mathcal{T}_p at height h . (From here it is elementary to deduce the well-known formula $\mathbb{E}[G(r_h(\mathcal{T}_p^*))] = \mathbb{E}[Z_h G(r_h(\mathcal{T}_p))]$, which also characterizes the distribution of Kesten’s tree.) We derive from (13) that, for a given $h \geq 0$, Kesten’s tree is distributed as a biased BGW tree with weight $Z_h^0/p(0)$ on which is grafted an independent Kesten’s tree at a uniformly chosen leaf of height h .

3. Definition of the distribution $\tilde{p}_{\theta, \beta}$

In this section we lay the groundwork for identifying the offspring distributions, as used in Theorem 4.1, such that the corresponding conditioned BGW trees have the same distribution. This leads us to consider exponentially tilted probability distributions (nondegenerate case) derived from an initial offspring distribution as well as some limiting (or degenerate) cases.

Let p be a probability distribution on \mathbb{N} satisfying (5). Recall that $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$ are pairwise disjoint nonempty subsets of $\text{supp}(p)$ and that A_0 is the complement of $\cup_{j \in \llbracket 1, J \rrbracket} A_j$ in $\text{supp}(p)$. Note that A_0 may be empty.

For a probability distribution q on \mathbb{N} , we consider the probabilities of the sets \mathcal{A} :

$$q(\mathcal{A}) = (q(A_1), \dots, q(A_J)) \in [0, 1]^J.$$

For a finite tree \mathbf{t} , let $L_{\mathcal{A}}(\mathbf{t})$ be the number of nodes in \mathbf{t} with out-degree in the finite collection \mathcal{A} of pairwise disjoint sets:

$$L_{\mathcal{A}}(\mathbf{t}) = (L_{A_1}(\mathbf{t}), \dots, L_{A_J}(\mathbf{t})) \in \mathbb{N}^J.$$

We extend the definition of the generic probability distribution with respect to a subset of \mathbb{N} in [4] to the generic probability distribution with respect to a family of subsets. Intuitively, an offspring distribution p is *generic* if there exists a critical offspring distribution p' such that conditional distributions of the BGW trees \mathcal{T}_p and $\mathcal{T}_{p'}$ are the same; and we see that the Kesten’s tree associated with p' appears then as the local limit of the BGW \mathcal{T}_p with some asymptotic conditioning. For this reason, we also investigate the family of *compatible* offspring distributions, that is, those offspring distributions for which the distribution of the conditioned BGW tree is identical.

Definition 3.1. (*Generic probability distribution.*) Let p be a probability distribution on \mathbb{N} satisfying (5). Let $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, be pairwise disjoint nonempty subsets of $\text{supp}(p)$.

- (i) *(p, A)-compatible probability distribution:* we say that a probability distribution p' is (p, \mathcal{A}) -compatible if p' satisfies (5), $\text{supp}(p') \subset \text{supp}(p)$, and for all $\mathbf{n} \in \mathbb{N}^J$ such that $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$ (and, thus, $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}) > 0$), we have

$$\text{dist}(\mathcal{T}_p \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) = \text{dist}(\mathcal{T}_{p'} \mid L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}). \tag{14}$$

- (ii) *Generic distribution:* we say that p is generic for \mathcal{A} in the direction $\alpha \in \Delta_J$ if there exists a critical (p, \mathcal{A}) -compatible probability distribution p' such that

$$p'(\mathcal{A}) = (1 - p'(A_0)) \alpha \quad \text{with } p'(A_0) < 1.$$

Remark 3.1. (*On the one-dimensional case.*) When $J = 1$, Definition 3.1(ii) reduces to the definition of the generic probability distribution with respect to the set $A_1 \subset \mathbb{N}$ with $\alpha = 1$.

Remark 3.2. (*Positivity of α .*) If p is generic for \mathcal{A} in the direction $\alpha \in \Delta_J$ with $0 \in A_j$ for some $j \in \llbracket 1, J \rrbracket$, then we have $\alpha_j > 0$. Indeed, the probability distribution p' from Definition 3.1(ii) is (p, \mathcal{A}) -compatible and, thus, $p'(A_j) = p'(A_0^c) \alpha_j$, with $p'(A_0^c) > 0$; since p' also satisfies (5), we deduce that $p'(A_j) \geq p'(0) > 0$, which then gives $\alpha_j > 0$.

We now introduce a set of parameters that will allow us to describe all the compatible probability distributions. We set

$$\mathcal{J}_{\infty} = \{j \in \llbracket 1, J \rrbracket : \sup A_j < \infty\}.$$

Definition 3.2. (*The set of parameters $\text{Pa}(p, \mathcal{A})$.*) Let p be a probability distribution on \mathbb{N} satisfying (5). Let $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, be pairwise disjoint nonempty subsets of $\text{supp}(p)$. For $\beta = (\beta_1, \dots, \beta_J) \in \mathbb{R}_+^J$, we set

$$\mathcal{J}_{\beta}^* = \{j \in \llbracket 1, J \rrbracket : \beta_j > 0\} \quad \text{and} \quad \mathcal{J}_{\beta} = \{0\} \cup \mathcal{J}_{\beta}^*. \tag{15}$$

The set $\text{Pa}(p, \mathcal{A}) \subset [0, +\infty] \times \mathbb{R}_+^J$ of parameters is defined as follows.

- (i) *Non degenerate case:* the parameter $(\theta, \beta) \in \mathbb{R}_+^* \times \mathbb{R}_+^J$ belongs to $\text{Pa}(p, \mathcal{A})$ if

$$\sum_{j \in \mathcal{J}_{\beta}} \beta_j g_{A_j}(\theta) = 1, \quad \text{where } \beta_0 = \theta^{-1}. \tag{16}$$

(ii) *Degenerate case:* the parameter (θ, β) , with $\theta \in \{0, +\infty\}$ and $\beta \in \mathbb{R}_+^J$, belongs to $\text{Pa}(p, \mathcal{A})$ if

$$\sum_{j \in \mathcal{J}_\beta} \beta_j = 1, \quad \text{where } \beta_0 = p(1) 1_{\{1 \in A_0\}}, \tag{17}$$

and

$$\text{if } \theta = 0 \text{ then } 0 \notin A_0, \tag{18}$$

$$\text{if } \theta = +\infty \text{ then } A_0 \subset \{0, 1\} \text{ and } \mathcal{J}_\beta^* \subset \mathcal{J}_\infty. \tag{19}$$

For simplicity, we write \mathcal{J}^* and \mathcal{J} for \mathcal{J}_β^* and \mathcal{J}_β . Note that \mathcal{J}^* might be empty in the nondegenerate case, and that \mathcal{J}^* is nonempty in the degenerate cases thanks to (5).

We now define the family of probability distributions indexed by the parameter (θ, β) . In the nondegenerate case the probability distribution puts weight on either all elements of A_j or none of them for each $j \in \mathcal{J}$; in the degenerate case the probability distribution puts weight either on the minimal element of A_j for some $j \in \mathcal{J}$ or on the maximal element of A_j for some $j \in \mathcal{J}$.

Definition 3.3. (*Probability distribution $\tilde{p}_{\theta, \beta}$.*) Let p be a probability distribution on \mathbb{N} satisfying (5). Let $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, be pairwise disjoint nonempty subsets of $\text{supp}(p)$. For $(\theta, \beta) \in \text{Pa}(p, \mathcal{A})$, we define the probability distribution $\tilde{p}_{\theta, \beta} = (\tilde{p}_{\theta, \beta}(n))_{n \in \mathbb{N}}$ as follows, where the unspecified probabilities are set to 0.

(i) *Non degenerate case:* if $\theta \in \mathbb{R}_+^*$ then we set

$$\tilde{p}_{\theta, \beta}(n) = \beta_j \theta^n p(n) \quad \text{for } j \in \mathcal{J} \text{ and } n \in A_j. \tag{20}$$

(ii) *Degenerate case at 0:* if $\theta = 0$ then we set

$$\tilde{p}_{0, \beta}(n) = \beta_j \quad \text{for } j \in \mathcal{J} \text{ and } n = \min A_j. \tag{21}$$

(iii) *Degenerate case at infinity:* if $\theta = +\infty$ then we set

$$\tilde{p}_{\infty, \beta}(n) = \beta_j \quad \text{for } j \in \mathcal{J} \text{ and } n = \max A_j. \tag{22}$$

Conditions (16) and (17) ensure that $\tilde{p}_{\theta, \beta}$ is indeed a probability distribution. In the next remark, we consider some particular cases of the probability distributions $\tilde{p}_{\theta, \beta}$. For convenience, for the BGW trees, we write

$$\mathcal{T}_{\theta, \beta} = \mathcal{T}_{\tilde{p}_{\theta, \beta}}.$$

Recall that $g_{A_0} = 0$ if $A_0 = \emptyset$. We consider the equation $g_{A_0}(\theta) = \theta$ on \mathbb{R}_+ which has at least one root and at most two. It is elementary to check that

$$\theta_{\min} = \min\{\theta \in \mathbb{R}_+ : g_{A_0}(\theta) = \theta\} \in [0, 1), \tag{23}$$

and that the second root, if it exists, belongs to $(1, +\infty)$. We deduce from (16) that, in the non-degenerate case, if the parameter (θ, β) belongs to $\text{Pa}(p, \mathcal{A})$ then $\theta > \theta_{\min}$. In the degenerate case, Lemma 3.1 states that if $(0, \beta)$ belongs to $\text{Pa}(p, \mathcal{A})$ then $\theta \geq \theta_{\min} = 0$. Conversely, we

indeed check that θ_{\min} is the infimum of the values θ for which there exists a β such that $\tilde{p}_{\theta,\beta}$ is a (p, \mathcal{A}) -compatible probability distribution (this a consequence of Theorem 4.1, (37), (39), and the comment thereafter).

Lemma 3.1. *We have $\theta_{\min} = 0$ if and only if $0 \notin A_0$.*

Remark 3.3. (On particular cases.) Recall $\mathbf{1}$ (respectively $\mathbf{0}$) denotes the vector of \mathbb{R}^J with all its coordinates equal to 1 (respectively 0).

- (i) Case $\theta = 1$ and $\beta = \mathbf{1}$: we trivially have $\tilde{p}_{(1,\mathbf{1})} = p$, and thus, $(1, \mathbf{1}) \in \text{Pa}(p, \mathcal{A})$. In particular, we get $\mathcal{T}_p = \mathcal{T}_{1,\mathbf{1}}$.
- (ii) Case $\theta = 0$: the support of $\tilde{p}_{0,\beta}$ is equal to $\{1\} \cup \{\min A_j : j \in \llbracket 1, J \rrbracket\}$ or to $\{\min A_j : j \in \llbracket 1, J \rrbracket\}$ according to 1 belonging to A_0 or not.
- (iii) Case $\theta = \infty$: the support of $\tilde{p}_{\infty,\beta}$ is equal to $\{1\} \cup \{\max A_j : j \in \mathcal{J}_\infty\}$ or $\{\max A_j : j \in \mathcal{J}_\infty\}$ according to 1 belonging to A_0 or not.
- (iv) Case $\beta = \mathbf{0}$: if $(\theta, \beta) \in \text{Pa}(p, \mathcal{A})$ then we have that $\beta = \mathbf{0}$ is equivalent to $\tilde{p}_{\theta,\mathbf{0}}(A_0^c) = 0$. In this case, we have $\theta \in (0, +\infty)$ (as (5), which implies that $p(1) < 1$, and (17) rule out the case $\theta \in \{0, +\infty\}$) and that θ is a root of $g_{A_0}(\theta) = \theta$ by (16). Thus, θ can take the value θ_{\min} (only if $\theta_{\min} > 0$ and then $0 \in A_0$), and possibly another value, say $\theta_M \in (1, +\infty)$. We also have that $\mu(\tilde{p}_{\theta,\mathbf{0}}) = g'_{A_0}(\theta)$, so that $\tilde{p}_{\theta_{\min},\mathbf{0}}$ is subcritical and, if θ_M exists, $\tilde{p}_{\theta_M,\mathbf{0}}$ is supercritical. In conclusion, for $\tilde{p}_{\theta,\mathbf{0}}$ not to be supercritical, we need $0 \in A_0$ and $\theta = \theta_{\min}$.

In order to (partially) remove the particular case $\beta = \mathbf{0}$, we set

$$\text{Pa}^{**}(p, \mathcal{A}) = \{(\theta, \beta) \in \text{Pa}(p, \mathcal{A}) : \beta \neq \mathbf{0}\}, \tag{24}$$

$$\text{Pa}^*(p, \mathcal{A}) = \begin{cases} \text{Pa}^{**}(p, \mathcal{A}) & \text{if } 0 \notin A_0, \\ \text{Pa}^{**}(p, \mathcal{A}) \cup \{(\theta_{\min}, \mathbf{0})\} & \text{if } 0 \in A_0. \end{cases} \tag{25}$$

Recall that we are interested in conditioning the BGW tree \mathcal{T}_p on the event $\{L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}\}$ for some $\mathbf{n} \in \mathbb{R}^J$ and that, under some mild assumptions, we prove in Theorem 4.1 that $\mathcal{T}_{\theta,\beta}$ similarly conditioned has the same distribution. For this to be meaningful, we shall at least check that the tree $\mathcal{T}_{\theta,\beta}$ conditioned to have a finite number of nodes in A_0^c is finite. The purpose of the next lemma is indeed to check this condition provided that (θ, β) belongs to the set $\text{Pa}^*(p, \mathcal{A})$.

Lemma 3.2. (Conditional finiteness of $\mathcal{T}_{\theta,\beta}$.) *Let $(\theta, \beta) \in \text{Pa}(p, \mathcal{A})$. We have*

$$\mathbb{P}(\mathcal{T}_{\theta,\beta} \notin \mathbb{T}_0, \sup_{j \in \mathcal{J}^*} L_{A_j}(\mathcal{T}_{\theta,\beta}) < +\infty) = 0 \quad \text{if and only if } (\theta, \beta) \in \text{Pa}^*(p, \mathcal{A}).$$

Proof. If $\beta \neq \mathbf{0}$ then there exists $j \in \llbracket 1, J \rrbracket$ such that $\beta_j > 0$. This gives that a.s. on the event $\{\mathcal{T}_{\theta,\beta} \notin \mathbb{T}_0\}$ we have $L_{A_j}(\mathcal{T}_{\theta,\beta}) = +\infty$ and, thus,

$$\mathbb{P}(\mathcal{T}_{\theta,\beta} \notin \mathbb{T}_0, \sup_{j \in \mathcal{J}^*} L_{A_j}(\mathcal{T}_{\theta,\beta}) < +\infty) = 0.$$

If $\beta = \mathbf{0}$, we deduce from Remark 3.3(iv) that if $(\theta, \mathbf{0}) \in \text{Pa}(p, \mathcal{A})$ then we have either $\theta = \theta_{\min} > 0$, $(\theta, \mathbf{0}) \in \text{Pa}^*(p, \mathcal{A})$, $\tilde{p}_{\theta, \mathbf{0}}$ is subcritical; or $\theta > \theta_{\min}$, $(\theta, \mathbf{0}) \notin \text{Pa}^*(p, \mathcal{A})$ and $\tilde{p}_{\theta, \mathbf{0}}$ is supercritical. In the former case, we get $\mathbb{P}(\mathcal{T}_{\theta, \mathbf{0}} \notin \mathbb{T}_0) = 0$ and in the latter case,

$$\mathbb{P}(\mathcal{T}_{\theta, \mathbf{0}} \notin \mathbb{T}_0, \sup_{j \in \mathcal{J}^*} L_{A_j}(\mathcal{T}_{\theta, \mathbf{0}}) < +\infty) = \mathbb{P}(\mathcal{T}_{\theta, \mathbf{0}} \notin \mathbb{T}_0) > 0.$$

This gives the result. □

We now restrict our study to the case where $\tilde{p}_{\theta, \beta}$ is nontrivial.

Definition 3.4. (*Compatible parameter.*) The parameter (θ, β) is (p, \mathcal{A}) -compatible if $(\theta, \beta) \in \text{Pa}^*(p, \mathcal{A})$ and the probability distribution $\tilde{p}_{\theta, \beta}$ satisfies condition (5).

We now characterize the (p, \mathcal{A}) -compatible parameters. For $\mathcal{J} \subset \llbracket 0, J \rrbracket$, we set

$$A_{\mathcal{J}} = \bigcup_{j \in \mathcal{J}} A_j.$$

Lemma 3.3. (*Characterization of the compatible parameters.*) *The parameter $(\theta, \beta) \in \text{Pa}^*(p, \mathcal{A})$ is (p, \mathcal{A}) -compatible if and only if*

- (i) for $\theta \in (0, +\infty)$, we have $0 \in A_{\mathcal{J}}$ and $p(A_{\mathcal{J}} \cap \{0, 1\}^c) > 0$;
- (ii) for $\theta = 0$, we have $0 \in A_{\mathcal{J}^*}$ and $\max_{j \in \mathcal{J}^*} \min(A_j) > 1$;
- (iii) for $\theta = +\infty$, we have $A_{j_0} = \{0\}$ for some $j_0 \in \mathcal{J}^*$ and $\max_{j \in \mathcal{J}^*} \max(A_j) > 1$ (and $A_0 \subset \{1\}$ and $\mathcal{J}^* \subset \mathcal{J}_{\infty}$).

Proof. By construction, we have $\text{supp}(\tilde{p}_{\theta, \beta}) \subset \text{supp}(p)$. Since $p(0) > 0$, there exists $j_0 \in \llbracket 0, J \rrbracket$ such that $0 \in A_{j_0}$.

Let $\theta \in (0, +\infty)$. The condition $0 \in A_{\mathcal{J}}$ in Lemma 3.3(i) is equivalent to $\beta_{j_0} > 0$ and, thus, to $\tilde{p}_{\theta, \beta}(0) > 0$. The condition $p(A_{\mathcal{J}} \cap \{0, 1\}^c) > 0$ is clearly equivalent to $\tilde{p}_{\theta, \beta}(\{0, 1\}^c) > 0$. Therefore, conditions in Lemma 3.3(i) are equivalent to $\tilde{p}_{\theta, \beta}$ satisfying the nondegeneracy condition (5).

Let $\theta = 0$ (and, thus, $0 \notin A_0$). The condition $0 \in A_{\mathcal{J}^*}$ in Lemma 3.3(ii) is equivalent to $\beta_{j_0} > 0$ and to $\tilde{p}_{\theta, \beta}(0) > 0$, as $\min A_{j_0} = 0$. Since $\{\min A_j : j \in \mathcal{J}^*\} \subset \text{supp}(\tilde{p}_{\theta, \beta}) \subset \{\min A_j : j \in \mathcal{J}^*\} \cup \{1\}$, we deduce that $\max_{j \in \mathcal{J}^*} \min(A_j) > 1$ is equivalent to $\tilde{p}_{\theta, \beta}(\{0, 1\}^c) > 0$. Thus, conditions in Lemma 3.3(ii) are equivalent to $\tilde{p}_{\theta, \beta}$ satisfying (5).

Let $\theta = +\infty$. The conditions $j_0 \in \mathcal{J}^*$ and $A_{j_0} = \{0\}$ in Lemma 3.3(iii) are equivalent to $\tilde{p}_{\theta, \beta}(0) > 0$ (note that, by (17), $\tilde{p}_{\theta, \beta}(A_0) > 0$ if and only if $1 \in A_0$, and then $0 \in A_0$ implies that $\tilde{p}_{\theta, \beta}(0) = 0$). Eventually the condition $\max_{j \in \mathcal{J}^*} \max(A_j) > 1$ is also clearly equivalent to $\tilde{p}_{\theta, \beta}(\{0, 1\}^c) > 0$. Thus, conditions in Lemma 3.3(iii) are equivalent to $\tilde{p}_{\theta, \beta}$ satisfying (5). □

4. The (p, \mathcal{A}) -compatible probability distributions

Recall that p is a probability distribution on \mathbb{N} satisfying (5) and $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, are pairwise disjoint nonempty subsets of $\text{supp}(p)$. This section is devoted to the following exhaustive characterization result that identify all the (p, \mathcal{A}) -compatible probability distributions; taking into account the degenerate cases adds extra technicalities.

Theorem 4.1. (Characterization of the compatible probability distributions.) *Let p be a probability distribution on \mathbb{N} satisfying (5). The probability distribution p' is (p, \mathcal{A}) -compatible if and only if $p' = \tilde{p}_{\theta, \beta}$ for some (p, \mathcal{A}) -compatible parameter (θ, β) .*

As being (p, \mathcal{A}) -compatible implies, by definition, that p is nontrivial (that is, condition (5) holds), we only consider the probability distributions $\tilde{p}_{\theta, \beta}$ satisfying the conditions of Lemma 3.3.

The end of this section is devoted to the proof of this result. We first prove that $\tilde{p}_{\theta, \beta}$ are (p, \mathcal{A}) -compatible, provided condition (5) holds; see Lemma 4.1 for the nondegenerate cases and Lemma 4.2 for the degenerate cases. Then we prove that all (p, \mathcal{A}) -compatible probability distributions are of the form $\tilde{p}_{\theta, \beta}$, distinguishing according to $0 \in A_0$ in Lemma 4.3 or $0 \notin A_0$ in Lemma 4.4, where the proof of the latter is more technical.

We set $\mathbf{e}^0 = \mathbf{0}$ and, for $j \in \llbracket 1, J \rrbracket$,

$$\mathbf{e}^j = (e_1^j, \dots, e_J^j) \in \mathbb{N}^J \quad \text{with } e_i^j = 1_{\{i=j\}}.$$

In particular, we have $\sum_{j=1}^J \mathbf{e}^j = \mathbf{1}$. Note that $\mathbb{P}(\mathcal{T}_p = \mathbf{t}) > 0$ implies that $\mathbf{t} \in \mathbb{T}_0$ and $k_u(\mathbf{t}) \in A_{\llbracket 0, J \rrbracket}$ for all $u \in \mathbf{t}$.

Lemma 4.1. *Under the assumptions of Theorem 4.1, if the parameter (θ, β) is (p, \mathcal{A}) -compatible with $\theta \in (0, +\infty)$ then the probability distribution $\tilde{p}_{\theta, \beta}$ is (p, \mathcal{A}) -compatible.*

Proof. We suppose that the parameter (θ, β) is (p, \mathcal{A}) -compatible (which implies that $\tilde{p}_{\theta, \beta}$ is nontrivial) and that $\theta \in (0, +\infty)$. We prove that the probability distribution $\tilde{p}_{\theta, \beta}$ is (p, \mathcal{A}) -compatible. Recall that $\mathcal{J} = \{0\} \cup \{j \in \llbracket 1, J \rrbracket : \beta_j > 0\}$.

Note that $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{\theta, \beta}) = \mathbf{n}) > 0$ implies that $n_j = 0$ for all $j \notin \mathcal{J}^*$, where $\mathbf{n} = (n_1, \dots, n_J)$. Using the definition of $\tilde{p}_{\theta, \beta}$, we obtain, for $\mathbf{t} \in \mathbb{T}_0$ such that $\mathbb{P}(\mathcal{T}_p = \mathbf{t}) > 0$ and $L_{A_j}(\mathbf{t}) = 0$ for $j \notin \mathcal{J}$,

$$\begin{aligned} \mathbb{P}(\mathcal{T}_{\theta, \beta} = \mathbf{t}) &= \prod_{u \in \mathbf{t}} \tilde{p}_{\theta, \beta}(k_u(\mathbf{t})) \\ &= \prod_{j \in \mathcal{J}} \prod_{u \in \mathcal{L}_{A_j}(\mathbf{t})} \beta_j \theta^{k_u(\mathbf{t})} p(k_u(\mathbf{t})) \\ &= \mathbb{P}(\mathcal{T}_p = \mathbf{t}) \theta^{\sharp \mathbf{t} - 1} \prod_{j \in \mathcal{J}} \beta_j^{L_{A_j}(\mathbf{t})} \\ &= \mathbb{P}(\mathcal{T}_p = \mathbf{t}) \theta^{-1} \prod_{j \in \mathcal{J}^*} (\beta_j \theta)^{L_{A_j}(\mathbf{t})}, \end{aligned}$$

where we used (6) for the third equality, and $\sum_{j \in \mathcal{J}} L_{A_j}(\mathbf{t}) = \sharp \mathbf{t}$ as well as $\beta_0 \theta = 1$ for the fourth equality. As $\mathbb{P}(\mathcal{T}_p = \mathbf{t}) = 0$ implies that $\mathbb{P}(\mathcal{T}_{\theta, \beta} = \mathbf{t}) = 0$, we deduce that, for $\mathbf{n} \in \mathbb{N}^J$,

$$\begin{aligned} \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{\theta, \beta}) = \mathbf{n}) &= \sum_{\mathbf{t} \in \mathbb{T}_0, L_{\mathcal{A}}(\mathbf{t}) = \mathbf{n}} \mathbb{P}(\mathcal{T}_{\theta, \beta} = \mathbf{t}) + \mathbb{P}(\mathcal{T}_{\theta, \beta} \notin \mathbb{T}_0, L_{\mathcal{A}}(\mathcal{T}_{\theta, \beta}) = \mathbf{n}) \\ &= \sum_{\mathbf{t} \in \mathbb{T}_0, L_{\mathcal{A}}(\mathbf{t}) = \mathbf{n}} \mathbb{P}(\mathcal{T}_{\theta, \beta} = \mathbf{t}) \\ &= \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) \theta^{-1} \prod_{j \in \mathcal{J}^*} (\beta_j \theta)^{n_j}, \end{aligned}$$

where we used Lemma 3.2 for the second equality.

So for every $\mathbf{n} \in \mathbb{N}^J$ such that $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{\theta, \beta}) = \mathbf{n}) > 0$, we have $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$, and for every $\mathbf{t} \in \mathbb{T}_0$ such that $L_{\mathcal{A}}(\mathbf{t}) = \mathbf{n}$, we get

$$\mathbb{P}(\mathcal{T}_{\theta, \beta} = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}_{\theta, \beta}) = \mathbf{n}) = \mathbb{P}(\mathcal{T}_p = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}).$$

Hence, the probability distribution $\tilde{p}_{\theta, \beta}$ is (p, \mathcal{A}) -compatible. □

Lemma 4.2. *Under the assumptions of Theorem 4.1, if the parameter (θ, β) is (p, \mathcal{A}) -compatible with $\theta \in \{0, +\infty\}$ then the probability distribution $\tilde{p}_{\theta, \beta}$ is (p, \mathcal{A}) -compatible.*

Proof. We first consider the case $\theta = +\infty$. For simplicity, we write $p' = \tilde{p}_{\infty, \beta}$ and $\mathcal{T}' = \mathcal{T}_{\infty, \beta}$. We suppose that the parameter (∞, β) is (p, \mathcal{A}) -compatible. As p' is nontrivial, it follows that $\beta \in \mathbb{R}'_+ \setminus \{0\}$ and, thus, $\mathcal{J}^* \neq \emptyset$. In particular, we have $A_0 \subset \{1\}$, thanks to (19) and Lemma 3.3(iii). We prove that the probability distribution p' is (p, \mathcal{A}) -compatible.

Note that $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}') = \mathbf{n}) > 0$ implies that $n_j = 0$ for all $j \notin \mathcal{J}^*$, where $\mathbf{n} = (n_1, \dots, n_J)$, and, by (6), that

$$\sum_{j \in \mathcal{J}^*} n_j (\max A_j - 1) = -1. \tag{26}$$

To simplify notation, recall that we write L_k for $L_{\{k\}}$ for $k \in \mathbb{N}$. Fix such $\mathbf{n} \in \mathbb{N}^J$ and consider the set $\mathbb{T}_{0, n} = \{\mathbf{t} \in \mathbb{T}_0 : L_{\mathcal{A}}(\mathbf{t}) = \mathbf{n} \text{ and } \mathbb{P}(\mathcal{T}' = \mathbf{t}) > 0\}$, which is clearly not empty. Using (6), it follows that, for $\mathbf{t} \in \mathbb{T}_{0, n}$, $\sum_{j \in \mathcal{J}^*} \sum_{k \in A_j} L_k(\mathbf{t})(k - 1) = -1$. We deduce from (26) that, for $k \notin A_0$, we have $L_k(\mathbf{t}) = n_j$ if $j \in \mathcal{J}^*$ and $k = \max A_j$, and $L_k(\mathbf{t}) = 0$ otherwise.

Let $\mathbf{t} \in \mathbb{T}_{0, n}$. We distinguish two cases according to 1 belonging to A_0 or not. First, we consider the case $1 \in A_0$, that is, $A_0 = \{1\}$, elementary computation gives

$$\mathbb{P}(\mathcal{T}' = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}') = \mathbf{n}) = c^{-1} p(1)^{L_1(\mathbf{t})},$$

with, as $\mathcal{J}^* \neq \emptyset$, c positive and finite given by

$$c = \sum_{\mathbf{t}' \in \mathbb{T}_0, L_{\mathcal{A}}(\mathbf{t}') = \mathbf{n}} p(1)^{L_1(\mathbf{t}')} + \mathbb{P}(\mathcal{T}' \notin \mathbb{T}_0, L_{\mathcal{A}}(\mathcal{T}') = \mathbf{n}) = \sum_{\mathbf{t}' \in \mathbb{T}_0, L_{\mathcal{A}}(\mathbf{t}') = \mathbf{n}} p(1)^{L_1(\mathbf{t}')}.$$

Similarly, we have

$$\mathbb{P}(\mathcal{T}_p = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) = c^{-1} p(1)^{L_1(\mathbf{t})}.$$

This readily implies that p' is (p, \mathcal{A}) -compatible.

Secondly, we consider the case $1 \notin A_0$, that is, $A_0 = \emptyset$. It follows that the set $\mathbb{T}_{0, n}$ is finite. Similarly to the first case, we show that, with $c = \text{Card}(\mathbb{T}_{0, n}) \geq 1$,

$$\mathbb{P}(\mathcal{T}' = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}') = \mathbf{n}) = \mathbb{P}(\mathcal{T}_p = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) = c^{-1}.$$

This readily also implies that p' is (p, \mathcal{A}) -compatible.

Eventually, the $\theta = 0$ case can be handled similarly using $\sum_{j \in \mathcal{J}^*} n_j (\min A_j - 1) = -1$ instead of (26). □

Lemma 4.3. *Under the assumptions of Theorem 4.1, if the probability distribution p' is (p, \mathcal{A}) -compatible and $0 \in A_0$ then we have $p' = \tilde{p}_{\theta, \beta}$ for some (p, \mathcal{A}) -compatible parameter (θ, β) .*

Proof. For $k \in \mathbb{N}$, let \mathbf{t}_k denote the tree with the root having k children, all of them being leaves (that is, $k_{\emptyset}(\mathbf{t}_k) = k$ and $\sharp \mathbf{t}_k = k + 1$).

We assume that $0 \in A_0$ and p' is (p, \mathcal{A}) -compatible. For simplicity, we write \mathcal{T}' for $\mathcal{T}_{p'}$. We have $p(0) > 0$ and $p'(0) > 0$ since p and p' satisfy (5). Let $j \in \llbracket 0, J \rrbracket$ be such that $p'(A_j) > 0$. There exists $k_j \in A_j$ such that $p'(k_j) > 0$. We have

$$\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}') = \mathbf{e}^j) \geq \mathbb{P}(\mathcal{T}' = \mathbf{t}_{k_j}) = p'(k_j)p'(0)^{k_j} > 0.$$

This also implies that $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{e}^j) > 0$ as $\text{supp}(p') \subset \text{supp}(p)$. In particular, thanks to (14), we have, for any $k \in A_j$,

$$\mathbb{P}(\mathcal{T}' = \mathbf{t}_k \mid L_{\mathcal{A}}(\mathcal{T}') = \mathbf{e}^j) = \mathbb{P}(\mathcal{T}_p = \mathbf{t}_k \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{e}^j).$$

This gives

$$\frac{p'(k)p'(0)^k}{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}') = \mathbf{e}^j)} = \frac{p(k)p(0)^k}{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{e}^j)},$$

that is,

$$p'(k) = \beta_j \theta^k p(k) \tag{27}$$

with $\theta = p(0)/p'(0) > 0$ and $\beta_j = \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}') = \mathbf{e}^j)/\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{e}^j) > 0$. Note that θ and β_j do not depend on $k \in A_j$. For $j = 0$, we have $p'(A_0) > 0$ as $p'(0) > 0$. For $k = 0 \in A_0$, we deduce from (27) that $\beta_0 = p'(0)/p(0) = 1/\theta$. For $j \in \llbracket 0, J \rrbracket$ such that $p'(A_j) = 0$, (27) also holds with $\beta_j = 0$. Note that when $\beta = \mathbf{0}$, we have $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}') = \mathbf{0}) = 1$ and $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{0}) < 1$, which entails the fact that $\beta_0 = \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}') = \mathbf{0})/\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{0}) > 1$, or equivalently, $\theta \in (0, 1)$. This proves that $p' = \tilde{p}_{\theta, \beta}$, the latter being defined in (20) as $\theta \in (0, +\infty)$.

To conclude, note that condition (16) holds as p' is a probability distribution and $\beta_0 = 1/\theta$. □

Lemma 4.4. *Under the assumptions of Theorem 4.1, if the probability distribution p' is (p, \mathcal{A}) -compatible and $0 \notin A_0$ then we have $p' = \tilde{p}_{\theta, \beta}$ for some (p, \mathcal{A}) -compatible parameter (θ, β) .*

The proof of this lemma is more technical and relies on the following result whose proof is postponed to the end of this section. We introduce the sets

$$\begin{aligned} \mathcal{J}^* &= \{j \in \llbracket 1, J \rrbracket : p'(A_j) > 0\}, \\ \mathcal{J}^{**} &= \{j \in \mathcal{J}^* : \text{Card}(A_j) \geq 2\}. \end{aligned}$$

Note that once Lemma 4.4 is proved, then \mathcal{J}^* coincides with \mathcal{J}_β^* as defined in (15).

Lemma 4.5. *Assume that $0 \notin A_0$ and let p' be a (p, \mathcal{A}) -compatible distribution. We have the following properties.*

1. *The map $\ell \mapsto (p'(\ell)/p(\ell))^{1/(\ell-1)}$ is constant over $\{\ell \in A_0 : \ell \geq 2\}$.*
2. *Assume that there exists $\ell \in A_0$ such that $\ell \geq 2$ and $k', k \in A_j$ with $j \in \mathcal{J}^{**}$ such that $k' > k \geq 0$. Then we have*

$$p'(k') > 0 \iff p'(k) > 0 \text{ and } p'(\ell) > 0.$$

Furthermore, if these conditions hold then we also have, with $\alpha = \ell - 1$ and $\beta = k' - k$,

$$\left(\frac{p'(k)}{p'(k')}\right)^\alpha p'(\ell)^\beta = \left(\frac{p(k)}{p(k')}\right)^\alpha p(\ell)^\beta. \tag{28}$$

3. Assume that there exist $i, j \in \mathcal{J}^{**}$, $\ell, k_i \in A_i$ such that $\ell > k_i \geq 0$ and $p'(k_i) > 0$, and $k_j, k \in A_j$ such that $k_j > k \geq 0$ and $p'(k_j) > 0$. Then we have $p'(\ell) > 0$ and $p'(k) > 0$ as well as, with $\alpha = \ell - k_i$ and $\beta = k_j - k$,

$$p'(k)^\alpha p'(\ell)^\beta = \left(\frac{p'(k_j)}{p(k_j)}\right)^\alpha \left(\frac{p'(k_i)}{p(k_i)}\right)^\beta p(k)^\alpha p(\ell)^\beta. \tag{29}$$

Proof of Lemma 4.4. We assume that $0 \notin A_0$, that is, without loss of generality, $0 \in A_1$, and that p' is (p, \mathcal{A}) -compatible. Since p and p' satisfy (5), we get $p(0) > 0$ and $p'(0) > 0$. So, we can define $\beta_1 = p'(0)/p(0) > 0$. By considering trees with vertices having zero or one child, we get

$$1 \in A_0 \implies p'(1) = p(1). \tag{30}$$

Recall the set $\mathcal{J}^* = \{j \in \llbracket 1, J \rrbracket : p'(A_j) > 0\}$. We set $\beta_j = 0$ for $j \in \llbracket 1, J \rrbracket \setminus \mathcal{J}^*$ in accordance with the definition of the $\tilde{p}_{\theta, \beta}$. Note that \mathcal{J}^* is nonempty as $1 \in \mathcal{J}^*$. For $j \in \mathcal{J}^*$, let k_j be an element of A_j such that $p'(k_j) > 0$. For $j \in \mathcal{J}^* \setminus \mathcal{J}^{**}$, we have $A_j = \{k_j\}$ and, whatever the value of θ , which will be defined later on, we set $\beta_j = \theta^{-k_j} p'(k_j)/p(k_j)$ if $\theta \in (0, +\infty)$ and $\beta_j = p'(k_j)/p(k_j)$ otherwise; this is also in accordance with the definition of the $\tilde{p}_{\theta, \beta}$. We now have to consider the value of $p'(\ell)$ for $\ell \in A_j$ and $j \in \{0\} \cup \mathcal{J}^{**}$.

We first consider the case $\mathcal{J}^{**} = \emptyset$. In particular, it follows that $A_1 = \{0\}$. Then we have either that $p'(A_0) = p'(A_0 \cap \{1\})$ or there exists $k_0 \in A_0$ such that $k_0 \geq 2$ and $p'(k_0) > 0$. In the former case, using (30) and $\beta_0 = p(1) 1_{\{1 \in A_0\}}$, the probability distribution can be written as a $\tilde{p}_{\theta, \beta}$ with $\theta = 0$. In the latter case, using Lemma 4.5(i) and (30), we deduce that (20) holds for all $\ell \in A_0$ with a common $\theta \in (0, +\infty)$ (given by the constant value of the map in Lemma 4.5(i) and $\beta_0 = 1/\theta$; thus, the probability distribution p' can be written as a $\tilde{p}_{\theta, \beta}$ with $\theta \in (0, +\infty)$).

We now assume that $\mathcal{J}^{**} \neq \emptyset$. From Lemma 4.5(iii), only the following three cases are possible.

- (C1) For all $j \in \mathcal{J}^{**}$, we have $k_j = \min A_j$, $p'(k_j) > 0$ and $p'(A_j \setminus \{k_j\}) = 0$.
- (C2) For all $j \in \mathcal{J}^{**}$, we have $k_j = \max A_j$, $p'(k_j) > 0$ and $p'(A_j \setminus \{k_j\}) = 0$.
- (C3) For all $j \in \mathcal{J}^{**}$, we have $p'(k) > 0$ for all $k \in A_j$.

We investigate each case separately. In (C1) we deduce from Lemma 4.5(ii) that $p'(\ell) = 0$ for all $\ell \in A_0$ with $\ell \geq 2$. Thus, using (30), it follows that the probability distribution p' can be written as a $\tilde{p}_{\theta, \beta}$ with $\theta = 0$.

In (C2) since $p'(0) > 0$, we deduce that $A_1 = \{0\}$ and that if j belongs to \mathcal{J}^{**} (and, thus, to \mathcal{J}^*) then $\sup A_j$ is finite. Then we use Lemma 4.5(ii) to deduce that there is no element $\ell \geq 2$ in A_0 , that is, $A_0 \subset \{1\}$. Thus, using (30), it follows that the probability distribution p' can be written as a $\tilde{p}_{\theta, \beta}$ with $\theta = +\infty$.

In (C3), to fix ideas, let $i = \min \mathcal{J}^{**}$ and consider that $k_i = \min A_i$ and $\ell = \min A_i \setminus \{k_i\}$. This uniquely determines β_i and the $\theta \in (0, +\infty)$ solution of, for $k' \in \{k_i, \ell\}$,

$$p'(k') = \beta_i \theta^{k'} p(k'). \tag{31}$$

Then for $k'' \in A_i$ larger than ℓ , we use Lemma 4.5(iii) and, in particular, (29), with $j = i$, $k_j = k''$ and $k = \ell$ to deduce that $p'(k'')$ can also be written as in (31). For $j \in \mathcal{J}^{**}$ with $j \neq i$, set $k = \min A_j$ and define β_j by

$$p'(k) = \beta_j \theta^k p(k). \tag{32}$$

Then we use Lemma 4.5(iii) and, in particular, (29), with $k_j > k$ (and $k_j \in A_j$), to deduce that (32) holds for k replaced by k_j . Then for $\ell \geq 2$ in A_0 , we use Lemma 4.5(ii) to obtain $p'(\ell) = \theta^{\ell-1}p(\ell)$ (which is also consistent with (30)). We deduce that the probability distribution p' can be written as a $\tilde{p}_{\theta, \beta}$ with $\theta \in (0, +\infty)$. This concludes the proof. \square

Proof of Lemma 4.5. Without loss of generality, we assume that $0 \in A_1$. In the following description of trees we do not give the precise number of leaves, as it is determined through (7). The argument is based on considering two well-chosen trees \mathbf{t} and \mathbf{t}' such that $L_{\mathcal{A}}(\mathbf{t}) = L_{\mathcal{A}}(\mathbf{t}')$. Recall that we write L_k for $L_{\{k\}}$ and $k \in \mathbb{N}$.

We prove part (i) of the lemma. Assume that there exist $k, k' \in A_0$ such that $\min(k, k') \geq 2$. Consider a tree \mathbf{t} having only leaves and $\alpha = k' - 1 > 0$ vertices of out-degree k (that is, $L_k(\mathbf{t}) = \alpha$) and a tree \mathbf{t}' having only leaves and $\beta = k - 1 > 0$ vertices of out-degree k' (that is, $L_{k'}(\mathbf{t}') = \beta$). Set $a = 1 + (k' - 1)(k - 1) > 1$. Thanks to (7), we get $L_0(\mathbf{t}) = L_0(\mathbf{t}') = a$, and thus,

$$L_{\mathcal{A}}(\mathbf{t}) = L_{\mathcal{A}}(\mathbf{t}') = \mathbf{n} \quad \text{with } \mathbf{n} = a \mathbf{e}^1.$$

Assume that $p'(k) > 0$. We have $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}) \geq \mathbb{P}(\mathcal{T}_{p'} = \mathbf{t}) = p'(k)^\alpha p'(0)^a > 0$, and thus, $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$. We set

$$c = \frac{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n})}{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n})} > 0.$$

Then, we apply (14) to the trees \mathbf{t} and \mathbf{t}' to obtain

$$p'(k)^\alpha p'(0)^a = cp(k)^\alpha p(0)^a \quad \text{and} \quad p'(k')^\beta p'(0)^a = cp(k')^\beta p(0)^a.$$

This readily implies that $p'(k')$ is positive (as $k' \in A_0$ implies that $p(k') > 0$) and that $(p'(k)/p(k))^{1/(k-1)} = (p'(k')/p(k'))^{1/(k'-1)}$. This gives Lemma 4.5(i).

We now prove part (ii) of the lemma. Assume that there exist $k' > k \geq 0$, which are elements of A_j with $j \in \mathcal{J}^{**}$ and $\ell \geq 2$, which is an element of A_0 . Note that j can possibly take the value 1 when $1 \in \mathcal{J}^{**}$ (that is, $\text{Card}(A_1) \geq 2$). Consider a tree \mathbf{t} having only leaves and $\alpha = \ell - 1 > 0$ vertices of out-degree k' (that is, $L_{k'}(\mathbf{t}) = \alpha$) and a tree \mathbf{t}' having only leaves, $\beta = k' - k > 0$ vertices of out-degree ℓ and, if $k > 0$, α vertices of out-degree k (that is, $L_\ell(\mathbf{t}') = \beta$ and, if $k > 0$, $L_k(\mathbf{t}') = \alpha$).

We first assume that $k \geq 1$ (which is automatically satisfied if $j \geq 2$). Thanks to (7), we get with $a = 1 + \alpha(k' - 1) = 1 + \alpha(k - 1) + \beta(\ell - 1)$ that $L_0(\mathbf{t}) = L_0(\mathbf{t}') = a$, and thus,

$$L_{\mathcal{A}}(\mathbf{t}) = L_{\mathcal{A}}(\mathbf{t}') = \mathbf{n} \quad \text{with } \mathbf{n} = a \mathbf{e}^1 + \alpha \mathbf{e}^k. \tag{33}$$

Assume that $p'(k') > 0$. We have $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}) \geq \mathbb{P}(\mathcal{T}_{p'} = \mathbf{t}) = p'(k')^\alpha p'(0)^a > 0$, and thus, $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$. We set

$$c = \frac{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n})}{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n})} > 0.$$

Then, we apply (14) to the trees \mathbf{t} and \mathbf{t}' to get

$$p'(k')^\alpha p'(0)^a = cp(k')^\alpha p(0)^a \quad \text{and} \quad p'(k)^\alpha p'(\ell)^\beta p'(0)^a = cp(k)^\alpha p(\ell)^\beta p(0)^a.$$

This readily implies that $p'(k)$ and $p'(\ell)$ are positive (as $k \in A_j$ and $\ell \in A_0$ imply that $p(k)$ and $p(\ell)$ are positive). Similarly, assuming that $p'(k)$ and $p'(\ell)$ are positive implies that $p'(k')$ is positive. Note also that (28) is obvious.

We now consider the case $k = 0$ and, thus, $j = 1$ (as $0 \in A_1$). We have $L_{A_1}(\mathbf{t}) = L_0(\mathbf{t}) + L_{k'}(\mathbf{t}) = 1 + \alpha(k' - 1) + \alpha = 1 + (\ell - 1)k'$ and $L_{A_1}(\mathbf{t}') = L_0(\mathbf{t}') = 1 + \beta(\ell - 1) = 1 + (\ell - 1)k'$, which implies that (33) still holds. We then conclude similarly as in the case $k > 0$. This gives Lemma 4.5(ii).

Finally, we prove part (iii) of the lemma. Assume that there exist $i, j \in \mathcal{J}^{**}$, $k_j > k \geq 0$ that are elements of A_j , $\ell > k_i \geq 0$ that are elements of A_i , with $p'(k_j) > 0$ and $p'(k_i) > 0$. Note that i and j can be possibly equal and can possibly take the value 1 when $1 \in \mathcal{J}^{**}$. Consider a tree \mathbf{t} having only leaves, $\alpha = \ell - k_i > 0$ vertices of out-degree k_j and, if $k_i > 0$, $\beta = k_j - k > 0$ vertices of out-degree k_i (that is, $L_{k_j}(\mathbf{t}) = \alpha$ and, if $k_i > 0$, $L_{k_i}(\mathbf{t}) = \beta$) and a tree \mathbf{t}' having only leaves, β vertices of out-degree ℓ , and, if $k > 0$, α vertices of out-degree k (that is, $L_\ell(\mathbf{t}') = \beta$ and, if $k > 0$, $L_k(\mathbf{t}') = \alpha$).

We assume that $k \geq 1$ and $k_i \geq 1$, and leave the cases $k = 0$ (and, thus, $j = 1$) and/or $k_i = 0$ (and, thus, $i = 1$) to the reader as the proof can be handled very similarly; see also the end of the proof of Lemma 4.5(ii). Thanks to (7), we get with $a = 1 + \alpha(k_j - 1) + \beta(k_i - 1) = 1 + \alpha(k - 1) + \beta(\ell - 1)$ that $L_0(\mathbf{t}) = L_0(\mathbf{t}') = a$, and thus,

$$L_{\mathcal{A}}(\mathbf{t}) = L_{\mathcal{A}}(\mathbf{t}') = \mathbf{n} \quad \text{with } \mathbf{n} = a \mathbf{e}^1 + \alpha \mathbf{e}^j + \beta \mathbf{e}^i.$$

We have $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n}) \geq \mathbb{P}(\mathcal{T}_{p'} = \mathbf{t}) = p'(k_j)^\alpha p'(k_i)^\beta p'(0)^a > 0$; thus, $\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}) > 0$. We set

$$c = \frac{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_{p'}) = \mathbf{n})}{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n})} > 0.$$

Then, we apply (14) to the trees \mathbf{t} and \mathbf{t}' to get

$$p'(k_j)^\alpha p'(k_i)^\beta p'(0)^a = cp(k_j)^\alpha p(k_i)^\beta p(0)^a$$

and

$$p'(k)^\alpha p'(\ell)^\beta p'(0)^a = cp(k)^\alpha p(\ell)^\beta p(0)^a.$$

This readily implies that $p'(k)$ and $p'(\ell)$ are positive (as $k \in A_j$ and $\ell \in A_i$ imply that $p(k)$ and $p(\ell)$ are positive). Equation (29) is then obvious. This gives Lemma 4.5(iii). □

5. Existence of a critical (p, \mathcal{A}) -compatible distribution

In Section 5.1 we first provide a new parametrization $(p_{\theta, \alpha})$ of the (p, \mathcal{A}) -compatible probability distributions $(\tilde{p}_{\theta, \beta})$ adapted to the direction α that appears in Theorem 1.1; this entails a discussion on the possible values of α . Then, in Section 5.2 we study the mean function $\theta \mapsto \mu_{\theta, \alpha} = \mu(p_{\theta, \alpha})$ for a given possible direction α ; and we provide in Section 5.3 necessary and sufficient conditions for the existence of (a unique) θ_α such that $p_{\theta_\alpha, \alpha}$ is critical, that is, p is generic. Recall Δ_J is the simplex of \mathbb{R}^J .

5.1. Parametrization of the (p, \mathcal{A}) -compatible probability distributions using their direction

We define the direction of a probability distribution p' with respect to \mathcal{A} .

Definition 5.1. (*Direction of compatible probability distributions.*) The direction $\alpha \in \Delta_J$ of a (p, \mathcal{A}) -compatible probability distribution p' such that $p'(A_0) < 1$ is defined by

$$\alpha = a^{-1} p'(\mathcal{A}) \quad \text{with } a = 1 - p'(A_0) > 0. \tag{34}$$

Because by definition p' is nontrivial, see (5), we see below that the set of possible directions is

$$\Delta_J^* = \Delta_J \setminus \Delta_J^o, \tag{35}$$

where the set of ineligible directions are given by

$$\Delta_J^o = \bigcup_{j=1}^J \{ \alpha \in \Delta_J : \alpha_j = 0 \text{ if } 0 \in A_j \text{ or } \alpha_j = 1 \text{ if } A_0 \cup A_j \subset \{0, 1\} \}.$$

We use a parametrization of the probability distribution $\tilde{p}_{\theta, \beta}$ using the parameter θ and its direction $\alpha \in \Delta_J^*$. Recall that $\tilde{p}_{\theta, \beta}(A_0) = 1$ if and only if $\beta = \mathbf{0}$; see Remark 3.3(iv). Recall the set of parameters $\text{Pa}^{**}(p, \mathcal{A})$ defined in (24), where the $\beta = \mathbf{0}$ cases are removed, and the set of (p, \mathcal{A}) -compatible parameters given in Definition 3.4. The direction map $\mathcal{D}: (\theta, \beta) \mapsto (\theta, \alpha)$ given by (34) is defined on the subset of $\text{Pa}^{**}(p, \mathcal{A})$ of (p, \mathcal{A}) -compatible parameters; and it is clearly injective.

Definition 5.2. (*Compatible parameters.*) The parameter $(\theta, \alpha) = \mathcal{D}(\theta, \beta)$ is (p, \mathcal{A}) -compatible if (θ, β) is (p, \mathcal{A}) -compatible (and $\beta \neq \mathbf{0}$).

We write $p_{\theta, \alpha}$ for $\tilde{p}_{\theta, \beta}$ when $(\theta, \alpha) = \mathcal{D}(\theta, \beta)$, and it satisfies (5) when (θ, β) is (p, \mathcal{A}) -compatible. Note that, for $\alpha \in \Delta_J^*$, the set

$$\mathcal{J}_\beta^* = \{j \in \llbracket 1, J \rrbracket : \beta_j > 0\} = \{j \in \llbracket 1, J \rrbracket : \alpha_j > 0\}$$

is by definition nonempty; we also denote it by \mathcal{J}_α^* or simply \mathcal{J}^* when there is no ambiguity on the parameter. For the convenience of the reader, we give explicit formulas for the probability distributions $p_{\theta, \alpha}$, using (34) and Lemma 3.3. To simplify the expression we set

$$q_1 = 1 - p(1)1_{\{1 \in A_0\}} > 0. \tag{36}$$

Lemma 5.1. (The probability distributions $p_{\theta, \alpha}$.) *Let $\alpha \in \Delta_J^*$ and $\theta \in [0, +\infty]$ be such that the parameter (θ, α) is (p, \mathcal{A}) -compatible. The nonzero terms of the probability distribution $p_{\theta, \alpha}$ are given as follows. (Recall that $0 \in A_{\mathcal{J}}$ and $p(A_{\mathcal{J}} \cap \{0, 1\}^c) > 0$.)*

(i) *If $\theta \in (0, +\infty)$ then we have $g_{A_0}(\theta) < \theta$, $g_{A_j}(\theta) < +\infty$ for $j \in \mathcal{J}^*$ and*

$$p_{\theta, \alpha}(k) = \theta^{k-1} p(k) \quad \text{for } k \in A_0,$$

$$p_{\theta, \alpha}(k) = \alpha_j \frac{\theta - g_{A_0}(\theta)}{g_{A_j}(\theta)} \theta^{k-1} p(k) \quad \text{for } k \in A_j \text{ and } j \in \mathcal{J}^*.$$

(ii) *If $\theta = 0$ then we have $0 \notin A_0$, $\max_{j \in \mathcal{J}^*} (A_j) > 1$, $p_{\theta, \alpha}(1) = p(1)$ if $1 \in A_0$ and*

$$p_{\theta, \alpha}(k) = \alpha_j q_1 \quad \text{for } k = \min A_j \text{ and } j \in \mathcal{J}^*.$$

(iii) *If $\theta = +\infty$ then we have $A_0 \subset \{1\}$, $A_{j_0} = \{0\}$ for some $j_0 \in \mathcal{J}^*$, $\mathcal{J}^* \subset \mathcal{J}_\infty$, $\max_{j \in \mathcal{J}^*} \max(A_j) > 1$, $p_{\theta, \alpha}(1) = p(1)$ if $1 \in A_0$ and*

$$p_{\theta, \alpha}(k) = \alpha_j q_1 \quad \text{for } k = \max A_j \text{ and } j \in \mathcal{J}^*.$$

For $\alpha \in \Delta_j^*$, it follows that $p_{1,\alpha} = \tilde{p}_{1,\beta}$ with $\beta_j = p(A_0^c) \alpha_j / p(A_j)$ satisfying (5), and thus, that $(1, \alpha)$ is (p, \mathcal{A}) -compatible. So Δ_j^* is indeed the set of all possible directions of (p, \mathcal{A}) -compatible probability distributions. We however complete this picture with the following result.

Proposition 5.1. (Possible directions.) *For every $\alpha \in \Delta_j^*$, there exists $\theta \in (0, 1)$ such that (θ, α) is (p, \mathcal{A}) -compatible, that is, such that (θ, β) is (p, \mathcal{A}) -compatible for some $\beta \in \mathbb{R}_+^J$ with $\beta \neq \mathbf{0}$ and α is the direction of $\tilde{p}_{\theta,\beta}$.*

Proof. We recall the convention $g_{A_0} = 0$ if $A_0 = \emptyset$. Let us first prove that there exists $\theta \in (0, 1)$ such that $\theta > g_{A_0}(\theta)$. If $0 \in A_0$ then $g_{A_0}(0) = p(0) > 0$ and $g_{A_0}(1) < 1$, so there exists $\theta_0 \in (0, 1)$ such that $\theta_0 = g_{A_0}(\theta_0)$, and then $g'_{A_0}(\theta_0) < 1$. This implies that $\theta > g_{A_0}(\theta)$ for $\theta \in (\theta_0, 1)$. If $0 \notin A_0$ then $g_{A_0}(0) = 0$ and $g_{A_0}(1) < 1$, so 0 is the only root of $\theta = g_{A_0}(\theta)$ in $[0, 1]$, and thus, $\theta > g_{A_0}(\theta)$ for $\theta \in (0, 1)$.

Let us now fix $\theta \in (0, 1)$ such that $\theta > g_{A_0}(\theta)$. We set, for all $j \in \llbracket 1, J \rrbracket$,

$$\beta_j = \alpha_j \frac{\theta - g_{A_0}(\theta)}{\theta g_{A_j}(\theta)} \geq 0$$

and $\beta_0 = \theta^{-1}$. Since $\sum_{j=1}^J \alpha_j = 1$, we have

$$\sum_{j \in \llbracket 0, J \rrbracket} \beta_j g_{A_j}(\theta) = \theta^{-1} g_{A_0}(\theta) + \sum_{j \in \llbracket 1, J \rrbracket} \alpha_j \theta^{-1} (\theta - g_{A_0}(\theta)) = 1.$$

Hence, condition (16) is satisfied. Moreover, we have

$$\tilde{p}_{\theta,\beta}(A_0^c) = 1 - \tilde{p}_{\theta,\beta}(A_0) = \theta^{-1} (\theta - g_{A_0}(\theta)) > 0.$$

Finally, $\alpha \in \Delta_j^*$ (and, thus, $\alpha \notin \Delta_j^?$) insures that $\tilde{p}_{\theta,\beta}$ satisfies Lemma 3.3(i), that is, $\tilde{p}_{\theta,\beta}$ is nontrivial. □

5.2. Properties of the mean of $p_{\theta,\alpha}$

For $\alpha \in \Delta_j^*$, we consider the following set:

$$I_\alpha = \{\theta \in [0, +\infty]: (\theta, \alpha) \text{ is } (p, \mathcal{A})\text{-compatible}\}. \tag{37}$$

Note that $1 \in I_\alpha$. Note $\rho_{\mathcal{J}}$ as the radius of convergence of $\sum_{j \in \mathcal{J}} g_{A_j}$ or, equivalently,

$$\rho_{\mathcal{J}} = \min_{j \in \mathcal{J}} \rho_{A_j} \in [1, +\infty]. \tag{38}$$

We define

$$\theta_{\min} = \inf I_\alpha \in [0, 1) \quad \text{and} \quad \theta_{\max} = \sup I_\alpha \in [1, \rho_{\mathcal{J}}]. \tag{39}$$

On the one hand, note that θ_{\min} is the only root of $g_{A_0}(\theta) = \theta$ in $[0, 1)$, so this definition is consistent with (23), and thus, θ_{\min} does not depend on α . On the other hand, we have that θ_{\max} depends on the support of α as

$$\theta_{\max} = \max(\rho_{\mathcal{J}^*}, \sup\{\theta \in [1, \rho_{A_0}): g_{A_0}(\theta) < \theta\}) \quad \text{with} \quad \rho_{\mathcal{J}^*} = \min_{j \in \mathcal{J}^*} \rho_{A_j}. \tag{40}$$

We also have that $(\theta_{\min}, \theta_{\max}) \subset I_\alpha$ and that $\theta \mapsto p_{\theta,\alpha}$ is continuous on I_α for the norm of the total variation. The next result is a direct consequence of Lemma 5.1 (the first point is also in Lemma 3.1).

Lemma 5.2. *We have*

- (i) $\theta_{\min} = 0$ if and only if $0 \notin A_0$,
- (ii) $\theta_{\min} \in I_\alpha$ if and only if $0 \notin A_0$ and $\max_{j \in \mathcal{J}^*} \min(A_j) > 1$,
- (iii) if $0 \in A_0$ then we have $I_\alpha \subset (0, +\infty)$.

In order to consider finite means, we set, for $\alpha \in \Delta_J^*$,

$$I_\alpha^f = \{\theta \in I_\alpha : \mu_{\theta, \alpha} < +\infty\}, \quad \text{where } \mu_{\theta, \alpha} = \mu(p_{\theta, \alpha}) \in [0, +\infty] \tag{41}$$

is the mean of $p_{\theta, \alpha}$. Note that $I_\alpha \subset I_\alpha^f \cup \{\theta_{\max}\}$.

We are now interested in the existence of a critical probability distribution among the (p, \mathcal{A}) -compatible probability distributions $p_{\theta, \alpha}$ with a given direction $\alpha \in \Delta_J^*$, that is, in the existence of $\theta \in I_\alpha$ such that $\mu_{\theta, \alpha} = 1$.

For $\alpha \in \Delta_J^*$ and $\theta \geq 0$ such that $\sum_{j \in \mathcal{J}^*} g_{A_j}(\theta) < +\infty$, we set

$$H_\alpha(\theta) = \sum_{j \in \mathcal{J}^*} \alpha_j h_j(\theta), \tag{42}$$

where for $j \in \mathcal{J}^*$, $h_j(0) = \min A_j$ and, for $\theta > 0$,

$$h_j(\theta) = \frac{\theta g'_{A_j}(\theta)}{g_{A_j}(\theta)} = \frac{\mathbb{E}[X\theta^X \mathbf{1}_{\{X \in A_j\}}]}{\mathbb{E}[\theta^X \mathbf{1}_{\{X \in A_j\}}]},$$

where X is distributed according to p . For $\theta \in I_\alpha \cap \mathbb{R}_+^*$, using Lemma 5.1(i), we have

$$\mu_{\theta, \alpha} = g'_{A_0}(\theta) + \frac{\theta - g_{A_0}(\theta)}{\theta} H_\alpha(\theta). \tag{43}$$

Recall q_1 from (36). Using Lemma 5.1, if $0 \in I_\alpha$, we have

$$\mu_{0, \alpha} = (1 - q_1) + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j \min A_j \tag{44}$$

and if $+\infty \in I_\alpha$, we have

$$\mu_{\infty, \alpha} = (1 - q_1) + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j \max A_j. \tag{45}$$

We now recall some elementary properties of the function h_j ; see also [16, Lemma 3.1] for a part of the proof. Note that if A_j is a singleton, say $\{k_j\}$, then the function h_j is constant and equal to k_j . Recall that ρ_{A_j} is the radius of convergence of g_{A_j} and that $\lim_{x \rightarrow \rho_{A_j}} g_{A_j}(x) = g_{A_j}(\rho_{A_j})$, with the limit being possibly infinite.

Lemma 5.3. *Let $j \in \llbracket 1, J \rrbracket$ with $\text{Card}(A_j) \geq 2$. The function h_j defined on $[0, \rho_{A_j})$ is \mathcal{C}^1 and increasing, with $h'_j > 0$ on $(0, \rho_{A_j})$. If $\rho_{A_j} = +\infty$ or if $g_{A_j}(\rho_{A_j}) = +\infty$, then we have $\lim_{\theta \rightarrow \rho_{A_j}} h_j(\theta) = \sup A_j$.*

As an immediate application, we get the following result (we only need to take care of the $\theta = 0$ case, where $H_\alpha(0) = \sum_{j \in \mathcal{J}^*} \alpha_j \min A_j$, and of the $\theta = +\infty$ case, where $H_\alpha(\infty) = \sum_{j \in \mathcal{J}^*} \alpha_j \sup A_j$ if $\min_{j \in \mathcal{J}^*} \rho_j = +\infty$).

Corollary 5.1. (Regularity of $\mu_{\theta, \alpha}$.) *The map $\theta \mapsto \mu_{\theta, \alpha}$ is continuous on I_α , finite on I_α^f , and \mathcal{C}^1 on $(\theta_{\min}, \theta_{\max})$ and also on $[0, \theta_{\max})$ if $0 \in I_\alpha$.*

5.3. Generic distributions

Note that $1 \in I_\alpha$; however, we do not assume *a priori* that $1 \in I_\alpha^f$ as $p_{1,\alpha}$ might have infinite mean. Recall that $\rho_{\mathcal{J}} = \min_{j \in \mathcal{J}} \rho_{A_j}$; see (38).

In the next lemmas we give preliminary results on the existence of $\theta \in I_\alpha$ for $p_{\theta,\alpha}$ to be subcritical/supercritical according to 0 belonging to A_0 or not.

Lemma 5.4. *Assume that $0 \in A_0$ and let $\alpha \in \Delta_j^*$.*

- (i) *There exists $\theta \in I_\alpha$ such that $\mu_{\theta,\alpha} < 1$.*
- (ii) *If $\rho_{\mathcal{J}} = +\infty$ then there exists $\theta \in I_\alpha$ such that $\mu_{\theta,\alpha} > 1$.*

Proof. As $0 \in A_0$, it follows that, using Lemma 5.2, $I_\alpha \cap [0, 1] = (\theta_{\min}, 1]$. By continuity, we deduce from (43) that

$$\lim_{\theta \downarrow \theta_{\min}} \mu_{\theta,\alpha} = g'_{A_0}(\theta_{\min}) < 1.$$

This gives Lemma 5.4(i).

We now prove part (ii) of the lemma. On the one hand, if there exists $k \in A_0$ such that $k \geq 2$, then the function g_{A_0} is strictly convex and $\lim_{\theta \rightarrow \infty} g_{A_0}(\theta)/\theta = +\infty$ as $\rho_{A_0} \geq \rho_{\mathcal{J}} = +\infty$. We deduce that $g_{A_0}(\theta_{\max}) = \theta_{\max}$ and, as $\theta_{\max} > 1$, that $g'_{A_0}(\theta_{\max}) > 1$. By continuity, we deduce from (43) that

$$\lim_{\theta \uparrow \theta_{\max}} \mu_{\theta,\alpha} = g'_{A_0}(\theta_{\max}) > 1.$$

On the other hand, if $A_0 \subset \{0, 1\}$ then it follows that $g_{A_0}(\theta) = p(0) + (1 - q_1)\theta$. Thus, there is no root of $g_{A_0}(\theta) = \theta$ on $(1, +\infty)$, but we have

$$\lim_{\theta \rightarrow \infty} \frac{\theta - g_{A_0}(\theta)}{\theta} = q_1 > 0.$$

Then, we use Lemma 5.3 and (43) to deduce that

$$\lim_{\theta \rightarrow +\infty} \mu_{\theta,\alpha} = (1 - q_1) + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j \sup A_j > (1 - q_1) + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j = 1,$$

where, for the inequality, we used the fact that $\sup_{j \in \mathcal{J}^*} \sup A_j > 1$ as $p_{\theta,\alpha}$ is nontrivial and $\sup A_0 \leq 1$. This gives Lemma 5.4(ii). □

Recall that $\mathcal{J}_\infty = \{j \in \llbracket 1, J \rrbracket : \sup A_j < \infty\}$.

Lemma 5.5. *Assume that $0 \notin A_0$ and let $\alpha \in \Delta_j^*$.*

- (i) *There exists $\theta \in I_\alpha$ such that $\mu_{\theta,\alpha} \leq 1$ if and only if*

$$\sum_{j \in \mathcal{J}^*} \alpha_j \min A_j \leq 1. \tag{46}$$

- (ii) *If $\rho_{\mathcal{J}} = +\infty$ then we have $\mu_{\theta,\alpha} < 1$ for all $\theta \in I_\alpha$ if and only if*

$$A_0 \subset \{1\}, \quad \mathcal{J}^* \subset \mathcal{J}_\infty, \quad \text{and} \quad \sum_{j \in \mathcal{J}^*} \alpha_j \max A_j < 1. \tag{47}$$

Proof. We prove part (i) of the lemma. We first assume that $\sum_{j \in \mathcal{J}^*} \alpha_j \min A_j \leq 1$. By assumption, $0 \notin A_0$ and $0 \in A_{\mathcal{J}^*}$ as $1 \in I_\alpha$ and $p_{1,\alpha}$ satisfies (5). According to Lemma 5.2, we have $\theta_{\min} = 0$.

If $0 \in I_\alpha$, we deduce from (44) that $\mu_{0,\alpha} = 1 - q_1 + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j \min A_j \leq 1$.

If $0 \notin I_\alpha$, that is, $\max_{j \in \mathcal{J}^*} \min(A_j) \leq 1$, it follows that $\sum_{j \in \mathcal{J}^*} \alpha_j \min A_j < 1$ as $0 \in A_{\mathcal{J}^*}$, that is, $\min A_j = 0$ for some $j \in \mathcal{J}^*$ and $\alpha_j > 0$. We then deduce from (43), using the fact that q_1 defined in (36) is positive, that

$$\lim_{\theta \downarrow 0} \mu_{\theta,\alpha} = (1 - q_1) + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j \min A_j < 1.$$

So by Corollary 5.1, there exists $\theta \in (0, 1]$ such that $\mu_{\theta,\alpha} < 1$.

In conclusion, condition (46) implies that $\mu_{\theta,\alpha} \leq 1$ for some $\theta \in I_\alpha$.

Let us now assume that $\sum_{j \in \mathcal{J}^*} \alpha_j \min A_j > 1$ (and, thus, $0 \in I_\alpha$ by Lemma 5.2). Using (43) and the fact that the h_j functions are nondecreasing (see Lemma 5.3 and the fact that h_j is constant and equal to k_j when A_j is reduced to the singleton $\{k_j\}$), it follows that, for $\theta \in I_\alpha \cap (0, +\infty)$,

$$\begin{aligned} \mu_{\theta,\alpha} &\geq g'_{A_0}(\theta) + \frac{\theta - g_{A_0}(\theta)}{\theta} \sum_{j \in \mathcal{J}^*} \alpha_j \min A_j \\ &> g'_{A_0}(\theta) + \frac{\theta - g_{A_0}(\theta)}{\theta} \\ &= 1 + \mathbb{E}[(X - 1)\theta^{X-1} 1_{\{X \in A_0\}}], \end{aligned}$$

where X has distribution p . Since $0 \notin A_0$, we deduce that $\mathbb{E}[(X - 1)\theta^{X-1} 1_{\{X \in A_0\}}]$ is nonnegative and, thus, $\mu_{\theta,\alpha} > 1$. Thanks to (44), we also have $\mu_{0,\alpha} > 1$, and thus, $\mu_{\theta,\alpha} > 1$ for all $\theta \in I_\alpha$ (note that, for $\theta = +\infty$, if it belongs to I_α , thanks to (45) we get $\mu_{\infty,\alpha} \geq \mu_{0,\alpha} > 1$). This completes the proof of Lemma 5.5(i).

We now prove part (ii) of the lemma. If $A_0 \subset \{1\}$, we get $\theta_{\min} = 0$ by Lemma 5.2 and $\theta_{\max} = +\infty$ as $g_{A_0}(\theta) = (1 - q_1)\theta < \theta$ and $\rho_{\mathcal{J}} = +\infty$; see (40). This gives us that $(0, +\infty) \subset I_\alpha$ and we have

$$\mu_{\theta,\alpha} = 1 - q_1 + q_1 H_\alpha(\theta) \quad \text{for } \theta \in I_\alpha. \tag{48}$$

So, if (47) holds, we have thanks to Lemma 5.3 that $\mu_{\theta,\alpha} < 1$ for $\theta \in (0, +\infty) \subset I_\alpha$. We also get $\mu_{0,\alpha} < 1$, respectively $\mu_{\infty,\alpha} < 1$, whenever $0 \in I_\alpha$, respectively $+\infty \in I_\alpha$. This gives us that $\mu_{\theta,\alpha} < 1$ for all $\theta \in I_\alpha$.

Let us assume that $A_0 \subset \{1\}$ and $\sup A_j = +\infty$ for some $j \in \mathcal{J}^*$ (that is, $\mathcal{J}^* \not\subset \mathcal{J}_\infty$). Since $\rho_{\mathcal{J}} = +\infty$, Lemma 5.3 gives us that $\lim_{\theta \rightarrow \infty} \mu_{\theta,\alpha} = +\infty$.

We now assume that $A_0 \subset \{1\}$, $\mathcal{J}^* \subset \mathcal{J}_\infty$, and $\sum_{j \in \mathcal{J}^*} \alpha_j \max A_j \geq 1$. Let $j_0 \in \mathcal{J}^*$ such that $0 \in A_{j_0}$. If $A_{j_0} = \{0\}$, it follows that $\theta_{\max} = +\infty$ belongs to I_α and (48) implies that $\mu_{\infty,\alpha} \geq 1$. If $\text{Card}(A_{j_0}) \geq 2$ then it follows that $\min_{j \in \mathcal{J}^*} \max A_j \geq 1$ and $\max_{j \in \mathcal{J}^*} \max A_j > 1$ (as $p_{\theta,\alpha}$ satisfies (5)), and thus, $\sum_{j \in \mathcal{J}^*} \alpha_j \max A_j > 1$. Using (48), it follows that $\lim_{\theta \rightarrow \infty} \mu_{\theta,\alpha} > 1$. In both cases, there exists $\theta \in I_\alpha$ such that $\mu_{\theta,\alpha} \geq 1$.

We now assume that $A_0 \not\subset \{1\}$. Then the function g_{A_0} is increasing and strictly convex. As $\rho_{\mathcal{J}} = \infty$, we deduce from (40) that $\theta_{\max} \in (1, +\infty)$ is the maximal root of $g_{A_0}(\theta) = \theta$, and thus, $g'_{A_0}(\theta_{\max}) > 1$. Then, it follows from (43) that $\lim_{\theta \uparrow \theta_{\max}} \mu_{\theta,\alpha} = g'_{A_0}(\theta_{\max}) > 1$.

In conclusion, if (47) does not hold then there exists $\theta \in I_\alpha$ such that $\mu_{\theta,\alpha} \geq 1$. □

Recall that $\mathcal{J}^{**} = \{j \in \mathcal{J}^* : \text{Card}(A_j) \geq 2\}$. We now consider the following condition:

$$A_0 \cap \{1\}^c \neq \emptyset \quad \text{or} \quad \mathcal{J}^{**} \neq \emptyset. \tag{49}$$

Proposition 5.2. (Monotonicity of $\theta \mapsto \mu_{\theta, \alpha}$.) *Let $\alpha \in \Delta_j^*$.*

- (i) *Assume that (49) does not hold. Then $I_\alpha = I_\alpha^f = [0, +\infty]$ and the map $\theta \mapsto p_{\theta, \alpha}$ (as well as the map $\theta \mapsto \mu_{\theta, \alpha}$) is constant.*
- (ii) *Assume that (49) holds. Then $\partial_\theta \mu_{\theta, \alpha} > 0$ on the interval $\{\theta \in I_\alpha^f : \mu_{\theta, \alpha} \leq 1\}$. If this set is not empty then its minimum is θ_{\min} . Furthermore, there exists at most one element $\theta \in I_\alpha$ such that $\mu_{\theta, \alpha} = 1$.*

Remark 5.1. ($\mu_{\theta, \alpha}$ independent of θ .) Recall that $q_1 = 1 - p(1)1_{\{1 \in A_0\}}$.

- (a) If (49) holds and the map $\theta \mapsto \mu_{\theta, \alpha}$ is constant, then $\mu_{\theta, \alpha} > 1$.
- (b) If (49) does not hold, that is, $A_0 \subset \{1\}$ and A_j is a singleton, say $\{k_j\}$ for all $j \in \mathcal{J}^*$ (with one of the k_j equal 0 and another larger than 1 in order for $p_{\theta, \alpha}$ to be nontrivial), then $I_\alpha = [0, +\infty]$ and $\mu_{\theta, \alpha} = (1 - q_1) + q_1 \sum_{j \in \mathcal{J}^*} \alpha_j k_j \in (0, +\infty)$. Thus, we recover Proposition 5.2(i).
- (c) Consider the example: $A_0 = \{1, k\}$ with $k \geq 2$, $\alpha \in \Delta_j^*$ and $A_j = \{k_j\}$ for $j \in \mathcal{J}^*$ (and, thus, $\mathcal{J}^{**} = \emptyset$) such that $k = \sum_{j \in \mathcal{J}^*} \alpha_j k_j$. Note that H_α is constant and equal to k and that the mean $\mu_{\theta, \alpha}$ is constant as

$$\mu_{\theta, \alpha} = p(1) + k\theta^{k-1}p(k) + (1 - p(1) - \theta^{k-1}p(k))H_\alpha(\theta) = 1 + (1 - p(1))(k - 1).$$

Thus, part (a) of this remark is not void.

Remark 5.2. (Is $\mu_{\theta, \alpha}$ monotone in θ ?) Consider the probability p defined by $p(2) = \frac{1}{2}$ and $p(0) = p(4) = \frac{1}{4}$; $J = 1$ (and, thus, $\alpha = 1$) and $A_1 = \{0, 4\}$ (and, thus, $A_0 = \{2\}$). It follows that $I_1 = (0, 2)$ as, by Lemma 5.2(ii), $\theta_{\min} = 0 \notin I_1$ and $\theta_{\max} = 2$ by (40). It is elementary to check that $\lim_{\theta \rightarrow 0^+} \mu_{\theta, 1} = 0$, $\mu_{\theta, 1} \simeq 1$ for $\theta \simeq 0.36$ and that $\partial_\theta \mu_{\theta, 1} < 0$ if and only if $\theta \in [\theta_0, \theta_1]$ with $\theta_0 \simeq 1.24$ and $\theta_1 \simeq 1.92$. This provides an example where the function $\theta \mapsto \mu_{\theta, \alpha}$ is not monotone.

Proof of Proposition 5.2. Let $\theta \in I_\alpha^f \cap \mathbb{R}_+^*$. Let X_θ be a random variable with distribution $p_{\theta, \alpha}$. We have

$$\mathbb{P}(X_\theta \in A_0) = \theta^{-1} g_{A_0}(\theta) \quad \text{and} \quad g'_{A_0}(\theta) = \mathbb{E}[X_\theta 1_{\{X_\theta \in A_0\}}].$$

We set

$$f(\theta) = \frac{\theta - g_{A_0}(\theta)}{\theta} = \mathbb{P}(X_\theta \notin A_0),$$

so that $\mu_{\theta, \alpha} = g'_{A_0}(\theta) + f(\theta)H_\alpha(\theta)$. We get

$$\partial_\theta \mu_{\theta, \alpha} = g''_{A_0}(\theta) + f(\theta)H'_\alpha(\theta) + f'(\theta)H_\alpha(\theta). \tag{50}$$

(This expression can possibly be equal to $+\infty$ if $\theta = \theta_{\max}$.) We have $H_\alpha(\theta) \in (0, +\infty]$ as $\theta > 0$. We have $H'_\alpha(\theta) = 0$ if $\mathcal{J}^{**} = \emptyset$ and, by Lemma 5.3, $H'_\alpha(\theta) \in (0, +\infty]$ otherwise. We

also have $g''_{A_0}(\theta) = 0$ if $A_0 \subset \{0, 1\}$ and $g''_{A_0}(\theta) \in (0, +\infty]$ otherwise. Finally, we have $f(\theta) > 0$ and

$$f'(\theta) = \frac{g_{A_0}(\theta) - \theta g'_{A_0}(\theta)}{\theta^2} = \theta^{-1} \mathbb{E}[(1 - X_\theta)1_{\{X_\theta \in A_0\}}], \tag{51}$$

which is finite as $\theta \in I_\alpha^f$. Note that $\mu_{\theta, \alpha} = \mathbb{E}[X_\theta]$, which is finite, and thus,

$$\theta f'(\theta) = \mathbb{E}[(X_\theta - 1)1_{\{X_\theta \notin A_0\}}] + (1 - \mu_{\theta, \alpha}). \tag{52}$$

Case $0 \in A_0$: we first assume that $0 \in A_0$, and thus, $I_\alpha \subset \mathbb{R}_+^*$ and $\theta_{\min} \notin I_\alpha$; see Lemma 5.2. We consider $\theta \in I_\alpha^f$ such that $\mu_{\theta, \alpha} \leq 1$. We deduce from (52) that $f'(\theta) = 0$ if \mathcal{J}^* is reduced to a singleton, say $\{j_0\}$, with $A_{j_0} = \{1\}$ and $\mu_{\theta, \alpha} = 1$, and that $f'(\theta) > 0$ otherwise (as $0 \in A_0$). Note that it is not possible to have $A_0 \subset \{0, 1\}$, $A_{j_0} = \{1\}$ and $\mathcal{J}^* = \{j_0\}$ together as $p_{\theta, \alpha}$ is nontrivial, so at least $g''_{A_0}(\theta) \in (0, +\infty]$ or $\mathbb{E}[(X_\theta - 1)1_{\{X_\theta \notin A_0\}}] > 0$. The latter implies that $f'(\theta) > 0$ as $\mu_{\theta, \alpha} \leq 1$. Since $f(\theta) > 0$ and $H_\alpha(\theta) > 0$, we deduce from (50) that $\partial_\theta \mu_{\theta, \alpha} > 0$ on $\{\theta \in I_\alpha^f : \mu_{\theta, \alpha} \leq 1\}$.

Case $0 \notin A_0$: we now assume that $0 \notin A_0$. This implies that $\theta_{\min} = 0$. The function f has a continuous extension at 0 given by $f(0) = q_1 > 0$; see (36). We distinguish according to A_0 being empty or reduced to $\{1\}$ and $A_0 \cap \{1\}^c \neq \emptyset$.

Subcase $A_0 \subset \{1\}$: we consider the subcase $A_0 \subset \{1\}$. The function f is constant and equal to q_1 , and $\mu_{\theta, \alpha} = 1 - q_1 + q_1 H_\alpha(\theta)$. If $\mathcal{J}^{**} = \emptyset$ and, thus, (49) does not hold, then we have $p_{\theta, \alpha}(k_j) = q_1 \alpha_j$ for $A_j = \{k_j\}$ and $j \in \mathcal{J}^*$. Thus, Proposition 5.2(i) is obvious.

If $\mathcal{J}^{**} \neq \emptyset$ then we deduce from Lemma 5.3 that the functions H_α and $\theta \mapsto \mu_{\theta, \alpha}$ are increasing on I_α^f .

Subcase $0 \notin A_0$ and $A_0 \cap \{1\}^c \neq \emptyset$: we get, in particular, that $\theta_{\max} < +\infty$. We introduce an auxiliary parameterized function defined on I_α^f for $\gamma > 0$ by

$$m_\gamma(\theta) = g'_{A_0}(\theta) + \gamma f(\theta). \tag{53}$$

Note that $m_\gamma(0) = (1 - q_1) + q_1 \gamma > 0$. On the one hand, direct computation yields

$$m_\gamma(\theta) = m_\gamma(0) + \sum_{k \in A_0 \cap \{1\}^c} (k - \gamma) \theta^{k-1} p(k). \tag{54}$$

We deduce that if $\gamma < \min(A_0 \cap \{1\}^c)$ and, in particular, if $\gamma < 2$, then $m'_\gamma > 0$ on I_α^f . (Note that m'_γ is finite on I_α^f except possibly at θ_{\max} in the case where it belongs to I_α^f .) On the other hand, if there exists $\theta_* \in I_\alpha^f$ such that $m_\gamma(\theta_*) \leq 1$, we deduce that $\gamma \leq 1$ if $\theta_* = 0$ and, from (53), that

$$\gamma \leq \frac{1 - g'_{A_0}(\theta_*)}{f(\theta_*)} = \frac{\theta_* - \theta_* g'_{A_0}(\theta_*)}{\theta_* - g_{A_0}(\theta_*)} < 1 \quad \text{if } \theta_* > 0.$$

In conclusion,

$$\text{there exists } \theta_* \in I_\alpha^f \text{ such that } m_\gamma(\theta_*) \leq 1 \implies m'_\gamma > 0. \tag{55}$$

Now, we go back to the function $\theta \mapsto \mu_{\theta, \alpha}$. Assume that there exists $\theta_* \in I_\alpha^f$ such that $\mu_{\theta_*, \alpha} \leq 1$. We set $\mu_* = \mu_{\theta_*, \alpha}$, $\gamma_* = H_\alpha(\theta_*)$, and $m_* = m_{\gamma_*}$. By construction, we have $m_*(\theta_*) = \mu_* \leq 1$ and, as H_α is nondecreasing, for $\theta \in I_\alpha^f$,

$$(\theta - \theta_*)(\mu_{\theta, \alpha} - m_*(\theta)) = (\theta - \theta_*)(H_\alpha(\theta) - H_\theta(\theta_*))f(\theta) \geq 0.$$

This implies that $\partial_{\theta=\theta_*} \mu_{\theta,\alpha} \geq m'_*(\theta_*)$, and thus, is positive thanks to (55). We find that $\partial_{\theta} \mu_{\theta,\alpha} > 0$ on $\{\theta \in I_{\alpha}^f : \mu_{\theta,\alpha} \leq 1\}$.

In conclusion, if $A_0 \subset \{1\}$ and $\mathcal{J}^{**} = \emptyset$ then $\theta \mapsto \mu_{\theta,\alpha}$ is constant; if $A_0 \cap \{1\}^c \neq \emptyset$ or $\mathcal{J}^{**} \neq \emptyset$, we have $\partial_{\theta} \mu_{\theta,\alpha} > 0$ on $\{\theta \in I_{\alpha}^f : \mu_{\theta,\alpha} \leq 1\}$, this set is either empty or equal to I_{α}^f or of the form $I_{\alpha}^f \cap [0, \theta']$, and it contains at most one element θ such that $\mu_{\theta,\alpha} = 1$. \square

Recall that $\rho_{\mathcal{J}}$ is the radius of convergence of the function $\sum_{j \in \mathcal{J}} g_{A_j}$ and I_{α} is an interval of $[0, +\infty]$ defined in (37). We have the following theorem on the characterization of the probability distribution p to be generic in the sense of Definition 3.1(ii). Recall that assuming Condition (49) is not very restrictive as otherwise the map $\theta \mapsto p_{\theta,\alpha}$ is constant; see Proposition 5.2.

Theorem 5.1 (Generic distribution.) *Let p be a probability distribution on \mathbb{N} satisfying (5). Let $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, be pairwise disjoint nonempty subsets of $\text{supp}(p)$. Let $\alpha \in \Delta_{\mathcal{J}}^*$ and assume that (49) holds. The distribution p is not generic for \mathcal{A} in the direction α if and only if one of the following conditions holds:*

(i) $\rho_{\mathcal{J}} < \infty$, $g'_{A_0}(\rho_{\mathcal{J}}) \leq 1$, $g_{\mathcal{J}}(\rho_{\mathcal{J}}) < \infty$, and

$$H_{\alpha}(\rho_{\mathcal{J}}) < \rho_{\mathcal{J}} \frac{1 - g'_{A_0}(\rho_{\mathcal{J}})}{\rho_{\mathcal{J}} - g_{A_0}(\rho_{\mathcal{J}})}; \tag{56}$$

- (ii) $0 \notin A_0$ and $\sum_{j \in \mathcal{J}^*} \alpha_j \min A_j > 1$;
 (iii) $A_0 \subset \{1\}$, $\rho_{\mathcal{J}} = \infty$ and $\sum_{j \in \mathcal{J}^*} \alpha_j \sup A_j < 1$.

Remark 5.3. (Direction and generic distribution.) A probability distribution might not be generic in all the directions; and it may happen that it is generic only in one direction.

(a) Suppose that $\text{supp}(p) = \llbracket 0, 3 \rrbracket$ and $\mathcal{A} = (\{0\}, \{2, 3\})$, and thus, $A_0 = \{1\}$. Consider the direction $\alpha = (\alpha_1, \alpha_2)$.

- The distribution p is generic for \mathcal{A} in the direction α if and only if $\alpha_2 \in [\frac{1}{3}, \frac{1}{2}]$.
- If $\alpha_2 \in (\frac{1}{2}, 1)$ then Theorem 5.1(ii) holds since $0 \notin A_0$ and $\sum_{j=1}^2 \alpha_j \min(A_j) = 2\alpha_2 > 1$; and in this case all the $p_{\theta,\alpha}$ are subcritical.
- If $\alpha_2 \in (0, \frac{1}{3})$ then Theorem 5.1(iii) holds since $A_0 \subset \{1\}$ and $\sum_{j=1}^2 \alpha_j \sup(A_j) = 3\alpha_2 < 1$; and in this case all the $p_{\theta,\alpha}$ are supercritical.

(b) Suppose that $\text{supp}(p) = \{0, 2\}$ and $\mathcal{A} = (\{0\}, \{2\})$. Note that (49) does not hold. In this example any probability distribution p' such that $\text{supp}(p') = \text{supp}(p)$ is (p, \mathcal{A}) -compatible. The direction of p' is given by $(p'(0), p'(2))$. We recover that, for all directions in Δ_2^* , there exists a (p, \mathcal{A}) -compatible probability distribution. However, the probability distribution p is generic for \mathcal{A} only in the direction $(\frac{1}{2}, \frac{1}{2})$.

Proof. We simply write M_{α} and m_{α} for the supremum and infimum of $\{\mu_{\theta,\alpha} : \theta \in I_{\alpha}\}$, respectively. Let \neg be the usual logical negation.

We first prove that

$$\rho_{\mathcal{J}} < +\infty \quad \text{and} \quad \neg(i) \implies M_{\alpha} \geq 1. \tag{57}$$

If $\rho_{\mathcal{J}} < \infty$ and $g'_{A_0}(\rho_{\mathcal{J}}) > 1$, then we have $A_0 \cap \{0, 1\}^c \neq \emptyset$, the function g_{A_0} is strictly convex. So there exists $\theta^* \in (\theta_{\min}, \rho_{\mathcal{J}})$ such that $g_{A_0}(\theta^*) < \theta^*$ and $g'_{A_0}(\theta^*) = 1$. Note that $\theta^* \in I_{\alpha}$ and use (43) to deduce that $\mu_{\theta^*,\alpha} \geq 1$.

If $\rho_{\mathcal{J}} < \infty$, $g'_{A_0}(\rho_{\mathcal{J}}) \leq 1$, and $g_{\mathcal{J}}(\rho_{\mathcal{J}}) = \infty$, then we have $\rho_{\mathcal{J}} > 1$ and $\rho_{\mathcal{J}} \notin I_{\alpha}$ (by Lemma 5.1(i)). Since $g'_{A_0}(\rho_{\mathcal{J}}) \leq 1$, we deduce that $g_{A_0}(\rho_{\mathcal{J}}) < \rho_{\mathcal{J}}$, and thus, $\theta_{\max} = \rho_{\mathcal{J}}$ by (40). Since $g_{\mathcal{J}}(\rho_{\mathcal{J}}) = \infty$, we deduce there exists $j \in \mathcal{J}^*$ such that $\rho_{A_j} = \rho_{\mathcal{J}}$, $g_{A_j}(\rho_{\mathcal{J}}) = \infty$ as well as $\sup A_j = +\infty$. This implies by Lemma 5.3 that $H_{\alpha}(\rho_{\mathcal{J}}) = +\infty$ and, hence, by continuity that $\lim_{\theta \uparrow \theta_{\max}} \mu_{\theta, \alpha} = +\infty$.

If $\rho_{\mathcal{J}} < \infty$, $g'_{A_0}(\rho_{\mathcal{J}}) \leq 1$, $g_{\mathcal{J}}(\rho_{\mathcal{J}}) < \infty$, and

$$H_{\alpha}(\rho_{\mathcal{J}}) \geq \rho_{\mathcal{J}} \frac{1 - g'_{A_0}(\rho_{\mathcal{J}})}{\rho_{\mathcal{J}} - g_{A_0}(\rho_{\mathcal{J}})}, \quad (58)$$

then we have $g_{A_0}(\rho_{\mathcal{J}}) < \rho_{\mathcal{J}}$, and thus, $\theta_{\max} = \rho_{\mathcal{J}}$ belongs to I_{α} . Using (43) and (58), we deduce that $\mu_{\theta_{\max}, \alpha} \geq 1$.

This proves that (57) holds.

We now consider the $0 \in A_0$ case. Thanks to Lemma 5.4(i), we have $m_{\alpha} < 1$. So p is generic in the direction α if and only if $M_{\alpha} \geq 1$. If Theorem 5.1(i) holds then $\theta_{\max} = \rho_{\mathcal{J}}$ belongs to I_{α} , and by (43) and (56) it follows that $\mu_{\theta_{\max}, \alpha} < 1$, which by Proposition 5.2 implies that $M_{\alpha} < 1$; thus, p is not generic in the direction α . Now assume that Theorem 5.1(i) does not hold. If $\rho_{\mathcal{J}} < +\infty$ then we use (57) to deduce that $M_{\alpha} \geq 1$ and that p is generic in the direction α . If $\rho_{\mathcal{J}} = +\infty$ then we use Lemma 5.4(ii) to show that $M_{\alpha} \geq 1$, and thus, p is also generic in the direction α . This proves the theorem in the $0 \in A_0$ case.

We now consider the $0 \notin A_0$ case. If Theorem 5.1(ii) holds, we deduce from Lemma 5.5(i) that $m_{\alpha} > 1$, and thus, p is not generic in the direction α .

We now assume that $\sum_{j \in \mathcal{J}^*} \alpha_j \min A_j \leq 1$. Thanks to Lemma 5.5(i), it follows that $m_{\alpha} \leq 1$. So p is generic in the direction α if and only if $M_{\alpha} \geq 1$. If $\rho_{\mathcal{J}} = \infty$, we deduce from Lemma 5.5(ii) that $A_0 \subset \{1\}$ and $\sum_{j \in \mathcal{J}^*} \alpha_j \sup A_j < 1$ are equivalent to $M_{\alpha} < 1$, that is, p is not generic in the direction α . We eventually assume that $\rho_{\mathcal{J}} < \infty$. If Theorem 5.1(i) holds then $\theta_{\max} = \rho_{\mathcal{J}}$ belongs to I_{α} , $\mu_{\theta_{\max}, \alpha} < 1$, and $M_{\alpha} < 1$ by Proposition 5.2, and thus, p is not generic in the direction α . If Theorem 5.1(i) does not hold then we use (57) to deduce that $M_{\alpha} \geq 1$ and that p is generic in the direction α . \square

6. Local limit of large Galton–Watson trees

We give the main theorem, see Theorem 6.1, on the local limit of conditioned BGW trees in the next section. Its proof relies on a transformation of BGW trees from Rizzolo (see Section 6.2) and a direct application of local limit theorems for multi-type BGW trees from [7] (see Section 6.3).

6.1. Main result

Let p be a probability distribution on \mathbb{N} satisfying (5), and let $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, be pairwise disjoint nonempty subsets of $\text{supp}(p)$. Let $\alpha \in \Delta_J^*$ be a possible direction; see (35). We assume that the distribution p is generic for \mathcal{A} in the direction α (recall Theorem 5.1). Thus, there exists a (unique) $\theta_{\alpha} \in I_{\alpha}$ such that

$$p_{\alpha} := p_{\theta_{\alpha}, \alpha} \quad \text{is critical.} \quad (59)$$

Recall that \mathcal{T}_q denotes a BGW tree with offspring distribution q and \mathcal{T}_q^* the corresponding Kesten's tree when $\mu(q) \leq 1$. Recall that $\|\mathbf{n}\|$ is the L^1 -norm of $\mathbf{n} \in \mathbb{N}^J$.

Lemma 6.1. *If p is generic for \mathcal{A} in the direction $\alpha \in \Delta_J^*$ then there exists a sequence $(\mathbf{n}^{(m)})_{m \in \mathbb{N}}$ in \mathbb{N}^J such that*

$$\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}^{(m)}) > 0, \tag{60}$$

$$\lim_{m \rightarrow \infty} |\mathbf{n}^{(m)}| = \infty, \quad \lim_{m \rightarrow \infty} \frac{\mathbf{n}^{(m)}}{|\mathbf{n}^{(m)}|} = \alpha, \tag{61}$$

and for all $m \in \mathbb{N}, j \in \llbracket 1, J \rrbracket$, with $\mathbf{n}^{(m)} = (n_1^{(m)}, \dots, n_j^{(m)})$,

$$\alpha_j = 0 \implies n_j^{(m)} = 0. \tag{62}$$

Proof. Since p is generic for \mathcal{A} in the direction $\alpha \in \Delta_J^*$, there exists a critical (p, \mathcal{A}) -compatible distribution p_α with direction α . Thus, by Definition 3.1 of the (p, \mathcal{A}) -compatible probability distribution, it is enough to find a sequence $(\mathbf{n}^{(m)})_{m \in \mathbb{N}}$ such that (60)–(62) hold, with \mathcal{T}_p replaced by \mathcal{T}_{p_α} .

Let $\mathcal{T}_{p_\alpha}^{(n)}$ be distributed as \mathcal{T}_{p_α} conditioned to have n vertices. Note that, for all finite $M > 0$, there exists $n > M$ such that the probability of \mathcal{T}_{p_α} having n vertices is positive. Recall that $L_k(\mathbf{t})$ denotes the number of vertices in \mathbf{t} with out-degree k . According to [16, Theorem 7.11], along the sequence $\{n \in \mathbb{N} : \mathbb{P}(\#\mathcal{T}_{p_\alpha} = n) > 0\}$, the following convergences hold in probability for all $k \in \mathbb{N}$:

$$n^{-1}L_k(\mathcal{T}_{p_\alpha}^{(n)}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} p_\alpha(k).$$

Since $(n^{-1}L_k(\mathcal{T}_{p_\alpha}^{(n)}))_{k \in \mathbb{N}}$ is a random probability distribution on \mathbb{N} , we also get

$$\max_{j \in \mathcal{J}} |n^{-1}L_{A_j}(\mathcal{T}_{p_\alpha}^{(n)}) - p_\alpha(A_j)| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0. \tag{63}$$

In particular, for $m \in \mathbb{N}^*$ large enough, we deduce that there exists a tree $\mathbf{t}^{[m]}$ such that $\#\mathbf{t}^{[m]} \geq m$; $\mathbb{P}(\mathcal{T}_{p_\alpha} = \mathbf{t}^{[m]}) > 0$; and, with $\mathbf{n}^{(m)} = (n_1^{(m)}, \dots, n_j^{(m)}) = L_{\mathcal{A}}(\mathbf{t}^{[m]})$, $n_j^{(m)} = 0$ if $\alpha_j = 0$ and, for all $j \in \mathcal{J}^*$,

$$\left| \frac{n_j^{(m)}}{|\mathbf{n}^{(m)}|} - \frac{p_\alpha(A_j)}{1 - p_\alpha(A_0)} \right| \leq \frac{1}{m}.$$

Recall also that $\mathbb{P}(\mathcal{T}_{p_\alpha} = \mathbf{t}) > 0$ implies that $\mathbb{P}(\mathcal{T}_p = \mathbf{t}) > 0$. This and the fact that α is the direction of p_α , that is, $p_\alpha(A_j)/(1 - p_\alpha(A_0)) = \alpha_j$, completes the proof. \square

Recall that

$$\mathcal{J}^* = \{j \in \llbracket 1, J \rrbracket : \alpha_j > 0\} \quad \text{and} \quad \mathcal{J} = \{0\} \cup \mathcal{J}^*.$$

For $A \subset \mathbb{N}$, we write $A - 1 = \{a - 1 : a \in A\}$ and $A - A = \{a - b : a, b \in A\}$. We define

$$\Gamma_\alpha = \bigcup_{j \in \mathcal{J}} A'_j, \quad \text{where } A'_0 = A_0 - 1 \text{ and } A'_j = A_j - A_j \text{ for } j \in \mathcal{J}^*. \tag{64}$$

Definition 6.1 (*Aperiodicity.*) Let p be generic for \mathcal{A} in the direction $\alpha \in \Delta_J^*$. The probability distribution p is aperiodic for \mathcal{A} in the direction α if $\theta_\alpha \in (0, +\infty)$ and the smallest subgroup of \mathbb{Z} that contains Γ_α is \mathbb{Z} .

Remark 6.1 (On aperiodicity.)

- (a) If $0 \in A_0$ then $-1 \in A'_0$, and thus, the distribution p is aperiodic for \mathcal{A} in the direction α .
- (b) If p is aperiodic for \mathcal{A} in the direction α then (49) holds.
- (c) Looking carefully at the proof of Lemma 6.3(iii), it would be more natural to consider Γ'_α defined as Γ_α in (64) but with A_j replaced by $A_j \cap \text{supp}(p_\alpha)$. For $\theta_\alpha \in (0, +\infty)$, this yields no modification as $\Gamma'_\alpha = \Gamma_\alpha$. However, for $\theta_\alpha \in \{0, \infty\}$ (which is ruled out in Definition 6.1), the set Γ'_α is reduced to $\{0\}$ (as $A_j \cap \text{supp}(p_\alpha)$ is either a singleton or empty). So, with the more natural definition that p is aperiodic for \mathcal{A} in the direction α if the smallest subgroup of \mathbb{Z} that contains Γ'_α is \mathbb{Z} , then we still have $\theta_\alpha \in (0, \infty)$.

We now state the main result. Note that we do not assume that p has a finite mean.

Theorem 6.1. (Local limit of conditioned BGW tree.) *Let p be a probability distribution on \mathbb{N} satisfying (5). Let $\mathcal{A} = (A_j)_{j \in \llbracket 1, J \rrbracket}$, with $J \in \mathbb{N}^*$, be a family of pairwise disjoint nonempty subsets of $\text{supp}(p)$ and consider the direction $\alpha \in \Delta_J^*$. We assume that p is generic and aperiodic for \mathcal{A} in the direction α (and, thus, $\theta_\alpha \in (0, +\infty)$). If $(\mathbf{n}^{(m)})_{m \in \mathbb{N}}$ is a sequence in \mathbb{N}^J satisfying (60)–(62) then we have*

$$\text{dist}(\mathcal{T}_p \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}^{(m)}) \xrightarrow{m \rightarrow \infty} \text{dist}(\mathcal{T}_{p_\alpha}^*).$$

We refer to Remark 1.1 and the details of its proof given in Section 6.5 for the usefulness of condition (62), that is, $\alpha_j = 0 \implies n_j^{(m)} = 0$ for the sequence $(\mathbf{n}^{(m)})_{m \in \mathbb{N}}$.

Remark 6.2. (Conditioning on the total population size and the number of leaves.) Note that conditioning on the total population size and the number of leaves, or equivalently, on the number of internal nodes and the number of leaves, corresponds to $A_1 = \mathbb{N}^* \cap \text{supp}(p)$ and $A_2 = \{0\}$. Note that Remark 6.1(b) that the case A_1 reduced to a singleton is excluded from Theorem 6.1, but then it is equivalent to conditioning on the number of leaves, as considered in [4].) Provided the assumptions on genericity and aperiodicity are satisfied, this case is included in Theorem 6.1 whereas it is excluded *a priori* in [7, Corollary 3.5].

Remark 6.3. (On the case $\mathcal{J}^{**} = \emptyset$ and $A_0 = \emptyset$ or $A_0 = \{1\}$.) Let $j_0 \in \mathcal{J}^*$ be such that $A_{j_0} = \{0\}$. Note that $\alpha_{j_0} \in (0, 1)$. Since $\mathcal{J}^{**} = \emptyset$ and $A_0 \subset \{1\}$, the aperiodic hypothesis is not satisfied. However, condition (7) implies no choice on L_0 , so that we can, without loss of generality, replace A_0 by $A'_0 = A_0 \cup \{0\}$ and remove j_0 from \mathcal{J}^* , as well as α by $\alpha' \in \mathbb{R}_+^{J-1}$ with $\alpha'_j = \alpha_j / (1 - \alpha_{j_0})$ for $j \in \llbracket 1, J \rrbracket \setminus \{j_0\}$. Then the conditioning is the same and the distribution p is aperiodic for $\mathcal{A}' = (A_j)_{j \in \llbracket 1, J \rrbracket \setminus \{j_0\}}$ in the direction α' .

Using this trick, we see that the local convergence of Theorem 6.1 holds in this case, even though p is not aperiodic.

The next two sections are devoted to the proof of the theorem. In Section 6.2, for a tree \mathbf{t} such that $\mathcal{L}_{\mathcal{A}}(\mathbf{t}) \neq \emptyset$, we describe a map from $\mathcal{L}_{\mathcal{A}}(\mathbf{t})$ onto a multi-type tree, which is a direct extension of [27]. Then, in Section 6.3 we use [7] on the local limit of multi-type BGW trees to conclude.

From now on the direction $\alpha \in \Delta_J^*$ is fixed and p_α , given by (59), is the unique critical probability distribution (p, \mathcal{A}) -compatible with the direction α . In particular, we have $\mu(p_\alpha) = 1$. By construction, we have $p_\alpha(A_j) = 0$ if $\alpha_j = 0$ for $j \in \llbracket 1, J \rrbracket$. Since only the indices j such that

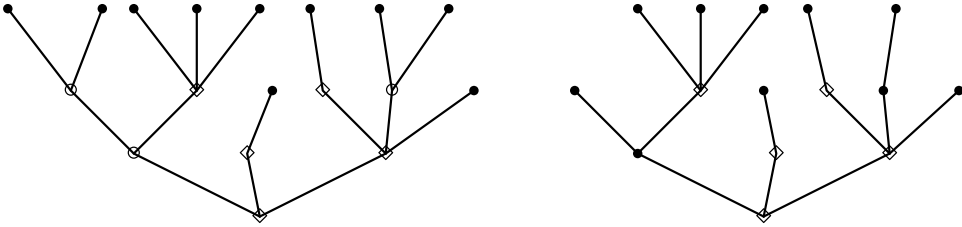


FIGURE 1. A tree \mathbf{t} on the left with $A_0 = \{2\}$, $A_1 = \{1, 3\}$, $A_2 = \{0\}$ and the tree \mathbf{t}^A after the map on the right. We represent type 0 with \circ , type 1 with \diamond , and type 2 with \bullet .

α_j is positive are pertinent, for a sequence $\mathbf{x} = (x_j)_{j \in \llbracket 1, J \rrbracket}$, we consider the subsequence

$$\mathbf{x}^* = (x_j)_{j \in \mathcal{J}^*}, \tag{65}$$

where we recall that $\mathcal{J}^* = \{j \in \llbracket 1, J \rrbracket : \alpha_j > 0\}$. For example, we write $L_{A^*}(\mathbf{t}) = (L_{A_j}(\mathbf{t}))_{j \in \mathcal{J}^*}$.

6.2. Extension of Rizzolo’s transformation

In the following, we use the framework for multi-type trees from [7, Section 2]. For a tree \mathbf{t} such that $\mathcal{L}_A(\mathbf{t}) \neq \emptyset$, we describe a map from $\mathcal{L}_A(\mathbf{t})$ onto a multi-type tree, which is a direct extension of [27].

The vertex $u \in \mathbf{t}$ is said to have type $j \in \mathcal{J}$, which we denote by $e_{\mathbf{t}}(u) = j$, if $k_u(\mathbf{t}) \in A_j$ so that $(\mathbf{t}, e_{\mathbf{t}})$ can be treated as a \mathcal{J} -type tree. Note that $\mathbf{t}^{[i]} = \{u \in \mathbf{t} : e_{\mathbf{t}}(u) = i\}$. In order to remove the 0-type vertices, following [27], we build a bijection ϕ (depending on $(\mathbf{t}, e_{\mathbf{t}})$) from $\mathbf{t} \setminus \mathbf{t}^{[0]}$ to a tree \mathbf{t}^A with a \mathcal{J}^* -type $e_{\mathbf{t}^A}$, which preserves the types, that is, $e_{\mathbf{t}^A}(\phi(u)) = e_{\mathbf{t}}(u) \in \mathcal{J}^*$. Furthermore, if $\mathbf{t}^{[0]}$ is empty (which is automatically the case if $A_0 = \emptyset$) then we have $\mathbf{t}^A = \mathbf{t}$ and ϕ is the identity map (and, thus, $e_{\mathbf{t}} = e_{\mathbf{t}^A}$).

Let $(\mathbf{t}, e_{\mathbf{t}})$ be a \mathcal{J} -type tree such that $\sharp(\mathbf{t} \setminus \mathbf{t}^{[0]}) = n \geq 1$. Following [4], we define recursively a sequence of growing \mathcal{J}^* -type trees $(\mathbf{t}_k, e_{\mathbf{t}_k})_{k \in \llbracket 1, n \rrbracket}$ and identify the last one as $(\mathbf{t}^A, e_{\mathbf{t}^A})$. The map ϕ is a by-product of this construction. Denote $<$ the lexicographic order on \mathcal{U} . Let $u_1 < \dots < u_n$ be the ordered list of vertices of $\mathbf{t} \setminus \mathbf{t}^{[0]}$. Then, we define recursively:

- $\phi(u_1) = \emptyset$, $\mathbf{t}_1 = \{\emptyset\}$, and $e_{\mathbf{t}_1}(\emptyset) = e_{\mathbf{t}}(u_1)$;
- for $1 < k \leq n$, let $M(u_{k-1}, u_k) \in \{u_1, \dots, u_{k-1}\}$ be the most recent common ancestor of u_{k-1} and u_k and \mathbf{s} the fringe subtree of \mathbf{t} above $M(u_{k-1}, u_k)$; see (8).

Note that $v = \min\{u \in \mathbf{s} : e_{\mathbf{t}}(u) \neq 0\}$ (for the lexicographic order). Then we set $\phi(u_k)$ as the concatenation of $\phi(v)$ and $(k_{\phi(v)}(\mathbf{t}_{k-1}) + 1)$ and consider the tree

$$\mathbf{t}_k = \mathbf{t}_{k-1} \cup \{\phi(u_k)\},$$

and the type map $e_{\mathbf{t}_k}$ coincide with $e_{\mathbf{t}_{k-1}}$ on \mathbf{t}_{k-1} and $e_{\mathbf{t}_k}(\phi(u_k)) = e_{\mathbf{t}}(u_k)$. (This ensures that ϕ preserves indeed the types.)

It is obvious that $(\mathbf{t}_k, e_{\mathbf{t}_k})_{k \in \llbracket 1, n \rrbracket}$ is a sequence of (increasing) multi-type trees. Let $(\mathbf{t}^A, e_{\mathbf{t}^A}) = (\mathbf{t}_n, e_{\mathbf{t}_n})$ and we view ϕ as a bijection from $\mathbf{t} \setminus \mathbf{t}^{[0]}$ to \mathbf{t}^A , which preserves the types. See Figure 1 for an example of \mathbf{t} and \mathbf{t}^A and their types.

Note that $L_{A_j}(\mathbf{t}) = \text{Card}\{u \in \mathbf{t}^A : e_{\mathbf{t}^A}(u) = j\}$ is the total population size of type j (which is equal to 0 if $j \notin \mathcal{J}^*$). For the \mathcal{J}^* -type tree $(\mathbf{t}^A, e_{\mathbf{t}^A})$, we denote by $\sharp \mathbf{t}^A$ the vector of the total

population size of each type in \mathcal{J}^* of \mathbf{t}^A :

$$\#\mathbf{t}^A = L_{\mathcal{A}^*}(\mathbf{t}) \in \mathbb{N}^{\mathcal{J}^*}. \tag{66}$$

Let \mathcal{T}_α be a BGW tree with critical offspring distribution p_α . Let $\mathcal{T}_{\alpha,*}$ be distributed as \mathcal{T}_α conditioned to have at least one vertex with out-degree in A_0^c , that is, on $\{L_{\mathcal{A}}(\mathcal{T}_\alpha) \neq \mathbf{0}\}$, and, with a slight abuse of notation, we set $(\mathcal{T}_\alpha^A, e_{\mathcal{T}_\alpha^A})$ as the \mathcal{J}^* -type tree associated with $\mathcal{T}_{\alpha,*}$ by the previous construction. The proof of the next result, which is left to the reader, is an adaptation of the proof of [27, Theorem 6].

Lemma 6.2. *The random tree $(\mathcal{T}_\alpha^A, e_{\mathcal{T}_\alpha^A})$ is a multi-type BGW tree with types in \mathcal{J}^* .*

The root of \mathcal{T}_α^A is of type $e_{\mathcal{T}_\alpha^A}(\emptyset) = j$ with probability α_j for $j \in \mathcal{J}^*$. Let $p^{A,\alpha} = (p_j^{A,\alpha})_{j \in \mathcal{J}^*}$ be the offspring distribution of the \mathcal{J}^* -type BGW tree $(\mathcal{T}_\alpha^A, e_{\mathcal{T}_\alpha^A})$, where $p_j^{A,\alpha}$, a probability distribution on $\mathbb{N}^{\mathcal{J}^*}$, is the offspring distribution of an individual of type j . To describe $p_j^{A,\alpha}$, we introduce several intermediate random variables.

- (A1) Let X be a random variable on \mathbb{N} distributed according to p_α .
- (A2) Let X^j be distributed as X conditionally on $\{X \in A_j\}$.
- (A3) Let $(X_i^0)_{i \in \mathbb{N}}$ be independent random variables distributed as $X - 1$ conditionally on $\{X \in A_0\}$.
- (A4) Let N be a geometric random variable with parameter $p_\alpha(A_0^c)$.

We assume that the random variables X^j , $(X_i^0)_{i \in \mathbb{N}}$, and N are independent. We adopt the convention $\inf \emptyset = +\infty$. [resume*]

- (A5) Set $T = \inf\{n \in \mathbb{N}^* : \sum_{i=1}^n X_i^0 = -1\}$.
- (A6) Set $Y_j = X^j + \sum_{i=1}^{N-1} X_i^0$ on the event $\{N \leq T\}$ and $Y_j = 0$ otherwise.
- (A7) Conditionally on the above random variables, let Z_j be a binomial random variable with parameters (Y_j, r) , where

$$r = \mathbb{P}(N \leq T) \in (0, 1]. \tag{67}$$

- (A8) Conditionally on the above random variables, let $X_j^A = (X_j^A(i))_{i \in \mathcal{J}^*}$ be a multinomial random variable with parameter (Z_j, α^*) .

Then, the probability distribution $p_j^{A,\alpha}$ is defined as the law of X_j^A conditionally on $\{N \leq T\}$.

Recall that $p^{A,\alpha}$ is said to be aperiodic, if the smallest subgroup of $\mathbb{Z}^{\mathcal{J}^*}$ that contains $\bigcup_{j \in \mathcal{J}^*} (\text{supp}(p_j^{A,\alpha}) - \text{supp}(p_j^{A,\alpha}))$ is $\mathbb{Z}^{\mathcal{J}^*}$. The mean matrix $M = (m_{j\ell})_{j,\ell \in \mathcal{J}^*}$ of $p^{A,\alpha}$ is defined by

$$m_{j\ell} = \mathbb{E}[X_j^A(\ell) \mid N \leq T]. \tag{68}$$

The offspring distribution $p^{A,\alpha}$ is critical if the spectral radius of the mean matrix M is one. We have the following properties. Recall that θ_α is the unique $\theta \in [0, +\infty]$ such that $p_{\theta,\alpha}$ is critical.

Lemma 6.3. (Properties of the offspring distribution $p^{A,\alpha}$.) *Let p be a probability distribution on \mathbb{N} satisfying (5), such that it is generic for \mathcal{A} in the direction $\alpha \in \Delta_j^*$.*

- (i) *The offspring distribution $p^{A,\alpha}$ is critical and α^* is the left eigenvector of the mean matrix associated with the eigenvalue 1.*
- (ii) *We have $m_{j\ell} > 0$ for all $j, \ell \in \mathcal{J}^*$ and $j \neq j_0$, where $j_0 \in \mathcal{J}$ is defined by $0 \in A_{j_0}$. If $j_0 \in \mathcal{J}^*$ then we have $m_{j_0\ell} = 0$ for all $\ell \in \mathcal{J}^*$ either if $A_{j_0} = \{0\}$ and $A_0 \subset \{1\}$ or if $\theta_\alpha = 0$, and that $m_{j_0\ell} > 0$ for all $\ell \in \mathcal{J}^*$ otherwise.*
- (iii) *The offspring distribution $p^{A,\alpha}$ is aperiodic if and only if p is aperiodic for \mathcal{A} in the direction α (and, thus, $\theta_\alpha \notin \{0, +\infty\}$).*

Proof. We prove part (i) of the lemma on the criticality of $p^{A,\alpha}$. By construction and (68), the entries of the mean matrix M are given by, for $j, \ell \in \mathcal{J}^*$,

$$m_{j\ell} = \mathbb{E}[X_j^A(\ell) \mid N \leq T] = \mathbb{E}[Z_j \mid N \leq T] \alpha_\ell = \mathbb{E}[Y_j] \alpha_\ell.$$

In particular, the mean matrix has rank one and α^* is the left eigenvector associated with the nonzero eigenvalue, say ρ . Since the mean matrix has nonnegative entries and α^* has positive entries, we also find that ρ is the Perron-Frobenius eigenvalue and, thus, the spectral radius of M . Since M has rank one, we also find that ρ is the trace of M :

$$\rho = \sum_{j \in \mathcal{J}^*} \mathbb{E}[Y_j] \alpha_j.$$

We now compute $r\mathbb{E}[Y_j]$ using (67):

$$\mathbb{E}[Y_j] = \mathbb{E} \left[\left(X^j + \sum_{i=1}^{N-1} X_i^0 \right) 1_{\{N \leq T\}} \right] = r\mathbb{E}[X^j] + \mathbb{E} \left[\sum_{i=1}^{N-1} X_i^0 \right] - \mathbb{E} \left[\sum_{i=1}^{N-1} X_i^0 1_{\{N \geq T\}} \right].$$

Using the strong Markov property of $(X_i^0)_{i \in \mathbb{N}^*}$ at the stopping time T and its definition (see (A5)), we get

$$\begin{aligned} \mathbb{E}[Y_j] &= r\mathbb{E}[X^j] + \mathbb{E} \left[\sum_{i=1}^{N-1} X_i^0 \right] - (1-r) \left(-1 + \mathbb{E} \left[\sum_{i=1}^{N-1} X_i^0 \right] \right) \\ &= 1 + r\mathbb{E} \left[X^j - 1 + \sum_{i=1}^{N-1} X_i^0 \right] \\ &= 1 + r \left(\frac{m_j}{p_\alpha(A_j)} - 1 + \frac{m_0 - p_\alpha(A_0)}{p_\alpha(A_0^c)} \right) \\ &= 1 + r \left(\frac{m_j}{p_\alpha(A_j)} + \frac{m_0 - 1}{p_\alpha(A_0^c)} \right), \end{aligned}$$

where $m_\ell = \sum_{k \in A_\ell} kp_\alpha(k)$ for $\ell \in \mathcal{J}$. As p_α is critical, it follows that $\sum_{j \in \mathcal{J}} m_j = 1$. Recall that $\alpha_j = p_\alpha(A_j)/p_\alpha(A_0^c)$ for $j \in \mathcal{J}^*$. Therefore, we obtain

$$\rho = \sum_{j \in \mathcal{J}^*} \alpha_j + r \sum_{j \in \mathcal{J}^*} \frac{m_j}{p_\alpha(A_0^c)} + r \frac{m_0 - 1}{p_\alpha(A_0^c)} \sum_{j \in \mathcal{J}^*} \alpha_j = 1 + r \frac{1 - m_0}{p_\alpha(A_0^c)} + r \frac{m_0 - 1}{p_\alpha(A_0^c)} = 1.$$

This ensures that $p^{A,\alpha}$ is critical.

We now prove part (ii) of the lemma on the positive entries of the mean matrix. Let $j \in \mathcal{J}^*$. We deduce from (68) and (A8) (where α^* has positive entries) and (A7) (where $r > 0$) that $(m_{j\ell})_{\ell \in \mathcal{J}^*}$ are all positive if $\mathbb{P}(Y_j = 0) < 1$ and all zero if $\mathbb{P}(Y_j = 0) = 1$. Note that $T = 1$ a.s. implies that $A_0 = \{0\}$, and thus, $Y_j > 0$ a.s. on $\{N \leq T\}$, so $\mathbb{P}(Y_j = 0) = 1$ implies that $\mathbb{P}(T \geq 2) > 0$. We deduce that $\mathbb{P}(Y_j = 0) = 1$ is equivalent to $\mathbb{P}(X^j = 0) = 1$ and $\mathbb{P}(N = 1) = 1$ or $\mathbb{P}(X_i^0 = 0) = 1$. Thus, $\mathbb{P}(Y_j = 0) = 1$ is equivalent to $0 \in A_j$, $p_\alpha(A_j \cap \{0\}^c) = 0$, and $p_\alpha(A_0) = 0$ or $p_\alpha(A_0 \cap \{1\}^c) = 0$, that is, $0 \in A_j$, $p_\alpha(A_j \cap \{0\}^c) = 0$, and $p_\alpha(A_0 \cap \{1\}^c) = 0$. To conclude, note that these conditions are equivalent to either $A_j = \{0\}$ and $A_0 \subset \{1\}$ or $0 \in A_j$ and $\theta_\alpha = 0$.

We now prove part (iii) of the lemma on the periodicity of $p^{A,\alpha}$. Thanks to (A8) and the fact that α^* has positive entries, we deduce that $p^{A,\alpha}$ is aperiodic (in $\mathbb{Z}^{\mathcal{J}^*}$) if and only if the smallest subgroup of \mathbb{Z} that contains $\Gamma = \bigcup_{j \in \mathcal{J}^*} (\text{supp}(\text{Law}(Z_j | N \leq T)) - \text{supp}(\text{Law}(Z_j | N \leq T)))$ is \mathbb{Z} . From the definition of the law of Z_j given in (A7), we consider the two cases $r = 1$ and $r < 1$. We also remark that $r = 1$ if and only if $T = +\infty$ a.s. or $N = 1$ a.s., which corresponds to $0 \notin A_0 \cap \text{supp}(p_\alpha)$, that is, $0 \notin A_0$ as $p_\alpha(0) > 0$.

In the easy $0 \in A_0$ case (and, thus, $\theta_\alpha \in (0, +\infty)$) we find on the one hand that $X^j > 0$ almost everywhere, and as $\mathbb{P}(N = 1) > 0$, we deduce that $\mathbb{P}(Y_j > 0) > 0$, and thus, that $\{0, 1\} \subset \text{supp}(\text{Law}(Z_j | N \leq T))$. This implies that $p^{A,\alpha}$ is aperiodic. On the other hand, we also find that p is (\mathcal{A}, α) -aperiodic; see Remark 6.1(a).

We now consider the $0 \notin A_0$ case, that is, $r = 1$ and, thus, $Z_j = Y_j$ and a.s. $T = +\infty$. If $p_\alpha(A_0) > 0$, we have $\mathbb{P}(N = k) > 0$ for all $k \in \mathbb{N}$, and if $p_\alpha(A_0) = 0$, we have a.s. $N = 1$ and $A_0 \cap \text{supp}(p_\alpha) \subset \{1\}$. In both cases, we deduce from (A6) that

$$\begin{aligned} \text{supp}(\text{Law}(Z_j | N \leq T)) &= \text{supp}(\text{Law}(Y_j | N \leq T)) \\ &= \text{supp}(\text{Law}(X^j | N \leq T)) + \mathbb{N}(A_0^\alpha - 1), \end{aligned}$$

where $\mathbb{N}B = \{nb : n \in \mathbb{N}, b \in B\}$ and $A_0^\alpha = A_0 \cap \text{supp}(p_\alpha)$. We set $A_j^\alpha = A_j \cap \text{supp}(p_\alpha)$ for $j \in \mathcal{J}^*$ and note that $A_j^\alpha = \text{supp}(\text{Law}(X^j | N \leq T))$. We then find that

$$\Gamma = \left(\bigcup_{j \in \mathcal{J}^*} (A_j^\alpha - A_j^\alpha) \right) \cup \mathbb{Z}(A_0^\alpha - 1).$$

If $\theta_\alpha \in \{0, +\infty\}$, we find that $A_0^\alpha \subset \{1\}$ and A_j^α are singletons for $j \in \mathcal{J}^*$, so that $\Gamma = \{0\}$. If $\theta_\alpha \in (0, +\infty)$, we find that $\Gamma = \Gamma_\alpha$, which is defined in (64).

Thus, $p^{A,\alpha}$ is aperiodic if and only if $\theta_\alpha \notin \{0, +\infty\}$ and the smallest subgroup in \mathbb{Z} containing Γ_α is \mathbb{Z} , that is, p is aperiodic for \mathcal{A} in the direction α . □

6.3 Proof of Theorem 6.1

Recall (10) on the characterization of the local convergence in distribution when the considered random trees are either finite or with a unique infinite spine. As p is generic for \mathcal{A} in the direction α , there exists by Lemma 6.1 a sequence $(\mathbf{n}^{(m)})_{m \in \mathbb{N}}$ satisfying (60)–(62). Since p_α is (p, \mathcal{A}) -compatible, we have, for $m \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}_0$, and $x \in \mathcal{L}_{\{0\}}(\mathbf{t})$, with $\mathcal{T}_\alpha = \mathcal{T}_{p_\alpha}$ and $\mathcal{T}_\alpha^* = \mathcal{T}_{p_\alpha}^*$, that

$$\mathbb{P}(\mathcal{T}_p \in \mathbb{T}(\mathbf{t}, x) \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}^{(m)}) = \mathbb{P}(\mathcal{T}_\alpha \in \mathbb{T}(\mathbf{t}, x) \mid L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)}).$$

We first prove that $\lim_{m \rightarrow \infty} \mathbb{P}(\mathcal{T}_p \in \mathbb{T}(\mathbf{t}, x) \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}^{(m)}) = \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x))$. For $j \in \llbracket 1, J \rrbracket$, recall that $\mathbf{e}^j \in \mathbb{N}^J$ is the vector with all its entries equal to 0 but the j th, which is equal to 1, and that $\mathbf{e}^0 = \mathbf{0}$. Set $j_0 \in \llbracket 0, J \rrbracket$ such that $0 \in A_{j_0}$ and set

$$\mathbf{b} = \mathbf{e}^{j_0}.$$

We have

$$\begin{aligned} &\mathbb{P}(\mathcal{T}_\alpha \in \mathbb{T}(\mathbf{t}, x), L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)}) \\ &= \sum_{\tilde{\mathbf{t}} \in \mathbb{T}_0} \mathbb{P}(\mathcal{T}_\alpha = \mathbf{t} \otimes (\tilde{\mathbf{t}}, x)) 1_{\{L_{\mathcal{A}}(\mathbf{t} \otimes (\tilde{\mathbf{t}}, x)) = \mathbf{n}^{(m)}\}} \\ &= \frac{1}{p_\alpha(0)} \sum_{\tilde{\mathbf{t}} \in \mathbb{T}_0} \mathbb{P}(\mathcal{T}_\alpha = \mathbf{t}) \mathbb{P}(\mathcal{T}_\alpha = \tilde{\mathbf{t}}) 1_{\{L_{\mathcal{A}}(\tilde{\mathbf{t}}) = \mathbf{n}^{(m)} - L_{\mathcal{A}}(\mathbf{t}) + \mathbf{b}\}} \\ &= \frac{1}{p_\alpha(0)} \mathbb{P}(\mathcal{T}_\alpha = \mathbf{t}) \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)} - L_{\mathcal{A}}(\mathbf{t}) + \mathbf{b}) \\ &= \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x)) \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)} - L_{\mathcal{A}}(\mathbf{t}) + \mathbf{b}), \end{aligned}$$

where we used the fact that p_α is critical (and, thus, \mathcal{T}_α is a.s. finite) for the first and third equalities and (12) for the last equality. Recall \mathbf{x}^* from (65), which is the restriction of the sequence \mathbf{x} indexed by $\llbracket 1, J \rrbracket$ to the indices \mathcal{J}^* , and that $L_{\mathcal{A}^*}(\mathbf{t}) = (L_{A_j}(\mathbf{t}))_{j \in \mathcal{J}^*}$. We get

$$\begin{aligned} &\mathbb{P}(\mathcal{T}_\alpha \in \mathbb{T}(\mathbf{t}, x) \mid L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)}) \\ &= \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x)) \frac{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)} - L_{\mathcal{A}}(\mathbf{t}) + \mathbf{b})}{\mathbb{P}(L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)})} \\ &= \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x)) \frac{\mathbb{P}(L_{\mathcal{A}^*}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)*} - L_{\mathcal{A}^*}(\mathbf{t}) + \mathbf{b}^*)}{\mathbb{P}(L_{\mathcal{A}^*}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)*})}, \end{aligned} \tag{69}$$

where we used the fact that $L_{A_j}(\mathcal{T}_\alpha) = 0$ for $j \notin \mathcal{J}$ (and, thus, $\mathbb{P}(\mathcal{T}_\alpha \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x)) = 0$ if $L_{A_j}(\mathbf{t}) \neq 0$ for some $j \notin \mathcal{J}$) as well as the fact that $(\mathbf{n}^{(m)})_{m \in \mathbb{N}}$ satisfies (62) for the second equality. Now we apply the extension of Rizzolo’s transformation for \mathcal{T}_α to get a \mathcal{J}^* -type BGW tree $\mathcal{T}_\alpha^{\mathcal{A}}$ such that $\sharp \mathcal{T}_\alpha^{\mathcal{A}} = L_{\mathcal{A}^*}(\mathcal{T}_\alpha)$ (see definition (66)). Hence, (69) is equivalent to

$$\mathbb{P}(\mathcal{T}_\alpha \in \mathbb{T}(\mathbf{t}, x) \mid L_{\mathcal{A}}(\mathcal{T}_\alpha) = \mathbf{n}^{(m)}) = \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x)) \frac{\mathbb{P}(\sharp \mathcal{T}_\alpha^{\mathcal{A}} = \mathbf{n}^{(m)*} - L_{\mathcal{A}^*}(\mathbf{t}) + \mathbf{b}^*)}{\mathbb{P}(\sharp \mathcal{T}_\alpha^{\mathcal{A}} = \mathbf{n}^{(m)*})}. \tag{70}$$

We consider the following condition that appears in Lemma 6.3(ii):

$$A_0 \subset \{1\} \quad \text{and} \quad A_{j_0} = \{0\} \quad \text{for some } j_0 \in \mathcal{J}^*. \tag{71}$$

We first assume that (71) does not hold. The hypothesis of Theorem 6.1 and Lemma 6.3 ensure that assumptions (H_1) (on the offspring distribution being critical and the mean matrix primitive) and (H_2) (on the aperiodicity of the offspring reproduction) hold in [7] and that α^* is the positive normalized left eigenvector of the mean matrix (see the hypothesis in [7, Lemma 3.11] where $a = \alpha^*$ and use the fact that $(\mathbf{n}^{(m)}, m \in \mathbb{N})$ is a sequence in \mathbb{N}^J satisfying (60)–(62)), so that the strong ratio theorem or, more precisely, [7, Equation (19)] holds, which entails that

$$\lim_{m \rightarrow \infty} \frac{\mathbb{P}(\sharp \mathcal{T}_\alpha^{\mathcal{A}} = \mathbf{n}^{(m)*} - L_{\mathcal{A}^*}(\mathbf{t}) + \mathbf{b}^*)}{\mathbb{P}(\sharp \mathcal{T}_\alpha^{\mathcal{A}} = \mathbf{n}^{(m)*})} = 1. \tag{72}$$

We deduce from (70) and (72) that, for all $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_{\{0\}}(\mathbf{t})$,

$$\lim_{m \rightarrow \infty} \mathbb{P}(\mathcal{T}_p \in \mathbb{T}(\mathbf{t}, x) \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}^{(m)}) = \mathbb{P}(\mathcal{T}_\alpha^* \in \mathbb{T}(\mathbf{t}, x)).$$

Second, for $\mathbf{t} \in \mathbb{T}_0$, it is obvious from (61) that

$$\lim_{m \rightarrow \infty} \mathbb{P}(\mathcal{T}_p = \mathbf{t} \mid L_{\mathcal{A}}(\mathcal{T}_p) = \mathbf{n}^{(m)}) = \mathbb{P}(\mathcal{T}_\alpha^* = \mathbf{t}) = 0.$$

The result thus follows from the fact that the family $\{(\mathbb{T}(\mathbf{t}, x), \mathbf{t} \in \mathbb{T}_0, x \in \mathcal{L}_0(\mathbf{t})) \cup \mathbb{T}_0$ is convergence determining for the local convergence in $\mathbb{T}_0 \cup \mathbb{T}_1$.

We now consider that (71) holds. We first check that it is enough to consider the $A_0 = \emptyset$ case, where the Rizzolo’s transformation is the identity map. Indeed, if $A_0 \neq \emptyset$, that is, $A_0 = \{1\}$, then the Rizzolo’s transformation corresponds to discarding individuals with only one child. This amounts to replacing the offspring distribution p (respectively $p_{\theta, \alpha}$) by p' (respectively $p'_{\theta, \alpha} = (p')_{\theta, \alpha}$), where $p'(k) = p(k)/q_1$ for $k \neq 1$ and $p'(k) = 0$ for $k = 1$. Then, note that θ such that $p'_{\theta, \alpha}$ is critical is exactly θ_α , so without confusion, we can also replace p_α by $p'_\alpha = (p')_\alpha$. In conclusion, using this modification amounts to only considering the following case:

$$A_0 = \emptyset \quad \text{and} \quad A_{j_0} = \{0\} \quad \text{for some } j_0 \in \mathcal{J}^*. \tag{73}$$

Note that this case is ruled out in [7, Corollary 3.5]. However, a slight modification of the proofs in [7], which we sketch in Section 6.4 (take $d = \text{Card}(\mathcal{J}^*)$ and $d = j_0$ therein), allows us to obtain (72), which we now read as, for a sequence $(\mathbf{n}^{(m)}, m \in \mathbb{N})$ in \mathbb{N}^J satisfying (60)–(62),

$$\lim_{m \rightarrow \infty} \frac{\mathbb{P}(\#\mathcal{T}_\alpha = \mathbf{n}^{(m)*} - L_{\mathcal{A}^*}(\mathbf{t}) + \mathbf{b}^*)}{\mathbb{P}(\#\mathcal{T}_\alpha = \mathbf{n}^{(m)*})} = 1, \tag{74}$$

where \mathcal{T}_α is seen as a multi-type BGW tree, where a node $u \in \mathcal{T}_\alpha$ as type $j \in \mathcal{J}^*$ if $k_u(\mathcal{T}_\alpha) \in A_j$. Note that the corresponding offspring distribution is $\mathbf{p} = (\mathbf{p}^{(i)})_{i \in \mathcal{J}^*}$, where $\mathbf{p}^{(i)} = (\mathbf{p}^{(i)}(\mathbf{k}))_{\mathbf{k} \in \mathbb{N}^{\mathcal{J}^*}}$ is a probability distribution on $\mathbb{N}^{\mathcal{J}^*}$ whose nonzero terms are given by

$$\mathbf{p}^{(i)}(\mathbf{k}) = \frac{p_\alpha(|\mathbf{k}|)}{p_\alpha(A_i)} \text{Mult}(\mathbf{k}, \alpha^*) \quad \text{for } |\mathbf{k}| \in A_i,$$

and $\text{Mult}(\mathbf{k}, \alpha^*)$ is the probability that a multinomial random variable with parameter $(|\mathbf{k}|, \alpha^*)$ takes the value \mathbf{k} . Furthermore, the type of the root is distributed as α^* . With this setting, we emphasize that (74) is exactly Equation (19) in [7], up to a relabeling. Once (74) is established, then we complete the proof as in the case where (71) does not hold.

We now give the properties of the offspring distribution \mathbf{p} and the type of the root (recall that (73) holds and that the Rizzolo’s transformation is the identity map); under assumption of Theorem 6.1, we have the following assertions.

1. The type of the root is distributed as α^* .
2. $\mathbf{p}^{(j_0)}(\mathbf{0}) = 1$.
3. By Lemma 6.3(ii), the mean matrix $M = (m_{ij})_{i, j \in \mathcal{J}^*}$ is such that, for $j \in \mathcal{J}^*$, we have $m_{ij} \in (0, +\infty)$ for $i \neq j_0$ and $m_{ij} = 0$ otherwise.
4. By Lemma 6.3(i), \mathbf{p} is critical and α^* is the left eigenvector with eigenvalue 1.
5. By Lemma 6.3(iii), \mathbf{p} is aperiodic.

- 6. Since p_α satisfies (5) and $A_0 = \emptyset$, we deduce from the definition of \mathbf{p} that there exists a type $j \in \mathcal{J}^*$ such that the individual of type j has two children or more with positive probability, that is, \mathbf{p} is nonsingular.

In particular, the offspring distribution \mathbf{p} satisfies hypothesis (75)–(78) from Section 6.4. To conclude, we refer to Section 6.4 on how to get [7, Equation (19)] under this set of hypotheses.

Remark 6.4. (*On related work.*) The case $A_0 = \emptyset$ and $\text{Card}(A_{j_0}) \geq 2$, where $j_0 \in \llbracket 1, J \rrbracket$ is such that $0 \in A_{j_0}$ (compare with condition (71)) could be handled using [25, Theorem 5.1] on multi-type BGW processes. (We also believe that condition (A5) there, which amounts to saying that, for each $j \in \llbracket 1, J \rrbracket$, there is $k \in A_j$ such that $k + 1 \in A_j$, could certainly be relaxed.) Note that the moments condition considered there does not allow us to consider directions α such that $\theta_\alpha = \theta_{\max}$ (this case might indeed exist). The possible vector $\mathbf{a} = (a_1, \dots, a_J)$ considered in [25] (which is associated with the critical BGW multi-type process) corresponds in our framework to $a_j = \theta_\alpha \alpha_j / p(A_j)$ for $j \in \llbracket 1, J \rrbracket$ and the direction $\bar{\mathbf{v}}$ that appears in [25, Equation (5.1)] corresponds to α . In our approach, we first fix the direction α and then give sufficient (and almost necessary) conditions for the existence and uniqueness of the critical parameter θ_α , and thus, how to choose the parameter \mathbf{a} given the direction α .

6.4. On the proof of (74)

In this section we quickly revisit the proof of Equation (19) in [7], using slightly different assumptions in order to take into account the particular case (71) from Section 6.3. In this section only, we stick to the notation introduced in [7]. Let $d \geq 2$ and set $[n] = \llbracket 1, n \rrbracket$ for $n \in \mathbb{N}^*$. Let $p = (p^{(i)}, i \in [d])$ with $p^{(i)} = (p^{(i)}(\mathbf{k}), \mathbf{k} \in \mathbb{N}^d)$ being probability distributions on \mathbb{N}^d . We assume that

$$p^{(d)}(\mathbf{0}) = 1. \tag{75}$$

For $i \in [d]$, let $X_i = (X_i^{(j)}, j \in [d])$ be a random variable on \mathbb{N}^d with probability distribution $p^{(i)}$. In particular, we have $X_d = \mathbf{0}$ a.s. We consider the generating function $f = (f^{(i)}, i \in [d])$ of p defined by

$$f^{(i)}(s) = \mathbb{E} \left[\prod_{j \in [d]} s_j^{X_i^{(j)}} \right], \quad \text{where } s = (s_j, j \in [d]) \in [0, 1]^d.$$

We consider the mean matrix $M = (m_{ij}; i, j \in [d])$ with $m_{ij} = \mathbb{E}[X_i^{(j)}]$. We assume that

$$m_{ij} \in (0, +\infty) \quad \text{for all } i \in [d - 1], j \in [d]; \tag{76}$$

note that $m_{dj} = 0$ for all $j \in [d]$. In particular, *the matrix M is not primitive*, as there is no $n \in \mathbb{N}^*$ such that M^n has only positive finite entries; note that M primitive is part of assumption (H_1) in [7] (this condition is mainly used to apply Perron-Frobenius theorem on the existence and uniqueness of a left and right eigenvector having nonnegative entries, and their corresponding eigenvalue is in fact the spectral radius of M). We recall that p is critical if the spectral radius of M is 1 and that p is nonsingular if $f(s) \neq Ms$. We assume that

$$p \text{ is critical and nonsingular}; \tag{77}$$

note that this is the other part of assumption (H_1) in [7]. We also assume that

$$p \text{ is aperiodic}, \tag{78}$$

that is, the smallest subgroup of \mathbb{Z}^d that contains $\bigcup_{i \in [d]} (\text{supp}(p^{(i)}) - \text{supp}(p^{(i)}))$ is \mathbb{Z}^d itself; this corresponds to hypothesis (H_2) in [7].

For $i \in [d]$, let \mathbf{e}_i denote the vector of \mathbb{R}^d with all its entries equal to 0 except the i th that is equal to 1. Using Perron-Frobenius theorem for the matrix M reduced to the first $d - 1$ lines and columns, and using the fact that the d th line of M is zero and the other entries are positive, we deduce that

- the eigenvalue 1 is simple;
- there exists two left eigenvectors with nonnegative entries: the vector \mathbf{e}_d with eigenvalue 0 and a vector $a \in \mathbb{R}^d$ having positive entries with eigenvalue 1;
- there exists a unique right eigenvector $a^* = (a^*(i), i \in [d])$ with eigenvalue 1, whose entries are positive except for the d th entry that is zero, i.e. $a^*(d) = 0$.

This result is the reason why we can remove the primitive assumption of M .

Then the results on the Dwass formula for BGW multi-type trees from [7, Section 3.2] also hold, as the expressions therein are algebraic in the entries of p . (For example, Lemma 3.8 holds, but note that both terms of the equality therein are zero if $r = d$.) Now Equation (19) of [7] is then a direct consequence of the Dwass formula and the technical Lemma 3.11 therein. This latter result, proved in Section 3.4, is also a direct consequence of Lemma 3.12, which asserts that an intermediate random variable Y on \mathbb{Z}^{2d-1} has an aperiodic distribution, and of Lemma 4.11, which is a variant of the strong ratio theorem for the random walk with increments distributed as Y . Now looking carefully at the proof of Lemma 3.12, we see that p is assumed to be aperiodic (this is (H_2) therein and (78) here) and that hypothesis (H_1) is only used at the end of the proof to show that $\mathbb{P}(X_d = \mathbf{0}) > 0$; but this is clearly the case if (75) holds. To conclude, note that Lemma 4.11 on the strong ratio theorem requires only that the law of Y is aperiodic (which is provided by Lemma 3.12) and that Y is integrable. By the construction of Y given in Section 3.4, we note that Y is integrable if and only if the mean matrix M has finite entries, which is hypothesis (76). In conclusion, we find that Equation (19) of [7] holds (note that the root has to be of type $r \neq d$ otherwise the numerator and denominator are both zero).

Remark 6.5. (On the extension to the main result of [7] under hypothesis (75)–(78).) We leave to the interested reader the construction of the corresponding Kesten’s tree (see [7, Section 2.6]), where here individuals on the infinite spine can not have type d (in particular, the root does not have type d). (For example, Lemma 2.9 therein holds provided i, r belong to $[d - 1]$.) Then, assuming that (75)–(78) hold, we have the analogue of Theorem 3.1 therein on the local convergence in distribution, towards the Kesten’s tree of the BGW multi-type tree (with the root not being a.s. of type d and with offspring distribution p) conditioned to have a population of type i equal to $k(n)_i$ for $i \in [d]$ with $\lim_{n \rightarrow \infty} k(n)_i / |k(n)| = a(i)$, where $|k(n)| = \sum_{j \in [d]} k(n)_j$ and $\lim_{n \rightarrow \infty} |k(n)| = \infty$.

6.5. Details for Remark 1.1 on the condition $n_j = 0$ if $\alpha_j = 0$

We consider the following example: a probability distribution p such that $\text{supp}(p)$ contains but not reduced to $\{0, 2\}$, $1 \notin \text{supp}(p)$, $J = 2$, $\mathcal{A} = (A_1, A_2)$ is a partition of $\text{supp}(p)$ (that is, $A_0 = \emptyset$) with $A_1 = \{0, 2\}$ and $A_2 \subset 3 + 2\mathbb{N}$. Note that $A_2 \neq \emptyset$. We set

$$a = \frac{\sqrt{p(0)}}{\sqrt{p(2)}} \in (0, +\infty). \quad (79)$$

We consider the direction $\alpha = (1, 0)$. It is elementary to check that p is generic for \mathcal{A} in the direction α and that $p_\alpha = (p_\alpha(n))_{n \in \mathbb{N}}$ is given by $p_\alpha(0) = p_\alpha(2) = \frac{1}{2}$. The distribution p is however not aperiodic for \mathcal{A} in the direction α , but thanks to Remark 1.2 we still have the convergence of \mathcal{T} conditionally on $L_{\mathcal{A}}(\mathcal{T}) = (n, 0)$, with n odd going to infinity, locally in distribution towards the Kesten’s tree \mathcal{T}_α^* .

For n odd going to infinity, we check that the distribution of \mathcal{T} conditionally on $L_{\mathcal{A}}(\mathcal{T}) = (n, 1)$ does not converge locally to the distribution of \mathcal{T}_α^* , and thus, condition (62) is required in general to get the local limit of conditioned BGW trees from Theorem 6.1. To do so, we simply check the positivity of the limit, for n odd going to infinity, of

$$\mathbb{P}(k_\emptyset(\mathcal{T}) \neq 2 \mid L_{\mathcal{A}}(\mathcal{T}) = (n, 1)) = \frac{B_1(n)}{B_2(n)},$$

where

$$B_1(n) = \mathbb{P}(k_\emptyset(\mathcal{T}) \neq 2, L_{\mathcal{A}}(\mathcal{T}) = (n, 1)) \quad \text{and} \quad B_2(n) = \mathbb{P}(L_{\mathcal{A}}(\mathcal{T}) = (n, 1)).$$

Before going further, we recall that the number of planar binary trees with n leaves is

$$f_{1,n} = \frac{1}{n} \binom{2n-2}{n-1},$$

(in particular, $f_{1,n+1}$ is the so-called n th Catalan’s number) and that $f_{1,n} = [z^n]zC = [z^{n-1}]C$, where we simply write C for $C(z) = (1 - \sqrt{1 - 4z})/2z$. Recall also that $zC^2 - C + 1 = 0$. We deduce that the number of planar forests with k binary trees and $n \geq k$ leaves is given by $f_{k,n} = [z^n]z^k C^k = [z^{n-k}]C^k$, and that according to [10, Equation (B.5)]:

$$f_{k,n} = \frac{k}{n} \binom{2n-k-1}{n-1} = \binom{2n-k-1}{n-1} - \binom{2n-k-1}{n} \quad \text{for } n \geq k \geq 1. \tag{80}$$

We set $f_{k,n} = 0$ if $k > n$.

Let $n \geq k$ be odd integers. On the event $k_\emptyset(\mathcal{T}) = k$ and $L_{\mathcal{A}}(\mathcal{T}) = (n, 1)$, it follows that \mathcal{T} can be seen as a forest of k trees grafted on the root and with the forest having $(n+k)/2$ leaves and $(n-k)/2$ internal nodes, all of them binary. We deduce that

$$\mathbb{P}(k_\emptyset(\mathcal{T}) = k, L_{\mathcal{A}}(\mathcal{T}) = (n, 1)) = p(k)p(0)^{(n+k)/2} p(2)^{(n-k)/2} f_{k,(n+k)/2},$$

and thus,

$$B_1(n) = (p(0)p(2))^{n/2} \sum_{k \in A_2, k \leq n} f_{k,(n+k)/2} p(k) a^k.$$

A tree \mathbf{t} such that $L_{\mathcal{A}}(\mathbf{t}) = (n, 1)$ can be decomposed as a binary tree with ℓ leaves, and on one of those ℓ leaves one grafts a forest with $k \in A_2$ (and $k \leq n$) binary trees with $(n+k)/2 + 1 - \ell$ leaves; and in total the tree \mathbf{t} has $(n+k)/2$ leaves, $(n-k)/2$ binary branching nodes, and one node with out-degree k . Thus, we have

$$\begin{aligned} B_2(n) &= \sum_{k \in A_2, k \leq n} p(k)p(0)^{(n+k)/2} p(2)^{(n-k)/2} \sum_{\ell=1}^{(n-k)/2+1} \ell f_{1,\ell} f_{k,(n+k)/2+1-\ell} \\ &= (p(0)p(2))^{n/2} \sum_{k \in A_2, k \leq n} F_{k,(n+k)/2} p(k) a^k, \end{aligned} \tag{81}$$

where, for $n \geq k \geq 1$,

$$F_{k,n} = \sum_{\ell=1}^{n-k+1} \ell f_{1,\ell} f_{k,n+1-\ell}.$$

We give an explicit formula of $F_{k,n}$.

Lemma 6.4. *We have*

$$F_{k,n} = \binom{2n-k}{n} = \frac{2n-k}{k} f_{k,n} \quad \text{for } n \geq k \geq 1.$$

Proof. We have

$$\begin{aligned} F_{k,n} &= [z^n](zC)' z^k C^k = [z^n] \frac{1}{k+1} (z^{k+1} C^{k+1})' \\ &= [z^{n+1}] \frac{n+1}{k+1} z^{k+1} C^{k+1} = \frac{n+1}{k+1} f_{k+1,n+1}. \end{aligned}$$

Then, we use (80) to conclude. □

We now consider that A_2 is unbounded. Let $\varepsilon \in (0, 1)$ and write

$$\begin{aligned} B_1(n, \varepsilon) &= \sum_{k \in A_2, k > \varepsilon n} f_{k,(n+k)/2} p(k) a^k, \\ B_3(n, \varepsilon) &= \sum_{k \in A_2, k \leq \varepsilon n} F_{k,(n+k)/2} p(k) a^k, \\ B_4(n, \varepsilon) &= \sum_{k \in A_2, k > \varepsilon n} F_{k,(n+k)/2} p(k) a^k, \end{aligned}$$

so that using the two latter terms, we can rewrite $B_2(n)$ as

$$B_2(n) = (p(0)p(2))^{n/2} (B_3(n, \varepsilon) + B_4(n, \varepsilon)). \tag{82}$$

For $k > \varepsilon n$, we have

$$F_{k,(n+k)/2} = \frac{n}{k} f_{k,(n+k)/2} \leq \frac{1}{\varepsilon} f_{k,(n+k)/2}.$$

This implies that

$$B_4(n, \varepsilon) \leq \frac{1}{\varepsilon} B_1(n, \varepsilon). \tag{83}$$

We now assume that $A_2 = 3 + 2\mathbb{N}$ and there exists $b \in (0, 1)$ and $M \geq 1$ finite such that $M^{-1} \leq p(k)b^{-k} \leq M$ for $k \in A_2$. Then, we have, with $2m = n + 3$ and $k = 3 + 2\ell$,

$$\begin{aligned} B_3(n, \varepsilon) &\leq M \sum_{k \in 3+2\mathbb{N}, k \leq \varepsilon n} F_{k,(n+k)/2} (ab)^k \\ &\leq M(ab)^3 \sum_{\ell \in \mathbb{N}, \ell \leq \varepsilon m} F_{3+2\ell,m+\ell} (ab)^{2\ell} \\ &= M(ab)^{-2m+3} (1 + (ab)^2)^{2m-3} \sum_{\ell \in \mathbb{N}, \ell \leq \varepsilon m} \binom{2m-3}{m+\ell} r^{m+\ell} (1-r)^{m-3-\ell}, \end{aligned}$$

with $r/(1-r) = (ab)^2$ and, thus, $r = (ab)^2/(1+(ab)^2)$. As $r < 1$, we deduce that

$$\sum_{\ell \in \mathbb{N}, \ell \leq \varepsilon m} \binom{2m-3}{m+\ell} r^{m+\ell} (1-r)^{m-3-\ell} = \mathbb{P}(m \leq X \leq (1+\varepsilon)m) \leq \mathbb{P}(X \leq (1+\varepsilon)m),$$

where X is binomial with parameter $(2m-3, r)$.

We now assume that $ab > 1$ and that ε is small enough so that $ab > (1+\varepsilon)/(1-\varepsilon)$. This yields $2r > 1+\varepsilon$, so that for m large enough, we have $(1+\varepsilon)m \leq r(2m-3)$. We deduce from [32, Theorem 2.1] that, with $j = \lfloor (1+\varepsilon)m \rfloor$ and $x = r(2m-3) - j + 1$,

$$\frac{\mathbb{P}(X \leq (1+\varepsilon)m)}{\mathbb{P}(X=j)} \leq 2-x + \sqrt{x^2 + 4(1-r)j}.$$

Since $\lim_{m \rightarrow \infty} x = +\infty$ and $\lim_{m \rightarrow \infty} (x+j)/j = 2r/(1+\varepsilon)$, we deduce that

$$\lim_{m \rightarrow \infty} \frac{\mathbb{P}(X \leq (1+\varepsilon)m)}{\mathbb{P}(X=j)} = c_0 \quad \text{with } c_0 = \frac{r(1-\varepsilon)}{2r - (1+\varepsilon)}.$$

Recall that $n = 2m - 3$. So for large enough n , we have, with $k' = 3 + 2\ell'$ and $\ell' = j - m = \lfloor \varepsilon m \rfloor$, and thus, $n + k' = 2j$,

$$\begin{aligned} B_3(n, \varepsilon) &\leq 2c_0 M (ab)^{-2m+3} \left(1 + (ab)^2\right)^{2m-3} \binom{2m-3}{j} r^j (1-r)^{2m-3-j} \\ &= 2c_0 M F_{k', (n+k')/2} (ab)^{k'}. \end{aligned}$$

Note that $j \geq (1+\varepsilon)m - 1$ and, thus, $k' \geq \varepsilon n + 1$, so that

$$B_3(n, \varepsilon) \leq 2c_0 M^2 B_4(n, \varepsilon). \quad (84)$$

Hence, using (83), we obtain

$$B_3(n, \varepsilon) + B_4(n, \varepsilon) \leq \frac{1 + 2c_0 M^2}{\varepsilon} B_1(n, \varepsilon).$$

From the definition of $B_1(n, \varepsilon)$, it follows that

$$(p(0)p(2))^{n/2} B_1(n, \varepsilon) = \mathbb{P}(k_\emptyset(\mathcal{T}) \geq \varepsilon n, L_{\mathcal{A}}(\mathcal{T}) = (n, 1)).$$

We deduce from (82) that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(k_\emptyset(\mathcal{T}) \geq \varepsilon n \mid L_{\mathcal{A}}(\mathcal{T}) = (n, 1)) = \liminf_{n \rightarrow \infty} \frac{B_1(n, \varepsilon)}{B_2(n)} \geq \frac{\varepsilon}{1 + 2c_0 M^2} > 0,$$

provided that there exists $b \in (0, 1)$ and $M \geq 1$ finite such that $M^{-1} \leq p(k)b^{-k} \leq M$ for $k \in A_2 = 3 + 2\mathbb{N}$ and $\varepsilon > 0$ is chosen so that $ab > (1+\varepsilon)/(1-\varepsilon)$, with a defined by (79).

In conclusion, under the above hypothesis, the distribution of \mathcal{T} conditionally on $\{L_{\mathcal{A}}(\mathcal{T}) = (n, 1)\}$ does not converge locally to the distribution of \mathcal{T}_α^* as n tends to infinity, whereas the distribution of \mathcal{T} conditionally on $\{L_{\mathcal{A}}(\mathcal{T}) = (n, 0)\}$ converges locally to the distribution of \mathcal{T}_α^* . Furthermore, conditioning on $\{L_{\mathcal{A}}(\mathcal{T}) = (n, 1)\}$ and letting n tend to infinity gives a condensation at the root with positive probability.

Acknowledgements

We thank the two referees for their precious work and whose comments allowed us to improve the presentation of the results. H. Bi would like to extend his gratitude to R. Abraham and J.-F. Delmas for their exceptional hospitality during his visit to CERMICS, where this work was carried out.

Funding information

There are no funding bodies to thank relating to this creation of this paper.

Competing interests

There were no competing interests to declare that arose during the preparation or publication process of this paper.

References

- [1] ABRAHAM, R., BOUAZIZ, A. AND DELMAS, J.-F. (2017). Local limits of Galton-Watson trees conditioned on the number of protected nodes. *J. Appl. Prob.* **54**, 55–65.
- [2] ABRAHAM, R., BOUAZIZ, A. AND DELMAS, J.-F. (2020). Very fat geometric Galton-Watson trees. *ESAIM: Prob. Statist.* **24**, 294–314.
- [3] ABRAHAM, R. AND DELMAS, J.-F. (2014). Local limits of conditioned Galton-Watson trees: the condensation case. *Electron. J. Prob.* **19**, 1–29.
- [4] ABRAHAM, R. AND DELMAS, J.-F. (2014). Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electron. J. Prob.* **19**, 1–19.
- [5] ABRAHAM, R. AND DELMAS, J.-F. (2015). An introduction to Galton-Watson trees and their local limits. Preprint. Available at <https://arxiv.org/abs/1506.05571>.
- [6] ABRAHAM, R. AND DELMAS, J.-F. (2019). Asymptotic properties of expansive Galton-Watson trees. *Electron. J. Prob.* **24**, 1–51.
- [7] ABRAHAM, R., DELMAS, J.-F. AND GUO, H. (2018). Critical multi-type Galton-Watson trees conditioned to be large. *J. Theoret. Prob.* **31**, 757–788.
- [8] ALDOUS, D. (1991). The continuum random tree. *I. Ann. Prob.* **19**, 1–28.
- [9] CURIEN, N. AND KORTCHEMSKI, I. (2014). Random non-crossing plane configurations: A conditioned Galton-Watson tree approach. *Random Struct. Algorithms* **45**, 236–260.
- [10] DEUTSCH, E. (1999). Dyck path enumeration. *Discrete Math.* **204**, 167–202.
- [11] DUQUESNE, T. AND LE GALL, J.-F. (2002). *Random trees, Lévy Processes and Spatial Branching Processes*. Astérisque vi+147.
- [12] DURHUUS, B. AND ÜNEL, M. (2023). Trees with exponential height dependent weight. *Prob. Theory Related Fields* **186**, 999–1043.
- [13] GEIGER, J. AND KAUFFMANN, L. (2004). The shape of large Galton-Watson trees with possibly infinite variance. *Random Struct. Algorithms* **25**, 311–335.
- [14] HE, X. (2017). Conditioning Galton-Watson trees on large maximal outdegree. *J. Theoret. Prob.* **30**, 842–851.
- [15] HE, X. (2022). Local convergence of critical random trees and continuous-state branching processes. *J. Theoret. Prob.* **35**, 685–713.
- [16] JANSON, S. (2012). Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Prob. Surv.* **9**, 103–252.
- [17] JANSON, S. (2016). Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. *Random Struct. Algorithms* **48**, 57–101.
- [18] JONSSON, T. AND STEFANSSON, S. Ö. (2011). Condensation in nongeneric trees. *J. Statist. Phys.* **142**, 277–313.
- [19] KARGIN, V. (2023). Scaling limits of slim and fat trees. *J. Theoret. Prob.* **36**, 2192–2228.
- [20] KENNEDY, D. P. (1975). The Galton-Watson process conditioned on the total progeny. *J. Appl. Prob.* **12**, 800–806.
- [21] KESTEN, H. (1986). Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Prob. Statist.* **22**, 425–487.
- [22] KORTCHEMSKI, I. AND MARZOUK, C. (2023). Large deviation local limit theorems and limits of biconditioned planar maps. *Ann. Appl. Prob.* **33**, 3755–3802.

- [23] MARZOUK, C. (2022). On scaling limits of random trees and maps with a prescribed degree sequence. *Ann. H. Lebesgue* **5**, 317–386.
- [24] NEVEU, J. (1986). Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré Prob. Statist.* **22**, 199–207.
- [25] PÉNISSON, S. (2016). Beyond the Q-process: various ways of conditioning the multitype Galton-Watson process. *Latin Amer. J. Prob. Math. Statist.* **13**, 223–237.
- [26] POUDEVIGNE, R. AND THÉVENIN, P. (2025). Existence of critical tiltings and local limits of general size-conditioned Bienaymé-Galton-Watson multitype trees. Preprint. Available at <https://arxiv.org/abs/2503.11501>.
- [27] RIZZOLO, D. (2015). Scaling limits of Markov branching trees and Galton-Watson trees conditioned on the number of vertices with out-degree in a given set. *Ann. Inst. H. Poincaré Prob. Statist.* **51**, 512–532.
- [28] STEPHENSON, R. (2018). Local convergence of large critical multi-type Galton-Watson trees and applications to random maps. *J. Theoret. Prob.* **31**, 159–205.
- [29] STUFLER, B. (2019). Local limits of large Galton-Watson trees rerooted at a random vertex. *Ann. Inst. H. Poincaré Prob. Statist.* **55**, 155–183.
- [30] THÉVENIN, P. (2020). Vertices with fixed outdegrees in large Galton-Watson trees. *Electron. J. Prob.* **25**, 1–25.
- [31] THÉVENIN, P. (2023). Critical exponential tiltings for size-conditioned multitype Bienaymé-Galton-Watson trees. Preprint. Available at <https://arxiv.org/abs/2310.12897>.
- [32] ZHU, H., LI, Z. AND HAYASHI, M. (2022). Nearly tight universal bounds for the binomial tail probabilities. Preprint. Available at <https://arxiv.org/abs/2211.01688>.