Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2012, Vol. 48, No. 3, 688–705
DOI: 10.1214/11-AIHP423

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Pruning Galton–Watson trees and tree-valued Markov processes[1]

Romain Abraham[a], Jean-François Delmas[b] and Hui He[c,d]

[a]*MAPMO, CNRS UMR 6628, Fédération Denis Poisson FR 2964, Université d'Orléans, B.P. 6759, 45067 Orléans cedex 2, France. E-mail: romain.abraham@univ-orleans.fr*
[b]*CERMICS, Université Paris-Est, 6-8 av. Blaise Pascal, Champs-sur-Marne, 77455 Marne La Vallée, France. E-mail: delmas@cermics.enpc.fr*
[c]*MAPMO, CNRS UMR 6628, Fédération Denis Poisson FR 2964, Université d'Orléans, B.P. 6759, 45067 Orléans cedex 2, France*
[d]*School of Mathematical Sciences, Beijing Normal University, Beijing 100875, P. R. China. E-mail: hehui@bnu.edu.cn*

Received 2 July 2010; revised 14 February 2011; accepted 23 February 2011

Abstract. We present a new pruning procedure on discrete trees by adding marks on the nodes of trees. This procedure allows us to construct and study a tree-valued Markov process $\{\mathcal{G}(u)\}$ by pruning Galton–Watson trees and an analogous process $\{\mathcal{G}^*(u)\}$ by pruning a critical or subcritical Galton–Watson tree conditioned to be infinite. Under a mild condition on offspring distributions, we show that the process $\{\mathcal{G}(u)\}$ run until its ascension time has a representation in terms of $\{\mathcal{G}^*(u)\}$. A similar result was obtained by Aldous and Pitman (*Ann. Inst. H. Poincaré Probab. Statist.* **34** (1998) 637–686) in the special case of Poisson offspring distributions where they considered uniform pruning of Galton–Watson trees by adding marks on the edges of trees.

Résumé. Nous présentons une nouvelle procédure d'élagage d'arbres discrets en ajoutant des marques sur les noeuds de l'arbre. Cette procédure nous permet de définir un processus de Markov $\{\mathcal{G}(u)\}$ à valeurs arbres en élaguant un arbre de Galton–Watson. Nous définissons également de manière analogue un processus $\{\mathcal{G}^*(u)\}$ en élaguant un arbre de Galton–Watson critique ou sous-critique conditionné à être infini. Sous de faibles hypothèses sur la loi de reproduction, nous montrons que le processus $\{\mathcal{G}(u)\}$ arrêté en son temps d'ascension admet une représentation en terme du processus $\{\mathcal{G}^*(u)\}$. Un résultat similaire a été obtenu par Aldous et Pitman (*Ann. Inst. H. Poincaré Probab. Statist.* **34** (1998) 637–686) dans le cas particulier de lois de reproductions poissoniennes en considérant un élagage uniforme sur les branches de l'arbre.

MSC: 05C05; 60J80; 60J27

Keywords: Pruning; Branching process; Galton–Watson process; Random tree; Ascension process

1. Introduction

Using percolation on the branches of a Galton–Watson tree, Aldous and Pitman constructed by time-reversal in [4] an inhomogeneous tree-valued Markov process that starts from the trivial tree consisting only of the root and ends at time 1 at the initial Galton–Watson tree. When the final Galton–Watson tree is infinite, they define the ascension time A as the first time where the tree becomes infinite. They also define another process by pruning at branches the tree conditioned on non-extinction and they show that, in the special case of Poisson offspring distribution, some connections exist between the first process up to the ascension time and the second process.

Using the same kind of ideas, continuum-tree-valued Markov processes are constructed in [2] and an analogous relation is exhibited between the process obtained by pruning the tree and the other one obtained by pruning the tree conditioned on non-extinction. However, in that continuous framework, such results hold under very general assumptions on the branching mechanism.

[1]This work is partially supported by the "Agence Nationale de la Recherche," ANR-08-BLAN-0190, and by NSFC (10525103 and 10721091).

Using the ideas of the pruning procedure [2] (which first appeared in [8] for a different purpose), we propose here to prune a Galton–Watson tree on the nodes instead of the branches so that the connections pointed out in [4] hold for any offspring distribution.

Let us first explain the pruning procedure. Given a probability distribution $p = \{p_n, n = 0, 1, \ldots\}$, let $\mathcal{G}_p$ be a Galton–Watson tree with offspring distribution p. Let $0 < u < 1$ be a constant. Then, if ν is an inner node of $\mathcal{G}_p$ that has n offsprings, we cut it (and discard all the sub-trees attached at this node) with probability u^{n-1} independently of the other nodes. The resulting tree is still a Galton–Watson tree with offspring distribution $p^{(u)}$ defined by:

$$p_n^{(u)} = u^{n-1} p_n \quad \text{for } n \geq 1 \tag{1.1}$$

and

$$p_0^{(u)} = 1 - \sum_{n=1}^{\infty} p_n^{(u)}. \tag{1.2}$$

This particular pruning is motivated by the following lemma whose proof is postponed to Section 5.

Lemma 1.1. *Let p and q be two offspring distributions. Let $\mathcal{G}_p$ and $\mathcal{G}_q$ be the associated Galton–Watson trees and let $\#\mathcal{L}_p$ and $\#\mathcal{L}_q$ denote the number of leaves of $\mathcal{G}_p$ and $\mathcal{G}_q$, respectively. Then we have that*

$$\forall N \geq 1, \quad \mathbb{P}(\mathcal{G}_p \in \cdot | \#\mathcal{L}_p = N) = \mathbb{P}(\mathcal{G}_q \in \cdot | \#\mathcal{L}_q = N) \tag{1.3}$$

if and only if

$$\exists u > 0, \forall n \geq 1 \quad q_n = u^{n-1} p_n.$$

This lemma can be viewed as the discrete analogue of Lemma 1.6 of [1] that explains the choice of the pruning parameters for the continuous case. In [3], a similar result for Poisson–Galton–Watson trees was obtained when conditioning by the total number of vertices, which explains why Poisson–Galton–Watson trees play a key role in [4].

Using the pruning at nodes procedure, given a critical offspring distribution p, we construct in Section 4 a tree-valued (inhomogeneous) Markov processes $(\mathcal{G}(u), 0 \leq u \leq \bar{u})$ for some $\bar{u} \geq 1$, such that

- the process is non-decreasing,
- for every u, $\mathcal{G}(u)$ is a Galton–Watson tree with offspring distribution $p^{(u)}$,
- the tree is critical for $u = 1$, sub-critical for $u < 1$ and super-critical for $1 < u \leq \bar{u}$.

Let us state the main properties that we prove for that process and let us compare them with the results of [4]. We write $(\mathcal{G}^{AP}(u))$ for the tree-valued Markov process defined in [4].

In Section 3, we compute the forward transition probabilities and the forward transition rates for that process and exhibit a martingale that will appear several times (see Corollary 3.4). For a tree $\mathbf{t}$, we set

$$M(u, \mathbf{t}) = \frac{(1 - \mu(u))\#\mathcal{L}(\mathbf{t})}{p_0^{(u)}}, \tag{1.4}$$

where $\#\mathcal{L}(\mathbf{t})$ denotes the number of leaves of $\mathbf{t}$ and $\mu(u)$ if the mean of offsprings in $\mathcal{G}(u)$. Then, the process

$$\big(M\big(u, \mathcal{G}(u)\big), 0 \leq u < 1\big)$$

is a martingale with respect to the filtration generated by $\mathcal{G}$. In [4], the martingale that appears (Corollary 23) for Poisson–Galton–Watson trees is $(1 - \mu(u))\#\mathcal{G}^{AP}(u)$.

When the tree $\mathcal{G}(u)$ is super-critical, it may be infinite. We define the ascension time A by:

$$A = \inf\big\{u \in [0, \bar{u}], \#\mathcal{G}(u) = \infty\big\}$$

with the convention $\inf\varnothing = \bar{u}$. We can then compute the joint law of A and $\mathcal{G}_{A-}$ (i.e. the tree just before it becomes infinite), see Proposition 4.6: we set $F(u)$ the extinction probability of $\mathcal{G}(u)$ and we have for $u \in [0, \bar{u})$

$$\mathbb{P}(A \le u) = 1 - F(u),$$
$$\mathbb{P}\big(\mathcal{G}(A-) = \mathbf{t} | A = u\big) = M(\hat{u}, \mathbf{t})\mathbb{P}\big(\mathcal{G}(\hat{u}) = \mathbf{t}\big)$$

with $\hat{u} = uF(u)$. These results are quite similar with those of Lemma 22 of [4]. They must also be compared to the continuous framework, Theorems 6.5 and 6.7 of [2].

When we have $p_0^{(\bar{u})} = 0$, then the final tree $\mathcal{G}(\bar{u})$ is a.s. infinite. We consider the tree $\mathcal{G}^*(1)$ which is distributed as the tree $\mathcal{G}(1)$ conditioned on non-extinction. From this tree, by the same pruning procedure, we construct a non-decreasing tree-valued process $(\mathcal{G}^*(u), 0 \le u \le 1)$. We then prove the following representation formula (Proposition 4.7):

$$\big(\mathcal{G}(u), 0 \le u < A\big) \stackrel{d}{=} \big(\mathcal{G}^*(u\gamma), 0 \le u < \bar{F}^{-1}(1-\gamma)\big),$$

where γ is a r.v., uniformly distributed on (0, 1), independent of $\{\mathcal{G}^*(\alpha)\colon 0 \le \alpha \le 1\}$ and $\bar{F}(u) = 1 - F(u)$. This result must also be compared to a similar result in [4], Proposition 26:

$$\big(\mathcal{G}^{AP}(u), 0 \le u < A\big) \stackrel{d}{=} \left(\mathcal{G}^{AP*}(u\gamma), 0 \le u < \frac{-\log\gamma}{(1-\gamma)}\right),$$

or to its continuous analogue, Corollary 8.2 of [2].

Let us stress again that, although the results are very similar, those in [4] only hold for Poisson–Galton–Watson trees whereas the results presented here hold for any offspring distribution.

The paper is organized as follows. In the next section, we recall some notation for trees and define the pruning procedure at nodes. In Section 3, we define the processes $\mathcal{G}$ and $\mathcal{G}^*$ and in Section 4 we state and prove the main results of the paper. Finally, we prove Lemma 1.1 in Section 5.

2. Trees and pruning

2.1. *Notation for trees*

We present the framework developed in [9] for trees, see also [7] or [4] for more notation and terminology. Introduce the set of labels

$$\mathcal{W} = \bigcup_{n=0}^{\infty} (\mathbb{N}^*)^n,$$

where $\mathbb{N}^* = \{1, 2, \dots\}$ and by convention $(\mathbb{N}^*)^0 = \{\varnothing\}$.

An element of $\mathcal{W}$ is thus a sequence $w = (w^1, \dots, w^n)$ of elements of $\mathbb{N}$, and we set $|w| = n$, so that $|w|$ represents the generation of w or the height of w. If $w = (w^1, \dots, w^m)$ and $v = (v^1, \dots, v^n)$ belong to $\mathcal{W}$, we write $wv = (w^1, \dots, w^m, v^1, \dots, v^n)$ for the concatenation of w and v. In particular $w\varnothing = \varnothing w = w$. The mapping $\pi : \mathcal{W} \setminus \{\varnothing\} \longrightarrow \mathcal{W}$ is defined by $\pi((w^1, \dots, w^n)) = (w^1, \dots, w^{n-1})$ if $n \ge 1$ and $\pi((w^1)) = \varnothing$, and we say that $\pi(w)$ is the father of w. We set $\pi^0(w) = w$ and $\pi^n(w) = \pi^{n-1}(\pi(w))$ for $1 \le n \le |w|$. In particular, $\pi^{|w|}(w) = \varnothing$.

A (finite or infinite) rooted ordered tree $\mathbf{t}$ is a subset of $\mathcal{W}$ such that

(1) $\varnothing \in \mathbf{t}$.
(2) $w \in \mathbf{t} \setminus \{\varnothing\} \Longrightarrow \pi(w) \in \mathbf{t}$.
(3) For every $w \in \mathbf{t}$, there exists a finite integer $k_w\mathbf{t} \ge 0$ such that, for every $j \in \mathbb{N}$, $wj \in \mathbf{t}$ if and only if $0 \le j \le k_w\mathbf{t}$ ($k_w\mathbf{t}$ is the number of children of $w \in \mathbf{t}$).

Let $\mathbf{T}^\infty$ denote the set of all such trees $\mathbf{t}$. Given a tree $\mathbf{t}$, we call an element in the set $\mathbf{t} \subset \mathcal{W}$ a node of $\mathbf{t}$. Denote the height of a tree $\mathbf{t}$ by $|\mathbf{t}| := \max\{|\nu|:\ \nu \in \mathbf{t}\}$. For $h \geq 0$, there exists a natural restriction map $r_h : \mathbf{T}^\infty \to \mathbf{T}^h$ such that $r_h\mathbf{t} = \{\nu \in \mathbf{t}:\ |\nu| \leq h\}$, where $\mathbf{T}^h := \{\mathbf{t} \in \mathbf{T}^\infty:\ |\mathbf{t}| \leq h\}$. In particular, $\mathbf{T}^0 = \{\varnothing\}$.

We denote by $\#\mathbf{t}$ the number of nodes of $\mathbf{t}$. Let

$$\mathbf{T} := \{\mathbf{t} \in \mathbf{T}^\infty:\ \#\mathbf{t} < \infty\}$$

be the set of all finite trees. Then $\mathbf{T} = \bigcup_{h=1}^\infty \mathbf{T}^h$.

We define the shifted subtree of $\mathbf{t}$ above ν by

$$T_\nu \mathbf{t} := \{w:\ \nu w \in \mathbf{t}\}.$$

For $n \geq 0$, let $\mathrm{gen}(n, \mathbf{t})$ be the nth generation of individuals in $\mathbf{t}$. That is

$$\mathrm{gen}(n, \mathbf{t}) := \{\nu \in \mathbf{t}:\ |\nu| = n\}.$$

We say that $w \in \mathbf{t}$ is a leaf of $\mathbf{t}$ if $k_w\mathbf{t} = 0$ and set

$$\mathcal{L}(\mathbf{t}) := \{w \in \mathbf{t} : k_w\mathbf{t} = 0\}.$$

So $\mathcal{L}(\mathbf{t})$ denotes the set of leaves of $\mathbf{t}$ and $\#\mathcal{L}(\mathbf{t})$ is the number of leaves of $\mathbf{t}$.

We say that $w \in \mathbf{t}$ is an inner node of $\mathbf{t}$ if it is not a leaf (i.e. $k_w\mathbf{t} > 0$) and we denote by $\mathbf{t}^i$ the set of inner nodes of $\mathbf{t}$ i.e.

$$\mathbf{t}^i = \mathbf{t} \setminus \mathcal{L}(\mathbf{t}).$$

Given a probability distribution $p = \{p_n, n = 0, 1, \ldots\}$ with $p_1 < 1$, following [4], call a random tree $\mathcal{G}_p$ a Galton–Watson tree with offspring distribution p if the number of children of $\varnothing$ has distribution p:

$$\mathbb{P}(k_\varnothing \mathcal{G}_p = n) = p_n \quad \forall n \geq 0$$

and for each $h = 1, 2, \ldots,$ conditionally given $r_h\mathcal{G} = \mathbf{t}^h \in \mathbf{T}^h$, for $\nu \in \mathrm{gen}(h, \mathbf{t}^h)$, $k_\nu\mathcal{G}_p$ are i.i.d. random variables distributed according to p. That means

$$\mathbb{P}(r_{h+1}\mathcal{G} = \mathbf{t} | r_h\mathcal{G} = r_h\mathbf{t}) = \prod_{\nu \in r_h\mathbf{t} \setminus r_{h-1}\mathbf{t}} p_{k_\nu \mathbf{t}}, \quad \mathbf{t} \in \mathbf{T}^{h+1},$$

where the product is over all nodes ν of $\mathbf{t}$ of height h. We have then

$$\mathbb{P}(\mathcal{G} = \mathbf{t}) = \prod_{\nu \in \mathbf{t}} p_{k_\nu \mathbf{t}}, \quad \mathbf{t} \in \mathbf{T}, \tag{2.1}$$

where the product is over all nodes ν of $\mathbf{t}$.

2.2. *Pruning at nodes*

Let $\mathcal{T}$ be a tree in $\mathbf{T}^\infty$. For $0 \leq u \leq 1$, a random tree $\mathcal{T}(u)$ is called a *node pruning* of $\mathcal{T}$ with parameter u if it is constructed as follows: conditionally given $\mathcal{T} = \mathbf{t}, \mathbf{t} \in \mathbf{T}^\infty$, for $0 \leq u \leq 1$, we consider a family of independent random variables $(\xi_\nu^u, \nu \in \mathbf{t})$ such that

$$P(\xi_\nu^u = 1) = 1 - P(\xi_\nu^u = 0) = \begin{cases} u^{k_\nu \mathbf{t} - 1}, & \text{if } k_\nu\mathbf{t} \geq 1, \\ 1, & \text{if } k_\nu\mathbf{t} = 0, \end{cases}$$

and define

$$\mathcal{T}(u) := \{\varnothing\} \cup \left\{\nu \in \mathbf{t} \setminus \{\varnothing\} : \prod_{n=1}^{|\nu|} \xi_{\pi^n(\nu)}^u = 1\right\}. \tag{2.2}$$

This means that if a node ν belongs to $\mathcal{T}(u)$ and $\xi_\nu^u = 1$, then νj, $j = 0, 1, \dots, k_\nu(\mathbf{t})$, all belong to $\mathcal{T}(u)$ and if $\xi_\nu^u = 0$, then all subsequent offsprings of ν will be removed with the subtrees attached to these nodes. Thus $\mathcal{T}(u)$ is a random tree, $\mathcal{T}(u) \subset \mathcal{T}$ and we have for every $\nu \in \mathcal{W}$,

$$\mathbb{P}\big(k_\nu \mathcal{T}(u) = n | \nu \in \mathcal{T}(u), \mathcal{T} = \mathbf{t}\big) = \begin{cases} u^{n-1} 1_{\{k_\nu \mathbf{t} = n\}}, & n \geq 1, \\ 1_{\{k_\nu \mathbf{t} = 0\}} + \big(1 - u^{k_\nu \mathbf{t} - 1}\big) 1_{\{k_\nu \mathbf{t} \geq 1\}}, & n = 0. \end{cases} \tag{2.3}$$

We also have that for $h \geq 1$ and $\mathbf{t} \in \mathbf{T}^\infty$,

$$\mathbb{P}\big(r_h \mathcal{T}(u) = r_h \mathbf{t} | \mathcal{T} = \mathbf{t}\big) = \mathbb{P}\left(\prod_{\nu \in n(h,\mathbf{t})} \xi_\nu^u = 1\right) = u^{\sum_{\nu \in n(h,\mathbf{t})} (k_\nu \mathbf{t} - 1)}, \tag{2.4}$$

where $n(h, \mathbf{t}) := \{\nu \in \mathbf{t}\colon\ k_\nu \mathbf{t} \geq 1 \text{ and } |\nu| < h\}$.

If $\mathcal{T}$ is a Galton–Watson tree, we have the following proposition.

Proposition 2.1. *If $\mathcal{T}$ is a Galton–Watson tree with offspring distribution $\{p_n, n \geq 0\}$, then $\mathcal{T}(u)$ is also a Galton–Watson tree with offspring distribution $\{p_n^{(u)}, n \geq 0\}$ defined by* (1.1) *and* (1.2).

Proof. By (2.3),

$$\mathbb{P}\big(k_\varnothing \mathcal{T}(u) = 0\big) = \mathbb{P}\big(\mathcal{T} = \{\varnothing\}\big) + \sum_{n=1}^{\infty} \big(1 - u^{n-1}\big) \mathbb{P}(k_\varnothing \mathcal{T} = n) = p_0 + \sum_{n=1}^{\infty} \big(1 - u^{n-1}\big) p_n,$$

which is equal to $p_0^{(u)}$. For $n \geq 1$,

$$\mathbb{P}\big(k_\varnothing \mathcal{T}(u) = n\big) = u^{n-1} \mathbb{P}(k_\varnothing \mathcal{T} = n) = u^{n-1} p_n.$$

The fact that $\{\xi_\nu^u\}$ are, conditionally on $\mathcal{T}$ independent random variables, gives that for each $h = 1, 2, \dots,$ conditionally given $r_h \mathcal{T}(u) = \mathbf{t}^h \in \mathbf{T}^h$, for $\nu \in \mathbf{t}^h$ with $|\nu| = h$, $k_\nu \mathcal{T}(u)$ are independent. Meanwhile, again by (2.3),

$$\begin{aligned} &\mathbb{P}\big(k_\nu \mathcal{T}(u) = n | r_h \mathcal{T}(u) = \mathbf{t}^h\big) \\ &\quad = \begin{cases} u^{n-1} \mathbb{P}(k_\nu \mathcal{T} = n) = p_n^{(u)}, & \text{if } n \geq 1, \\ \mathbb{P}(k_\nu \mathcal{T} = 0) + \sum_{k \geq 1} \big(1 - u^{k-1}\big) P(k_\nu \mathcal{T} = k) = p_0^{(u)}, & \text{if } n = 0. \end{cases} \end{aligned}$$

Then the desired result follows readily. □

3. A tree-valued Markov process

3.1. *A tree-valued process given the terminal tree*

Let $\mathcal{T}$ be a tree in $\mathbf{T}^\infty$. We want to construct a $\mathbf{T}^\infty$-valued stochastic process $\{\mathcal{T}(u)\colon\ 0 \leq u \leq 1\}$ such that

- $\mathcal{T}(1) = \mathcal{T}$,
- for every $0 \leq u_1 < u_2 \leq 1$, $\mathcal{T}(u_1)$ is a node pruning of $\mathcal{T}(u_2)$ with pruning parameter u_1/u_2.

Recall that $\mathcal{T}^i$ is the set of the inner nodes of $\mathcal{T}$. Let $(\xi_\nu, \nu \in \mathcal{T}^i)$ be a family of independent random variables such that, for every $\nu \in \mathcal{G}^i$,

$$\mathbb{P}(\xi_\nu \leq u) = u^{k_\nu \mathcal{T} - 1}.$$

Then, for every $u \in [0, 1]$, we set

$$\mathcal{T}(u) = \big\{\nu \in \mathcal{T}, \forall 1 \leq n \leq |\nu|, \xi_{\pi^n(\nu)} \leq u\big\}.$$

We call the process $(\mathcal{T}(u), 0 \le u \le 1)$ a pruning process associated with $\mathcal{T}$. Let us remark that, contrary to the process of [4], the tree $\mathcal{T}(0)$ may not be reduced to the root as the nodes with one offspring are never pruned. More precisely, if we denote by $(1)^n$ the n-uple $(1, 1, \dots, 1) \in (\mathbb{N}^*)^n$ with the convention $(1)^0 = \varnothing$, we have

$$\mathcal{T}(0) = \big\{(1)^n, n \le \sup\{k, \forall l < k, k_{(1)^l}\mathcal{G} = 1\}\big\}$$

with the convention $\sup \varnothing = 0$.

We deduce from Formula (2.4) the following proposition:

Proposition 3.1. *We have that*

$$\lim_{u \to 1} \mathcal{T}(u) = \mathcal{T} \quad a.s., \tag{3.1}$$

where the limit means that for almost every ω in the basic probability space, for each h there exists a $u(h, \omega) < 1$ such that $r_h\mathcal{T}(u, \omega) = r_h\mathcal{T}(\omega)$ for all $u(h, \omega) < u \le 1$.

3.2. *Pruning Galton–Watson trees*

Let $p = \{p_n, n = 0, 1, \dots\}$ be an offspring distribution. Let $\mathcal{G}$ be a Galton–Watson tree with offspring distribution p. Then we consider the process $(\mathcal{G}(u), 0 \le u \le 1)$ such that, conditionally on $\mathcal{G}$, the process is a pruning process associated with $\mathcal{G}$.

Then for each $u \in [0, 1]$, $\mathcal{G}(u)$ is a Galton–Watson tree with offspring distribution $p^{(u)}$. Let $g(s)$ denote the generating function of p. Then the distribution of $\mathcal{G}(u)$ is determined by the following generating function

$$g_u(s) = 1 - g(u)/u + g(us)/u, \quad 0 < u \le 1. \tag{3.2}$$

3.3. *Forward transition probabilities*

Let $\mathcal{L}(u)$ be the set of leaves of $\mathcal{G}(u)$. Fix α and β with $0 \le \alpha \le \beta \le 1$. Let us define

$$p_{\alpha,\beta}(k) = \frac{(1 - (\alpha/\beta)^{k-1})p_k^{(\beta)}}{p_0^{(\alpha)}} \quad \text{for } k \ge 1 \quad \text{and} \quad p_{\alpha,\beta}(0) = \frac{p_0^{(\beta)}}{p_0^{(\alpha)}}.$$

We define a modified Galton–Watson tree in which the size of the first generation has distribution $p_{\alpha,\beta}$, while these and all subsequent individuals have offspring distribution $p^{(\beta)}$. More precisely, let N be a random variable with law $p_{\alpha,\beta}$ and let $(\mathcal{T}_k, k \in \mathbb{N}^*)$ be a sequence of i.i.d. Galton–Watson trees with offspring distribution $p^{(\beta)}$ independent of N. Then we define the modified Galton–Watson tree $\mathcal{G}_{\alpha,\beta}$ by

$$\mathcal{G}_{\alpha,\beta} = \{\varnothing\} \cup \{k, 1 \le k \le N\} \cup \bigcup_{k=1}^{N} \{kw, w \in \mathcal{T}_k\}. \tag{3.3}$$

Let $(\mathcal{G}_{\alpha,\beta}^{\nu}, \nu \in \mathcal{L}(\alpha))$ be, conditionally given $\mathcal{G}(\alpha)$, i.i.d. modified Galton–Watson trees. We set

$$\hat{\mathcal{G}}(\beta) = \mathcal{G}(\alpha) \cup \bigcup_{\nu \in \mathcal{L}(\alpha)} \big\{\nu w \colon w \in \mathcal{G}_{\alpha,\beta}^{\nu}\big\}. \tag{3.4}$$

That is $\hat{\mathcal{G}}(\beta)$ is a random tree obtained by adding a modified Galton–Watson tree $\mathcal{G}_{\alpha,\beta}^{\nu}$ on each leaf ν of $\mathcal{G}(\alpha)$. The following proposition, which implies the Markov property of $\{\mathcal{G}(u), 0 \le u \le 1\}$, describes the transition probabilities of that tree-valued process.

Proposition 3.2. *For every* $0 \le \alpha \le \beta \le 1$, $(\mathcal{G}(\alpha), \mathcal{G}(\beta)) \stackrel{d}{=} (\mathcal{G}(\alpha), \hat{\mathcal{G}}(\beta))$.

Proof. Let $\alpha < \beta$, let $h \in \mathbb{N}^*$ and let $\mathbf{s}$ and $\mathbf{t}$ be two trees of $\mathbf{T}^h$ such that $\mathbf{s}$ can be obtained from $\mathbf{t}$ by pruning at nodes. Then, by definition of the pruning procedure, we have

$$\begin{aligned}\mathbb{P}\big(r_h\mathcal{G}(\alpha)=\mathbf{s}, r_h\mathcal{G}(\beta)=\mathbf{t}\big) &= \prod_{\nu\in r_{h-1}\mathbf{t}} p^{(\beta)}_{k_\nu\mathbf{t}} \prod_{\nu\in\mathbf{s}^i}\left(\frac{\alpha}{\beta}\right)^{k_\nu\mathbf{t}-1} \prod_{\nu\in\mathcal{L}(\mathbf{s})\setminus\mathcal{L}(\mathbf{t})}\left(1-\left(\frac{\alpha}{\beta}\right)^{k_\nu\mathbf{t}-1}\right)\\
&= \prod_{\nu\in\mathbf{s}^i} p^{(\alpha)}_{k_\nu\mathbf{t}} \prod_{\nu\in r_{h-1}\mathbf{t}\setminus\mathbf{s}^i} p^{(\beta)}_{k_\nu\mathbf{t}} \prod_{\nu\in\mathcal{L}(\mathbf{s}),|\nu|<h}\left(1-\left(\frac{\alpha}{\beta}\right)^{k_\nu\mathbf{t}-1}\mathbf{1}_{\{k_\nu\mathbf{t}>0\}}\right)\\
&= \prod_{\nu\in\mathbf{s}^i} p^{(\alpha)}_{k_\nu\mathbf{t}} \prod_{\nu\in r_{h-1}\mathbf{t}\setminus\mathbf{s}} p^{(\beta)}_{k_\nu\mathbf{t}} \prod_{\nu\in\mathcal{L}(\mathbf{s}),|\nu|<h} p^{(\beta)}_{k_\nu\mathbf{t}}\left(1-\left(\frac{\alpha}{\beta}\right)^{k_\nu\mathbf{t}-1}\mathbf{1}_{\{k_\nu\mathbf{t}>0\}}\right)\\
&= \prod_{\nu\in r_{h-1}\mathbf{s}} p^{(\alpha)}_{k_\nu\mathbf{s}} \prod_{\nu\in r_{h-1}\mathbf{t}\setminus\mathbf{s}} p^{(\beta)}_{k_\nu\mathbf{t}} \prod_{\nu\in\mathcal{L}(\mathbf{s}),|\nu|<h} \frac{p^{(\beta)}_{k_\nu\mathbf{t}}}{p^{(\alpha)}_0}\left(1-\left(\frac{\alpha}{\beta}\right)^{k_\nu\mathbf{t}-1}\mathbf{1}_{\{k_\nu\mathbf{t}>0\}}\right)\\
&= \prod_{\nu\in r_{h-1}\mathbf{s}} p^{(\alpha)}_{k_\nu\mathbf{s}} \prod_{\nu\in r_{h-1}\mathbf{t}\setminus\mathbf{s}} p^{(\beta)}_{k_\nu\mathbf{t}} \prod_{\nu\in\mathcal{L}(\mathbf{s}),|\nu|<h} p_{\alpha,\beta}(k_\nu\mathbf{t}).\end{aligned}$$

The definition of $\hat{\mathcal{G}}(\beta)$ readily implies

$$\mathbb{P}\big(r_h\mathcal{G}(\alpha)=\mathbf{s}, r_h\hat{\mathcal{G}}(\beta)=\mathbf{t}\big) = \prod_{\nu\in r_{h-1}\mathbf{s}} p^{(\alpha)}_{k_\nu\mathbf{s}} \prod_{\nu\in r_{h-1}\mathbf{t}\setminus\mathbf{s}} p^{(\beta)}_{k_\nu\mathbf{t}} \prod_{\nu\in\mathcal{L}(\mathbf{s}),|\nu|<h} p_{\alpha,\beta}(k_\nu\mathbf{t})$$

which ends the proof. □

Let $\#\mathcal{L}(u)$ denote the number of leaves of $\mathcal{G}(u)$. The latter proposition together with the description of $\hat{\mathcal{G}}$ readily imply

$$\big(\mathcal{G}(\alpha), \#\mathcal{L}(\beta)\big) \stackrel{d}{=} \left(\mathcal{G}(\alpha), \sum_{\nu\in\mathcal{L}(\alpha)} \#\mathcal{L}\big(\mathcal{G}^\nu_{\alpha,\beta}\big)\right). \tag{3.5}$$

We can also describe the forward transition rates when the trees are finite. If $\mathbf{s}$ and $\mathbf{t}$ are two trees and if $\nu \in \mathcal{L}(\mathbf{s})$, we define the tree obtained by grafting $\mathbf{t}$ on ν by

$$\mathbf{r}(\mathbf{s}, \nu; \mathbf{t}) := \mathbf{s} \cup \{\nu w, w \in \mathbf{t}\}.$$

We also define, for $s \in \mathbf{T}$, $\nu \in \mathcal{L}(\mathbf{s})$ and $k \in \mathbb{N}^*$, the set of trees

$$\mathbf{r}\big(\mathbf{s}, \nu; \mathbf{t}_k(\infty)\big) := \big\{\mathbf{r}(\mathbf{s}, \nu; \mathbf{t}), k_\varnothing\mathbf{t} = k, \#\mathbf{t} = \infty\big\}.$$

Corollary 3.3. *Let $\mathbf{s} \in \mathbf{T}$, $\mathbf{t} \in \mathbf{T}$, $\mathbf{t} \neq \{\varnothing\}$ and let $\nu \in \mathcal{L}(\mathbf{s})$. Then the transition rate at time u from $\mathbf{s}$ to $\mathbf{r}(\mathbf{s}, \nu; \mathbf{t})$ is given by*

$$q_u\big(\mathbf{s} \to \mathbf{r}(\mathbf{s}, \nu; \mathbf{t})\big) := \frac{k_\varnothing\mathbf{t} - 1}{u}\frac{\mathbb{P}(\mathcal{G}(u) = \mathbf{t})}{p^{(u)}_0}. \tag{3.6}$$

Let $\mathbf{s} \in \mathbf{T}$, $\nu \in \mathcal{L}(\mathbf{s})$ and $k \geq 1$. Then the transition rate at time u from $\mathbf{s}$ to the set $\mathbf{r}(\mathbf{s}, \nu; \mathbf{t}_k(\infty))$ is

$$q_u\big(\mathbf{s} \to \mathbf{r}\big(\mathbf{s}, \nu; \mathbf{t}_k(\infty)\big)\big) := \frac{k-1}{u}\frac{1 - F(u)^k}{p^{(u)}_0} p^{(u)}_k, \tag{3.7}$$

and no other transitions are allowed.

Proof. Let $\mathbf{s} \in \mathbf{T}$, $\mathbf{t} \in \mathbf{T}$, $\mathbf{t} \neq \{\varnothing\}$ and let $\nu \in \mathcal{L}(\mathbf{s})$. By Proposition 3.2, we have

$$\begin{aligned}\mathbb{P}\big(\mathcal{G}(u) = \mathbf{s}, \mathcal{G}(u + \mathrm{d}u) = \mathbf{r}(\mathbf{s}, \nu; \mathbf{t})\big) &= \mathbb{P}\big(\mathcal{G}(u) = \mathbf{s}, \hat{\mathcal{G}}(u + \mathrm{d}u) = \mathbf{r}(\mathbf{s}, \nu; \mathbf{t})\big) \\ &= \mathbb{P}\big(\mathcal{G}(u) = \mathbf{s}\big)\mathbb{P}\big(\mathcal{G}^{\nu}_{u,u+\mathrm{d}u} = \mathbf{t}\big) \prod_{\tilde{\nu} \in \mathcal{L}(\mathbf{s}) \setminus \{\nu\}} \mathbb{P}\big(\mathcal{G}^{\tilde{\nu}}_{u,u+\mathrm{d}u} = \{\varnothing\}\big) \\ &= \mathbb{P}\big(\mathcal{G}(u) = \mathbf{s}\big)\mathbb{P}\big(\mathcal{G}^{\nu}_{u,u+\mathrm{d}u} = \mathbf{t}\big) p_{u,u+\mathrm{d}u}(0)^{\#\mathcal{L}(\mathbf{s})-1}.\end{aligned}$$

Using (3.3), we get

$$\begin{aligned}&\mathbb{P}\big(\mathcal{G}(u + \mathrm{d}u) = \mathbf{r}(\mathbf{s}, \nu; \mathbf{t}) | \mathcal{G}(u) = \mathbf{s}\big) \\ &\quad= \mathbb{P}\big(\mathcal{G}(u + \mathrm{d}u) = \mathbf{t}\big) \frac{p_{u,u+\mathrm{d}u}(k_\varnothing \mathbf{t})}{p^{(u+\mathrm{d}u)}_{k_\varnothing \mathbf{t}}} p_{u,u+\mathrm{d}u}(0)^{\#\mathcal{L}(\mathbf{s})-1} \\ &\quad= \mathbb{P}\big(\mathcal{G}(u + \mathrm{d}u) = \mathbf{t}\big) \frac{1}{p^{(u)}_0} \left(1 - \left(\frac{u}{u + \mathrm{d}u}\right)^{k_\varnothing \mathbf{t} - 1}\right) \left(\frac{p^{(u+\mathrm{d}u)}_0}{p^{(u)}_0}\right)^{\#\mathcal{L}(\mathbf{s})-1} \\ &\quad\underset{\mathrm{d}u \to 0}{\sim} \mathbb{P}\big(\mathcal{G}(u) = \mathbf{t}\big) \frac{1}{p^{(u)}_0} (k_\varnothing \mathbf{t} - 1) \frac{\mathrm{d}u}{u} \cdot\end{aligned}$$

This gives Formula (3.6).

A similar computation replacing $\mathbb{P}(\mathcal{G}(u) = \mathbf{t})$ by $\mathbb{P}(\#\mathcal{G}(u) = \infty | k_\varnothing \mathcal{G}(u) = k) = 1 - F(u)^k$ gives Formula (3.7).

Another similar computation gives that, if $\mathbf{t}$ is obtained by grafting two trees (or more) on the leaves of $\mathbf{s}$,

$$\mathbb{P}\big(\mathcal{G}(u + \mathrm{d}u) = \mathbf{t} | \mathcal{G}(u) = \mathbf{s}\big) = \mathrm{o}(\mathrm{d}u)$$

and it is clear by construction that, in all the other cases,

$$\mathbb{P}\big(\mathcal{G}(u + \mathrm{d}u) = \mathbf{t} | \mathcal{G}(u) = \mathbf{s}\big) = 0. \qquad \square$$

Let us define

$$\mu(u) := \sum_{k=1}^{\infty} k p^{(u)}_k \tag{3.8}$$

if it exists the mean of $p^{(u)}$. We set $u_1 = \sup\{u \in [0, 1], \mu(u) \leq 1\}$. Recall the definition of function M in (1.4).

Corollary 3.4. *The process*

$$\big(M\big(u, \mathcal{G}(u)\big), 0 \leq u < u_1\big) \tag{3.9}$$

is a martingale with respect to the filtration generated by $\{\mathcal{G}(u), 0 \leq u < u_1\}$.

Proof. First, by the branching property of Galton–Watson process, for each $n \geq 1$, $0 \leq u < u_1$ and $\ell \geq n$,

$$\mathbb{P}\big(\#\mathcal{L}(u) = \ell | k_\varnothing \mathcal{G}(u) = n\big) = \mathbb{P}\left(\sum_{i=1}^{n} L_i = \ell\right),$$

where $L_1, L_2, \ldots$ are i.i.d. copies of $\#\mathcal{L}(u)$. This gives

$$\mathbb{E}\big[\#\mathcal{L}(u)\big] = p^{(u)}_0 + \mathbb{E}\big[k_\varnothing \mathcal{G}(u)\big]\mathbb{E}\big[\#\mathcal{L}(u)\big]$$

which implies

$$\mathbb{E}\big[\#\mathcal{L}(u)\big]=\frac{p_0^{(u)}}{1-\mu(u)}. \tag{3.10}$$

A straightforward computation gives that the mean of the offspring distribution $p_{\alpha,\beta}$ is

$$\mu_{\alpha,\beta}:=\frac{\mu(\beta)-\mu(\alpha)}{p_0^{(\alpha)}}.$$

By the same reasoning, (3.3) and (3.5) imply, for $0\le\alpha\le\beta<u_1$,

$$\begin{aligned}\mathbb{E}\big[\#\mathcal{L}(\beta)|\mathcal{G}(\alpha)\big]&=\#\mathcal{L}(\alpha)\mathbb{E}\big[\#\mathcal{L}\big(\mathcal{G}^{\nu}_{\alpha,\beta}\big)\big]=\#\mathcal{L}(\alpha)\big(p_{\alpha,\beta}(0)+\mu_{\alpha,\beta}\mathbb{E}\big[\#\mathcal{L}(\beta)\big]\big)\\&=\#\mathcal{L}(\alpha)\left(p_{\alpha,\beta}(0)+\mu_{\alpha,\beta}\frac{p_0^{(\beta)}}{1-\mu(\beta)}\right)\end{aligned}$$

by (3.10). Then the martingale property of (3.9) follows from a simple calculation. □

3.4. *Pruning a Galton–Watson tree conditioned on non-extinction*

Let p be a critical or sub-critical offspring distribution with mean μ such that $p_0<1$. We define the size-biased probability distribution p^* of p by

$$p_k^*=\frac{kp_k}{\mu},\quad k\ge 0.$$

Let $\mathcal{G}$ be a Galton–Watson tree with offspring distribution p. For a tree $\mathbf{t}$, we denote by $Z_n\mathbf{t}=\#\mathrm{gen}(n,\mathbf{t})$ the number of individuals in the nth generation of $\mathbf{t}$. We first recall a result in [6].

Proposition 3.5 (Kesten [6], Aldous and Pitman [4]).

(i) *The conditional distribution of $\mathcal{G}$ given $\{Z_n\mathcal{G}>0\}$ converges, as n tends to $+\infty$, toward the law of a random family tree $\mathcal{G}^\infty$ specified by*

$$\mathbb{P}\big(r_h\mathcal{G}^\infty=\mathbf{t}\big)=\mu^{-h}(Z_h\mathbf{t})\mathbb{P}(r_h\mathcal{G}=\mathbf{t})\quad\forall\mathbf{t}\in\mathbf{T}^{(h)},h\ge 0.$$

(ii) *Almost surely $\mathcal{G}^\infty$ contains a unique infinite path $(\varnothing=V_0,V_1,V_2,\dots)$ such that $\pi(V_{h+1})=V_h$ for every* $h=0,1,2,\dots$.

(iii) *The joint distribution of $(V_0,V_1,V_2,\dots)$ and $\mathcal{G}^\infty$ is determined recursively as follows: for each* $h=0,1,2,\dots$, *given $(V_0,V_1,V_2,\dots,V_h)$ and $r_h\mathcal{G}^\infty$, the numbers of children $(k_v\mathcal{G}^\infty,v\in\mathrm{gen}(h,\mathcal{G}^\infty))$ are independent with distribution p for $v\ne V_h$, and with the size-biased distribution p^* for $v=V_h$; given also the numbers of children $k_v\mathcal{G}^\infty$ for $v\in\mathrm{gen}(h,\mathcal{G}^\infty)$, the vertex V_{h+1} has uniform distribution on the set* $\{(V_h,i),1\le i\le k_{V_h}\mathcal{G}^\infty\}$.

We say that $\mathcal{G}^\infty$ is the Galton–Watson tree associated with p conditioned on non-extinction. We then define the process $(\mathcal{G}^*(u),0\le u\le 1)$ as a pruning process associated with $\mathcal{G}^\infty$. By Proposition 3.1, $\mathcal{G}^*(1-)=\mathcal{G}^*(1)=\mathcal{G}^\infty(1)$ almost surely. And since there exists a unique infinite path, we get that $\mathcal{G}^*(u)$ is finite almost surely for all $0\le u<1$.

The distribution of $\mathcal{G}^*(u)$ for fixed u is given in the following proposition. Let us recall that $\mu(u)$ is the mean of $p^{(u)}$ defined in (3.8).

Proposition 3.6. *For each* $0\le u<1$,

$$\mathbb{P}\big(\mathcal{G}^*(u)=\mathbf{t}\big)=\left(\sum_{\nu\in\mathcal{L}(\mathbf{t})}\frac{1}{\mu(1)^{|\nu|+1}}\right)\frac{\mu(1)-\mu(u)}{p_0^{(u)}}\mathbb{P}\big(\mathcal{G}(u)=\mathbf{t}\big),\quad \mathbf{t}\in\mathbf{T}. \tag{3.11}$$

Proof. We prove (3.11) inductively. First, note that

$$\mathbb{P}\big(k_\varnothing \mathcal{G}^\infty(1) = n\big) = p_n^* = np_n/\mu(1).$$

Then

$$\mathbb{P}\big(\mathcal{G}^*(u) = \{\varnothing\}\big) = \sum_{n\geq 1}\big(1-u^{n-1}\big)\mathbb{P}\big(k_\varnothing \mathcal{G}^\infty(1) = n\big) = \big(\mu(1)-\mu(u)\big)/\mu(1).$$

Since $\mathbb{P}(\mathcal{G}(u) = \{\varnothing\}) = p_0^{(u)}$, (3.11) holds for $\mathbf{t} = \{\varnothing\}$.

On the other hand, by Proposition 3.5, we have

$$\mathbb{P}\big(T_\nu \mathcal{G}^\infty(1) = \mathbf{t} | \nu = V_{|\nu|}\big) = \mathbb{P}\big(\mathcal{G}^\infty = \mathbf{t}\big)$$

and

$$\mathbb{P}\big(T_\nu \mathcal{G}^\infty(1) = \mathbf{t} | \nu \neq V_{|\nu|}\big) = \mathbb{P}(\mathcal{G} = \mathbf{t})$$

which gives

$$\mathbb{P}\big(T_\nu \mathcal{G}^*(u) = \mathbf{t} | \nu \in \mathcal{G}^*(u), \nu = V_{|\nu|}\big) = \mathbb{P}\big(\mathcal{G}^*(u) = \mathbf{t}\big) \tag{3.12}$$

and

$$\mathbb{P}\big(T_\nu \mathcal{G}^*(u) = \mathbf{t} | \nu \in \mathcal{G}^*(u), \nu \neq V_{|\nu|}\big) = \mathbb{P}\big(\mathcal{G}(u) = \mathbf{t}\big), \tag{3.13}$$

respectively. Meanwhile, since $\mathcal{G}(u)$ is a Galton–Watson tree,

$$\mathbb{P}\big(T_\nu \mathcal{G}(u) = \mathbf{t} | \nu \in \mathcal{G}(u)\big) = \mathbb{P}\big(\mathcal{G}(u) = \mathbf{t}\big),$$

which implies

$$\mathbb{P}\big(\mathcal{G}(u) = \mathbf{t}\big) = p_{k_\varnothing \mathbf{t}}^{(u)} \prod_{1\leq j\leq k_\varnothing \mathbf{t}} \mathbb{P}\big(\mathcal{G}(u) = T_{(j)}\mathbf{t}\big). \tag{3.14}$$

For some $h \geq 0$, assume that (3.11) holds for all trees in $\mathbf{T}^h$. By (3.12) and (3.13), we have for $\mathbf{t} \in \mathbf{T}^{h+1} \setminus \mathbf{T}^h$,

$$\begin{aligned}
\mathbb{P}\big(\mathcal{G}^*(u) = \mathbf{t}\big) &= \mathbb{P}\big(k_\varnothing \mathcal{G}^*(u) = k_\varnothing \mathbf{t}\big) \sum_{i=1}^{k_\varnothing \mathbf{t}} \mathbb{P}\big(\forall 1 \leq j \leq k_\varnothing \mathbf{t}, T_{(j)}\mathcal{G}^*(u) = T_{(j)}\mathbf{t} | V_1 = i\big)\mathbb{P}(V_1 = i) \\
&= u^{k_\varnothing \mathbf{t}-1} p_{k_\varnothing \mathbf{t}}^* \frac{1}{k_\varnothing \mathbf{t}} \sum_{i=1}^{k_\varnothing \mathbf{t}} \Big(\mathbb{P}\big(\mathcal{G}^*(u) = T_{(i)}\mathbf{t}\big) \cdot \prod_{j\neq i, 1\leq j\leq k_\varnothing \mathbf{t}} \mathbb{P}\big(\mathcal{G}(u) = T_{(j)}\mathbf{t}\big)\Big) \\
&= \frac{p_{k_\varnothing \mathbf{t}}^{(u)}}{\mu(1)} \sum_{i=1}^{k_\varnothing \mathbf{t}} \sum_{\nu \in \mathcal{L}(T_{(i)}\mathbf{t})} \Big(\frac{1}{\mu(1)^{|\nu|}} \frac{\mu(1)-\mu(u)}{p_0^{(u)}} \cdot \prod_{1\leq j\leq k_\varnothing \mathbf{t}} \mathbb{P}\big(\mathcal{G}(u) = T_{(j)}\mathbf{t}\big)\Big) \\
&= \Big(\sum_{\nu' \in \mathcal{L}(\mathbf{t})} \frac{1}{\mu(1)^{|\nu'|+1}}\Big) \frac{\mu(1)-\mu(u)}{p_0^{(u)}} \mathbb{P}\big(\mathcal{G}(u) = \mathbf{t}\big),
\end{aligned}$$

where the last equality follows from (3.14) and the facts

$$\bigcup_{1\leq i\leq k_\varnothing \mathbf{t}} \big\{i\nu \colon \nu \in \mathcal{L}(T_{(i)}\mathbf{t})\big\} = \mathcal{L}(\mathbf{t})$$

and $|i\nu| = |\nu| + 1$. Since $\mathbf{T} = \bigcup_{h=1}^\infty \mathbf{T}^h$, (3.11) follows inductively. □

Remark 3.7. *If $\mathcal{G}$ is a critical Galton–Watson tree ($\mu(1)=1$) with $p_1<1$, Formula (3.11) becomes*

$$\mathbb{P}\big(\mathcal{G}^*(u)=\mathbf{t}\big)=\frac{\#\mathcal{L}(\mathbf{t})(1-\mu(u))}{p_0^{(u)}}\mathbb{P}\big(\mathcal{G}(u)=\mathbf{t}\big)=M(u,\mathbf{t})\mathbb{P}\big(\mathcal{G}(u)=\mathbf{t}\big).$$

In other words, the law of $\mathcal{G}^*(u)$ is absolutely continuous with respect to the law of $\mathcal{G}(u)$ with density the martingale of Corollary 3.4.

4. The ascension process and its representation

In this section, we consider a critical offspring distribution p with $p_1<1$ and set

$$I=\left\{u\geq 0,\sum_{k=1}^{+\infty}u^{k-1}p_k\leq 1\right\}.$$

Remark 4.1. *We have $I=[0,\bar{u}]$ with $\bar{u}\geq 1$. Indeed, as the all the coefficients of the sum are non-negative, either the sum converges at the radius of convergence R and is continuous on $[0,R]$, or it tends to infinity when $u\to R$. In the latter case, there is a unique $\bar{u}<R$ for which the generating function takes the value 1 (by continuity). In the former case, either the value of the generating function at R is greater than 1 and the previous argument also applies, or the value of the generating function at R is less than 1 and $\bar{u}=R$.*

Let us give some examples:

Example 4.2 (The binary case). *We consider the critical offspring distribution p defined by $p_0=p_2=1/2$ (each individual dies out or gives birth to two children with equal probability). In that case, we have*

$$\sum_{k=1}^{+\infty}u^{k-1}p_k=\frac{1}{2}u$$

and hence we have $\bar{u}=2$, $I=[0,2]$ and $p_0^{(2)}=0$.

Example 4.3 (The geometric case). *We consider the critical offspring distribution p defined by*

$$\begin{cases} p_k=\alpha\beta^{k-1} & \text{for } k\geq 1,\\ p_0=1-\frac{\alpha}{1-\beta}. \end{cases}$$

Then, for every u, $p^{(u)}$ is still of that form. As the offspring distribution p is critical, we must have $\alpha=(1-\beta)^2$, $0<\beta<1$. In that case, we have

$$\sum_{k=1}^{+\infty}p_k u^{k-1}=\frac{(1-\beta)^2}{1-\beta u}$$

and hence we have $\bar{u}=2-\beta$, $I=[0,\bar{u}]$, $p_0^{(\bar{u})}=0$.

Example 4.4. *We consider the critical offspring distribution p defined by*

$$\begin{cases} p_k=\frac{6}{\pi^2}\frac{1}{k^3} & \text{for } k\geq 1,\\ p_0=1-\sum_{k=1}^{+\infty}p_k, \end{cases}$$

then $\bar{u}=1$ and $p_0^{(\bar{u})}=p_0>0$.

For $u \in I$, let us define

$$\begin{cases} p_k^{(u)} = u^{k-1} p_k, & k \geq 1, \\ p_0^{(u)} = 1 - \sum_{k=1}^{+\infty} p_k^{(u)}. \end{cases} \tag{4.1}$$

Then, for $u \in I$, $p^{(u)}$ is still an offspring distribution, it is sub-critical for $u < 1$, critical for $u = 1$ and super-critical for $u > 1$.

We consider a tree-valued process $(\mathcal{G}(u), u \in I)$ such that the process $(\mathcal{G}(t\bar{u}), t \in [0, 1])$ is a pruning process associated with $\mathcal{G}(\bar{u})$. Then this process $\mathcal{G}$ satisfies the following properties:

- for every $u \in I$, $\mathcal{G}(u)$ is a Galton–Watson tree with offspring distribution $p^{(u)}$,
- for every $\alpha, \beta \in I$, $\alpha < \beta$, $\mathcal{G}(\alpha)$ is a pruning of $\mathcal{G}(\beta)$.

We now consider $\{\mathcal{G}(u), u \in I\}$ as an *ascension process* with the *ascension time*

$$A := \inf\{u \in I, \#\mathcal{G}(u) = \infty\}$$

with the convention $\inf \varnothing = \bar{u}$.

The state in the ascension process at time u is $\mathcal{G}(u)$ if $0 \leq u < A$ and $\mathbf{t}(\infty)$ if $A \leq u$ where $\mathbf{t}(\infty)$ is a state representing any infinite tree. Then the ascension process is still a Markov process with countable state-space $\mathbf{T} \cup \{\mathbf{t}(\infty)\}$, where $\mathbf{t}(\infty)$ is an absorbing state.

Denote by $F(u)$ the extinction probability of a Galton–Watson process with offspring distribution $p^{(u)}$, which is the least non-negative root of the following equation with respect to s

$$s = g_u(s) = 1 - g_1(u)/u + g_1(us)/u, \tag{4.2}$$

where g_1 is the generating function associated with the offspring distribution p and g_u is the generating function associated with $p^{(u)}$.

We set

$$\bar{F}(u) = 1 - F(u). \tag{4.3}$$

Thus

$$u\bar{F}(u) + g_1\big(u - u\bar{F}(u)\big) = g_1(u). \tag{4.4}$$

The distribution of the ascension process is determined by the transition rates (3.6) and

$$q_u\big(\mathbf{s} \to \mathbf{t}(\infty)\big) = \frac{\#\mathcal{L}(\mathbf{s})}{up_0^{(u)}} \sum_{k=2}^{\infty} (k-1) p_k^{(u)} \big(1 - F(u)^k\big). \tag{4.5}$$

Define the conjugate $\hat{u}$ by

$$\hat{u} = uF(u) \quad \text{for } u \in I. \tag{4.6}$$

In particular, for $u \leq 1$, $F(u) = 1$ and consequently $\hat{u} = u$. On the contrary, for $u > 1$, Proposition 4.5 shows that $\hat{u} \leq 1$.

We can restate Eq. (4.4) into

$$g_1(\hat{u}) - g_1(u) = \hat{u} - u. \tag{4.7}$$

Notice that this equation with the condition $\hat{u} \leq 1$ characterizes $\hat{u}$.

We first prove the following result which is already well known, see for instance [5], p. 52. We just restate this property in terms of our pruning parameter.

Proposition 4.5. *For any $u \in I$, $u \geq 1$*

$$\mathbb{P}\big(\mathcal{G}(u)=\mathbf{t}\big)=F(u)\mathbb{P}\big(\mathcal{G}(\hat{u})=\mathbf{t}\big), \quad \mathbf{t}\in\mathbf{T}. \tag{4.8}$$

In other words, the law of $\mathcal{G}(u)$ conditioned to be finite is $\mathcal{G}(\hat{u})$, which explains the term *conjugate* for $\hat{u}$.

Proof of Proposition 4.5. By (2.1), we have

$$\mathbb{P}\big(\mathcal{G}(u)=\mathbf{t}\big)=\prod_{v\in\mathbf{t}\setminus\mathcal{L}(\mathbf{t})} p^{(u)}_{k_v\mathbf{t}}\cdot\prod_{v\in\mathcal{L}(\mathbf{t})} p^{(u)}_0$$

and by (4.6),

$$\mathbb{P}\big(\mathcal{G}(u)=\mathbf{t}\big)=F(u)^{-(\sum_{v\in\mathbf{t}\setminus\mathcal{L}(\mathbf{t})}(k_v\mathbf{t}-1))}\left(\frac{p^{(u)}_0}{p^{(\hat{u})}_0}\right)^{\#\mathcal{L}(\mathbf{t})}\mathbb{P}\big(\mathcal{G}(\hat{u})=\mathbf{t}\big).$$

We also have

$$p^{(\hat{u})}_0=1-\sum_{k=1}^{\infty}F(u)^{k-1}p^{(u)}_k=1+p^{(u)}_0/F(u)-g_u\big(F(u)\big)/F(u)=p^{(u)}_0/F(u). \tag{4.9}$$

Then the desired result follows from the fact that given a tree $\mathbf{t}\in\mathbf{T}$,

$$\#\mathcal{L}(\mathbf{t})=1+\sum_{v\in\mathbf{t}\setminus\mathcal{L}(\mathbf{t})}(k_v\mathbf{t}-1).$$

□

In what follows, we will often suppose that

$$p^{(\bar{u})}_0=0, \tag{4.10}$$

which is equivalent to the condition

$$\sum_{k=1}^{+\infty}\bar{u}^{k-1}p_k=1$$

and which implies that $\mathcal{G}(\bar{u})$ is infinite.

We can however give the law of A and of the tree $\mathcal{G}(A-)$ just before the ascension time in general, this is the purpose of the next proposition.

Proposition 4.6. *For $u\in[1,\bar{u})$ and $\mathbf{t}\in\mathbf{T}$,*

$$\mathbb{P}(A\leq u)=\bar{F}(u), \tag{4.11}$$

$$\mathbb{P}\big(\mathcal{G}(A-)=\mathbf{t}|A=u\big)=M(\hat{u},\mathbf{t})\mathbb{P}\big(\mathcal{G}(\hat{u})=\mathbf{t}\big), \tag{4.12}$$

$$\mathbb{P}\big(\#\mathcal{G}(A-)<+\infty|A=u\big)=1. \tag{4.13}$$

Furthermore, under assumption (4.10),

$$\mathbb{P}(A<\bar{u})=1 \tag{4.14}$$

and

$$\left(A,\frac{\hat{A}}{A}\right)=\big(A,F(A)\big)\stackrel{d}{=}\big(\bar{F}^{-1}(1-\gamma),\gamma\big), \tag{4.15}$$

where $\bar{F}^{-1}:[0,1]\to[1,\bar{u}]$ is the inverse function of $\bar{F}$ and γ is a r.v. uniformly distributed on $(0,1)$.

Proof. We have

$$\mathbb{P}(A \le u) = \mathbb{P}\big(\#\mathcal{G}(u) = \infty\big) = \bar{F}(u)$$

which gives (4.11).

By the definition of $F(u)$ in Formula (4.2) and the implicit function theorem, function F is differentiable on $(1, \bar{u})$. This gives for $u \in (1, \bar{u})$

$$\mathbb{P}(A \in \mathrm{d}u) = -F'(u)\,\mathrm{d}u$$

and differentiating (4.2) gives

$$u\big(1 - g_1'(\hat{u})\big)F'(u) = 1 - g_1'(u) - F(u)\big(1 - g_1'(\hat{u})\big). \tag{4.16}$$

By (4.5), we have for $\mathbf{t} \in \mathbf{T}$

$$\mathbb{P}\big(\mathcal{G}(A-) = \mathbf{t}, A \in \mathrm{d}u\big) = \frac{\#\mathcal{L}(\mathbf{t})\mathbb{P}(\mathcal{G}(u) = \mathbf{t})}{u p_0^{(u)}} \sum_{k=2}^{\infty} (k-1) p_k^{(u)} \big(1 - F(u)^k\big)\,\mathrm{d}u. \tag{4.17}$$

Now, using (4.8), we have

$$\mathbb{P}\big(\mathcal{G}(A-) = \mathbf{t}, A \in \mathrm{d}u\big) = \frac{\#\mathcal{L}(\mathbf{t})\mathbb{P}(\mathcal{G}(\hat{u}) = \mathbf{t})F(u)}{u p_0^{(u)}} \sum_{k=2}^{\infty} (k-1) p_k^{(u)} \big(1 - F(u)^k\big)\,\mathrm{d}u.$$

Easy computations give

$$\begin{aligned}\sum_{k=2}^{\infty} (k-1) p_k^{(u)} \big(1 - F(u)^k\big) &= g_1'(u) - \frac{g_1(u)}{u} - F(u) g_1'(\hat{u}) + \frac{g(\hat{u})}{u} \\ &= -F'(u)u\big(1 - g_1'(\hat{u})\big) + 1 - F(u) + \frac{g_1(\hat{u}) - g_1(u)}{u} \\ &= -F'(u)u\big(1 - g_1'(\hat{u})\big)\end{aligned}$$

using first Eq. (4.16) and then Eq. (4.7).

This finally gives

$$\begin{aligned}\mathbb{P}\big(\mathcal{G}(A-) = \mathbf{t}, A \in \mathrm{d}u\big) &= -\frac{\#\mathcal{L}(\mathbf{t})\mathbb{P}(\mathcal{G}(\hat{u}) = \mathbf{t})F(u)}{p_0^{(u)}} F'(u)\big(1 - g_1'(\hat{u})\big)\,\mathrm{d}u \\ &= -\frac{\#\mathcal{L}(\mathbf{t})\mathbb{P}(\mathcal{G}(\hat{u}) = \mathbf{t})}{p_0^{(\hat{u})}} F'(u)\big(1 - \mu(\hat{u})\big)\,\mathrm{d}u\end{aligned}$$

by Eq. (4.9), which yields (4.12).

Summing (4.12) over all finite trees $\mathbf{t}$ gives

$$\mathbb{P}\big(\#\mathcal{G}(A-) < +\infty | A = u\big) = \mathbb{E}\big[M\big(\hat{u}, \mathcal{G}(\hat{u})\big)\big] = 1$$

by the martingale property, which is (4.13).

Finally, (4.11) gives

$$\mathbb{P}(A = \bar{u}) = \mathbb{P}(\forall u < \bar{u}, A > u) = \lim_{u \to \bar{u}} 1 - \bar{F}(u) = F(\bar{u}-).$$

As F is non-increasing, $F(\bar{u}-)$ indeed exists. Moreover, we have by taking the limit in (4.2)

$$F(\bar{u}-) = g_{\bar{u}}\big(F(\bar{u}-)\big)$$

and, by assumption (4.10), the only fixed points of $g_{\bar{u}}$ are 0 and 1, which gives (4.14).

This together with Formula (4.11) also give

$$A \stackrel{d}{=} \bar{F}^{-1}(\gamma) \stackrel{d}{=} \bar{F}^{-1}(1-\gamma).$$

Thus $(A, \bar{F}(A)) \stackrel{d}{=} (\bar{F}^{-1}(1-\gamma), 1-\gamma)$. So we have

$$\big(A, F(A)\big) = \big(A, 1-\bar{F}(A)\big) \stackrel{d}{=} \big(\bar{F}^{-1}(1-\gamma), \gamma\big),$$

which is just (4.15). □

With Remark 3.7 and Proposition 4.6 in hand, we have the following representation of the ascension process $\{\mathcal{G}(\alpha):\ 0 \le \alpha < A\}$ under assumption (4.10).

Proposition 4.7. *Under assumption* (4.10), *we have*

$$\big\{\mathcal{G}(u), 0 \le u < A\big\} \stackrel{d}{=} \big\{\mathcal{G}^*(u\gamma):\ 0 \le u < \bar{F}^{-1}(1-\gamma)\big\}, \tag{4.18}$$

where γ is a r.v. with uniform distribution on $(0,1)$, *independent of* $\{\mathcal{G}^*(u):\ 0 \le u \le 1\}$.

Proof. Let $\{\mathcal{G}^*(u):\ 0 \le u \le 1\}$ be independent of A. Then by Remark 3.7,

$$\mathbb{P}\big(\mathcal{G}^*(\hat{A}) = \mathbf{t} | A = a\big) = \mathbb{P}\big(\mathcal{G}^*(\hat{a}) = \mathbf{t}\big) = M(\hat{a}, \mathbf{t})\mathbb{P}\big(\mathcal{G}(\hat{a}) = \mathbf{t}\big).$$

Thus it follows from (4.12) that $(A, \mathcal{G}(A-)) \stackrel{(d)}{=} (A, \mathcal{G}^*(\hat{A}))$. On the other hand, by the definition of node-pruning, for every $\mathbf{t} \in \mathbf{T}$, $0 \le \alpha < 1$ and $0 \le \beta < \bar{u}$,

$$\mathbb{P}\big(\big(\mathcal{G}(s\beta), 0 \le s \le 1\big) \in \cdot | \mathcal{G}(\beta) = \mathbf{t}\big) = \mathbb{P}\big(\big(\mathcal{G}^*(s\alpha), 0 \le s \le 1\big) \in \cdot | \mathcal{G}^*(\alpha) = \mathbf{t}\big).$$

Thus conditioning on the terminal value implies

$$\big\{\mathcal{G}(u), 0 \le u < A\big\} \stackrel{d}{=} \big\{\mathcal{G}^*(\hat{A}u/A):\ 0 \le u < A\big\}.$$

Then (4.18) follows from (4.15). □

Example 4.8 (Binary case). *If $p = \frac{1}{2}\delta_2 + \frac{1}{2}\delta_0$, then $\bar{u} = 2$ and $\mathcal{G}(u)$ is a Galton–Watson tree with binary offspring distribution $\frac{u}{2}\delta_2 + (1-\frac{u}{2})\delta_0$ for $0 \le u \le 2$. In this case, we have*

$$F(u) = \frac{2}{u} - 1$$

and the ascension time A is distributed as

$$\bar{F}^{-1}(1-\gamma) = \frac{2}{1+\gamma},$$

where γ is a uniform random variable on $(0,1)$. A's density is given by

$$f(t) = -F'(t) = \frac{2}{t^2}\mathbf{1}_{[1,2]}(t).$$

Example 4.9 (Geometric case). *We suppose that the critical offspring distribution p is of the form*

$$p_k = (1-\beta)^2\beta^{k-1} \quad \textit{for } k \ge 1, \qquad p_0 = \beta.$$

In that case, we have

$$\begin{cases} p_k^{(u)} = (1-\beta)^2(u\beta)^{k-1} & \text{for } k \geq 1, \\ p_0^{(u)} = 1 - \frac{(1-\beta)^2}{1-u\beta}, \end{cases}$$

$\bar{u} = 2 - \beta$, *and assumption* (4.10) *is statisfied.*

We then get

$$F(u) = \frac{2-u-\beta}{1-u\beta}\frac{1}{u}$$

and the ascension time A *has density*

$$\left(\frac{2-\beta}{u^2} + \frac{(1-\beta)^2\beta}{(1-u\beta)^2}\right)\mathbf{1}_{[1,2-\beta]}(u).$$

5. Proof of Lemma 1.1

Let $\mathcal{G}_p$ be a Galton–Watson tree with offspring distribution p such that $p_1 < 1$.

If $\mathbf{t}$ is a tree, we denote by $(a_1, a_2, \dots, a_m)$ the numbers of offsprings of the inner nodes. Its number of leaves is then

$$a_1 + \cdots + a_m - m + 1.$$

If $\mathbf{t}$ is a tree with n leaves, we have

$$\mathbb{P}(\mathcal{G}_p = \mathbf{t}) = p_{a_1} \cdots p_{a_m} p_0^n$$

and therefore

$$\mathbb{P}(\#\mathcal{L}_p = n) = C_p(n) p_0^n$$

with

$$C_p(n) = \sum_{\mathbf{t}, \#\mathcal{L}(\mathbf{t})=n} p_{a_1} \cdots p_{a_m}.$$

Then we have, for every n such that $C_p(n) \neq 0$,

$$\mathbb{P}(\mathcal{G}_p = \mathbf{t} | \#\mathcal{L}_p = n) = \mathbb{P}(\mathcal{G}_q = \mathbf{t} | \#\mathcal{L}_q = n) \quad \iff \quad \frac{p_{a_1} \cdots p_{a_m}}{C_p(n)} = \frac{q_{a_1} \cdots q_{a_m}}{C_q(n)}. \tag{5.1}$$

First, let us suppose that

$$\mathbb{P}(\mathcal{G}_p = \mathbf{t} | \#\mathcal{L}_p = n) = \mathbb{P}(\mathcal{G}_q = \mathbf{t} | \#\mathcal{L}_q = n).$$

For $n = 1$, all the trees with one leaf are those with one offspring at each generation until the last individual dies. Therefore,

$$\begin{aligned} \mathbb{P}(\mathcal{G}_p = \mathbf{t} | \#\mathcal{L}_p = 1) = \mathbb{P}(\mathcal{G}_q = \mathbf{t} | \#\mathcal{L}_q = 1) \quad &\iff \quad \forall k \geq 0, \quad p_1^k(1-p_1) = q_1^k(1-q_1) \\ &\iff \quad p_1 = q_1. \end{aligned}$$

We set $n_0 = \inf\{n \geq 2, p_n > 0\}$. We then set $u = (q_{n_0}/p_{n_0})^{1/(n_0-1)}$.

If the only non-zero terms of p are p_0, p_1 and p_{n_0}, the relation

$$q_n = u^{n-1} p_n$$

is trivally true for every $n \geq 1$.

In the other cases, let $n > n_0$ such that $p_n > 0$ and let N be the integer defined by:

$$N = 2(n-1)(n_0-1).$$

Let us consider first a tree $\mathbf{t}$ that has $N+1$ leaves, $n-1$ inner nodes with n_0 offsprings and n_0-1 inner nodes with n offsprings. Applying (5.1) to that tree gives

$$\frac{p_{n_0}^{n-1} p_n^{n_0-1}}{C_p(N+1)} = \frac{q_{n_0}^{n-1} q_n^{n_0-1}}{C_q(N+1)}.$$

Then, let us consider another tree with $N+1$ leaves composed of $2(n-1)$ inner nodes with n_0 offsprings. For that new tree, (5.1) gives

$$\frac{p_{n_0}^{2(n-1)}}{C_p(N+1)} = \frac{q_{n_0}^{2(n-1)}}{C_q(N+1)}.$$

Dividing the two latter equations gives

$$q_n = u^{n-1} p_n.$$

It remains to remark that this identity also holds when $n = n_0$ and when $p_n = 0$.

Conversely, let us suppose that $q_n = u^{n-1} p_n$ for every $n \geq 1$. Let n such that $C_p(n) \neq 0$. Then, for every $\mathbf{t}$ with n leaves, we have

$$\begin{aligned} q_{a_1} \cdots q_{a_m} &= u^{a_1-1} p_{a_1} \cdots u^{a_m-1} p_{a_m} \\ &= u^{a_1+\cdots+a_m-m} p_{a_1} \cdots p_{a_m} \\ &= u^{n-1} p_{a_1} \cdots p_{a_m}. \end{aligned}$$

We then have $C_q(n) = u^{n-1} C_p(n)$ and

$$\frac{q_{a_1} \cdots q_{a_m}}{C_q(n)} = \frac{u^{n-1} p_{a_1} \cdots p_{a_m}}{u^{n-1} C_p(n)} = \frac{p_{a_1} \cdots p_{a_m}}{C_p(n)},$$

that is

$$\mathbb{P}(\mathcal{G}_p = \mathbf{t} | \#\mathcal{L}_p = n) = \mathbb{P}(\mathcal{G}_q = \mathbf{t} | \#\mathcal{L}_q = n).$$

Acknowlegments

The authors wish to thank an anonymous referee for his remarks that improved significantly the presentation of the paper.

Hui He also wants to thank the laboratory MAPMO for his pleasant stay at Orléans where this work was done.

References

[1] R. Abraham and J.-F. Delmas. Fragmentation associated with Lévy processes using snake. *Probab. Theory Related Fields* **141** (2008) 113–154. MR2372967

[2] R. Abraham and J.-F. Delmas. A continuum-tree-valued Markov process. *Ann. Probab.* **40** (2012) 1167–1211.
[3] D. Aldous. The continuum random tree I. *Ann. Probab.* **19** (1991) 1–28. MR1085326
[4] D. Aldous and J. Pitman. Tree-valued Markov chains derived from Galton–Watson processes. *Ann. Inst. H. Poincaré Probab. Statist.* **34** (1998) 637–686. MR1641670
[5] K. B. Athreya and P. E. Ney. *Branching Processes*. *Die Grundlehren der Mathematischen Wissenschaften* **196**. Springer, New York, 1972. MR0373040
[6] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.* **22** (1987) 425–487. MR0871905
[7] J.-F. Le Gall. Random trees and applications. *Probab. Surv.* **2** (2005) 245–311. MR2203728
[8] G. Miermont. Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields* **131** (2005) 341–375. MR2123249
[9] J. Neveu. Arbres et processus de Galton–Watson. *Ann. Inst. H. Poincaré Probab. Statist.* **22** (1986) 199–207. MR0850756