ALEA, Lat. Am. J. Probab. Math. Stat. 12 (1), 451–476 (2015)

# $\beta$-coalescents and stable Galton-Watson trees

**Romain Abraham and Jean-François Delmas**

Laboratoire MAPMO, CNRS, UMR 7349, Fédération Denis Poisson, FR 2964,
Université d'Orléans,
B.P. 6759,
45067 Orléans cedex 2, France.

Université Paris-Est, École des Ponts, CERMICS,
6-8 av. Blaise Pascal,
Champs-sur-Marne, 77455 Marne La Vallée, France.
*E-mail address*: romain.abraham@univ-orleans.fr, delmas@cermics.enpc.fr

**Abstract.** Representation of coalescent processes using pruning of trees has been used by Goldschmidt and Martin for the Bolthausen-Sznitman coalescent and by Abraham and Delmas for the $\beta(3/2, 1/2)$-coalescent. By considering a pruning procedure on stable Galton-Watson tree with $n$ labeled leaves, we give a representation of the discrete $\beta(1+\alpha, 1-\alpha)$-coalescent, with $\alpha \in [1/2, 1)$ starting from the trivial partition of the $n$ first integers. The construction can also be made directly on the stable continuum Lévy tree, with parameter $1/\alpha$, simultaneously for all $n$. This representation allows to use results on the asymptotic number of coalescence events to get the asymptotic number of cuts in stable Galton-Watson tree (with infinite variance for the offspring distribution) needed to isolate the root. Using convergence of the stable Galton-Watson tree conditioned to have infinitely many leaves, one can get the asymptotic distribution of blocks in the last coalescence event in the $\beta(1+\alpha, 1-\alpha)$-coalescent.

## 1. Introduction

1.1. *Framework.* The idea of constructing coalescent processes by pruning discrete trees arises first in Goldschmidt and Martin (2005) where the Bolthausen-Sznitman coalescent is constructed by a uniform pruning of the branches of a random recursive tree, see also Schweinsberg (2012) and Freund and Siri-Jégousse (2014) for applications of such a representation. The same kind of ideas has been used in Abraham and Delmas (2013a) to construct a $\beta(3/1, 1/2)$-coalescent process using the pruning at node of a uniform random binary tree. This construction is also

*Received by the editors January 7, 2015; accepted June 10, 2015.*
2010 *Mathematics Subject Classification.* 60J80, 60G18.
*Key words and phrases.* Galton-Watson trees, stable offspring distribution, coalescent processes.

closely related to Aldous's continuum random tree. The goal of this paper is to extend this result by applying a pruning at nodes (introduced in Abraham and Delmas (2008) in a continuous setting and in Abraham et al. (2012) in a discrete setting) to a stable Lévy tree, obtaining a $\beta(1+\alpha, 1-\alpha)$-coalescent process, with $1/2 \leq \alpha < 1$.

Let $\Lambda$ be a finite measure on $[0,1]$. A $\Lambda$-coalescent $(\Pi(t), t \geq 0)$ is a Markov process which takes values in the set of partitions of $\mathbb{N}^* = \{1, 2, \ldots\}$ introduced in Pitman (1999) for coalescent processes with possible multiple collisions. It is defined via the transition rates of its restriction $\Pi^{[n]} = (\Pi^{[n]}(t), t \geq 0)$ to the $n$ first integers: if $\Pi^{[n]}(t)$ is composed of $b$ blocks, then $k$ $(2 \leq k \leq b)$ fixed blocks coalesce at rate:

$$\lambda_{b,k} = \int_0^1 u^{k-2}(1-u)^{b-k}\Lambda(du). \tag{1.1}$$

In particular a coalescence event happens at rate:

$$\lambda_b = \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}. \tag{1.2}$$

We take the convention $\lambda_1 = 0$. We also define the discrete process $\Pi^{[n]}_{\text{dis}} = (\Pi^{[n]}_{\text{dis}}(k), k \in \mathbb{N})$ as the different successive states of the process $\Pi^{[n]}$ until it reaches the absorbing state (which is the trivial partition consisting in one block) and afterward the discrete process remains constant.

As examples of $\Lambda$-coalescents, let us mention:

- the Kingman's coalescent with $\Lambda(dx) = \delta_0(dx)$, see Kingman (1982),
- the Bolthausen-Sznitman coalescent with $\Lambda(dx) = \mathbf{1}_{(0,1)}(x)dx$, see Bolthausen and Sznitman (1998),
- the $\beta$-coalescents where $\Lambda(dx)$ is (up to a multiplicative constant) the $\beta(a,b)$ distribution. In the case of the $\beta(1+\alpha, 1-\alpha)$-coalescent, that is $\Lambda(dx) = (x/(1-x))^\alpha \, dx$, see Birkner et al. (2005); Berestycki et al. (2007) for $-1 < \alpha < 0$. The case $\alpha = 0$ corresponds to the Bolthausen-Sznitman coalescent, while the limit case $\alpha = -1$ formally corresponds to the Kingman's coalescent. For the $\beta(1+\alpha, -\alpha)$-coalescent, with $-1 < \alpha < 0$ see Foucart and Hénard (2013).

We refer to the survey Berestycki (2009) for further results on coalescent processes.

Let $\alpha \in [1/2, 1)$. We consider a critical Galton-Watson (GW) tree $T$ with offspring distribution characterized by its generating function for $r \in [0,1]$:

$$g(r) = r + \alpha(1-r)^{1/\alpha}. \tag{1.3}$$

This GW tree arises as the shape of the sub-tree of a stable Lévy tree with index $\gamma = 1/\alpha$ generated by leaves chosen in a Poissonian manner, see Duquesne and Le Gall (2002), Theorem 3.2.1. We shall call these random trees the stable GW trees with parameter $\gamma$. We denote by $\mathbf{P}$ the distribution of $T$. If $x$ is a node of $T$ we denote by $k_x(T)$ the number of offsprings of $x$. If $k_x(T) = 0$ (resp. $k_x(T) > 0$), then $x$ is called a leaf (resp. an internal node) of $T$. We denote by $L(T)$ the number of leaves of the tree $T$. Since $g'(0) = 0$, we get that a.s. $k_x(T) \neq 1$ for all $x \in T$. We denote by $\mathbf{P}_n$ the law of $T$ conditioned to have exactly $n$ leaves. Under $\mathbf{P}_n$, we label the leaves of $T$ from 1 to $n$ uniformly at random, independently of $T$, and then we consider the following pruning procedure which is derived from Abraham et al. (2015), see Section 2.2. Choose an internal node $x_1$ (which has at least 2 children) at random with probability:

$$\frac{k_{x_1}(T)-1}{L(T)-1}.$$

This internal node separates the tree into two subtrees: the fringe sub-tree $T_{x_1}$ rooted at $x_1$ that consists of all nodes of $T$ that have $x_1$ on their lineage to the root (including $x_1$), and the set $T \setminus T_{x_1}$ which is still a tree. We set $T_{(1)} = (T \setminus T_{x_1}) \cup \{x_1\}$ which is the new tree we work with. All the leaves of $T_{(1)}$ except $x_1$ are leaves of $T$ and they keep their label. Notice that $x_1$ is a new leaf of $T_{(1)}$ and we label it by the block (i.e. the sequence) of labels of the leaves of $T_{x_1}$. We then iterate the procedure on the tree $T_{(1)}$ and so on until the root is chosen (see Figure 1.1).

This pruning procedure defines a discrete time process $\Pi_{\mathrm{GW}}^{[n]} = (\Pi_{\mathrm{GW}}^{[n]}(k), k \in \mathbb{N})$ taking values in the set of partitions of the $n$ first integers, $\Pi_{\mathrm{GW}}^{[n]}(k)$ being the set of labels of the leaves of the tree $T_{(k)}$ obtained after the $k$-th cut.

1.2. *Main result.* The process $\Pi_{\mathrm{GW}}^{[n]}$ is then a coalescent process starting from the trivial partition consisting of singletons and blocks merge together as time goes by. Its law is given in the next theorem.

**Theorem 1.1.** *We set $\alpha = \frac{1}{\gamma} \in [1/2, 1)$. The process $\Pi_{GW}^{[n]}$ is distributed under $\mathbf{P}_n$ as $\Pi_{dis}^{[n]}$ for the $\beta(1+\alpha, 1-\alpha)$-coalescent with coalescent measure:*

$$\Lambda(dx) = \left(\frac{x}{1-x}\right)^{\alpha} dx. \tag{1.4}$$

*Remark* 1.2. Notice that the process $\Pi_{\mathrm{dis}}^{[n]}$ is discrete in time and thus characterizes the coalescent measure up to a multiplicative constant. It is possible to construct the continuous-time coalescent process $\Pi^{[n]}$ associated with the measure $\Lambda$ given by Equation (1.4) from the process $\Pi_{\mathrm{GW}}^{[n]}$ by adding exponential times between the successive states of this process. More precisely, recall the definitions of the transitions rates $\lambda_{b,k}$ of Equation (1.1) and of the jump rates $\lambda_b$ of Equation (1.2). Let $(\tau_k)_{k\in\mathbb{N}}$ be a sequence of independent random variables such that, conditionally given the process $\Pi_{\mathrm{GW}}^{[n]}$, the random variable $\tau_k$ is exponentially distributed with parameter $\lambda_{\ell_k}$ where $\ell_k$ is the number of blocks of the partition $\Pi_{\mathrm{GW}}^{[n]}(k)$, with the convention that $\tau_k = +\infty$ if $\ell_k = 1$. Then we set

$$\tilde{\Pi}^{[n]}(t) = \Pi_{\mathrm{GW}}^{[n]}(k) \quad \text{if } \sum_{i=0}^{k-1} \tau_i \le t < \sum_{i=0}^{k} \tau_i.$$

As a direct consequence of Theorem 1.1 and the definition of a $\Lambda$-coalescent, we get that the processes $\Pi^{[n]}$ and $\tilde{\Pi}^{[n]}$ have the same distribution.

One major drawback of this construction is that we define the process for fixed $n$ and not simultaneously for all $n$. However, as in Abraham and Delmas (2013a), we can construct directly the process $(\Pi(\theta), \theta \ge 0)$ taking values in the set of partitions of the integers using the pruning of a Lévy continuum random tree. More precisely, we consider the weighted stable Lévy tree $(\mathcal{T}, d, \mathbf{m}^{\mathcal{T}})$ associated with the branching mechanism $\psi(\lambda) = \lambda^{\gamma}$ for $\gamma \in (1, 2)$ (the case $\gamma = 2$ is studied in Abraham and

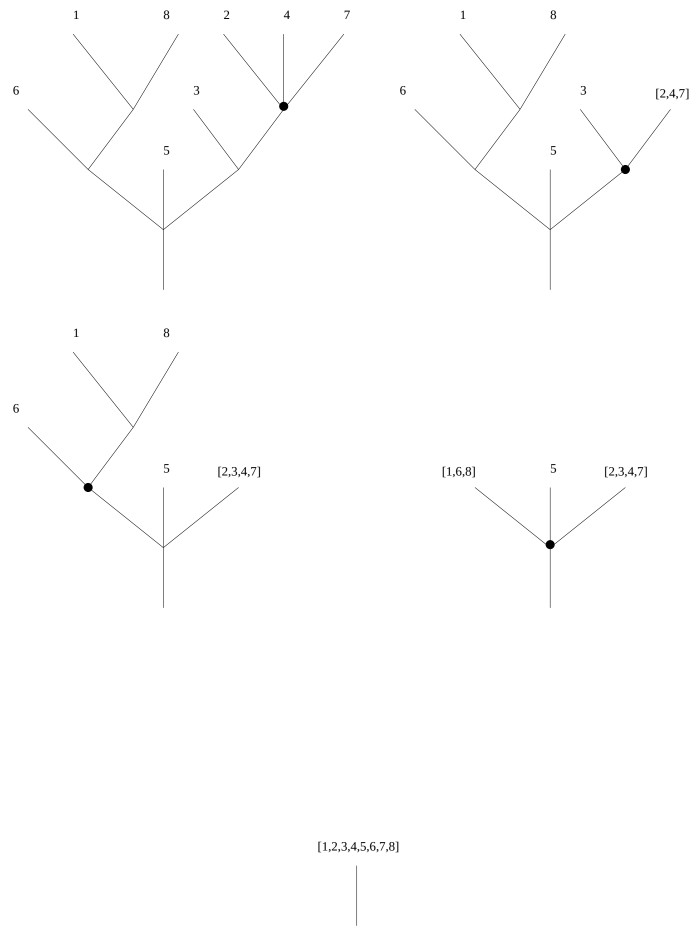


Figure 1.1. The pruning at node of a given tree. The bold internal node corresponds to the next chosen node.

Delmas (2013a) and requires a different pruning). We recall that $\mathcal{T}$ is a real tree and that $\mathbf{m}^{\mathcal{T}}$ corresponds to a uniform measure on the leaves of $\mathcal{T}$, see Duquesne and Le Gall (2002, 2005) and also Abraham et al. (2014) more specifically for the space of weighted real trees. We work under the so-called normalized excursion measure $\mathbb{N}^{(1)}$ under which $\mathbf{m}^{\mathcal{T}}$ is a probability measure. We consider given $\mathcal{T}$ the pruning defined in Abraham and Delmas (2008): to each branching point $x$ of $\mathcal{T}$ we can associate a "mass" $\Delta_x$ of this node, which intuitively represents the size of its progeny, and a random variable $E_x$ which is exponentially distributed with parameter $\Delta_x$. This random variable represents the time at which the node $x$ is cut. When we cut such a node, we remove the sub-tree above it. Let $\mathcal{T}_\theta$ denote the continuum random sub-tree obtained at time $\theta \geq 0$. We define a partition-valued process using the usual paintbox procedure. Let $(U_i, i \in \mathbb{N}^*)$ be independent random variables with distribution $\mathbf{m}^{\mathcal{T}}$ under $\mathbb{N}^{(1)}$. We define a partition of $\mathbb{N}^*$ at time $\theta$, $\Pi_{\text{Lévy}}(\theta)$ by saying that two integers $i$ and $j$ belong to the same block of $\Pi_{\text{Lévy}}(\theta)$ if and only if the random variables $U_i$ and $U_j$ have a leaf of $\mathcal{T}_\theta$ as a common ancestor. Intuitively this means that $U_i$ and $U_j$ belong to the same sub-tree attached above $\mathcal{T}_\theta$. This defines a coalescent process $\Pi_{\text{Lévy}} = (\Pi_{\text{Lévy}}(\theta), \theta \geq 0)$. We are now interested in its discrete (in time) restriction to the $n$ first integers. Let $\Pi^{[n]}_{\text{Lévy}} = (\Pi^{[n]}_{\text{Lévy}}(k), k \in \mathbb{N})$ be the discrete process associated with $\Pi_{\text{Lévy}}$ restricted to the $n$ first integers until it reaches the absorbing state (which is the trivial partition consisting in one block) and which afterward remains constant.

By construction, and thanks to Theorem 3.2.1 in Duquesne and Le Gall (2002), we can deduce that under $\mathbb{N}^{(1)}$, the discrete coalescent process $\Pi^{[n]}_{\text{Lévy}}$ is distributed as $\Pi^{[n]}_{\text{GW}}$ under $\mathbf{P}_n$. In fact, we have the following stronger result.

**Theorem 1.3.** *We set $\alpha = \frac{1}{\gamma} \in (1/2, 1)$. Under $\mathbb{N}^{(1)}$, the processes $(\Pi^{[n]}_{Lévy}, n \in \mathbb{N}^*)$ associated with the Lévy tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ are distributed as $(\Pi^{[n]}_{dis}, n \in \mathbb{N}^*)$ associated with the Lévy measure $\Lambda(dx) = (x/1-x)^\alpha \, dx$.*

*Remark* 1.4. Although the process $\Pi_{\text{Lévy}}$ is a continuous-time process like $\Pi_{\text{GW}}$, it is not a coalescent process under $\mathbb{N}^{(1)}$ as for instance the time of the first coalescence event in $\Pi^{[n]}_{\text{Lévy}}$ is not exponentially distributed, see Corollary 4.5.

We conjecture that there exists a random time-change $(R(t), t \geq 0)$ such that the process $(\Pi_{\text{Lévy}}(R(t)), t \geq 0)$ is indeed under $\mathbb{N}^{(1)}$ a $\beta(1+\alpha, 1-\alpha)$-coalescent, but we have no guess on what this time change could be.

*Remark* 1.5. Let us remark that the $\beta(1+\alpha, 1-\alpha)$-coalescent we obtain is also a $\beta(2-a, a)$-coalescent (with $a = 1-\alpha$) as in Berestycki et al. (2007) but with a different range for $a$. The difference between the two cases is that in Berestycki et al. (2007) $\alpha \in (-1, 0)$ and the coalescent process comes down from infinity (i.e. for every positive time $\theta$, the partition $\Pi(\theta)$ contains only a finite number of blocks) whereas in our case $\alpha \in (1/2, 1)$ the process always contains an infinite number of singletons (also called "dust").

*Remark* 1.6. Let us remark that the pruning procedure described above is the same as in Miermont (2005) used to construct Miermont's self-similar fragmentation process (see also Abraham and Delmas (2008)). However, the time reversal of the process $\Pi_{\text{Lévy}}$ is not Miermont's fragmentation as once a sub-tree is cut and discarded, it is no more considered in our construction whereas it undergoes some other fragmentations in Miermont's construction. There are still some strong connections. For instance, the tree $\mathcal{T}_\theta$ is linked with a tagged fragment in the fragmentation, see Abraham and Delmas (2008) Theorem 1.5 and Proposition 1.7 for the distribution of the tree $\mathcal{T}_\theta$ and for the distribution of a tagged fragment in Miermont's fragmentation.

1.3. *Number of cuts needed to isolate the root in a stable GW tree.* Using the above link between Galton-Watson trees and $\beta$-coalescents, known results in one field translate immediately in the other field giving sometimes new results. In that direction, we first focus on how known asymptotics on the number of coalescence events yield new results on the number of cuts needed to isolate the root in a stable GW tree with $n$ leaves.

The original problem of cutting randomly a rooted tree arises first in Meir and Moon Meir and Moon (1970). Given a rooted tree $T_n$ with $n$ edges, select an edge uniformly at random (notice that this is not exactly our pruning procedure) and delete the subtree not containing the root attached to this edge. On the remaining tree, iterate this procedure until only the edge attached to the root is left. We denote by $\tilde{Z}_n$ the number of edge-removals needed to isolate the root. The problem is then to study asymptotics of this random number $\tilde{Z}_n$, depending on the law of the initial tree $T_n$.

In the original paper Meir and Moon (1970), Meir and Moon considered Cayley trees and obtained asymptotics for the first two moments of $X_n$. Limits in distribution were then obtained, see for instance Panholzer (2006) for some simply generated trees, Drmota et al. (2009) for random recursive trees, Holmgren (2010) for binary search trees, Bertoin (2012) for Cayley trees. In Janson (2006), Janson focuses on conditioned Galton-Watson trees associated with critical offspring distributions with finite variance and proves that

$$\tilde{Z}_n/\sqrt{n} \xrightarrow[n\to+\infty]{(d)} \tilde{Z},$$

where the random variable $\tilde{Z}$ has Rayleigh distribution with density $x\,\mathrm{e}^{-x^2/2}\,\mathbf{1}_{\{x>0\}}$, and can be explicitly constructed using a pruning procedure on the Brownian continuum random tree (which corresponds to the cases $\gamma = 2$ in our setting), see Abraham and Delmas (2013c). In particular $\tilde{Z}$ is distributed as the height of a random leaf of the Brownian continuum random tree. See also Addario-Berry et al. (2014); Bertoin and Miermont (2013) for further work on cutting randomly rooted trees.

Notice that the reproduction law for stable GW trees has an infinite variance for $\alpha \in (1/2, 1)$, and the uniform pruning does not seem to be adapted to isolate the root. For this reason, we consider the pruning procedure developed in Section 1.1 to tackle the infinite variance case. So, let $Z_n$ be the number of cuts, using this procedure, needed to isolate the root of a stable GW tree:

$$Z_n = \inf\{k;\ \Pi^{[n]}_{\mathrm{GW}}(k) = \{\{1, \dots, n\}\}\}.$$

Notice that for $r$-ary trees, since all the internal nodes have the same degree, the cutting procedure given in Section 1.1 corresponds to choose an internal node uniformly.

We immediately deduce from asymptotics of the number of coalescence events in $\beta$-coalescents (see Corollary 1 Haas and Miermont (2011), see also Gnedin et al. (2014), Table 1 for a summary of all the results concerning $\beta$-coalescents), the following result which extends part of the result in Janson (2006) to GW tree with infinite variance of the reproduction law.

**Corollary 1.7.** *Let $\alpha = 1/\gamma \in [1/2, 1)$. We have the following convergence in distribution:*

$$n^{\alpha-1} Z_n \xrightarrow[n\to+\infty]{(d)} Z,$$

*with the distribution of $Z$ characterized by, for $n \in \mathbb{N}^*$:*

$$\mathbb{E}\left[Z^n\right] = \alpha^n \frac{\Gamma(n+1)\Gamma(1-\alpha)}{\Gamma((n+1)(1-\alpha))}.$$

Let us insist on the fact that this corollary does not need any proof as this is just a translation of known results on $\beta$-coalescents using our links with GW trees, only the moment computation needs some explanations and is done in Section 5, The distribution of $Z$ corresponds to the expected limit distribution in the Conjecture that is stated at the end of the introduction in Abraham and Delmas (2013b) for the number of cuts needed to isolate the root in general GW trees. (Notice that in the conjecture, one choose an internal node $x \in T$ with probability proportional to $k_x(T)$ whereas in Section 1.1 one choose an internal node $x \in T$ with probability proportional to $k_x(T)-1$.) In particular, $Z$ is distributed as the height of a random leaf of the normalized Lévy tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$.

1.4. *Number of blocks in the last coalescence event.* On the other hand, we can use results on GW trees conditioned to have an infinite number of leaves (which is very close to Kesten's result on GW tree conditionally on the non extinction, see Curien and Kortchemski (2014) Theorem 3.1 or Abraham and Delmas (2014) Proposition 4.6) to get asymptotics on the number $B_n$ of blocks involved in the last coalescence event of $\Pi^{[n]}$.

The proof of the following Proposition is given in Section 6.

**Proposition 1.8.** *Let $\alpha = 1/\gamma \in [1/2, 1)$. We have the following convergence in distribution:*

$$B_n \xrightarrow[n\to+\infty]{(d)} B,$$

*with the distribution of $B$ given by its generating function $\varphi_\alpha(r) = \mathbb{E}\left[r^B\right]$, with for $r \in [0,1]$:*

$$\varphi_\alpha(r) = (1-\alpha)r \int_0^1 \frac{dx}{1-(1-x)^\alpha} \left( \frac{1}{(1-rx)^\alpha} - 1 \right). \tag{1.5}$$

See also Abraham and Delmas (2013a) for more results in this direction when $\alpha = 1/2$ including the number of singletons involved in the last coalescence event as well as a closed form for $\varphi_{1/2}$.

*Remark* 1.9. After we first posted this paper on arXiv, Hénard proved in Henard (2015) Theorem 3.5 that Equation (1.5) remains valid for all $\beta(1+\alpha, 1-\alpha)$-coalescents with $\alpha \in (-1, 1)$ (taking the limit when $\alpha = 0$).

For $\alpha = 0$, the $\beta(1+\alpha, 1-\alpha)$-coalescent corresponds to the Bolthausen-Sznitman coalescent, and thus $\varphi_0$ is the generating function of the asymptotic number of blocks of the last coalescence event in the Bolthausen-Sznitman coalescent whose distribution is given in Theorem 3.1 and Proposition 3.2 of Goldschmidt and Martin (2005).

As $\alpha$ goes down to $-1$, we recover the Kingman's coalescent as a limit. We also get $\varphi_{-1}(r) = r^2$ and notice that $\varphi_{-1}$ is trivially the generating function of the number of blocks of the last (in fact all) coalescence event in the Kingman's coalescent, as all the coalescence events are binary.

1.5. *Organization of the paper.* Section 2 gives a representation of the pruning at node procedure for GW tree in continuous time motivated by Abraham et al. (2015). This procedure corresponds in fact to the one presented in Introduction, Section 1.1. Section 3 is devoted to the proof of Theorem 1.1. Section 4 devoted to the proof of Theorem 1.3 is more technical as it relies on continuum random Lévy trees and the pruning of such trees as developed in Abraham and Delmas (2008). Eventually Sections 5 and 6 are devoted to the proofs of Propositions 1.7 and 1.8.

## 2. Pruning at node of discrete GW trees

2.1. *Discrete trees.* Let us recall here the formalism for ordered discrete trees. We set

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

the set of finite sequences of positive integers with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. For $u \in \mathcal{U}$ let $|u|$ be the length or generation of $u$ defined as the integer $n$ such that $u \in (\mathbb{N}^*)^n$. If $u$ and $v$ are two sequences of $\mathcal{U}$, we denote by $uv$ the concatenation of the two sequences, with the convention that $uv = u$ if $v = \emptyset$ and $uv = v$ if $u = \emptyset$. The set of ancestors of $u$ is the set:

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

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

- $\emptyset \in \mathbf{t}$,
- If $u \in \mathbf{t}$, then $A_u \subset \mathbf{t}$.
- For every $u \in \mathbf{t}$, there exists a non-negative integer $k_u(\mathbf{t})$ such that, for all positive integers $i$, $ui \in \mathbf{t}$ iff $1 \leq i \leq k_u(\mathbf{t})$.

The integer $k_u(\mathbf{t})$ represents the number of offsprings of the node $u$ in the tree $\mathbf{t}$. We define $\mathcal{L}(\mathbf{t})$ the set of leaves of $\mathbf{t}$ and $\mathcal{N}(\mathbf{t})$ the set of internal nodes of $\mathbf{t}$ by:

$$\mathcal{L}(\mathbf{t}) = \{u \in \mathbf{t},\ k_u(\mathbf{t}) = 0\} \quad \text{and} \quad \mathcal{N}(\mathbf{t}) = \mathbf{t} \setminus \mathcal{L}(\mathbf{t}).$$

Let $L(\mathbf{t}) = \text{Card}\,(\mathcal{L}(\mathbf{t}))$ be the number of leaves of the tree $\mathbf{t}$, and notice that:

$$L(\mathbf{t}) - 1 = \sum_{u \in \mathcal{N}(\mathbf{t})} (k_u(\mathbf{t}) - 1). \tag{2.2}$$

We denote by $\mathbb{T}$ the set of discrete trees and by $\mathbb{T}_n = \{\mathbf{t} \in \mathbb{T}; L(\mathbf{t}) = n\}$ the set of discrete trees with $n$ leaves.

2.2. *A discrete tree-valued process.* We consider the pruning procedure developed in Abraham et al. (2012). Let $\mathbf{t} \in \mathbb{T}$. Under some probability measure $\mathbf{P}^{\mathbf{t}}$, we consider a family of marks $(\xi_u, u \in \mathcal{U})$ of independent non-negative real random variables (possibly infinite) such that:

- $\mathbf{P}^{\mathbf{t}}$-a.s. $\xi_u = +\infty$ if $u \notin \mathbf{t}$ or if $u \in \mathbf{t}$ and $k_u(\mathbf{t}) \in \{0, 1\}$,
- $\mathbf{P}^{\mathbf{t}}(\xi_u \geq \theta) = (1 + \theta)^{1 - k_u(\mathbf{t})}$ if $u \in \mathbf{t}$ and $k_u(\mathbf{t}) \geq 2$.

At time $\theta$, we define the pruned tree $\mathcal{P}_\theta(\mathbf{t})$ as the sub-tree given by:

$$\mathcal{P}_\theta(\mathbf{t}) = \{u \in \mathbf{t};\ \xi_v > \theta \text{ for all } v \in A_u,\ v \neq u\}.$$

In particular, we always have $\emptyset \in \mathcal{P}_\theta(\mathbf{t})$.

For $u \in \mathcal{N}(\mathbf{t})$, let $D_u$ be the event that $u$ is marked first, that is:

$$D_u = \{\xi_u = \min_{v \in \mathcal{N}(\mathbf{t})} \xi_v\}.$$

**Lemma 2.1.** *We suppose that $L(\mathbf{t}) \neq 1$. Let $u \in \mathcal{N}(\mathbf{t})$. We have:*

$$\mathbf{P}^{\mathbf{t}}(D_u) = \frac{k_u(\mathbf{t}) - 1}{L(\mathbf{t}) - 1} \cdot$$

This lemma implies that the cutting procedure given in Section 1.1, corresponds to the successive states of the process $(\mathcal{P}_\theta(\mathbf{t}), \theta \geq 0)$.

*Proof*: We have, using (2.2) for the last equality:

$$\begin{aligned}
\mathbf{P}^{\mathbf{t}}(D_u) &= \mathbf{P}^{\mathbf{t}}(\xi_u \le \xi_v \quad \forall v \neq u, v \in \mathcal{N}(\mathcal{T})) \\
&= \mathbf{E}^{\mathbf{t}}\left[(1+\xi_u)^{-\sum_{v\neq u, v\in\mathcal{N}(\mathbf{t})}(k_v(\mathbf{t})-1)}\right] \\
&= (k_u(\mathbf{t})-1)\int_{[0,+\infty)} (1+\theta)^{-\sum_{v\in\mathcal{N}(\mathbf{t})}(k_v(\mathbf{t})-1)-1}\, d\theta \\
&= \frac{k_u(\mathbf{t})-1}{\sum_{v\in\mathcal{N}(\mathbf{t})}(k_v(\mathbf{t})-1)} \\
&= \frac{k_u(\mathbf{t})-1}{L(\mathbf{t})-1}.
\end{aligned}$$

□

2.3. *Construction of the partition-valued process* $\Pi^{[n]}_{GW}$. Let $\alpha \in [1/2, 1)$. Recall that the function $g$ defined by (1.3) is the generating function of a probability measure $\nu_g$ on $\mathbb{N}$. We denote by $G_g(dT)$ the distribution on $\mathbb{T}$ of the critical GW tree with offspring distribution $\nu_g$. We will denote by $\mathbf{P}$ the probability measure on $\mathbb{T} \times [0, +\infty]^{\mathcal{U}}$:

$$\mathbf{P}(dT, d\xi) = G_g(dT)\mathbf{P}^T(d\xi).$$

Under $\mathbf{P}$, the random tree $T$ is a GW tree whose offspring distribution $\nu_g$ has generating function $g$ given by (1.3). According to Propositions 2.1 and 3.2 in Abraham et al. (2015), $(\mathcal{P}_\theta(T), \theta \ge 0)$ is a Markov process and $\mathcal{P}_\theta(T)$ is a GW tree whose reproduction law has generating function $g_\theta$, with:

$$g_\theta(r) = 1 + (1+\theta)\left[g\left(\frac{r}{1+\theta}\right) - g\left(\frac{1}{1+\theta}\right)\right].$$

Notice that:

$$g_\theta(r) = r + \alpha\frac{(1-r+\theta)^\gamma - \theta^\gamma}{(1+\theta)^{(\gamma)-1}} \tag{2.3}$$

with $\gamma = 1/\alpha$.

For every positive integer $n$, we set:

$$\mathbf{P}_n(\bullet) = \mathbf{P}(\bullet \mid L(T) = n).$$

Under $\mathbf{P}_n$, the distribution of the tree $T$ is given by the following formula (see Duquesne and Le Gall (2002), Theorem 3.3.3, or Marchal (2008)), for $\mathbf{t} \in \mathbb{T}_n$:

$$\mathbf{P}_n(T = \mathbf{t}) = n!\left(\prod_{v\in\mathcal{N}(\mathbf{t})} \frac{p_{k_v(\mathbf{t})}}{k_v(\mathbf{t})!}\right)\frac{\alpha^{n-1}\Gamma(1-\alpha)}{\Gamma(n-\alpha)} \tag{2.4}$$

where $p_1 = 0$ and, for $k \ge 2$, $p_k = |(1-\gamma)(2-\gamma)\cdots(k-\gamma)|$.

Let $n \in \mathbb{N}^*$. Let $T$ be a random tree distributed according to $\mathbf{P}_n$. Conditionally on $T$, we define a uniform random labeling $U_1, \dots, U_n$ of the leaves of $T$, independently of the variables $(\xi_u, u \in T)$. Recall the set of ancestors defined in (2.1) and the pruning procedure $\mathcal{P}_\theta$ introduced in Section 2.2. We define the equivalence relation $\mathcal{R}^{[n]}_\theta$ on $\{1, 2, \dots, n\}$ by: $i\mathcal{R}^{[n]}_\theta j$ if $A_{U_i} \bigcap A_{U_j} \bigcap \mathcal{L}(\mathcal{P}_\theta(T))$ is non empty, that is $U_i$ and $U_j$ have a leaf of $\mathcal{P}_\theta(T)$ as common ancestor. Then, for every $\theta \ge 0$, let $\hat{\Pi}^{[n]}_{\mathrm{GW}}(\theta)$ be the equivalence classes of the equivalence relation $\mathcal{R}^{[n]}_\theta$ of the $n$ first integers. Let $\Pi^{[n]}_{\mathrm{GW}} = (\Pi^{[n]}_{\mathrm{GW}}(k), k \in \mathbb{N})$ be the discrete process associated with

$\hat{\Pi}^{[n]}_{\mathrm{GW}} = (\hat{\Pi}^{[n]}_{\mathrm{GW}}(\theta), \theta \geq 0)$ until it reaches the absorbing state (which is the trivial partition consisting in one block) and afterward the discrete process remains constant.

We end this section with an elementary lemma which will be used in the proof of Theorem 1.1.

**Lemma 2.2.** *We have for $n \geq 2$:*

$$\mathbf{E}_n\left[k_\emptyset(T) - 1\right] = \frac{1-\alpha}{\alpha} \frac{\Gamma(1-\alpha)}{\Gamma(\alpha)} \frac{\Gamma(n-1+\alpha)}{\Gamma(n-\alpha)}. \tag{2.5}$$

*Proof*: We consider the generating function of $(k_\emptyset(T), L(T))$ under $\mathbf{P}$, that is $H(s,t) = \mathbf{E}\left[s^{k_\emptyset(T)} t^{L(T)}\right]$. Using the branching property of GW trees, we have:

$$H(s,t) = \mathbf{E}\left[s^{k_\emptyset(T)} \mathbf{E}[t^{L(T)}]^{k_\emptyset(T)} \mathbf{1}_{\{k_\emptyset(T) \neq 0\}}\right] + t\mathbf{P}(k_\emptyset(T) = 0). \tag{2.6}$$

Notice that $g(s) = \mathbf{E}\left[s^{k_\emptyset(T)}\right] = H(s,1)$. We set $h(t) = H(1,t) = \mathbf{E}\left[t^{L(T)}\right]$ the generating function of $L(T)$. So that (2.6) becomes:

$$H(s,t) = g(s\,h(t)) - g(0)(1-t). \tag{2.7}$$

Taking $s = 1$ in (2.7), we get:

$$g(h(t)) - h(t) = g(0)(1-t). \tag{2.8}$$

Using expression (1.3), we get:

$$h(t) = 1 - (1-t)^\alpha \quad \text{and} \quad H(s,t) = s\,h(t) + \alpha(1 - s\,h(t))^{1/\alpha} - \alpha(1-t).$$

We deduce that:

$$\begin{aligned}
\mathbf{E}\left[k_\emptyset(T) t^{L(T)}\right] = \frac{\partial H}{\partial s}(1,t) &= h(t) - h(t)(1-h(t))^{(1/\alpha)-1} \\
&= \mathbf{E}\left[t^{L(T)}\right] - \left[1 - (1-t)^\alpha\right](1-t)^{1-\alpha} \\
&= \mathbf{E}\left[t^{L(T)}\right] - (1-t)^{1-\alpha} + 1 - t.
\end{aligned}$$

This gives:

$$\mathbf{E}\left[(k_\emptyset(T) - 1) t^{L(T)}\right] = -(1-t)^{1-\alpha} + 1 - t.$$

For $n \geq 2$, we get:

$$\begin{aligned}
\mathbf{E}\left[(k_\emptyset(T) - 1)\mathbf{1}_{\{L(T)=n\}}\right] &= \frac{1}{n!}\left(\frac{d^n}{dt^n}\mathbf{E}\left[(k_\emptyset(T) - 1) t^{L(T)}\right]\right)_{|t=0} \\
&= \frac{1}{n!}(1-\alpha)\prod_{k=0}^{n-2}(\alpha + k) \\
&= \frac{1}{n!}(1-\alpha)\frac{\Gamma(n-1+\alpha)}{\Gamma(\alpha)}.
\end{aligned}$$

We also get for $n \geq 2$:

$$\mathbf{P}(L(T) = n) = \frac{1}{n!} h^{(n)}(0) = \frac{1}{n!}\alpha\prod_{k=1}^{n-1}(k-\alpha) = \frac{1}{n!}\alpha\frac{\Gamma(n-\alpha)}{\Gamma(1-\alpha)}.$$

We deduce that:

$$\mathbf{E}_n\left[k_\emptyset(T) - 1\right] = \frac{\mathbf{E}\left[(k_\emptyset(T) - 1)\mathbf{1}_{\{L(T)=n\}}\right]}{\mathbf{P}(L(T) = n)} = \frac{1-\alpha}{\alpha}\frac{\Gamma(1-\alpha)}{\Gamma(\alpha)}\frac{\Gamma(n-1+\alpha)}{\Gamma(n-\alpha)}.$$

$\square$

## 3. Proof of Theorem 1.1

Let $\alpha \in [1/2, 1)$ and $\Lambda$ given by (1.4). Notice that the probability that the first coalescence event for $\Pi_{\text{dis}}^{[n]}$ corresponds to the collision of $k$ given blocks is $\lambda_{n,k}/\lambda_n$, with $\lambda_{n,k}$ and $\lambda_n$ given respectively by (1.1) and (1.2).

Theorem 1.1 is a direct consequence of Lemma 3.3 which states that the probability that the first coalescence event for $\Pi_{\text{GW}}^{[n]}$ corresponds to the collision of $k$ given blocks is $\lambda_{n,k}/\lambda_n$, and of Lemma 3.4, which states that after the first coalescence event, the law of the pruned tree under $\mathbf{P}_n$ conditionally given that it has $k$ leaves is exactly $\mathbf{P}_k$.

The proof of Lemma3.3 (resp. 3.4) is given in Section 3.1 (resp. 3.2).

3.1. *Computation of the coalescence rates.* We first give an intermediate lemma. For $\alpha \in (0, 1)$ and $\lambda > \alpha - 1$, we set:

$$\phi_{1+\alpha,1-\alpha}(\lambda) = \int_0^1 \left(1 - (1-x)^\lambda\right) x^{\alpha-2}(1-x)^{-\alpha}\, dx. \tag{3.1}$$

**Lemma 3.1.** *For $\alpha \in (0, 1)$ and $\lambda > \alpha - 1$, we have:*

$$\phi_{1+\alpha,1-\alpha}(\lambda) = \lambda \frac{\Gamma(\alpha)\Gamma(\lambda + 1 - \alpha)}{(1-\alpha)\Gamma(\lambda + 1)}. \tag{3.2}$$

Notice that for $\lambda > 0$, (3.2) reduces to:

$$\phi_{1+\alpha,1-\alpha}(\lambda) = \frac{\Gamma(\alpha)\Gamma(\lambda + 1 - \alpha)}{(1-\alpha)\Gamma(\lambda)}. \tag{3.3}$$

*Proof*: We set:

$$I = \int_0^1 \left((1-u)^{-\alpha} - 1\right)\, u^{\alpha-2}\, du.$$

Notice that $I$ is finite and $\phi_{1+\alpha,1-\alpha}(\alpha) = I$. For $\lambda > \alpha$, using an integration by part, we have:

$$\begin{aligned}
\phi_{1+\alpha,1-\alpha}(\lambda) &= \int_0^1 \left(1 - (1-x)^\lambda\right) x^{\alpha-2}(1-x)^{-\alpha}\, dx \\
&= \int_0^1 \left((1-x)^{-\alpha} - 1\right) x^{\alpha-2}\, dx + \int_0^1 \left(1 - (1-x)^{\lambda-\alpha}\right) x^{\alpha-2} dx \\
&= I - \frac{1}{1-\alpha} + \frac{\lambda - \alpha}{1-\alpha} \int_0^1 (1-x)^{\lambda-\alpha-1} x^{\alpha-1}\, dx \\
&= I - \frac{1}{1-\alpha} + \frac{\Gamma(\alpha)\Gamma(\lambda+1-\alpha)}{(1-\alpha)\Gamma(\lambda)}.
\end{aligned}$$

We now compute $I$. Remark first that, by (3.3) for $\lambda = 1$, we have:

$$\phi_{1+\alpha,1-\alpha}(1) = \Gamma(\alpha)\Gamma(1-\alpha).$$

We deduce that:

$$I - \frac{1}{1-\alpha} + \frac{\Gamma(\alpha)\Gamma(2-\alpha)}{(1-\alpha)\Gamma(1)} = \phi_{1+\alpha,1-\alpha}(1) = \Gamma(\alpha)\Gamma(1-\alpha).$$

This readily implies that $I = 1/(1-\alpha)$ and thus (3.2) holds for $\lambda \geq \alpha$. Then uses that the right-hand sides of (3.1) and (3.2) are analytic for $\lambda > \alpha - 1$ to get that (3.2) also holds for $\lambda > \alpha - 1$. □

Recall that $\lambda_{n,k}$ and $\lambda_n$ are given respectively by (1.1) and (1.2), for $\Lambda$ given by (1.4).

**Lemma 3.2.** *Let $\alpha \in [1/2, 1)$. We have for $2 \leq k \leq n$:*

$$\frac{\lambda_{n,k}}{\lambda_n} = \frac{1-\alpha}{\Gamma(\alpha+1)} \frac{\Gamma(k+\alpha-1)\Gamma(n-k-\alpha+1)}{\Gamma(n-\alpha)} \frac{1}{n-1}. \tag{3.4}$$

*Proof*: We have

$$\begin{aligned}
\lambda_{n,k} &= \int_0^1 u^{k-2}(1-u)^{n-k}\Lambda(du) \\
&= \int_0^1 u^{k-2+\alpha}(1-u)^{n-k-\alpha} du \\
&= \beta(k+\alpha-1, n-k-\alpha+1) \\
&= \frac{\Gamma(k+\alpha-1)\Gamma(n-k-\alpha+1)}{\Gamma(n)},
\end{aligned}$$

and

$$\lambda_n = \sum_{k=2}^{n} \binom{n}{k} \lambda_{n,k} = \int_0^1 (1-(1-u)^n - nu(1-u)^{n-1})u^{-2}\Lambda(du).$$

Then using notations (3.1) and (3.3), we deduce that:

$$\begin{aligned}
\lambda_n &= \phi_{1+\alpha,1-\alpha}(n) - n\int_0^1 u^{\alpha-1}(1-u)^{n-1-\alpha}\, du \\
&= \frac{\Gamma(\alpha)\Gamma(n+1-\alpha)}{(1-\alpha)\Gamma(n)} - n\frac{\Gamma(\alpha)\Gamma(n-\alpha)}{\Gamma(n)} \\
&= \left(\frac{n-\alpha}{1-\alpha} - n\right)\frac{\Gamma(\alpha)\Gamma(n-\alpha)}{\Gamma(n)} \\
&= (n-1)\frac{\alpha}{1-\alpha}\frac{\Gamma(\alpha)\Gamma(n-\alpha)}{\Gamma(n)}.
\end{aligned}$$

The expression obtained for $\lambda_{n,k}$ then gives the result. □

If $\mathbf{t}_1$ and $\mathbf{t}_2$ are two discrete trees and $u \in \mathcal{L}(\mathbf{t}_1)$ is a leaf of $\mathbf{t}_1$, we shall denote by $\mathbf{t}_1 \circledast_u \mathbf{t}_2$ the tree obtained by grafting the tree $\mathbf{t}_2$ on the leaf $u$ of $\mathbf{t}_1$, that is:

$$\mathbf{t}_1 \circledast_u \mathbf{t}_2 = \mathbf{t}_1 \cup \{uv,\ v \in \mathbf{t}_2\}. \tag{3.5}$$

**Lemma 3.3.** *Let $\alpha \in [1/2, 1)$. The probability under $\mathbf{P}_n$ that the first coalescence event in $\Pi_{GW}^{[n]}$ is the coalescence of $k$ given integers into one block is $\lambda_{n,k}/\lambda_n$.*

*Proof*: Let $A_k$ be the event that the first coalescence event corresponds to the $k$ first integers merging together. By exchangeability, the lemma is proved as soon as we check that $\mathbf{P}_n(A_k) = \lambda_{n,k}/\lambda_n$.

The event $A_k$ is realized, if and only if:

- The initial tree $T$ is of the form $\mathbf{t}_1 \circledast_u \mathbf{t}_2$ for some $\mathbf{t}_2 \in \mathbb{T}_k$ and $\mathbf{t}_1 \in \mathbb{T}_{n-k+1}$ and $u \in \mathcal{L}(\mathbf{t}_1)$.

- The leaves of $\mathbf{t}_2$ are labeled from 1 to $k$ (and therefore, the leaves of $\mathbf{t}_1$ except $u$ are labeled from $k+1$ to $n$). This occurs with probability $\frac{k!(n-k)!}{n!}$.
- The first chosen node of $\mathbf{t}_1 \circledast_u \mathbf{t}_2$ is $u$. This occurs according to Lemma 2.1 with probability $\frac{k_\emptyset(\mathbf{t}_2)-1}{n-1}$.

Thus, using (2.4) for the probability of having a given tree, we have:

$$\begin{aligned}
\mathbf{P}_n(A_k) &= \sum_{\substack{\mathbf{t}_1\in\mathbb{T}_{n-k+1}\\ \mathbf{t}_2\in\mathbb{T}_k\\ u\in\mathcal{L}(\mathbf{t}_1)}} \mathbf{P}_n(T=\mathbf{t}_1\circledast_u \mathbf{t}_2)\frac{k!(n-k)!}{n!}\frac{k_\emptyset(\mathbf{t}_2)-1}{n-1}\\
&= \sum_{\substack{\mathbf{t}_1\in\mathbb{T}_{n-k+1}\\ \mathbf{t}_2\in\mathbb{T}_k\\ u\in\mathcal{L}(\mathbf{t}_1)}} n!\left(\prod_{v\in\mathcal{N}(\mathbf{t}_1\circledast_u\mathbf{t}_2)}\frac{p_{k_v(\mathbf{t}_1\circledast_u\mathbf{t}_2)}}{k_v(\mathbf{t}_1\circledast_u\mathbf{t}_2)!}\right)\\
&\qquad\qquad\times\frac{\alpha^{n-1}\Gamma(1-\alpha)}{\Gamma(n-\alpha)}\frac{k!(n-k)!}{n!}\frac{k_\emptyset(\mathbf{t}_2)-1}{n-1}\\
&=(n-k+1)\sum_{\substack{\mathbf{t}_1\in\mathbb{T}_{n-k+1}\\ \mathbf{t}_2\in\mathbb{T}_k}}\frac{n!}{k!(n-k+1)!}\mathbf{P}_{n-k+1}(T=\mathbf{t}_1)\mathbf{P}_k(T=\mathbf{t}_2)\\
&\quad\times\frac{\alpha^{n-1}\Gamma(1-\alpha)}{\Gamma(n-\alpha)}\frac{\Gamma(n-k-\alpha+1)}{\alpha^{n-k}\Gamma(1-\alpha)}\frac{\Gamma(k-\alpha)}{\alpha^{k-1}\Gamma(1-\alpha)}\frac{k!(n-k)!}{n!}\frac{k_\emptyset(\mathbf{t}_2)-1}{n-1}\\
&=\frac{\Gamma(n-k-\alpha+1)\Gamma(k-\alpha)}{\Gamma(n-\alpha)\Gamma(1-\alpha)}\frac{1}{n-1}\mathbf{E}_k\left[k_\emptyset(T)-1\right].
\end{aligned}$$

We then use Lemma 2.2 and Lemma 3.2 to conclude. □

3.2. *Law of the tree after the first coalescence event.* Let $S$ be the time of the first coalescence event and recall that $\mathcal{P}_S(T)$ denote the pruned tree at the first coalescence event.

**Lemma 3.4.** *Let* $\mathbf{t}\in\mathbb{T}_k$. *We have:*

$$\mathbf{P}_n(\mathcal{P}_S(T)=\mathbf{t}\,\big|\,L(\mathcal{P}_S(T))=k)=\mathbf{P}_k(T=\mathbf{t}). \tag{3.6}$$

*Proof*: Let $\mathbf{t}\in\mathbb{T}_k$. We obtain $\mathbf{t}$ just after the first coalescence event if $T$ is of the form $\mathbf{t}\circledast_u\mathbf{s}$ for some $s\in\mathbb{T}_{n-k+1}$, $u\in\mathcal{L}(\mathbf{t})$ and $u$ is the first chosen internal node.

This gives:

$$\begin{aligned}
\mathbf{P}_n(\mathcal{P}_S(T)=\mathbf{t}) &= \sum_{\substack{u\in\mathcal{L}(\mathbf{t})\\ \mathbf{s}\in\mathbb{T}_{n-k+1}}} \mathbf{P}_n(T=\mathbf{t}\circledast_u \mathbf{s})\frac{k_\emptyset(\mathbf{s})-1}{n-1}\\
&= k\sum_{\mathbf{s}\in\mathbb{T}_{n-k+1}} n!\left(\prod_{v\in\mathcal{N}(\mathbf{t})}\frac{p_{k_v(\mathbf{t})}}{k_v(\mathbf{t})!}\prod_{v\in\mathcal{N}(\mathbf{s})}\frac{p_{k_v(\mathbf{s})}}{k_v(\mathbf{s})!}\right)\\
&\qquad\qquad\times\frac{\alpha^{n-1}\Gamma(1-\alpha)}{\Gamma(n-\alpha)}\frac{k_\emptyset(\mathbf{s})-1}{n-1}\\
&= k\sum_{\mathbf{s}\in\mathbb{T}_{n-k+1}}\frac{n!}{k!(n-k+1)!}\mathbf{P}_k(T=\mathbf{t})\mathbf{P}_{n-k+1}(T=\mathbf{s})\\
&\quad\times\frac{\alpha^{n-1}\Gamma(1-\alpha)}{\Gamma(n-\alpha)}\frac{\Gamma(k-\alpha)}{\alpha^{k-1}\Gamma(1-\alpha)}\frac{\Gamma(n-k+1-\alpha)}{\alpha^{n-k}\Gamma(1-\alpha)}\frac{k_\emptyset(\mathbf{s})-1}{n-1}\\
&= \frac{n!}{(k-1)!(n-k+1)!}\frac{\Gamma(n-k+1-\alpha)\Gamma(k-\alpha)}{\Gamma(n-\alpha)\Gamma(1-\alpha)}\\
&\qquad\qquad\frac{1}{n-1}\mathbf{E}_{n-k+1}[k_\emptyset(T)-1]\mathbf{P}_k(T=\mathbf{t}).
\end{aligned}$$

As the term in front of $\mathbf{P}_k(T=\mathbf{t})$ does not depend on $\mathbf{t}$, it has to be equal to $\mathbf{P}_n(L(\mathcal{P}_S(T))=k)$ and therefore (3.6) holds. □

## 4. Pruning of rooted real trees and proof of Theorem 1.3

The aim of this section is to use the pruning procedure for Lévy trees developed in Abraham and Delmas (2008) to give a consistent representation of the family of coalescent processes $(\hat{\Pi}^{[n]}_{\mathrm{GW}}, n\in\mathbb{N}^*)$, see Corollary 4.4 and thus deduce Theorem 1.3.

### 4.1. *The CRT framework.*

4.1.1. *Real trees.* Real trees have been introduced first in the field of geometric group theory (see for instance Dress et al. (1996)) and then used later for defining continuum random trees (the framework first appeared in Evans et al. (2006)). A real tree is a metric space $(\mathcal{T},d)$ satisfying the following two properties for every $x,y\in\mathcal{T}$:

- (unique geodesic) There is a unique isometric map $f_{x,y}$ from $[0,d(x,y)]$ into $\mathcal{T}$ such that $f_{x,y}(0)=x$ and $f_{x,y}(d(x,y))=y$.
- (no loop) If $\varphi$ is a continuous injective map from $[0,1]$ into $\mathcal{T}$ such that $\varphi(0)=x$ and $\varphi(1)=y$, then
$$\varphi([0,1])=f_{x,y}([0,d(x,y)]).$$

A rooted real tree is a real tree with a distinguished vertex denoted $\emptyset$ and called the root.

For every $x,y\in\mathcal{T}$, we denote by $[\![x,y]\!]$ the range of the map $f_{x,y}$ (i.e. the only path in the tree that links $x$ to $y$) and we set $[\![x,y[\![=[\![x,y]\!]\setminus\{y\}$.

If $\mathcal{T}$ is a rooted real tree, for $x\in\mathcal{T}$, we define the degree of $x$, denoted by $n_x$, as the number of connected components of $\mathcal{T}\setminus\{x\}$. The set of leaves of $\mathcal{T}$ is

$\mathcal{L}(\mathcal{T}) = \{x \in \mathcal{T} \setminus \{\emptyset\};\ n_x = 1\}$. If $n_x \geq 3$, we say that $x$ is a branching point of $\mathcal{T}$. We denote by $\mathcal{B}_{\mathrm{br}}(\mathcal{T})$ the set of branching points of $\mathcal{T}$. The height of $\mathcal{T}$ is $H_{\max}(\mathcal{T}) = \sup\{d(\emptyset, x);\ x \in \mathcal{T}\}$. Let $(x_i, i \in I)$ be a family of elements of $\mathcal{T}$, we define their most recent common ancestor denoted by $\mathrm{MRCA}(x_i, i \in I)$ as the element $x$ of $\mathcal{T}$ such that $[\![\emptyset, x]\!] = \bigcap_{i\in I} [\![\emptyset, x_i]\!]$.

A weighted rooted real tree $(\mathcal{T}, d, \mathbf{m})$ is a rooted real tree $(\mathcal{T}, d)$ endowed with a $\sigma$-finite measure $\mathbf{m}$ called the mass measure.

4.1.2. *Stable Lévy tree.* We now always let $\psi(\lambda) = \lambda^\gamma$ with $\gamma \in (1, 2)$. We refer to Duquesne and Le Gall (2005) and Abraham et al. (2014) for the existence of a measure $\mathbb{N}[d\mathcal{T}]$ on the set of weighted locally compact rooted real trees such that $\mathcal{T}$ is under $\mathbb{N}[d\mathcal{T}]$ a Lévy tree associated with the branching mechanism $\psi$. For the Lévy tree $(\mathcal{T}, d, \mathbf{m})$, $\mathbb{N}[d\mathcal{T}]$ -a.e., the mass measure has support $\mathcal{L}(\mathcal{T})$ and has no atom. Furthermore, $\mathbb{N}[d\mathcal{T}]$-a.e., all the branching points of the tree are of infinite degree. Following Duquesne and Le Gall (2005), there exists a local time process $(\ell^a, a \geq 0)$ with values on finite measures on $\mathcal{T}$, which is cdlg for the weak topology on finite measures on $\mathcal{T}$ and such that $\mathbb{N}[d\mathcal{T}]$-a.e.:

$$\mathbf{m}(dx) = \int_0^\infty \ell^a(dx)\, da,$$

$\ell^0 = 0$, $\inf\{a > 0; \ell^a = 0\} = \sup\{a \geq 0; \ell^a \neq 0\} = H_{\max}(\mathcal{T})$ and for every fixed $a \geq 0$, $\mathbb{N}[d\mathcal{T}]$-a.e. the measure $\ell^a$ is supported on $\{x \in \mathcal{T}; d(\emptyset, x) = a\}$ and the real valued process $(\langle \ell^a, 1\rangle, a \geq 0)$ is distributed as a continuous state branching process (CSBP) with branching mechanism $\psi$ under its canonical measure. In particular, as the total size of a critical CSBP is finite, we get that $\mathbb{N}$-a.e. $\sigma = \mathbf{m}(\mathcal{T})$ is finite.

The set $\{d(\emptyset, x);\ x \in \mathcal{B}_{\mathrm{br}}(\mathcal{T})\}$ coincides $\mathbb{N}$-a.e. with the set of discontinuity times of the mapping $a \mapsto \ell^a$. Moreover, $\mathbb{N}$-a.e., for every such discontinuity time $b$, there is a unique $x \in \mathcal{B}_{\mathrm{br}}(\mathcal{T})$ such that $d(\emptyset, x) = b$ and $\Delta_x > 0$, such that:

$$\ell^b = \ell^{b-} + \Delta_x \delta_x,$$

where $\Delta_x > 0$ is called the mass of the node $x$. Intuitively $\Delta_x$ represents the size of the progeny of $x$.

The scaling property of the stable Lévy tree implies that there exists a well defined probability measure $\mathbb{N}^{(1)}$ defined as the measure $\mathbb{N}$ conditioned on $\{\sigma = 1\}$. The probability measure is also referred to as the normalized excursion measure for Lévy trees.

4.2. *The partition-valued process.*

4.2.1. *Pruning of the stable Lévy tree.* We consider the pruning procedure introduced in Abraham and Delmas (2008) (this procedure is defined when there is no Brownian part in the Lévy process with index given by the branching mechanism $\psi$). Under $\mathbb{N}$ or $\mathbb{N}^{(1)}$, conditionally given $\mathcal{T}$, we consider a family $(E_x, x \in \mathcal{B}_{\mathrm{br}}(\mathcal{T}))$ of independent real random variables such that the random variable $E_x$ is exponentially distributed with parameter $\Delta_x$. This random variable represents the time at which the branching point $x$ is marked. For every $\theta > 0$, we set

$$\mathcal{T}_\theta = \{x \in \mathcal{T},\ \forall y \in [\![\emptyset, x[\![,\ E_y \geq \theta\}.$$

The set $\mathcal{T}_\theta$ is still a real tree which represents the tree $\mathcal{T}$ pruned at time $\theta$: we cut $\mathcal{T}$ at the points that are marked before time $\theta$ and keep the connected component of the tree that contains the root. We set $\mathcal{T}_0 = \mathcal{T}$. By Abraham and Delmas (2008), Theorem 1.5, the tree $\mathcal{T}_\theta$ is distributed under $\mathbb{N}$ as a Lévy tree with branching mechanism $\psi_\theta$ defined by:

$$\psi_\theta(\lambda) = \psi(\lambda + \theta) - \psi(\theta).$$

Moreover, by Abraham and Delmas (2012), the process $(\mathcal{T}_\theta, \theta \geq 0)$ is under $\mathbb{N}$ a Markov process.

4.2.2. *Definition of the partition-valued process.* Under $\mathbb{N}$ or $\mathbb{N}^{(1)}$, conditionally on $\mathcal{T}$, let $(F_i, i \in \mathbb{N}^*)$ be independent random variables on $\mathcal{T}$ distributed according to the probability mass measure $\mathbf{m}/\mathbf{m}(\mathcal{T})$, and independent of the marks $(E_x, x \in \mathcal{B}_{\mathrm{br}}(\mathcal{T}))$. Notice that $\mathbb{N}$-a.e. or $\mathbb{N}^{(1)}$-a.s. $(F_i, i \in \mathbb{N}^*)$ are leaves of $\mathcal{T}$. For $\theta \geq 0$, we define the equivalence relation $\mathcal{R}_\theta^{\text{Lévy}}$ on $\mathbb{N}^*$ by: $i\mathcal{R}_\theta^{\text{Lévy}}j$ if $[\![\emptyset, F_i]\!] \bigcap [\![\emptyset, F_j]\!] \bigcap \mathcal{L}(\mathcal{T}_\theta)$ is non empty, that is $F_i$ and $F_j$ have a leaf of $\mathcal{T}_\theta$ as common ancestor. This is very close to the definition of the equivalence relation $\mathcal{R}_\theta^{[\mathrm{n}]}$ defined in Section 2.3. We denote by $\Pi_{\text{Lévy}}(\theta)$ the partition of $\mathbb{N}^*$ formed by the equivalence classes of $\mathcal{R}_\theta^{\text{Lévy}}$ and set $\Pi_{\text{Lévy}} = (\Pi_{\text{Lévy}}(\theta), \theta \geq 0)$.

4.3. *Lévy sub-trees.*

4.3.1. *Skeleton of a finite real tree.* Let $\hat{\mathbf{t}}$ be a real tree with finite height and a finite number of leaves, such that the leaves $(f_i, i \in I(\hat{\mathbf{t}}))$ are indexed by a totally ordered set $I(\hat{\mathbf{t}})$. We define the skeleton $\tilde{\mathbf{t}}$ of the tree $\hat{\mathbf{t}}$ as the discrete tree (belonging to $\mathbb{T}$) where we forget the edge lengths. As the trees in $\mathbb{T}$ are ordered, we must be a bit more rigorous for the definition of $\tilde{\mathbf{t}}$.

The skeleton $\tilde{\mathbf{t}}$ of the real tree with ordered leaves $\big(\hat{\mathbf{t}}, (f_i, i \in I(\hat{\mathbf{t}}))\big)$ is defined recursively as follows. We define $k_\emptyset(\tilde{\mathbf{t}})$ as the degree of $\mathrm{MRCA}(f_i, i \in I(\hat{\mathbf{t}}))$ the ancestor of all the leaves of $\hat{\mathbf{t}}$. If $k_\emptyset(\tilde{\mathbf{t}}) = 0$, then $\tilde{\mathbf{t}}$ is reduced to $\emptyset$. In this case $\hat{\mathbf{t}}$ has one leaf, let $f$ be its label, and the discrete tree $\tilde{\mathbf{t}}$ has thus one leaf to which we give the label $f$. If $k_\emptyset(\tilde{\mathbf{t}}) > 0$, then we consider the $k_\emptyset(\tilde{\mathbf{t}})$ connected components of $\hat{\mathbf{t}} \setminus \{\mathrm{MRCA}(f_i, i \in I(\hat{\mathbf{t}}))\}$ that do not contain the root and label them from 1 to $k_\emptyset(\tilde{\mathbf{t}})$ according to the lowest label of the leaves of $\hat{\mathbf{t}}$ which belongs to them. This gives an ordered family $(\hat{\mathbf{t}}_k, k \in \{1, \ldots, k_\emptyset(\tilde{\mathbf{t}})\})$ of real trees, and let $\mathrm{MRCA}(f_i, i \in I(\hat{\mathbf{t}}))$ be the root of each one. For $k \in \{1, \ldots, k_\emptyset(\tilde{\mathbf{t}})\}$, let $I(\hat{\mathbf{t}}_k) = \{i \in I(\hat{\mathbf{t}}); f_i \in \hat{\mathbf{t}}_k\}$ be the labels of the leaves of $\hat{\mathbf{t}}_k$ and the discrete tree $\tilde{\mathbf{t}}_k$ is the skeleton of $\big(\hat{\mathbf{t}}_k, (f_i, i \in I(\hat{\mathbf{t}}_k))\big)$.

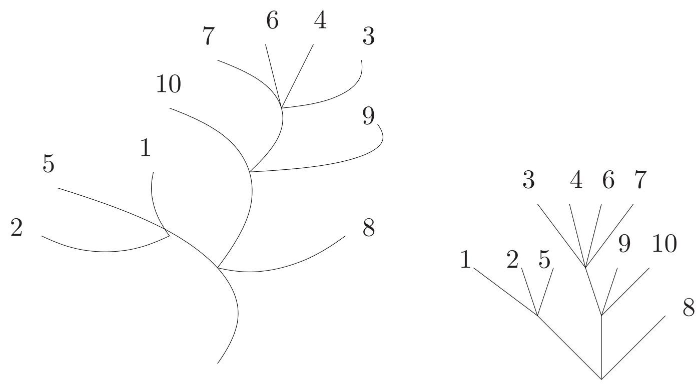


FIGURE 4.2. A real tree $\hat{\mathbf{t}}$ with labeled leaves and the associated skeleton $\tilde{\mathbf{t}}$ (with the same labels)

Notice that $\tilde{\mathbf{t}}$ is finite, $k_u(\tilde{\mathbf{t}}) \neq 1$ for all $u \in \tilde{\mathbf{t}}$, and $\hat{\mathbf{t}}$ and $\tilde{\mathbf{t}}$ have the same number of leaves. In the previous construction to a leaf $f_i$ of $\hat{\mathbf{t}}$ with label $i$ corresponds a unique leaf $e_i$ of $\tilde{\mathbf{t}}$ with label $i$. For $u \in \tilde{\mathbf{t}}$, we define $\tilde{\mathbf{t}}_u$ the sub-tree of $\tilde{\mathbf{t}}$ attached to the node $u$, i.e.

$$\tilde{\mathbf{t}}_u = \{w \in \mathcal{U};\ uw \in \tilde{\mathbf{t}}\}, \tag{4.1}$$

and let $I_u = \{i; e_i \in \tilde{\mathbf{t}}_u\}$. Define $\hat{\mathbf{t}}_u$ as $\mathbf{s}_u = \hat{\mathbf{t}} \backslash \bigcup_{i \notin I_u} [\![\emptyset, f_i]\!]$ to which we add the root $\emptyset_u = \overline{\mathbf{s}_u} \backslash \mathbf{s}_u$, and $I(\hat{\mathbf{t}}_u) = \{i; e_i \in \tilde{\mathbf{t}}_u\}$. Notice that by construction $\tilde{\mathbf{t}}_u$ is the skeleton of $\big(\hat{\mathbf{t}}_u, (f_i; i \in I(\hat{\mathbf{t}}_u))\big)$. We say that $u \in \tilde{\mathbf{t}}$ are the individuals of $\hat{\mathbf{t}}$, and define their lifetime as the length $h_u$ of the geodesic $B(u) = [\![\emptyset_u, \mathrm{MRCA}(f_i, i \in I(\hat{\mathbf{t}}_u))]\!]$. We say the corresponding node in $\hat{\mathbf{t}}$ of $u \in \tilde{\mathbf{t}}$ is $C(u) = \mathrm{MRCA}(f_i, i \in I(\hat{\mathbf{t}}_u))$.

Notice it is easy to reconstruct $\hat{\mathbf{t}}$ from $\tilde{\mathbf{t}}$ and the family of lifetimes $(h_u, u \in \tilde{\mathbf{t}})$.

4.3.2. *Coalescence of Lévy tree and GW tree.* Let $(F_i, i \in \mathbb{N}^*)$ be defined as in Section 4.2.2. Let $M$ be, under $\mathbb{N}$ or $\mathbb{N}^{(1)}$ conditionally on $\mathcal{T}$, a Poisson random variable with finite mean $\sigma = \mathbf{m}(\mathcal{T})$, independant of the $F_i$'s. We shall work on $\{M \geq 1\}$. On $\{M \geq 1\}$, let $\hat{T}_0$ be the real sub-tree of $\mathcal{T}$ generated by the root and $(F_i, 1 \leq i \leq M)$:

$$\hat{T}_0 = \bigcup_{1 \leq i \leq M} [\![\emptyset, F_i]\!].$$

Since $\mathbf{m}$ has support $\mathcal{L}(\mathcal{T})$ and has no atom, we deduce that $(F_i, 1 \leq i \leq M)$ are distinct and are the leaves of $\hat{T}_0$.

We denote by $\tilde{T}_0$ the skeleton of $\hat{T}_0$ with the labeled leaves $(F_i, 1 \leq i \leq M)$. According to Duquesne and Le Gall (2002), Theorem 3.2.1, the tree $\hat{T}_0$ is distributed under $\mathbb{N}[\,\cdot \mid M \geq 1]$ as a continuous GW tree (i.e. a GW tree with edge-lengths) such that

- The discrete tree $\tilde{T}_0$ is a GW tree with offspring distribution characterized by its generating function $g$ defined by (1.3) with $\alpha = 1/\gamma$.
- Lifetimes of individuals $(h_u, u \in \tilde{T}_0)$ are independent random variables with exponential distribution with parameter $\gamma$.

We must first prove the following lemma which will be a key point in the sequel. Its proof relies on the scaling property of the Lévy tree.

**Lemma 4.1.** *The distributions of $\tilde{T}_0$ under $\mathbb{N}[\,\cdot \mid M = n]$ and under $\mathbb{N}^{(1)}[\,\cdot \mid M = n]$ are the same.*

*Proof*: For a tree $\mathcal{T}$ and points $x_1, \ldots, x_n$ of $\mathcal{T}$, let us denote by $T(\mathcal{T}, x_1, \ldots, x_n)$ the tree spanned by the points $(x_i)$ and the root of the tree and $\tilde{T}(\mathcal{T}, x_1, \ldots, x_n)$ the associated discrete tree so that under $\mathbb{N}[\,\cdot \mid M = n]$ or $\mathbb{N}^{(1)}[\,\cdot \mid M = n]$, we have

$$\tilde{T}_0 = \tilde{T}(\mathcal{T}, F_1, \ldots, F_n).$$

Then, for every bounded measurable function $\phi$, we have

$$\mathbb{N}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\mathbf{1}_{\{M=n\}}\right] = \mathbb{N}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\frac{\sigma^n}{n!}\,\mathrm{e}^{-\sigma}\right].$$

Let $\nu$ be the distribution of $\sigma$ under $\mathbb{N}$ i.e. the only measure $\nu$ such that for every $\lambda > 0$,

$$\int_0^{+\infty} (1 - \mathrm{e}^{-\lambda u})\nu(du) = \lambda^\alpha.$$

Then we have

$$\mathbb{N}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\mathbf{1}_{\{M=n\}}\right] = \int_0^{+\infty} \mathbb{N}^{(u)}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\right]\frac{u^n}{n!}\,\mathrm{e}^{-u}\,\nu(du).$$

Using the scaling property of the stable Lévy tree (see Duquesne and Le Gall (2002) Section 3.3), we have that the law of the tree $\mathcal{T}$ under $\mathbb{N}^{(u)}$ is the same as the law of $u^{1-\alpha}\mathcal{T}$ under $\mathbb{N}^{(1)}$ where the notation $\lambda\mathcal{T}$ means that we multiply the distance that defines $\mathcal{T}$ by the factor $\lambda$ (i.e. we scale all the edge lengths by $\lambda$). Moreover, as we only look at discrete trees, this factor does not modify the tree $\tilde{T}_0$. Therefore, we get:

$$\begin{aligned}\mathbb{N}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\mathbf{1}_{\{M=n\}}\right] &= \int_0^{+\infty} \mathbb{N}^{(1)}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\right]\frac{u^n}{n!}\,\mathrm{e}^{-u}\,\nu(du)\\ &= \mathbb{N}^{(1)}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\right]\mathbb{N}[M=n].\end{aligned}$$

We deduce:

$$\begin{aligned}\mathbb{N}[\phi(\tilde{T}_0) \mid M = n] &= \mathbb{N}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big) \mid M = n\right]\\ &= \mathbb{N}^{(1)}\left[\phi\big(\tilde{T}(\mathcal{T}, F_1, \ldots, F_n)\big)\right] = \mathbb{N}^{(1)}[\phi(\tilde{T}_0) \mid M = n]\end{aligned}$$

since $\mathcal{T}$ and $M$ are independent under $\mathbb{N}^{(1)}$. □

We now consider the marks that define the pruned tree $\mathcal{T}_\theta$ and we define on the event $\{M \geq 1\}$ the tree $\hat{T}_\theta$ as the tree $\hat{T}_0$ pruned on the same marks, in other words, we set

$$\hat{T}_\theta = \hat{T}_0 \cap \mathcal{T}_\theta.$$

Let $\hat{\Pi}^{[n]}_{\text{Lévy}}$ be the restriction of $\Pi_{\text{Lévy}}$ to the $n$ first integers. By construction, if $C_\theta$ is an element of $\hat{\Pi}^{[n]}_{\text{Lévy}}(\theta)$, then there exists a leaf $x$ of $\hat{T}_\theta$ such that $x$ belongs to the sub-tree $\bigcup_{i\in C_\theta}[\![\emptyset, F_i]\!]$, and $x$ is the only leaf of $\hat{T}_\theta$ with this property. We set $C_\theta$ for the label of $x$, and we consider the order of the elements of $\tilde{\Pi}^{[n]}_{\text{Lévy}}$ given by the order of their smallest integer. We set $I_\theta = I(\hat{T}_\theta)$ for the labels of the leaves of $\hat{T}_\theta$ and $(F_i^\theta, i \in I_\theta)$ for the leaves of $\hat{T}_\theta$.

We denote by $\tilde{T}_\theta$ the skeleton of $\hat{T}_\theta$ with the labeled leaves $(F_i^\theta, i \in I_\theta)$. According to Abraham et al. (2015), Proposition 4.1, the tree $\hat{T}_\theta$ is distributed under $\mathbb{N}[\,\cdot \mid M \geq 1]$ as a continuous GW tree such that

- $\tilde{T}_\theta$ is a GW tree with offspring distribution characterized by its generating function $g_\theta$ given in (2.3) with $\alpha = 1/\gamma$.
- The lifetimes of individuals $(h_u, u \in \hat{T}_\theta)$ are independent random variable with exponential distribution with parameter $\psi'_\theta(1) = \gamma(1+\theta)^{\gamma-1}$.

The following lemma is a consequence of Theorem 6.1 of Abraham et al. (2015).

**Lemma 4.2.** *The process $(\tilde{T}_\theta, \theta \geq 0)$ is distributed under $\mathbb{N}[\,\cdot \mid M \geq 1]$ as the process $(\mathcal{P}_\theta(T), \theta \geq 0)$ under $\mathbf{P}$.*

*Proof*: Let $\theta > 0$. Theorem 6.1 of Abraham et al. (2015) describes how $\hat{T}_\theta$ is obtained from $\hat{T}_0$:

- A branching point $x$ of $\hat{T}_0$ with $k_x = k_x(\hat{T}_0)$ children is marked at time $\tau_x$ with distribution given by:
$$\mathbb{N}[\tau_x \geq \theta \mid \hat{T}_0] = -\int_\theta^{+\infty} \frac{\psi^{(k_x+1)}(1+z)}{\psi^{(k_x)}(1)} dz = \frac{\psi^{(k_x)}(1+\theta)}{\psi^{(k_x)}(1)} = \left(\frac{1}{1+\theta}\right)^{k_x-\gamma}.$$
- A branch $B$ of length $h$ is marked at time $\tau_B$ with distribution given by:
$$\mathbb{N}[\tau_B \geq \theta \mid \hat{T}_0] = \exp\left(-h \int_0^\theta \psi''(1+z)dz\right) = \mathrm{e}^{-\left(\psi'(1+\theta)-\psi'(1)\right)h}.$$

Then the tree $\hat{T}_0$ is cut according to the marks present at time $\theta$ and the tree $\hat{T}_\theta$ is the connected component that contains the root. Therefore, the tree $\tilde{T}_\theta$ is obtained from the tree $\tilde{T}_0$ by a pruning at node. A node $u \in \tilde{T}_0$ is marked if the corresponding node $C(u) \in \hat{T}_0$ is marked at time $\theta$ in the previous procedure OR the branch $B(u)$ with length $h_u$ is marked. So the node $u$ of $\tilde{T}_0$ is marked at time $\zeta_u = \tau_{C(u)} \wedge \tau_{B(u)}$ and using that the edge lengths of $\hat{T}_0$ are independent and exponentially distributed with parameter $\gamma = \psi'(1)$, we have with $k_u = k_u(\hat{T}_0)$:

$$\begin{aligned}
\mathbb{N}[\zeta_u \geq \theta \mid \tilde{T}_0] &= \mathbb{N}[\tau_{C(u)} \geq \theta \mid \tilde{T}_0]\, \mathbb{N}[\tau_{B(u)} \geq \theta \mid \tilde{T}_0] \\
&= \left(\frac{1}{1+\theta}\right)^{k_u-\gamma} \int_0^{+\infty} dh\, \gamma\, \mathrm{e}^{-\gamma h}\, \mathrm{e}^{-\left(\psi'(1+\theta)-\gamma\right)h} \\
&= \left(\frac{1}{1+\theta}\right)^{k_u-\gamma} \left(\frac{1}{1+\theta}\right)^{\gamma-1} \\
&= \left(\frac{1}{1+\theta}\right)^{k_u-1}.
\end{aligned}$$

Since the cutting time $\tau_{C(u)}$ and $\tau_{B(u)}$ are independent for all internal nodes $u$, we recover the discrete pruning procedure that defines the process $(\mathcal{P}_\theta(T), \theta \geq 0)$ under $\mathbf{P}$. To conclude notice that $\tilde{T}_0$ and $T$ are GW trees with offspring distributions characterized by their generating function $g$. □

4.4. *Proof of Theorem 1.3.* The next corollary states that the pruning procedure for stable GW trees developed in Abraham et al. (2012) and the pruning procedure for Lévy trees developed in Abraham and Delmas (2008) and applied in Abraham et al. (2015) to sub-trees with finite number of leaves coincide.

**Corollary 4.3.** *Let $n \in \mathbb{N}$. The process $(\tilde{T}_\theta, \theta \geq 0)$ is distributed under $\mathbb{N}[\,\cdot \mid M = n]$ as the process $(\mathcal{P}_\theta(T), \theta \geq 0)$ under $\mathbf{P}_n$.*

*Proof*: This is a direct consequence of Lemma 4.2 and the fact that $M = L(\tilde{T}_0)$. □

Theorem 1.3 follows directly from Theorem 1.1 and from the following corollary, which is a direct consequence of Corollary 4.3. Recall that $\hat{\Pi}^{[n]}_{\text{Lévy}}$ is the restriction of $\Pi_{\text{Lévy}}$ defined in Section 4.2.2 to the $n$ first integers.

**Corollary 4.4.** *The process $\hat{\Pi}^{[n]}_{Lévy}$ is under $\mathbb{N}^{(1)}$ distributed as $\hat{\Pi}^{[n]}_{GW}$ under $\mathbf{P}_n$.*

Using Lemma 4.2, we also have the following corollary which shows that the first coalescent event in $\hat{\Pi}^{[n]}_{\text{Lévy}}$ is not exponentially distributed.

**Corollary 4.5.** *Let $\tau_1^{(n)}$ be the first coalescent event in $\hat{\Pi}^{[n]}_{Lévy}$. Then we have for $\theta \geq 0$:*

$$\mathbb{N}^{(1)}[\tau_1^{(n)} \geq \theta] = \left(\frac{1}{1+\theta}\right)^{n-1}.$$

*Proof*: We keep the notations of the proof of Lemma 4.2. We have:

$$\begin{aligned}
\mathbb{N}^{(1)}[\tau_1^{(n)} \geq \theta] &= \mathbb{N}\left[\mathbb{N}\left[\inf_{u \in \mathcal{N}(\tilde{T}_0)} \zeta_u \geq \theta \,\Big|\, \tilde{T}_0\right] \,\Big|\, M = n\right] \\
&= \mathbb{N}\left[\prod_{u \in \mathcal{N}(\tilde{T}_0)} \left(\frac{1}{1+\theta}\right)^{k_u(\tilde{T}_0)-1} \,\Big|\, M = n\right] \\
&= \mathbb{N}\left[\left(\frac{1}{1+\theta}\right)^{M-1} \,\Big|\, M = n\right] \\
&= \left(\frac{1}{1+\theta}\right)^{n-1},
\end{aligned}$$

using (2.2) for the third equality. □

## 5. Proof of Corollary 1.7

We recall results from Haas and Miermont (2011), Corollary 1. Let $X_n$ be the number of coalescence events for a $\beta(a,b)$-coalescent. For $1 < a < 2$ and $b > 0$, we have that:

$$\frac{2-a}{\Gamma(a)} n^{a-2} X_n$$

converges in distribution towards

$$W_{a,b} = \int_0^\infty dt \;\mathrm{e}^{-(2-a)S_{a,b}(t)},$$

where $S_{a,b}$ is a subordinator with Laplace exponent $\phi_{a,b}$ given by:

$$\phi_{a,b}(\lambda) = \int_0^1 \left(1 - (1-x)^\lambda\right) x^{a-3}(1-x)^{b-1}\, dx.$$

Notice that this notation is consistent with (3.1). Since $Z_n$ is distributed as $X_n$ with $a = 1+\alpha$ and $b = 1-\alpha$, we deduce that:

$$n^{\alpha-1} Z_n \xrightarrow[n\to+\infty]{(d)} Z,$$

with $Z$ distributed as $\frac{\Gamma(1+\alpha)}{1-\alpha} W_{1+\alpha,1-\alpha}$.

Using Lemma 3.1, we compute the moments of $Z$:

$$\begin{aligned}
\mathbb{E}\left[W^n_{1+\alpha,1-\alpha}\right] &= n! \int_{0\le t_1 \le \cdots \le t_n} \mathbb{E}\left[\mathrm{e}^{-(1-\alpha)\sum_{k=1}^n S_{1-\alpha,1+\alpha}(t_k)}\right] dt_1 \cdots dt_n \\
&= n! \int_{0\le r_1,\cdots,0\le r_n} \prod_{k=1}^n \mathbb{E}\left[\mathrm{e}^{-(1-\alpha)k S_{1-\alpha,1+\alpha}(r_k)}\right] dr_1 \cdots dr_n \\
&= \frac{n!}{\prod_{k=1}^n \phi_{1+\alpha,1-\alpha}(k(1-\alpha))} \\
&= \left(\frac{1-\alpha}{\Gamma(\alpha)}\right)^n \frac{\Gamma(n+1)\Gamma(1-\alpha)}{\Gamma((n+1)(1-\alpha))}.
\end{aligned}$$

We deduce that:

$$\mathbb{E}\left[Z^n\right] = \alpha^n \frac{\Gamma(n+1)\Gamma(1-\alpha)}{\Gamma((n+1)(1-\alpha))}.$$

## 6. Number of blocks in the last coalescence event

We consider the number of blocks $B_n$ involved in the last coalescence event of $\Pi^{[n]}_{\text{dis}}$. In order to stress the dependence in $n$, we shall denote by $T_n$ the GW tree $T$ under $\mathbf{P}_n$. We also write $\xi_u(T_n)$ for $\xi_u$ to stress the dependence of the marks introduced in Section 2.2 as a function of the underlying tree $T_n$. Notice that the time $\xi_\emptyset(T_n)$ at which the root of $T_n$ is marked corresponds to the last coalescence event associated with $T_n$. Thanks to Theorem 1.1, $B_n$ is distributed as the number of leaves of the pruned tree obtained from $T_n$ just before the last coalescence event, that is:

$$B_n \overset{(\mathrm{d})}{=} L(\mathcal{P}_{\xi_\emptyset(T_n)-}(T_n)). \tag{6.1}$$

6.1. *Local limit.* The method used in Abraham and Delmas (2013a) when $\alpha = 1/2$ relies on Aldous's CRT, which is the (global) limit of $T_n$ when the length of the branches of $T_n$ are rescaled by $1/\sqrt{n}$, see Duquesne (2003). Since Lévy's trees are more difficult to handle, we choose here to use the local limit of $T_n$, which is Kesten's tree $T^*$, according to Curien and Kortchemski (2014) Theorem 3.1 or Abraham and Delmas (2014) Proposition 4.6.

Recall that $\nu_g$ is the distribution with generating function $g$ given in (1.3) and that $\nu_g$ is critical as $g'(1) = 1$. We recall the distribution of Kesten's tree $T^*$ associated with the critical reproduction law $\nu_g$, see Kesten (1986). Let $\nu_g^*$ be the corresponding size-biased distribution: $\nu_g^*(k) = k\nu_g(k)$ for all $k \in \mathbb{N}$. For $h \in \mathbb{N}$, we consider the truncation operator $r_h$ on $\mathbb{T}$ defined as:

$$r_h \mathbf{t} = \{u \in \mathbf{t}; |u| \le h\}.$$

The distribution of $T^*$ is as follows. Almost surely, $T^*$ contains a unique infinite path i.e. a unique infinite sequence $(V_k, k \in \mathbb{N}^*)$ of positive integers such that, for every $h \in \mathbb{N}$, $V_1 \cdots V_h \in T^*$, with the convention that $V_1 \cdots V_h = \emptyset$ if $h = 0$. The joint distribution of $(V_k, k \in \mathbb{N}^*)$ and $T^*$ is determined recursively as follows: for each $h \in \mathbb{N}$, conditionally given $(V_1, \ldots, V_h)$ and $r_h T^*$, we have:

- The numbers of children $(k_v(T^*),\ v \in T^*,\ |v| = h)$ are independent and distributed according to $\nu_g$ if $v \neq V_1 \cdots V_h$ and according to $\nu_g^*$ if $v = V_1 \cdots V_h$.
- Given also the numbers of children $(k_v(T^*),\ v \in T^*,\ |v| = h)$, the vertex $V_{h+1}$ is uniformly distributed on the set of integers

$$\left\{1, \ldots, \sum_{v \in T^*,\ |v|=h} k_v(T^*)\right\}.$$

We denote by $\mathbb{P}$ the distribution of $T^*$.

Recall that the height of a discrete tree $\mathbf{t} \in \mathbb{T}$ is $H_{\max}(\mathbf{t}) = \sup\{|u|, u \in \mathbf{t}\}$. The local limit convergence of critical GW trees, see Abraham and Delmas (2014), implies that, for all $h \in \mathbb{N}^*$, $\mathbf{t} \in \mathbb{T}$ with height $h$:

$$\lim_{n \to +\infty} \mathbf{P}_n(r_h T_n = \mathbf{t}) = \mathbb{P}(r_h T^* = \mathbf{t}).$$

Notice that $\mathcal{P}_\theta(T^*)$ is a.s. finite for any $\theta > 0$. By construction of the marks, we easily get that the local limit of $(\mathcal{P}_\theta(T_n), \theta \geq 0)$ is given by $(\mathcal{P}_\theta(T^*), \theta \geq 0)$. Since $k_\emptyset(T_n)$ converges in distribution to $k_\emptyset(T^*)$ (with distribution $\nu_g^*$), we deduce the convergence in distribution of the mark $\xi_\emptyset(T_n)$ to $\xi_\emptyset^*$ distributed under $\mathbb{P}$ as:

$$\mathbb{P}(\xi_\emptyset^* \geq \theta | T^*) = (1+\theta)^{1-k_\emptyset(T^*)}.$$

We deduce that the local limit in distribution of $\mathcal{P}_{\xi_\emptyset(T_n)-}(T_n)$ is given by $\mathcal{P}_{\xi_\emptyset^*-}(T^*)$.

This and the definition of $T^*$ gives the following lemma. For $\mathbf{t} \in \mathbb{T}$, and $u \in \mathbf{t}$, recall the notation $\mathbf{t}_u$ for the sub-tree attached at $u$, see (4.1).

**Lemma 6.1.** *We have, for all* $\mathbf{t} \in \mathbb{T}$*:*

$$\lim_{n \to +\infty} \mathbf{P}_n(\mathcal{P}_{\xi_\emptyset(T_n)-}(T_n) = \mathbf{t}) = \mathbb{P}(\bar{T} = \mathbf{t}),$$

*where* $\bar{T}$ *is such that:*

- $k_\emptyset(\bar{T})$ *has distribution* $\nu_g^*$.
- *Conditionally on* $k_\emptyset(\bar{T})$, $\xi$ *is a random variable such that* $\mathbb{P}(\xi \geq \theta) = (1+\theta)^{1-k_\emptyset(\bar{T})}$ *for all* $\theta \geq 0$.
- *Conditionally on* $k_\emptyset(\bar{T})$ *and* $\xi$, $V_1$ *is a uniform random variable on* $\{1, \ldots, k_\emptyset(\bar{T})\}$.
- *Conditionally on* $k_\emptyset(\bar{T})$, $\xi$ *and* $V_1$, $(\bar{T}_u, u \in \{1, \ldots, k_\emptyset(\bar{T})\})$ *are independent random trees distributed such that for* $u \neq V_1$, $T_u$ *is distributed as* $\mathcal{P}_\xi(T)$ *with* $T$ *a GW tree with offspring distribution* $\nu_g$*, and* $T_{V_1}$ *is distributed as* $\mathcal{P}_\xi(T^*)$*, with* $T^*$ *distributed as Kesten's tree associated with the reproduction law* $\nu_g$.

Notice that by construction, $\bar{T}$ is finite.

6.2. *Proof of Proposition 1.8.* We deduce from (6.1), Lemma 6.1 and the fact that $\bar{T}$ is a.s. finite, that $B_n$ converge in distribution to $B = L(\bar{T})$. From Lemma 6.1, we have that $B$ is distributed as

$$L(\mathcal{P}_\xi(T^*)) + \sum_{k=1}^{k_\emptyset - 1} L(\mathcal{P}_\xi(T_k)),$$

where $k_\emptyset$ has distribution $\nu_g^*$, $\xi$ has density $(k_\emptyset - 1)(1+\theta)^{-k_\emptyset}\mathbf{1}_{\{\theta\geq 0\}}$, $T^*$ is independent and distributed as Kesten's tree associated with $\nu_g$, and $(T_k, k\in\mathbb{N}^*)$ are independent and distributed as a Galton-Watson tree $T$ with offspring distribution $\nu_g$. We deduce that:

$$\mathbb{E}\left[r^B\right] = \mathbb{E}\left[N(N-1)\int_0^{+\infty}(1+\theta)^{-N}d\theta\,\mathbb{E}\left[r^{L_\theta}\right]^{N-1}\mathbb{E}\left[r^{L_\theta^*}\right]\right],$$

where $N$ has distribution $\nu_g$, $L_\theta$ is the number of leaves of $\mathcal{P}_\theta(T)$ and $L_\theta^*$ is the number of leaves of $\mathcal{P}_\theta(T^*)$.

Let $h_\theta$ be the generating function of $L_\theta$ and $h_\theta^*$ be the generating function of $L_\theta^*$. We have:

$$\mathbb{E}\left[r^B\right] = \int_0^{+\infty}\frac{d\theta}{(1+\theta)^2}\,g''\left(\frac{h_\theta(r)}{1+\theta}\right)h_\theta(r)h_\theta^*(r).$$

Recall that $\mathcal{P}_\theta(T)$ is a GW tree whose reproduction law has generating function $g_\theta$ given by (2.3). Similar arguments as in the proof of (2.8), yield that:

$$g_\theta(h_\theta(r)) - h_\theta(r) = g_\theta(0)(1-r). \tag{6.2}$$

We deduce from (2.3) that:

$$g_\theta''(r) = g''\left(\frac{r}{1+\theta}\right)\frac{1}{1+\theta}.$$

We deduce from (6.2) that:

$$(1-g_\theta'(h_\theta(r))) = \frac{g_\theta(0)}{h_\theta'(r)} \quad\text{and}\quad g_\theta''(h_\theta(r)) = (1-g_\theta'(h_\theta(r)))\frac{h_\theta''(r)}{(h_\theta'(r))^2}. \tag{6.3}$$

We obtain:

$$g''\left(\frac{h_\theta(r)}{1+\theta}\right)\frac{1}{1+\theta} = g_\theta(0)\frac{h_\theta''(r)}{(h_\theta'(r))^3}.$$

We now compute $h_\theta^*$. According to Remark 3.7 in Abraham et al. (2015), we have for $\mathbf{t}\in\mathbb{T}$:

$$\mathbf{P}(\mathcal{P}_\theta(T^*)=\mathbf{t}) = L(\mathbf{t})\,\frac{1-g_\theta'(1)}{g_\theta'(0)}\,\mathbf{P}(\mathcal{P}_\theta(T)=\mathbf{t}).$$

We deduce that:

$$\begin{aligned} h_\theta^*(r) = \mathbf{E}\left[r^{L_\theta^*}\right] &= \sum_{\mathbf{t}\in\mathbb{T}} r^{L(\mathbf{t})}\mathbf{P}(\mathcal{P}_\theta(T^*)=\mathbf{t}) \\ &= \frac{1-g_\theta'(1)}{g_\theta(0)}\sum_{\mathbf{t}\in\mathbb{T}} L(\mathbf{t})r^{L(\mathbf{t})}\mathbf{P}(\mathcal{P}_\theta(T)=\mathbf{t}) \\ &= r\,\frac{h_\theta'(r)}{h_\theta'(1)}, \end{aligned}$$

where we used the first equality in (6.3) with $r=1$ and $h_\theta(1)=1$. We get:

$$\mathbb{E}\left[r^B\right] = r\int_0^{+\infty}\frac{d\theta}{1+\theta}\,\frac{g_\theta(0)}{h_\theta'(1)}\,\frac{h_\theta''(r)}{(h_\theta'(r))^2}\,h_\theta(r). \tag{6.4}$$

We have from (2.3) that:

$$g_\theta(0) = \alpha(1+\theta)\left[1-\left(\frac{\theta}{1+\theta}\right)^{1/\alpha}\right].$$

We deduce from (6.2) that:

$$h_\theta(r) = (1+\theta)\left[1 - \left\{1 - r\left[1 - \left(\frac{\theta}{1+\theta}\right)^{1/\alpha}\right]\right\}^{\alpha}\right].$$

Then, the change of variable $x = 1 - (\theta/(1+\theta))^{1/\alpha}$ in (6.4) gives that $\varphi_\alpha$, given in (1.5), is the generating function of $B$.

## References

R. Abraham and J.-F. Delmas. Fragmentation associated with Lévy processes using snake. *Probab. Theory Related Fields* **141** (1-2), 113–154 (2008). MR2372967.

R. Abraham and J.-F. Delmas. A continuum-tree-valued Markov process. *Ann. Probab.* **40** (3), 1167–1211 (2012). MR2962090.

R. Abraham and J.-F. Delmas. A construction of a $\beta$-coalescent via the pruning of binary trees. *J. Appl. Probab.* **50** (3), 772–790 (2013a). MR3102514.

R. Abraham and J.-F. Delmas. The forest associated with the record process on a Lévy tree. *Stochastic Process. Appl.* **123** (9), 3497–3517 (2013b). MR3071387.

R. Abraham and J.-F. Delmas. Record process on the continuum random tree. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** (1), 225–251 (2013c). MR3083925.

R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electron. J. Probab.* **19** (2), 1–19 (2014). DOI: 10.1214/EJP.v19-2747.

R. Abraham, J.-F. Delmas and H. He. Pruning Galton-Watson trees and tree-valued Markov processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (3), 688–705 (2012). MR2976559.

R. Abraham, J.-F. Delmas and H. He. Pruning of CRT-sub-trees. *Stochastic Process. Appl.* **125** (4), 1569–1604 (2015). MR3310357.

R. Abraham, J.-F. Delmas and P. Hoscheit. Exit times for an increasing Lévy tree-valued process. *Probab. Theory Related Fields* **159** (1-2), 357–403 (2014). MR3201925.

L. Addario-Berry, N. Broutin and C. Holmgren. Cutting down trees with a Markov chainsaw. *Ann. Appl. Probab.* **24** (6), 2297–2339 (2014). MR3262504.

J. Berestycki, N. Berestycki and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *Ann. Probab.* **35** (5), 1835–1887 (2007). MR2349577.

N. Berestycki. *Recent progress in coalescent theory*, volume 16 of *Ensaios Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, Rio de Janeiro (2009). ISBN 978-85-85818-40-1. MR2574323.

J. Bertoin. Fires on trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (4), 909–921 (2012). MR3052398.

J. Bertoin and G. Miermont. The cut-tree of large Galton-Watson trees and the Brownian CRT. *Ann. Appl. Probab.* **23** (4), 1469–1493 (2013). MR3098439.

M. Birkner, J. Blath, M. Capaldo, A. Etheridge, M. Möhle, J. Schweinsberg and A. Wakolbinger. Alpha-stable branching and beta-coalescents. *Electron. J. Probab.* **10**, no. 9, 303–325 (electronic) (2005). MR2120246.

E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197** (2), 247–276 (1998). MR1652734.

N. Curien and I. Kortchemski. Random non-crossing plane configurations: a conditioned Galton-Watson tree approach. *Random Structures Algorithms* **45** (2), 236–260 (2014). MR3245291.

A. Dress, V. Moulton and W. Terhalle. $T$-theory: an overview. *European J. Combin.* **17** (2-3), 161–175 (1996). Discrete metric spaces (Bielefeld, 1994). MR1379369.

M. Drmota, A. Iksanov, M. Moehle and U. Roesler. A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms* **34** (3), 319–336 (2009). MR2504401.

T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.* **31** (2), 996–1027 (2003). MR1964956.

T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque* (281), vi+147 (2002). MR1954248.

T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** (4), 553–603 (2005). MR2147221.

S.N. Evans, J. Pitman and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields* **134** (1), 81–126 (2006). MR2221786.

C. Foucart and O. Hénard. Stable continuous-state branching processes with immigration and Beta-Fleming-Viot processes with immigration. *Electron. J. Probab.* **18**, no. 23, 21 (2013). MR3035751.

F. Freund and A. Siri-Jégousse. Minimal clade size in the Bolthausen-Sznitman coalescent. *J. Appl. Probab.* **51** (3), 657–668 (2014). MR3256218.

A. Gnedin, A. Iksanov, A. Marynych and M. Möhle. On asymptotics of the beta coalescents. *Adv. in Appl. Probab.* **46** (2), 496–515 (2014). MR3215543.

C. Goldschmidt and J.B. Martin. Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **10**, no. 21, 718–745 (electronic) (2005). MR2164028.

B. Haas and G. Miermont. Self-similar scaling limits of non-increasing Markov chains. *Bernoulli* **17** (4), 1217–1247 (2011). MR2854770.

O. Henard. The fixation line (2015). To appear in Ann of Appl. Probab.

C. Holmgren. Random records and cuttings in binary search trees. *Combin. Probab. Comput.* **19** (3), 391–424 (2010). MR2607374.

S. Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms* **29** (2), 139–179 (2006). MR2245498.

H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.* **22** (4), 425–487 (1986). MR871905.

J.F.C. Kingman. The coalescent. *Stochastic Process. Appl.* **13** (3), 235–248 (1982). MR671034.

P. Marchal. A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 489–499. Assoc. Discrete Math. Theor. Comput. Sci., Nancy (2008). MR2508809.

A. Meir and J.W. Moon. Cutting down random trees. *J. Austral. Math. Soc.* **11**, 313–324 (1970). MR0284370.

G. Miermont. Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields* **131** (3), 341–375 (2005). MR2123249.

A. Panholzer. Cutting down very simple trees. *Quaest. Math.* **29** (2), 211–227 (2006). MR2233368.

J. Pitman. Coalescents with multiple collisions. *Ann. Probab.* **27** (4), 1870–1902 (1999). MR1742892.

J. Schweinsberg. Dynamics of the evolving Bolthausen-Sznitman coalecent. *Electron. J. Probab.* **17**, no. 91, 50 (2012). MR2988406.