THE EXISTENCE OF A GIANT CLUSTER FOR PERCOLATION ON LARGE CRUMP–MODE–JAGERS TREES

G. BERZUNZA,* *Department of Mathematics, Uppsala University*

Abstract

In this paper we consider random trees associated with the genealogy of Crump–Mode–Jagers processes and perform Bernoulli bond-percolation whose parameter depends on the size of the tree. Our purpose is to show the existence of a giant percolation cluster for appropriate regimes as the size grows. We stress that the family trees of Crump–Mode–Jagers processes include random recursive trees, preferential attachment trees, binary search trees for which this question has been answered by Bertoin [7], as well as (more general) *m*-ary search trees, fragmentation trees, and median-of-($2\ell + 1$) binary search trees, to name a few, where to our knowledge percolation has not yet been studied.

Keywords: Random tree; percolation; giant component; Crump–Mode–Jagers processes

2010 Mathematics Subject Classification: Primary 60J80
Secondary 60K35; 05C05

1. Introduction and main results

Consider a graph G_n of large but finite size $n \in \mathbb{N}$ and perform Bernoulli bond-percolation with parameter $p_n \in (0, 1)$ that depends on the size of G_n (typically the size of a graph refers to its number of vertices but not necessarily). This means we first pick a finite graph and then remove each edge with probability $1 - p_n$, independently of the other edges, inducing a partition of its set of vertices into connected clusters. A natural problem in this setting is to show the existence of a giant cluster for appropriate regimes of the percolation parameter p_n, when the size of the graph grows. More precisely, we are interested in finding a supercritical p_n such that, with high probability as $n \to \infty$, there exists a cluster that is of a size comparable to the entire graph. Let us recall some known answers to this question in some important instances.

In the case of the complete graph with n vertices, a classical result due to Erdős and Rényi (see e.g. [12]) shows that for $p_n \sim c/n$ as $n \to \infty$ with $c > 1$ fixed, with high probability, there is a unique giant cluster of size close to $\theta(c)n$ where $\theta(c)$ is the unique strictly positive solution to the equation $x + \mathrm{e}^{-cx} = 1$; recall that for two sequences of real numbers $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$, we write $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to \infty$. Second, consider a uniform Cayley tree with n vertices (i.e. a tree picked uniformly at random among the n^{n-2} trees on a set of n labeled vertices). Pitman [30, 31] showed that for $1 - p_n \sim c/\sqrt{n}$ as $n \to \infty$ with a fixed $c > 0$, the sequence of sizes of the percolation clusters ranked in decreasing order and renormalized by a factor $1/n$ converges weakly as $n \to \infty$ to a random mass partition which can be described explicitly in

terms of a conditioned Poisson measure. Finally, Bertoin [7] has shown that for fairly general families of trees with n vertices, the supercritical regime corresponds to percolation parameters of the form $1 - p_n \sim c/\ell(n)$ as $n \to \infty$, where $c > 0$ fixed and $\ell(n)$ is an estimate of the height of a typical vertex in the tree structure. Roughly speaking, Bertoin [7] established that under the previous regime the size of the cluster containing the root is of order n as $n \to \infty$. The latter result includes, for instance, some important families of random trees, such as random recursive trees, preferential attachment trees, binary search trees, etc., where it is well known that $\ell(n) = \ln n$; see [15] and [16, Section 4.4].

The main purpose of this work is to investigate the same question for large random Crump–Mode–Jagers trees, or CMJ-trees for short. More precisely, CMJ-trees are the family trees (or genealogical trees) of Crump–Mode–Jagers processes, also referred to as general or age-dependent branching processes; for further details we refer to the classic book of Jagers [20]. These are general branching population models where the number of individuals can be measured or counted in many different ways: those born, those alive, or in some sub-phase of life, for instance. More generally, we can assign random characteristics or weights to each of the individuals and measure the size of the population according to those characteristics (for instance, special choices of reproduction point process and counting yield the classical Galton–Watson or Bellman–Harris processes). We postpone its formal definition to later in this work and continue by informally describing our main results. Loosely speaking, we study Bernoulli bond-percolation on CMJ-trees at the time when the total weight ('size') of the underlying CMJ-process reaches n in three different regimes:

- weakly supercritical, $1/(\ln n) \ll 1 - p_n \ll 1$,

- supercritical, $1 - p_n \sim c/(\ln n)$ for some $c > 0$ fixed, and

- strongly supercritical, $0 < 1 - p_n \ll 1/(\ln n)$.

Recall that for two sequences of real numbers $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$, we write $a_n \ll b_n$ or $b_n \gg a_n$ if and only if $a_n/b_n \to 0$ as $n \to \infty$. We show that, under standard conditions on the underlying CMJ-process, the root cluster is of order $n^{\kappa(p_n)/\alpha}$, where $\kappa(p_n) > 0$ is a function of the percolation parameter and $\alpha > 0$ is the so-called *Malthusian parameter*. We have used the same terminology as in [4], where only random recursive trees are studied. In Section 3 we shall see that several important families of random trees can be constructed as family trees of a CMJ-process stopped at a suitable time: for example, random recursive trees, preferential attachment trees, and binary search trees where the existence of a giant cluster has been shown by Bertoin [7]. On the other hand, the general nature of the CMJ-processes will allow us to provide new results on percolation for (more general) m-ary search trees [27], fragmentation trees [22], median-of-$(2\ell + 1)$ binary search trees [13], and so-called splitting trees introduced in [17], to name a few.

In the rest of the introduction we are going to describe our setting more precisely and give the exact definition of CMJ-trees. This will enable us to state our main result in Section 1.2.

1.1. Crump–Mode–Jagers trees

We start by recalling the definition of Crump–Mode–Jagers processes (CMJ-processes) whose associated family trees are called CMJ-trees. Following Jagers [20], we present a CMJ-process as a general branching process that starts with a single individual born at time 0. We use the usual Ulam–Harris notation and introduce the set of labels $\mathbb{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$, with the convention $\mathbb{N}^0 = \{\varnothing\}$. The ancestor has label $\varnothing$. An individual with label $u = (u_1, \dots, u_n) \in \mathbb{U}$

belongs to the *n*th generation and is understood to be the u_nth descendant of $(u_1, \ldots, u_{n-1})$, which is the u_{n-1}th descendant of $(u_1, \ldots, u_{n-2})$ and so on. The initial individual has a random number N of children, born at random times $(\xi_i)_{i=1}^N$ where $0 \leq N \leq \infty$ and $0 \leq \xi_1 \leq \xi_2 \leq \cdots \leq \xi_N$. Formally, we describe the birth times $(\xi_i)_{i=1}^N$ as a point process Ξ on $[0, \infty)$, that is, $\Xi = \sum_{i=1}^N \delta_{\xi_i}$ is an integer-valued random measure, where δ_t is a point mass (Dirac measure) at time $t \geq 0$; see e.g. [23]. We let $\mu(\cdot) := \mathbb{E}[\Xi(\cdot)]$ denote the intensity measure of Ξ, and write $\mu(t) := \mu([0, t]) = \mathbb{E}[\Xi([0, t])]$. In particular, we have $N = \Xi([0, \infty))$, and thus $\mu(\infty) = \mathbb{E}[N]$. Every child that is born evolves in the same way, that is, every individual u has its own copy Ξ_u of Ξ (where now ξ_i means the age of the mother when child i is born); these copies are assumed to be independent and identically distributed. We denote the time an individual u is born by σ_u. We also assume that each individual has a random lifetime $\lambda \in [0, \infty]$ (for several of our applications we assume $\lambda \equiv \infty$). Formally, we assign to each possible individual u a copy $(\Omega_u, \mathcal{F}_u, \nu_u)$ of some generic probability space $(\Omega, \mathcal{F}, \nu)$ on which we define Ξ, and possibly a random characteristic or weight ϕ. The general branching process is then defined on the product $\prod_u (\Omega_u, \mathcal{F}_u, \nu_u)$ of these probability spaces.

The simplest way to measure or monitor the evolution of the CMJ-process is to consider the process $Z = (Z(t), t \geq 0)$ of the total number of individuals that have been born up to time $t \geq 0$, i.e. the number of births in $[0, t]$. More precisely,

$$Z(t) = \sum_u \mathbf{1}_{\{\sigma_u \leq t\}}, \quad t \geq 0;$$

see e.g. [20] and [21]. Following Jagers' work on CMJ-processes (see e.g. [20], [21], [28], and [29]), it is going to be relevant to monitor the evolution of individuals that satisfy some random property, instead of the total number of births in some fixed time interval. This random property or characteristic of an individual might be unrelated or heavily dependent on its reproduction behavior. More precisely, a characteristic or weight of an individual is a random function $\phi \colon \mathbb{R}_+ \to \mathbb{R}_+$ that assigns the value $\phi(t)$ when the individual's age is $t \geq 0$. We assume that ϕ is càdlàg (we may extend ϕ to $\mathbb{R}$ by setting $\phi(t) = 0$ for $t < 0$). We assume that each individual u has its own copy ϕ_u and thus we associated each of them with a triple $(\Xi_u, \lambda_u, \phi_u)$. These triples for all individuals are independent and identically distributed. We then define the ϕ-counted process $Z^\phi = (Z^\phi(t), t \geq 0)$ by

$$Z^\phi(t) := \sum_{u \colon \sigma_u \leq t} \phi_u(t - \sigma_u), \quad t \geq 0,$$

and say that Z_t^ϕ is the total weight at time t of all individuals that have been born so far (recall that u is born at time σ_u and thus has age $t - \sigma_u$ at time t). If $\phi \equiv 1$, we have $Z^\phi = Z$. On the other hand, the characteristic $\phi = \mathbb{1}_{[0, \lambda)}$ yields the number $Z^\phi(t) = \sum_u \mathbb{1}_{\{\sigma_u \leq t < \sigma_u + \lambda_u\}}$ of individuals alive at time $t \geq 0$.

Following [19], we let $T(\infty)$ be the family tree of the entire CMJ-process or (complete) CMJ-tree. This tree is obtained from the general branching process described at the beginning of this section by ignoring the time structure. Specifically, the individuals in the population are seen as vertices where the initial individual is the root. The children of a vertex in the tree are the same as the children in the general branching process. The tree $T(\infty)$ may be infinite, which happens when the process does not die out, i.e. $Z(\infty) = \infty$. For $t \geq 0$, we let $T(t)$ be the CMJ-tree consisting of all individuals born up to time t. Note that the number of vertices at time $t \geq 0$ is given by $Z(t)$. Clearly, $T(t)$ is an unordered tree for $t > 0$. However, one could get an ordered tree by adding an additional ordering of the children of each individual. This can

be done by taking the children in order of birth, or by choosing a random order; we refer to
[19, Remark 5.1] for further details. Finally, observe that the random tree $T(t)$ has a random
size for $t > 0$ (possibly infinite). In this work we shall be mainly interested in CMJ-trees with a
given number of vertices or when some random property is fulfilled. More precisely, we have
the following definition.

Definition 1. Fix a random characteristic or weight ϕ. For $n \in \mathbb{N}$, we let

$$\tau^\phi(n) := \inf\{t \geq 0 \colon Z^\phi(t) \geq n\}$$

be the first time the total weight is at least n (as usual $\inf \varnothing = \infty$). We exclude the case $\phi \equiv 0$,
which would yield $\tau^\phi(n) = \infty$ almost surely (a.s). We then define

$$T_n^\phi := T(\tau^\phi(n)),$$

the CMJ-tree at the time when the total weight or 'size' reaches n (provided this ever happens).

Random trees T_n^ϕ defined in this way, for some CMJ-process and some weight ϕ, are the
focus of the present paper. From now on we usually refer to T_n^ϕ as the CMJ-tree, whose size is
given by $|T_n^\phi| := Z(\tau^\phi(n))$. If $\phi \equiv 1$, then T_n^ϕ is the family tree of a CMJ-process stopped when
its number of vertices is greater than n. In particular, if the birth times have continuous dis-
tributions and there are no twins, then a.s. no two vertices are born simultaneously. Therefore
$|T_n^\phi| = n$.

Note that T_n^ϕ could be an infinite random tree, or also the time $\tau^\phi(n)$ could be infinite. In
order to avoid such possibilities, we only study cases where $Z^\phi(t) < \infty$ for every finite $t \geq 0$,
but $Z(\infty) = \infty$. In this direction, we define the Laplace transform of a function f on $[0, \infty)$ by

$$\hat{f}(\theta) = \theta \int_0^\infty e^{-\theta t} f(t)\, dt, \quad \theta > 0,$$

and the Laplace transform of a measure ν on $[0, \infty)$ by

$$\hat{\nu}(\theta) = \int_0^\infty e^{-\theta t} \nu(dt), \quad -\infty < \theta < \infty. \tag{1}$$

Throughout this work and unless we specify otherwise, we make the following assumptions.

Assumption 1. *We consider CMJ-processes that satisfy the following.*

(A1) $\mu(\{0\}) = \mathbb{E}[\Xi(\{0\})] < 1$. *This excludes a trivial case with explosions at the start (in our
examples, $\mu(\{0\}) = 0$).*

(A2) μ *is not concentrated on any lattice* $h\mathbb{Z}$, $h > 0$.

(A3) $N \geq 1$ *a.s. (in this case every individual has at least one child, so the process never dies
out and $Z(\infty) = \infty$).*

(A4) *There exists $\theta_1 > 0$ such that $\hat{\mu}(\theta_1) \in (1, \infty)$. Note that $\hat{\mu}(\cdot)$ is monotone decreasing on
(θ_1, ∞), and that $\hat{\mu}(\theta) \to 0$, as $\theta \to \infty$, by the dominated convergence theorem. Thus
there exists a real number $\alpha > 0$ (the Malthusian parameter) such that $\hat{\mu}(\alpha) = 1$.*

(A5) *For θ_1 as in (A4), we have $\mathrm{Var}\,(\hat{\Xi}(\theta_1)) < \infty$.*

(A6) *The random variable* $\sup_{t \geq 0} (e^{-\theta_2 t} \phi(t))$ *has finite expectation for some* $0 < \theta_2 < \alpha$.

(A7) Var $(\phi(t))$ *is bounded in finite intervals. Furthermore, there exists* $0 < \theta_3 \leq 2\alpha$ *such that* $\lim_{t \to \infty} e^{-\theta_3 t} \text{Var}(\phi(t)) = 0$.

Observe that **(A1)**–**(A5)** are conditions on the general branching process, while **(A6)**–**(A7)** are conditions on the characteristic ϕ. The following result shows that T_n^ϕ is well-defined.

Proposition 1. *Under assumptions* **(A1)**–**(A4)** *and for any characteristic* ϕ *satisfying* **(A6)**, *we have the following*.

(i) $\lim_{t \to \infty} Z^\phi(t) = \infty$ *almost surely. Thus a.s.* $\tau^\phi(n) < \infty$ *for every* $n \geq 0$ *and* T_n^ϕ *is a well-defined finite random tree.*

(ii) $\lim_{n \to \infty} |T_n^\phi|/n = 1/\mathbb{E}[\hat{\phi}(\alpha)] \in (0, \infty)$ *a.s., and* $\lim_{n \to \infty} \tau_n^\phi/(\ln n) = 1/\alpha$ *almost surely.*

Proof. See [19, Theorem 5.12]. □

We end this section by making a few remarks on our assumptions. Note that **(A4)** implies that $\mathbb{E}[N] > 1$ (this is known as the supercritical case). Note also that **(A4)** implies that $\mu(t) < \infty$ for every $0 \leq t < \infty$. However, $\mu(\infty) = \mathbb{E}[N]$ may be infinite. Furthermore, this condition also implies that $Z(t)$ and $\mathbb{E}[Z(t)]$ are finite for every $0 \leq t < \infty$; see e.g. [20, Theorem 6.3.3]. Finally, we do not really need the assumption $N \geq 1$ in **(A3)**; it suffices that $\mathbb{E}[N] > 1$. In this case the extinction probability $\mathbb{P}(Z(\infty) < \infty) < 1$, so there is a positive probability that the process is infinite, and Proposition 1 and the results below hold conditioned on the event $\{Z(\infty) = \infty\}$ (this is a standard setting in [21], [28], and [29]).

1.2. Main results

We now consider Bernoulli bond-percolation with parameter $p_n \in (0, 1)$ on the CMJ-tree T_n^ϕ with given weight ϕ (recall Definition 1). Following the idea of [10], we incorporate Bernoulli bond-percolation on the growth algorithm of the random tree process $(T(t), t \geq 0)$ in a dynamic way and stop at the time $\tau^\phi(n)$. This will lead us to interpret Bernoulli bond-percolation in terms of neutral mutations which are superposed to the structure of the CMJ-process and that appear at the birth events. More precisely, at each birth event, independently of all other individuals, the newborn is a clone of its parent with probability p_n or a mutant with probability $1 - p_n$. The mutations are considered to be neutral, that is, the behavior (reproduction laws and lifetimes) of the individuals is the same regardless of whether they are clones or mutants. A mutation event corresponds to the insertion of an edge in $T(t)$ that is immediately destroyed. This creates a new percolation cluster (with one vertex or individual) that grows following the same dynamic. We write $T^{(p_n)}(t)$ for the resulting combinatorial structure at time $t \geq 0$. That is, $T^{(p_n)}(t)$ has the same set of vertices as $T(t)$ and its set of intact edges is a subset of the edges of $T(t)$. Thus the connected clusters of $T^{(p_n)}(t)$ are the subtrees of $T(t)$ formed by the subsets of vertices which can be connected by a path of intact edges.

In this work we are interested in the evolution of the percolation cluster that contains the root. In this direction, we write $T_{\varnothing}^{(p_n)}(t)$ for the subtree of $T(t)$ at time $t \geq 0$ that contains the progenitor of the entire population at time 0. It should be clear that the sub-population with

the ancestral type is a CMJ-process whose generic birth process, denoted by $\Xi^{(p_n)}$, has intensity measure given by

$$\mu^{(p_n)}(\mathrm{d}t) := p_n \mu(\mathrm{d}t), \tag{2}$$

where μ is the intensity measure of the birth process Ξ of the original CMJ-process. This is a consequence of the thinning property of point processes. For a characteristic ϕ, we let

$$Z_{\varnothing}^{(p_n),\phi} = (Z_{\varnothing}^{(p_n),\phi}(t), t \geq 0)$$

denote the ϕ-counted process associated with the (clonal) CMJ-process of the sub-population bearing the same type as the initial individual. In particular, if $\phi \equiv 1$, then $Z_{\varnothing}^{(p_n),\phi} = Z_{\varnothing}^{(p_n)} = (Z_{\varnothing}^{(p_n)}(t), t \geq 0)$ counts the number of vertices in the root cluster. Clearly, if $p_n \equiv 1$, we recover the original CMJ-process. Let

$$T_{\varnothing}^{n,\phi} := T_{\varnothing}^{(p_n)}(\tau^{\phi}(n))$$

be the sub-tree which contains the original root of the CMJ-tree T_n^{ϕ} (associated with the weight ϕ) after performing percolation of parameter $p_n \in (0, 1)$. Recall that $\tau^{\phi}(n) := \inf\{t \geq 0 : Z^{\phi}(t) \geq n\}$ is the first time that the total weight or 'size' of the tree process $(T(t), t \geq 0)$ is at least n. Therefore the size of the root percolation cluster is given by

$$|T_{\varnothing}^{n,\phi}| := Z_{\varnothing}^{(p_n)}(\tau^{\phi}(n)).$$

We now turn to the statement of our main result, Theorem 1. Recall that we consider the regimes weakly supercritical, supercritical, and strongly supercritical of $p_n \in (0, 1)$, with $p_n \to 1$ as $n \to \infty$. First, we need to introduce some notation that we will use in the rest of the work. Note that (**A4**) implies that there exists $n^* \in \mathbb{N}$ such that for $n \geq n^*$ there is $\alpha_{p_n} > 0$ (the Malthusian parameter of $\mu^{(p_n)}$) such that $\hat{\mu}^{(p_n)}(\alpha_{p_n}) = 1$. We write

$$\bar{\mu}(\alpha) := \int_0^{\infty} t\, e^{-\alpha t} \mu(\mathrm{d}t), \tag{3}$$

which is finite and strictly positive due to our assumptions; see Remarks 1 and 2 below.

Theorem 1. *Let ϕ be any characteristic that does not depend on p_n. Under assumptions* (**A1**)–(**A7**), *we have*

$$\lim_{n \to \infty} n^{-\alpha_{p_n}/\alpha} |T_{\varnothing}^{n,\phi}| = \mathbb{E}[\hat{\phi}(\alpha)], \quad \text{in probability},$$

where $\alpha - \alpha_{p_n} \sim (1 - p_n)\bar{\mu}(\alpha)^{-1}$ as $n \to \infty$. In particular:

(i) *in the weakly supercritical regime, $\lim_{n\to\infty} n^{-1}|T_{\varnothing}^{n,\phi}| = 0$ in probability,*

(ii) *in the supercritical regime, $\lim_{n\to\infty} n^{-1}|T_{\varnothing}^{n,\phi}| = e^{-c/(\alpha\bar{\mu}(\alpha))}\mathbb{E}[\hat{\phi}(\alpha)]$ in probability,*

(iii) *in the strongly supercritical regime, $\lim_{n\to\infty} n^{-1}|T_{\varnothing}^{n,\phi}| = \mathbb{E}[\hat{\phi}(\alpha)]$ in probability.*

The parameters α and α_{p_n} are difficult to compute explicitly. Nevertheless, in the supercritical and strongly supercritical regimes, we note that $n^{\alpha_{p_n}\alpha^{-1}} \sim n^{1-(1-p_n)(\alpha\bar{\mu}(\alpha))^{-1}}$ as $n \to \infty$. We also note that in the weakly supercritical regime, the size of the root cluster is $o(n)$ whereas in the other regimes it is of order n.

On the other hand, in the supercritical regime (i.e. $1 - p_n \sim c/\ln n$ as $n \to \infty$, with $c > 0$ fixed), we find that Theorem 1 (ii) agrees with [7, Theorem 1]. However, it must be pointed out

that Theorem 1 (ii) cannot always be deduced from [7, Theorem 1]. More precisely, note that [7, Theorem 1] considers size as the number of vertices in the tree structure, whereas the size of a CMJ-tree depends on the underlying characteristic or weight (recall Definition 1) that does not always coincide with the number of vertices, for instance m-ary search trees, fragmentation trees, or the so-called splitting trees where the notion of 'size' is different; see Section 3 below for details. Then the arguments used in [7] do not always apply in our setting except in particular cases, for example random recursive trees, preferential attachment trees, or binary search trees that agree with a type of CMJ-tree when the correct birth process, lifespan, and characteristic are chosen. Moreover, Bertoin [7] did not treat the strongly and weakly supercritical regimes as we will do in this work. Therefore Theorem 1 may be seen as complementary to (or an extension of) the results in [7].

Inspired by Bertoin and Uribe Bravo [10], our approach relies on the connection between CMJ-processes with neutral mutations and Bernoulli bond-percolation. This leads us to investigate the asymptotic behavior of a CMJ-process with neutral mutations up to a large random time, in certain regimes when the small mutation parameter is related to the size of the total population. In [10], the authors connected Bernoulli bond-percolation in preferential attachment trees with a Markovian system of branching processes with neutral mutations. This is clearly not the case here, since it is well known that CMJ-processes are not always Markovian. Thus we have to use different tools, although some guidelines are similar to [10]. We stress that similar connections with systems of (Markovian) branching processes have been used before to study percolation on random recursive trees [4, 5] and m-ary random increasing trees [11].

This work leaves open some natural questions that we plan to investigate in the future. One can consider estimating the sizes of the largest clusters which do not contain the root. In this work we restrict ourselves to the root cluster because the absence of the Markov property makes the analysis much harder. This is not the case in [10] and [4], where the connection with a Markovian branching system with neutral mutations is used to answer this question for random recursive trees and preferential attachment trees. We refer also to Bertoin [9], where this question has been answered for random recursive trees by using a different approach. The second direction of future work would be to analyze the fluctuations of the giant component that we expect to be non-Gaussian as for random recursive trees [8], preferential attachment trees, and m-ary random increasing trees [11]. Finally, it would be interesting to estimate the size of the largest percolation clusters in the sub-critical regime, i.e. $1 - p_n \gg c / \ln n$ as $n \to \infty$ and $c > 0$ is fixed; see e.g. [5], where the case of the random recursive tree has been studied.

The rest of this paper is organized as follows. In Section 2 we prove our main result. Section 3 is devoted to the application of Theorem 1 to important families of random trees that can be constructed via CMJ-processes. Finally, the key results used in the proof of Theorem 1 are proved in Section A. More precisely, we investigate the asymptotic behavior of CMJ-processes with mutations and deduce some crucial properties that may be of independent interest.

2. Proof of Theorem 1

In this section we prove our main result, Theorem 1. Recall that the size of the root cluster is related to the clonal CMJ-process with generic birth process $\Xi^{(p_n)}$ whose intensity measure $\mu^{(p_n)}$ is given in (2). The starting point is to investigate the asymptotic behavior of the ϕ-counted clonal process $Z_{\varnothing}^{(p_n),\phi} = (Z_{\varnothing}^{(p_n),\phi}(t), t \geq 0)$ as $p_n \to 1$ and $t \to \infty$. The approach relies crucially on the use of a remarkable martingale that can be found in the work of Nerman [28]. More precisely, we improve the results in [28] and [21] on the convergence of Nerman's

martingale in order to hold uniformly in the percolation parameter (see Lemmas 2 and 3 below). Then these results and further remarks are put together to conclude with the proof of Theorem 1.

Recall that (**A4**) implies that there exists $n^* \in \mathbb{N}$ such that for $n \geq n^*$ there is $\alpha_{p_n} > 0$ (the Malthusian parameter of $\mu^{(p_n)}$) such that $\hat{\mu}^{(p_n)}(\alpha_{p_n}) = 1$. This implies that $\mu^{(p_n)}(\infty) = p_n \mathbb{E}[N] > 1$ (i.e. the clonal process is supercritical). Since $\alpha_{p_n} \to \alpha$ as $p_n \to 1$, we choose n^* large enough such that $0 < \theta_1 < \inf_{n \geq n^*} \alpha_{p_n}$, where θ_1 satisfies (**A4**)–(**A5**). Moreover, consider n^* even larger such that $0 < \theta_2 < \inf_{n \geq n^*} \alpha_{p_n}$ and $0 < \theta_3 < 2 \inf_{n \geq n^*} \alpha_{p_n}$, where θ_2 and θ_3 satisfy (**A6**) and (**A7**), respectively.

Remark 1. Note for future reference that $\mu_\alpha^{(p_n)}(\mathrm{d}t) := \mathrm{e}^{-t\alpha_{p_n}} \mu^{(p_n)}(\mathrm{d}t)$, for $n \geq n^*$, is a probability measure concentrated on $(0, \infty)$. Moreover, condition (**A4**) implies that

$$\bar{\mu}^{(p_n)}(\alpha_{p_n}) := \int_0^\infty t \mu_\alpha^{(p_n)}(\mathrm{d}t) < \infty. \tag{4}$$

We write $W_\varnothing^{(p_n),\phi} = (W_\varnothing^{(p_n),\phi}(t), t \geq 0)$ for the process given by

$$W_\varnothing^{(p_n),\phi}(t) := \mathrm{e}^{-t\alpha_{p_n}} Z_\varnothing^{(p_n),\phi}(t), \quad t \geq 0.$$

For $p_n \equiv 1$, we sometimes remove the superscript (p_n) and the subscript $\varnothing$ from the previous notations. That is, we write $W^\phi = (W^\phi(t), t \geq 0)$ for the process given by $W^\phi(t) := \mathrm{e}^{-t\alpha} Z^\phi(t)$. For $n \geq n^*$, consider the characteristic

$$\psi^{(p_n)}(t) = \mathbb{1}_{\{t \geq 0\}} \, \mathrm{e}^{t\alpha_{p_n}} \int_t^\infty \mathrm{e}^{-s\alpha_{p_n}} \, \Xi^{(p_n)}(\mathrm{d}s), \quad t \geq 0. \tag{5}$$

The next lemma shows that $\psi^{(p_n)}$ satisfies the conditions (**A6**)–(**A7**).

Lemma 1. *Assume that conditions* (**A1**)–(**A5**) *are fulfilled. Then we have the following.*

  (i) $\sup_{n \geq n^*} \sup_{t \geq 0} \mathrm{e}^{-\theta_1 t} \psi^{(p_n)}(t)$ *has finite expectation, for θ_1 as in* (**A4**).

  (ii) $\sup_{n \geq n^*} \mathrm{Var}\,(\psi^{(p_n)}(t))$ *is bounded in finite intervals. Furthermore, there exists $0 < \theta \leq 2 \inf_{n \geq n^*} \alpha_{p_n}$ such that $\lim_{t \to \infty} \sup_{n \geq n^*} \mathrm{e}^{-\theta t} \mathrm{Var}\,(\psi^{(p_n)}(t)) = 0$.*

Proof. For $t \geq 0$ and $0 < \theta < \alpha_{p_n}$, we note that $\psi^{(p_n)}(t) \leq \mathrm{e}^{t\theta} \int_t^\infty \mathrm{e}^{-s\theta} \Xi(\mathrm{d}s) \leq \mathrm{e}^{t\theta} \hat{\Xi}(\theta)$. This inequality and conditions (**A4**)–(**A5**) imply our claim. □

Henceforth, and for the sake of simplicity, we omit the superscript (p_n) from $\psi^{(p_n)}$ and only write ψ for the characteristic defined in (5).

It is well known that the process $W_\varnothing^{(p_n),\psi}$ is a nonnegative square-integrable martingale whose terminal value will be denoted by $W_\varnothing^{(p_n),\psi}(\infty)$. Furthermore, $W_\varnothing^{(p_n),\psi}(\infty) \geq 0$ almost surely (see [28, Proposition 2.4] for the proof of the martingale property and [21, Theorem 4.1 and Corollary 4.2] for the convergence result). In particular, if $p_n \equiv 1$, [21, Corollary 4.2] and condition (**A3**) imply that $W^{(p_n),\psi}(\infty) = W^\psi(\infty) > 0$ almost surely. An important result established by Nerman [28] and Jagers and Nerman [21, Theorem 4.3] (see also Jagers [20, Section 6.10]) implies that, for $n \geq n^*$,

$$\lim_{t \to \infty} W_\varnothing^{(p_n),\phi}(t) = m_\infty^{(p_n),\phi} W_\varnothing^{(p_n),\psi}(\infty), \tag{6}$$

almost surely and in $L_2(\mathbb{P})$, where

$$m_\infty^{(p_n),\phi} := \frac{\mathbb{E}[\hat{\phi}(\alpha_{p_n})]}{\alpha_p \bar{\mu}^{(p_n)}(\alpha_{p_n})} < \infty, \tag{7}$$

with $\bar{\mu}^{(p_n)}(\alpha_{p_n})$ defined in (4).

Remark 2. Note that the previous convergence implies that $\bar{\mu}^{(p_n)}(\alpha_{p_n}) > 0$. Furthermore, $m_\infty^{(p_n),\phi} > 0$ (or equivalently, $\mathbb{E}[\hat{\phi}(\alpha_{p_n})] > 0$) whenever $\phi \not\equiv 0$ almost surely.

Remark 3. In particular, a simple computation shows that $m_\infty^{(p_n),\psi} = 1$, where ψ is defined in (5); see e.g. [21, Theorem 4.1].

The next two results are the key ingredients in the proof of Theorem 1. The first lemma shows that the $L_2(\mathbb{P})$ convergence of the square-integrable martingale $W_\varnothing^{(p_n),\psi}$ holds uniformly for $n \geq n^*$. Furthermore, it shows that the $L_2(\mathbb{P})$ convergence in (6) also holds uniformly for $n \geq n^*$. The second lemma establishes an even stronger convergence result by showing that the convergence remains true as $t \to \infty$ and $p_n \to 1$.

Lemma 2. *Assume that conditions* (**A1**)–(**A5**) *are fulfilled. We have*

$$\lim_{t \to \infty} \sup_{n \geq n^*} \mathbb{E}\left[\sup_{s \geq t} |W_\varnothing^{(p_n),\psi}(s) - W_\varnothing^{(p_n),\psi}(\infty)|^2 \right] = 0,$$

where ψ is defined in (5). *Furthermore, if ϕ is a characteristic that does not depend on p_n and that satisfies* (**A6**)–(**A7**), *then we have*

$$\lim_{t \to \infty} \sup_{n \geq n^*} \mathbb{E}[|W_\varnothing^{(p_n),\phi}(t) - m_\infty^{(p_n),\phi} W_\varnothing^{(p_n),\psi}(\infty)|^2] = 0.$$

Lemma 3. *Assume that conditions* (**A1**)–(**A5**) *are fulfilled. For a characteristic ϕ that does not depend on p_n and that satisfies* (**A6**)–(**A7**), *we have*

$$\lim_{n \to \infty, t \to \infty} \mathbb{E}[|W_\varnothing^{(p_n),\phi}(t) - m_\infty^\phi W^\psi(\infty)|^2] = 0,$$

where ψ is defined in (5) *with $p_n \equiv 1$ (the limit must be understood as a double limit).*

The proofs of these lemmas are rather technical and it is convenient to postpone their proofs to the Appendix. We will then finish the proof of Theorem 1, but first we need the next result.

Lemma 4. *Assume that conditions* (**A1**)–(**A4**) *are fulfilled. We have $\alpha - \alpha_{p_n} \sim (1 - p_n) \bar{\mu}(\alpha)^{-1}$, as $n \to \infty$, where $\bar{\mu}(\alpha)$ is defined in* (3).

Proof. Recall the definition of $\hat{\mu}(\cdot)$ in (1). Recall also that (**A4**) implies that $\hat{\mu}(\cdot)$ is a continuous monotone decreasing function on (θ_1, ∞), where θ_1 is defined in (**A4**). Furthermore, (**A4**) and the dominated convergence theorem show that $\hat{\mu}(\cdot)$ is differentiable with continuous derivative given by

$$\hat{\mu}'(\theta) := \frac{\mathrm{d}}{\mathrm{d}s}\hat{\mu}(s)\bigg|_{s=\theta} = -\int_0^\infty t\,\mathrm{e}^{-\theta t}\mu(\mathrm{d}t) \quad \text{for } \theta \in (\theta_1, \infty). \tag{8}$$

Since $p_n \in (0, 1)$, we have $\alpha_{p_n} < \alpha$ for $n \geq n^*$. Then, for $n \geq n^*$, the mean value theorem implies that there exists $\varepsilon_n \in (\alpha_{p_n}, \alpha)$ such that

$$\hat{\mu}'(\varepsilon_n) = \frac{\hat{\mu}(\alpha) - \hat{\mu}(\alpha_{p_n})}{\alpha - \alpha_{p_n}}.$$

Recall that (**A4**) implies that $\hat{\mu}(\alpha) = 1$ and $\hat{\mu}(\alpha_{p_n}) = 1/p_n$. Moreover, we have $0 < \bar{\mu}(\alpha) < -\hat{\mu}'(\varepsilon_n) < \infty$, where $\bar{\mu}(\alpha)$ is defined in (3). Hence

$$\alpha - \alpha_{p_n} = -\hat{\mu}'(\varepsilon_n)^{-1} \frac{1 - p_n}{p_n}.$$

Finally, we use the continuity of $\hat{\mu}'(\cdot)$, since $\alpha_{p_n} \to \alpha$, as $n \to \infty$, and $-\hat{\mu}'(\alpha) = \bar{\mu}(\alpha)$. □

Proof of Theorem 1. We deduce from Lemmas 2 and 3 that

$$\lim_{n \to \infty} e^{-\alpha \tau_n^\phi} Z(\tau_n^\phi) = \lim_{n \to \infty} e^{-\alpha_{p_n} \tau_n^\phi} Z_\varnothing^{(p_n)}(\tau_n^\phi) = m_\infty W^\psi(\infty),$$

in probability, where ψ is defined in (5) with $p_n \equiv 1$ and $m_\infty = (\alpha \bar{\mu}(\alpha))^{-1} \in (0, \infty)$; see Remark 2. Proposition 1 implies that $\lim_{n \to \infty} n^{-1} Z(\tau_n^\phi) = \mathbb{E}[\hat{\phi}(\alpha)]$ a.s. Then

$$\lim_{n \to \infty} n^{\alpha_{p_n}/\alpha} e^{-\alpha_{p_n} \tau_n^\phi} = m_\infty \mathbb{E}^{-1}[\hat{\phi}(\alpha)] W^\psi(\infty).$$

Since $W^\psi(\infty) > 0$ almost surely (by (**A3**) together with [21, Corollary 4.2]), we obtain that

$$\lim_{n \to \infty} n^{-\alpha_{p_n}/\alpha} Z_\varnothing^{(p_n)}(\tau_n^\phi) = \mathbb{E}[\hat{\phi}(\alpha)],$$

in probability. Therefore our claim follows by Lemma 4. □

3. Applications

In this section we use Theorem 1 to deduce known and new results on the existence of a giant percolation cluster in the supercritical regime for several families of trees. We begin in Subsection 3.1 with some examples where the result has been established in [7]. In Subsections 3.2, 3.3, 3.4, and 3.5 we present several new examples.

3.1. General preferential attachment trees

We introduce the procedure studied by Rudas, Tóth, and Valkó [35] and Rudas and Tóth [34] to grow a so-called general preferential attachment tree. Fix a sequence of nonnegative weights $\mathbf{w} = (w_k)_{k=0}^\infty$ with $w_0 > 0$. Start the construction from a unique tree with a single vertex and build a random tree $T_n^{(\mathbf{w})}$ with n vertices recursively as follows. Suppose that $T_n^{(\mathbf{w})}$ has been constructed for $n \geq 1$, and for every vertex $v \in T_n^{(\mathbf{w})}$ let $d_n^+(v)$ denote its outdegree. Given $T_n^{(\mathbf{w})}$, the tree $T_{n+1}^{(\mathbf{w})}$ is derived from $T_n^{(\mathbf{w})}$ by incorporating a new vertex u and creating an edge between u and a vertex $v_n \in T_n^{(\mathbf{w})}$ chosen at random according to the law

$$\mathbb{P}(v_n = v \mid T_n^{(\mathbf{w})}) = w_{d_n^+(v)} \left(\sum_{v'} w_{d_n^+(v')} \mathbb{1}_{\{v' \in T_n^{(\mathbf{w})}\}} \right)^{-1} \quad \text{for } v \in T_n^{(\mathbf{w})}.$$

We point out that different choices of the sequence $\mathbf{w}$ yield well-known families of trees. Some of these families are summarized in Table 1; see [36], [1], [26], and [32] for background.

TABLE 1: Examples of general preferential attachment trees.

	$\mathbf{w} = (w_k)_{k=0}^{\infty}$
Random recursive tree	$w_k = 1$ for all $k \geq 0$
Binary search tree	$w_0 = 2$, $w_1 = 1$ and $w_k = 0$ for $k \geq 2$
m-ary increasing tree ($m \geq 2$)	$w_k = m - k$, for $k = 0, 1, \ldots, m - 1$, and $w_k = 0$ for $k \geq m$
Linear preferential attachment	$w_k = \beta k + \rho$, where $\beta \in \{-1, 0, 1\}$, and $\rho \in \mathbb{R}_+ \setminus \{0\}$
Binary pyramid	$w_0 = w_1 = 1$ and $w_k = 0$ for $k \geq 2$.

Preferential attachment trees can be constructed via CMJ-processes as in Definition 1. More precisely, consider the characteristic (or weight) $\phi \equiv 1$ and a CMJ-process with birth times $\xi_i = \sum_{k=1}^{i} X_k$ for $i \in \mathbb{N} \cup \{0\}$ (with the convention $\xi_0 = \sum_{k=1}^{0} X_k = 0$), where $X_i = \xi_i - \xi_{i-1}$, for $i \in \mathbb{N}$, are independent random variables and distributed according as an exponential random variable of parameter w_{i-1}. In other words, we find that the process $(\Xi([0, t]), t \geq 0)$ is a pure birth process starting at 0 with birth rate w_k when the state is $k \in \mathbb{N} \cup \{0\}$. In this example, the lifetime of the individuals $\lambda \equiv \infty$. Henceforth we assume that the pure birth process $(\Xi([0, t]), t \geq 0)$ is non-explosive, that is,

$$\sum_{k=0}^{\infty} \frac{1}{w_k} = \infty; \tag{9}$$

see [2] for details. This implies that each individual in the CMJ-process has a.s. a finite number of children in each finite interval. Furthermore, note that $\hat{\Xi}(\theta) = \sum_{k=1}^{\infty} e^{-\theta \xi_k}$ for $\theta > 0$. Then

$$\hat{\mu}(\theta) = \sum_{k=1}^{\infty} \mathbb{E}[e^{-\theta \xi_k}] = \sum_{k=1}^{\infty} \prod_{i=1}^{k} \mathbb{E}[e^{-\theta X_i}] = \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{1}{1 + \theta/w_i}.$$

Assume that

$$\text{there exists } \varepsilon_1 > 0 \text{ such that } 1 < \hat{\mu}(\varepsilon_1) < \infty. \tag{E1}$$

This implies that $w_1 > 0$ and that the condition of non-explosion (9) is fulfilled. At the same time, it should be plain that condition (**A4**) is verified. Thus the Malthusian parameter exists, that is, there exists $\alpha > \varepsilon_1$ such that $\hat{\mu}(\alpha) = 1$. Finally, we further assume that

$$\text{Var} \left(\hat{\Xi}(\varepsilon_1) \right) < \infty. \tag{E2}$$

Hence the conditions (**A1**)–(**A7**) are satisfied and Theorem 1 implies the following result.

Corollary 1. *In the supercritical regime, under* (**E1**)–(**E2**)*, we have*

$$\lim_{n \to \infty} n^{-1} |T_n^{(\mathbf{w})}| = e^{-c/(\alpha \bar{\mu}(\alpha))} \quad \text{in probability.}$$

Finally, Corollary 1 allows us to recover some of the cases studied in [7]; see [19, Section 6] for details of the calculations. See Table 2.

3.2. The m-ary search tree

The m-ary search trees, where $m \geq 2$ is a fixed integer, were first introduced in [27]. In particular, $m = 2$ corresponds to the binary search tree described in Table 1 (Section 3.1). An m-ary search tree is an m-ary tree constructed recursively from a sequence of keys (real

TABLE 2: Some values of $\hat{\mu}(\theta)$, α, and $\bar{\mu}(\alpha)$ for several distributions.

	$\hat{\mu}(\theta)$	α	$\bar{\mu}(\alpha)$
Random recursive tree	$\dfrac{1}{\theta}, \theta > 0$	1	1
Binary search tree	$\dfrac{2}{\theta + 1}, \theta > -1$	1	$\dfrac{1}{2}$
m-ary increasing tree ($m \geq 2$)	$\dfrac{m}{\theta + 1}, \theta > -1$	$m - 1$	$\dfrac{1}{m}$
Linear preferential attachment	$\dfrac{\rho}{\theta - \beta}, \theta > \beta$	$\beta + \rho$	$\dfrac{1}{\rho}$
Binary pyramid	$\dfrac{1}{1 + \theta} + \dfrac{1}{(1 + \theta)^2}, \theta > -1$	$\dfrac{\sqrt{5} - 1}{2}$	$\dfrac{4\sqrt{5}}{(1 + \sqrt{5})^2}$

numbers), where each vertex stores up to $m - 1$ keys. More precisely, we start from a tree containing just an empty vertex (the root). Assume that the keys are i.i.d. random variables with a continuous distribution on $\mathbb{R}$. Then add keys one by one until the $(m - 1)$th key is placed in the root (i.e. the root becomes full) and add m new empty vertices as children of the root. Furthermore, the $m - 1$ keys in the root divide the set of real numbers into m intervals $I_1, \ldots, I_m$ that we associate with each of the m children of the root. Then each further key is passed to one of the children of the root depending on which interval it belongs to, that is, a key in I_i is stored in the ith child. Finally, we continue by iterating this procedure in an obvious way every time a vertex becomes full. This construction yields the extended m-ary search tree. In this setting, the vertices containing at least one key are called internal and the empty vertices are called external. In this work we have decided to eliminate the external vertices and consider the tree consisting of the internal nodes only. This is the m-ary search tree (in any case our results also apply to extended m-ary search tree). We also consider m-ary search trees with a fixed number of keys, say $n \in \mathbb{N}$. In other words, we stop the previous procedure at the time when the nth key is added and let $T_n^{(m)}$ denote the resulting (random) m-ary search tree with n keys. Note that the number of vertices of $T_n^{(m)}$ is actually random.

Following [19, Section 7.2], we can construct $T_n^{(m)}$ as the family tree of a CMJ-process. Consider a continuous-time version of the construction procedure of an m-ary search tree and start with one vertex (the root) with a single key. The root acquires more keys after successive independent waiting times $Y_2, \ldots, Y_{m-1}$, where Y_i is an exponential random variable of parameter $i \in \{2, \ldots, m\}$. At the arrival of the $(m - 1)$th key, at time $\sum_{i=2}^{m-1} Y_i$ (with the convention that the sum is equal to 0 when $m = 2$), the root gets m children with one key for each of them, marked by $1, \ldots, m$, with the child k born after a further waiting time X_k, i.e. at time $\sum_{i=2}^{m-1} Y_i + X_k$, where $X_1, \ldots, X_m$ are independent and exponentially distributed random variables of parameter 1. Finally, we continue growing the tree in an obvious way. Clearly, the CMJ-process associated with an m-ary search tree possesses birth times $\xi_k = \sum_{i=2}^{m-1} Y_i + X_k$, for $k = 1, \ldots, m$ (note that $N = m$ is non-random), and lifetime of each individual $\lambda \equiv \infty$. By considering the characteristic

$$\phi_m(t) = k \quad \text{for} \quad \sum_{i=2}^{k} Y_i \leq t < \sum_{i=2}^{k+1} Y_i, \quad \text{and} \quad k = 1, 2, \ldots, m - 1, \; t \geq 0,$$

we see that $T_n^{\phi_m}$, in Definition 1, is a random m-ary search tree with n keys. Note that

$$\hat{\Xi}(\theta) = \sum_{k=1}^{m} \exp\left(-\theta\left(\sum_{i=2}^{m-1} Y_i + X_k\right)\right).$$

Hence a simple computation implies that

$$\hat{\mu}(\theta) = \sum_{k=1}^{m} \mathbb{E}\left[\exp\left(-\theta\left(\sum_{i=2}^{m-1} Y_i + X_k\right)\right)\right] = m! \prod_{i=1}^{m-1} (i+\theta)^{-1}, \quad \theta > -1. \quad (10)$$

In particular, we see that the Malthusian parameter is $\alpha = 1$. Furthermore,

$$\mathbb{E}[\hat{\Xi}(\theta)^2] = \hat{\mu}(2\theta)\left(1 + \frac{(m-1)(m-2)}{m}\frac{1+2\theta}{(1+\theta)^2}\right), \quad \theta > -1/2,$$

which implies that $\mathrm{Var}\,(\hat{\Xi}(\theta)) < \infty$ for $\theta > -1/2$. Thus it should be clear that conditions (**A1**)–(**A7**) are satisfied. Consequently, Theorem 1 implies the following result.

Corollary 2. *In the supercritical regime, we have* $\lim_{n\to\infty} n^{-1}|T_n^{(m)}| = 2(H_m - 1)e^{-c/(H_m - 1)}$ *in probability, where* $H_m = \sum_{i=1}^{m} i^{-1}$.

   Proof. The result follows from Theorem 1 by computing $\bar{\mu}(1)$ and $\mathbb{E}[\hat{\phi}_m(1)]$. Note that (8) and (10) show that $\bar{\mu}(1) = -\hat{\mu}'(1) = H_m - 1$. Note also that

$$\hat{\phi}_m(1) = \int_0^{\infty} e^{-t}\phi_m(t)\,dt = 1 + \sum_{k=2}^{m-1} e^{-\sum_{i=2}^{k} Y_i}.$$

Therefore a direct computation shows that

$$\mathbb{E}[\hat{\phi}_m(1)] = 1 + \sum_{k=2}^{m-1} \mathbb{E}[e^{-\sum_{i=2}^{k} Y_i}] = 1 + 2\sum_{k=2}^{m-1} \frac{1}{k+1} = 2(H_m - 1),$$

which concludes the proof.   □

3.3. Median-of-$(2\ell + 1)$ binary search tree

   The random median-of-$(2\ell + 1)$ binary search tree, for $\ell \in \mathbb{N}$ (see e.g. [13]), is a modification of the binary search tree (or 2-ary search tree), where each internal vertex contains exactly one key, but each external vertex can contain up to 2ℓ keys (recall that keys are real numbers).
   This tree is constructed recursively from an initial tree with a single external vertex without any keys. Then we add keys one by one until the $(2\ell + 1)$th key is placed at this first external vertex (or to another external vertex later in the process). The vertex becomes an internal one with two new external vertices as its children, say v_L and v_R. Immediately, the median of the $2\ell + 1$ keys at the vertex is computed and put at the external vertex, while the ℓ keys that are smaller than the median are put in the left child v_L and the ℓ keys that are larger than the median are put in the right child v_R. We continue by adding new keys to the root and send them to the left or to the right whenever they are smaller or larger than the median of the first $2\ell + 1$ keys. Finally, we iterate this procedure in an obvious way in order to grow the tree until $n \in \mathbb{N}$ keys

have been added. We let $T_n^{(\ell)}$ denote the random median-of-$(2\ell + 1)$ binary search tree with n keys.

Following [19, Section 8], we can construct a median-of-$(2\ell + 1)$ binary search tree via a CMJ-process. Start with one vertex (the root) with ℓ keys (note that this is not a problem because the first ℓ keys always go there). Then each external vertex will contain between ℓ and 2ℓ keys, throughout the process. On the other hand, a vertex acquires $\ell + 1$ additional keys after successive independent waiting times $Y_1, \ldots, Y_{\ell+1}$, where Y_i has exponential distribution of parameter $\ell + i$, for $i = 1, \ldots, \ell + 1$. At the time when the $(\ell + 1)$th key arrives, the vertex immediately gets two children, each containing ℓ keys. Therefore it should be clear that the CMJ-process related to the median-of-$(2\ell + 1)$ binary search tree has birth times $\xi_1 = \xi_2 = \sum_{i=1}^{\ell+1} Y_i$ (in distribution) and where each individual has lifetime $\lambda \equiv \infty$. In this case $N = 2$. Define the characteristic

$$\phi_\ell(t) = \begin{cases} \ell + k & \sum_{i=1}^{k} Y_i \leq t < \sum_{i=1}^{k+1} Y_i, \ 0 \leq k \leq \ell, \\ 1 & \sum_{i=1}^{\ell+1} Y_i \leq t, \end{cases}$$

for $t \geq 0$ (with the convention $\sum_{i=1}^{0} Y_i = 0$). Therefore the tree $T_n^{\phi_\ell}$ in Definition 1 is a median-of-$(2\ell + 1)$ binary search tree with n keys. In this example, note that $\hat{\Xi}(\theta) = 2 \exp(-\theta \sum_{i=1}^{\ell+1} Y_i)$. Hence

$$\hat{\mu}(\theta) = \mathbb{E}\left[2 \exp\left(-\theta \sum_{i=1}^{\ell+1} Y_i\right)\right] = 2 \prod_{i=1}^{\ell+1} \frac{\ell + i}{\ell + i + \theta}, \quad \theta > -(\ell + 1). \tag{11}$$

We deduce that the Malthusian parameter is $\alpha = 1$. Furthermore, it is not difficult to see that

$$\mathrm{Var}(\hat{\Xi}(\theta)) = 2\hat{\mu}(2\theta) - \hat{\mu}(\theta)^2 < \infty, \quad \theta > -(\ell + 1)/2.$$

Thus it should be plain that all the conditions (**A1**)–(**A7**) are satisfied. Consequently, Theorem 1 implies the following result.

Corollary 3. *In the supercritical regime, we have*

$$\lim_{n \to \infty} n^{-1} |T_n^{(\ell)}| = (\ell + 1)(H_{2\ell+2} - H_{\ell+1}) e^{-c(H_{2\ell+2} - H_{\ell+1})^{-1}}, \quad \text{in probability.}$$

Proof. By Theorem 1, we only need to compute $\bar{\mu}(1)$ and $\mathbb{E}[\hat{\phi}_\ell(1)]$. Note that (8) and (11) show that $\bar{\mu}(1) = -\hat{\mu}'(1) = H_{2\ell+2} - H_{\ell+1}$. Note also that

$$\hat{\phi}_\ell(1) = \int_0^\infty e^{-t} \hat{\phi}_\ell(t) \, dt = \sum_{k=0}^{\ell} (e^{-t \sum_{i=1}^{k} Y_i} - e^{-t \sum_{i=1}^{k+1} Y_i}) + e^{-t \sum_{i=1}^{\ell+1} Y_i}.$$

Therefore a simple but tedious computation implies that $\mathbb{E}[\hat{\phi}_\ell(1)] = (\ell + 1)(H_{2\ell+2} - H_{\ell+1})$. $\square$

3.4. Fragmentation trees

In this section we study family trees induced by fragmentation processes. These processes were introduced in [24]; see also [6] and [22] for general background and further references.

Fix $b \geq 2$ and consider a random vector $\mathbf{V} = (V_1, \ldots, V_b)$; this is the so-called dislocation law. Assume that $0 \leq V_i < 1$ a.s., for $i = 1, \ldots, b$, and $\sum_{i=1}^{b} V_i = 1$, that is, $\mathbf{V}$ belongs to the standard simplex.

We then describe the construction of the fragmentation tree. Start with a vertex (the root) with mass x_0. This vertex has b children with masses $x_0 V_1, \ldots, x_0 V_b$, that is, we break x_0 into b pieces with masses driven by the dislocation law $\mathbf{V}$. Consider a threshold $x_1 \in (0, x_0]$, and continue recursively with each vertex that has mass larger than x_1, using new (independent) copies of $\mathbf{V}$ each time. The process terminates a.s. after a finite number of steps, leaving a finite set of vertices (or fragments) with masses smaller than x_1. We regard the vertices of mass larger than x_1 that occur during this process as the internal vertices and the resulting vertices of mass strictly less than x_1 as external. Note that the fragmentation tree depends only on the ratio x_0/x_1, so we denote it by $T^{\mathrm{f}}_{x_0/x_1}$.

One can relate the fragmentation process to a CMJ-process by regarding a fragment of mass x as born at time $\log(x_0/x)$. In other words, a vertex will have b children that are born at times $\xi_i = -\log V_i$, for $i = 1, \ldots, b$ (observe that $N = b$ in this case). If $V_i = 0$, we have $\xi_i = \infty$, meaning that this child is not born, and thus a vertex has fewer than b children. Note also that the lifetime $\lambda = \infty$. Finally, it is easy to see that the fragmentation tree $T^{\mathrm{f}}_{x_0/x_1}$ is the same as the family tree of this CMJ-process at time $\log(x_0/x_1)$, i.e. $T(\log(x_0/x_1))$ in the notation of Section 1.1.

By using the characteristic $\phi \equiv 1$, we define the fragmentation tree $T^{\mathrm{f}}_n := T(\tau^{\phi}_n)$ of fixed size $n \in \mathbb{N}$ as in Definition 1. This means that we choose a threshold $x_1 > 0$ to be the mass of the nth largest fragment in the process, so that there will be exactly n fragments of size x_1 (unless there is a tie). In this case, note that $\hat{\Xi}(\theta) = \sum_{i=1}^{b} e^{-\theta \xi_i} = \sum_{i=1}^{b} V_i^{\theta}$. Then

$$\hat{\mu}(\theta) = \sum_{i=1}^{b} \mathbb{E}[V_i^{\theta}], \quad \theta \geq 0. \tag{12}$$

Thus we conclude that the Malthusian parameter is $\alpha = 1$. This has the consequence that the martingale W^{ψ} (with $p_n \equiv 1$), in Section 2, is constant 1. Then we find that its limit $W^{\psi}(\infty) \equiv 1$. Furthermore, we also find that $\mathrm{Var}(\hat{\Xi}(\theta)) < \infty$, for $\theta \geq 0$. Thus the conditions (A1)–(A7) are satisfied and Theorem 1 implies the following result.

Corollary 4. *In the supercritical regime, we have* $\lim_{n \to \infty} n^{-1} |T^{\mathrm{f}}_n| = e^{-c/\beta}$, *in probability, where* $\beta := \sum_{i=1}^{b} \mathbb{E}[V_i \log(1/V_i)]$.

Proof. The result follows from Theorem 1. Since $\phi \equiv 1$, we need only compute $\bar{\mu}(1)$, but note that (8) and (12) imply that $\bar{\mu}(1) = -\hat{\mu}'(1) = \beta$. $\qquad\square$

Example 1. (*Binary splitting.*) Consider $b = 2$ and $\mathbf{V} = (V_1, V_2) = (V_1, 1 - V_1)$, where V_1 is a uniform random variable on $(0, 1)$. Thus, at each fragmentation event, the fragment is split into two parts, with uniformly random sizes. In the corresponding CMJ-process the birth times ξ_1 and ξ_2 are exponential random variables with parameter 1, where one of them determines the other by $e^{-\xi_1} + e^{-\xi_2} = 1$. Furthermore, $\hat{\mu}(\theta) = (1 + \theta)^{-1}$ for $\theta \geq 0$, and $\bar{\mu}(1) = 1/2$. Finally, note also that there are similarities with the CMJ-process associated with the binary search tree in Table 1 (Section 3.1); the difference is that there ξ_1 and ξ_2 are independent, while here they are dependent.

Finally, we should mention that the split trees defined by Devroye [14] are related to fragmentation trees. A split tree is a b-ary tree defined using a number of balls that enter the root

and are distributed (randomly and recursively) to the subtrees of the root and further down in the tree according to certain rules that are based on a splitting law **V**; see [14] for details. For instance, binary search trees are a particular type of split tree. Percolation on split trees will be studied in another paper.

3.5. Homogeneous CMJ-trees

Let Λ be a finite positive measure on $(0, \infty]$ with total mass $b := \Lambda((0, \infty])$ such that $m := \int_{(0,\infty]} t\Lambda(\mathrm{d}t)$ satisfies

$$1 < m < \infty. \tag{E3}$$

Consider a CMJ-process with birth process Ξ whose intensity measure is given by $\mu(\mathrm{d}t) = \mathrm{d}t\Lambda((t, \infty])$. Moreover, each individual has lifetime given by a random variable λ with distribution $\Lambda(\cdot)/b$. Note that, conditional on the birth time and lifetime, the birth point process of an individual is distributed as a Poisson point process during his life. We can say that the CMJ-process is homogeneous (constant birth rate) and binary (birth occurs singly). Set

$$\Psi(\theta) = \theta - \int_{(0,\infty]} (1 - \mathrm{e}^{-\theta t})\Lambda(\mathrm{d}t), \quad \theta \geq 0,$$

and observe that Ψ is a convex function such that $\Psi(0+) = 0$ and $\Psi'(0+) = 1 - m < 0$. Hence there exists a unique $\alpha > 0$ such that $\Psi(\alpha) = 0$.

In this section we are interested in studying Bernoulli bond-percolation on the family tree of this CMJ-process. More precisely, we consider the characteristic $\phi \equiv 1$ and define the CMJ-tree $T_n^{\mathrm{hom}} := T(\tau_n^\phi)$ as in Definition 1. The family tree of this particular CMJ-process was called a splitting tree in [17], [18], and [25]. Further, these papers studied the case when Λ is not necessary finite, but for simplicity we have decided to restrict ourselves to finite measures. Nevertheless, our results can be applied to the general case.

In [33, Chapter 3], the author studied this CMJ-process under neutral rare mutations, where mutations occur at birth with probability $1 - p \in [0, 1]$. The difference from our setting is that in [33, Chapter 3] the probability of mutation (or percolation parameter) does not depend on the 'size' of the tree. Therefore the result in this section may be of independent interest.

We now check that assumptions (**A1**)–(**A7**) are satisfied. Clearly, (**A1**)–(**A2**) are fulfilled by the assumptions made on Λ. Note that in this case (**A3**) is not fulfilled. Nevertheless, [33, Proposition 2.1; see also (2.1)] shows that the extinction probability $\mathbb{P}(Z(\infty) < \infty) = 1 - \alpha/b < 1$. This has the consequence that the limit of the martingale W^ψ (with $p_n \equiv 1$), in Section 2, is strictly positive (i.e. $W^\psi(\infty) > 0$ a.s.) on the event $\{Z(\infty) = \infty\}$; see [21, Corollary 4.2]. Thus our results hold conditioned on the event $\{Z(\infty) = \infty\}$ since (**A3**) is only needed to guarantee that $W^\psi(\infty) > 0$ almost surely. On the other hand,

$$\hat{\mu}(\theta) = \int_0^\infty \int_{(t,\infty]} \mathrm{e}^{-\theta t}\,\mathrm{d}t\Lambda(\mathrm{d}u) = \int_{(0,\infty]} \frac{1 - \mathrm{e}^{-\theta t}}{\theta}\Lambda(\mathrm{d}t) = \frac{\theta - \Psi(\theta)}{\theta}, \quad \theta \geq 0,$$

which clearly shows that (**A4**) is satisfied. Moreover, the Malthusian parameter is given by $\alpha > 0$. Note also that Campbell's formula implies that

$$\mathrm{Var}\,(\hat{\Xi}(\theta)) = \int_0^\infty \int_{(t,\infty]} \mathrm{e}^{-2\theta t}\,\mathrm{d}t\Lambda(\mathrm{d}u) = \frac{2\theta - \Psi(2\theta)}{2\theta} < \infty, \quad \theta \geq 0,$$

which implies (**A5**). Finally, (**A6**)–(**A7**) follow immediately since $\phi \equiv 1$. Therefore Theorem 1 implies the following result.

Corollary 5. *In the supercritical regime and under* (**E3**), *conditional on the event* $\{Z(\infty) = \infty\}$, *we have* $\lim_{n\to\infty} n^{-1}|T_n^{\mathrm{hom}}| = \mathrm{e}^{-c/(\Psi'(\alpha))}$, *in probability.*

Proof. The result follows from Theorem 1 by computing $\bar{\mu}(\alpha)$ since $\phi \equiv 1$. This follows from (3):

$$
\bar{\mu}(\alpha) = \int_0^\infty t\,\mathrm{e}^{-\alpha t}\mu(\mathrm{d}t)
$$

$$
= \int_0^\infty \int_{(t,\infty]} t\,\mathrm{e}^{-\alpha t}\,\mathrm{d}t\Lambda(\mathrm{d}u)
$$

$$
= \int_{(0,\infty]} \left(-\frac{t\,\mathrm{e}^{-\alpha t}}{\alpha} + \frac{1 - \mathrm{e}^{-\alpha t}}{\alpha^2} \right)\Lambda(\mathrm{d}t)
$$

$$
= \frac{\Psi'(\alpha)}{\alpha}. \qquad\qquad \square
$$

Appendix A. Proofs of Lemmas 2 and 3

In this section we establish some general results on the long time behavior of the CMJ-process with neutral mutations that may be of independent interest. It will be helpful to write p rather than p_n, omitting the integer n from the notation. To be more precise, we consider that the percolation parameter is a real number $p \in [0, 1]$ and study the behavior of the ϕ-counted clonal process $Z_\varnothing^{(p),\phi} = (Z_\varnothing^{(p),\phi}(t), t \geq 0)$ as $p \to 1$ and $t \to \infty$. Recall that $\Xi^{(p)}$ denotes the generic birth process of the clonal CMJ-process whose intensity measure $\mu^{(p)}$ is defined in (2).

Note that (**A4**) implies that there exists $p^* \in (0, 1)$ such that for $p \in [p^*, 1]$ there exists $\alpha_p > 0$ (the Malthusian parameter of $\mu^{(p)}$) such that $\hat{\mu}^{(p)}(\alpha_p) = 1$. This implies that $p\mathbb{E}[N] > 1$ (i.e. the clonal process is supercritical). Since $\alpha_p \to \alpha$ as $p \to 1$, we choose p^* such that $0 < \theta_1 < \alpha_{p^*}$ where θ_1 satisfies (**A4**)–(**A5**). Furthermore, we also consider p^* such that $0 < \theta_2 < \alpha_{p^*}$ and $0 < \theta_3 < 2\alpha_{p^*}$, where θ_2 and θ_3 satisfy (**A6**) and (**A7**), respectively.

We start by recalling some well-known results about the moments of $Z_\varnothing^{(p),\phi}$. For $k \in \mathbb{N} \cup \{0\}$, we let ν^{*k} denote the k-fold convolution of a measure ν on $[0, \infty)$ (here ν^{*0} is a unit point mass at 0).

Theorem 2. *Assume that conditions* (**A1**)–(**A4**) *are fulfilled. For a characteristic ϕ that may depend on p, we have*

$$
\mathbb{E}[Z_\varnothing^{(p),\phi}(t)] = \int_0^t \mathbb{E}[\phi(t - s)] \sum_{k=0}^\infty (\mu^{(p)})^{*k}(\mathrm{d}s), \quad t \geq 0.
$$

Furthermore, if $\mathbb{E}[\phi(t)]$ is bounded on finite intervals,

$$
\mathrm{Var}\,(Z_\varnothing^{(p),\phi}(t)) = \int_0^t h_\varnothing(t - s) \sum_{k=0}^\infty (\mu^{(p)})^{*k}(\mathrm{d}s), \quad t \geq 0,
$$

with

$$
h_\varnothing^{(p)}(t) = \mathrm{Var}\left(\phi_\varnothing(t) + \int_0^t \mathbb{E}[Z_\varnothing^{(p),\phi}(t - s)]\Xi_\varnothing^{(p)}(\mathrm{d}s) \right), \quad t \geq 0, \tag{13}
$$

and where $(\Xi_{\varnothing}^{(p)}, \phi_{\varnothing})$ *is the birth process and weight associated with the progenitor of the population.*

Proof. The first claim is a consequence of [21, Theorem 3.1]. Note that (**A1**) and (**A4**) imply that the measure $\sum_{k=0}^{\infty} (\mu^{(p)})^{*k}(\mathrm{d}t)$ is finite. Moreover, $\mathbb{E}[Z_{\varnothing}^{(p),\phi}(t)] < \infty$, for $t \geq 0$, since $\mathbb{E}[\phi(t)]$ is bounded on finite intervals. Therefore the second claim follows from [21, Theorem 3.2]. $\qquad\square$

We next provide an improvement of the results of [20, Theorem 6.9.2] and [21, Theorem 3.5] on the asymptotic behavior of the first and second moment of the clonal CMJ-process. Recall that we write $W_{\varnothing}^{(p),\phi} = (W_{\varnothing}^{(p),\phi}(t), t \geq 0)$ for the process given by $W_{\varnothing}^{(p),\phi}(t) := \mathrm{e}^{-t\alpha_p} Z_{\varnothing}^{(p),\phi}(t)$, for $t \geq 0$. Set

$$m_t^{(p),\phi} = \mathbb{E}[W_{\varnothing}^{(p),\phi}(t)], \quad t \geq 0.$$

Recall also that for $p \equiv 1$ we sometimes remove the superscript (p) and the subscript $\varnothing$ from the previous notations. That is, we write $W^{\phi} = (W^{\phi}(t), t \geq 0)$ for the processes given by $W^{\phi}(t) := \mathrm{e}^{-t\alpha} Z^{\phi}(t)$ and $m_t^{\phi} = \mathbb{E}[W^{\phi}(t)]$.

Proposition 2. *Assume that conditions* (**A1**)–(**A4**) *are fulfilled.*

(i) *For any characteristic ϕ that does not depend on p and that satisfies* (**A6**), *we have*

$$\lim_{t \to \infty} \sup_{p \in [p^*, 1]} |m_t^{(p),\phi} - m_{\infty}^{(p),\phi}| = 0,$$

where $m_{\infty}^{(p),\phi}$ is defined in (7). *In particular,* $\sup_{t \geq 0} \sup_{p \in [p^*, 1]} m_t^{(p),\phi} < \infty$.

(ii) *For the characteristic ψ defined in* (5), *we have*

$$\lim_{t \to \infty} \sup_{p \in [p^*, 1]} |m_t^{(p),\psi} - m_{\infty}^{(p),\psi}| = 0.$$

In particular, $\sup_{t \geq 0} \sup_{p \in [p^*, 1]} m_t^{(p),\psi} < \infty$.

Proof. We only prove (i). The proof of (ii) follows from exactly same argument and Lemma 1 (i). Recall from Remark 1 that $\mu_{\alpha}^{(p)}(\mathrm{d}t) := \mathrm{e}^{-t\alpha_p} \mu^{(p)}(\mathrm{d}t)$, for $p \in [p^*, 1]$, is a probability measure on $(0, \infty)$. Theorem 2 implies that

$$m_t^{(p),\phi} = \int_0^t \mathrm{e}^{-(t-s)\alpha_p} \mathbb{E}[\phi(t-s)] \sum_{k=0}^{\infty} (\mu_{\alpha}^{(p)})^{*k}(\mathrm{d}s), \quad t \geq 0,$$

where we have used that

$$\mathrm{e}^{-s\alpha_p} \sum_{k=0}^{\infty} (\mu^{(p)})^{*k}(\mathrm{d}s) = \sum_{k=0}^{\infty} (\mu_{\alpha}^{(p)})^{*k}(\mathrm{d}s).$$

From (**A6**), we deduce that the family of functions $t \mapsto \mathrm{e}^{-t\alpha_p} \mathbb{E}[\phi(t)]$, for $p \in [p^*, 1]$, it is uniformly directly Riemann integrable (see [37, Definition 2.8]). Furthermore, (**A4**) implies that

the family of probability measures $\{\mu_\alpha^{(p)} : p \in [p^*, 1]\}$ is weakly compact (treated as a set of measure on $(0, \infty)$), and uniformly integrable, that is,

$$\lim_{a \to \infty} \sup_{p \in [p^*, 1]} \int_{[a, \infty)} t \mu_\alpha^{(p)}(\mathrm{d}t) = 0.$$

Therefore the point (i) is a consequence of the uniform version of the key renewal theorem [37, Theorem 2.12] since μ satisfies (**A2**). The second claim in (i) follows from the definition of $m_\infty^{(p),\phi}$ by noting that

$$\alpha_{p^*} \leq \alpha_p \leq \alpha, \quad \bar{\mu}^{(p)}(\alpha_p) \geq p^* \bar{\mu}^{(1)}(\alpha) \quad \text{and} \quad \mathbb{E}[\hat{\phi}(\alpha_p)] \leq \frac{\alpha}{\alpha_{p^*}} \mathbb{E}[\hat{\phi}(\alpha_{p^*})];$$

we also need to recall that $\bar{\mu}^{(1)}(\alpha) > 0$ by Remark 2. $\square$

By Proposition 2,

$$v_t^{(p),\phi} = \mathrm{Var}\,(W_\varnothing^{(p),\phi}(t)), \quad t \geq 0,$$

is well-defined. For $p \equiv 1$, we sometimes write $v_t^\phi = \mathrm{Var}\,(W^\phi(t))$, for $t \geq 0$.

Proposition 3. *Assume that conditions* (**A1**)–(**A5**) *are fulfilled.*

(i) *For any characteristic ϕ that does not depend on p and that satisfies* (**A6**)–(**A7**), *we have*

$$\lim_{t \to \infty} \sup_{p \in [p^*, 1]} |v_t^{(p),\phi} - v_\infty^{(p),\phi}| = 0,$$

where

$$v_\infty^{(p),\phi} := (m_\infty^{(p),\phi})^2 \frac{\mathrm{Var}\,(\hat{\Xi}^{(p)}(\alpha_p))}{1 - \hat{\mu}^{(p)}(2\alpha_p)}.$$

(ii) *For the characteristic ψ defined in* (5), *we have*

$$\lim_{t \to \infty} \sup_{p \in [p^*, 1]} |v_t^{(p),\psi} - v_\infty^{(p),\psi}| = 0,$$

where $v_\infty^{(p),\psi} := \mathrm{Var}\,(\hat{\Xi}^{(p)}(\alpha_p))(1 - \hat{\mu}^{(p)}(2\alpha_p))^{-1}$.

(iii) *For the characteristic $\phi' = \phi + m_\infty^{(p),\phi} \psi$, we have*

$$\lim_{t \to \infty} \sup_{p \in [p^*, 1]} |v_t^{(p),\phi'} - v_\infty^{(p),\phi'}| = 0,$$

where $v_\infty^{(p),\phi'} = 4 v_\infty^{(p),\phi}$.

Proof. We only prove (i). The proofs of (ii) and (iii) follow similarly by using Lemma 1 and Remark 3. Theorem 2 implies that

$$v_t^{(p),\phi} = \int_0^t \mathrm{e}^{-2(t-s)\alpha_p} h_\varnothing^{(p)}(t-s)\, \mathrm{e}^{-2s\alpha_p} \sum_{k=0}^\infty (\mu^{(p)})^{*k}(\mathrm{d}s), \quad t \geq 0,$$

where the function $h_{\varnothing}^{(p)}$ is defined in (13). First, note the identity

$$\int_0^\infty e^{-2s\alpha_p} \sum_{k=0}^\infty (\mu^{(p)})^{*k}(ds) = \sum_{k=0}^\infty \hat{\mu}^{(p)}(2\alpha_p)^k = \frac{1}{1 - \hat{\mu}^{(p)}(2\alpha_p)} < \infty, \qquad (14)$$

since $\hat{\mu}^{(p)}(2\alpha_p) < \hat{\mu}^{(p)}(\alpha_p) = 1$. Then the triangle inequality implies that

$$|v_t^{(p),\phi} - v_\infty^{(p),\phi}|$$

$$\leq \int_0^t e^{-2(t-s)\alpha_p} \, \text{Var}\,(\phi_{\varnothing}(t-s)) \, e^{-2s\alpha_p} \sum_{k=0}^\infty (\mu^{(p)})^{*k}(ds)$$

$$+ 2 \int_0^t e^{-(t-s)\alpha_p} \left| \text{Cov}\left(\phi_{\varnothing}(t-s), \int_0^{t-s} m_{t-s-u}^{(p),\phi} \, e^{-u\alpha_p} \, \Xi_{\varnothing}^{(p)}(du)\right)\right| e^{-2s\alpha_p} \sum_{k=0}^\infty (\mu^{(p)})^{*k}(ds)$$

$$+ \int_0^\infty g^{(p)}(t-s) \, e^{-2s\alpha_p} \sum_{k=0}^\infty (\mu^{(p)})^{*k}(ds), \qquad (15)$$

where $(\Xi_{\varnothing}^{(p)}, \phi_{\varnothing})$ is the birth process and weight associated with the progenitor of the population, and

$$g^{(p)}(t-s) = \left| \mathbb{1}_{\{s \in [0,t]\}} \, \text{Var}\left(\int_0^{t-s} m_{t-s-u}^{(p),\phi} \, e^{-u\alpha_p} \, \Xi_{\varnothing}^{(p)}(du)\right) - (m_\infty^{(p),\phi})^2 \, \text{Var}\,(\hat{\Xi}_{\varnothing}^{(p)}(\alpha_p)) \right|.$$

Let $I_1^{(p)}(t)$, $I_2^{(p)}(t)$, and $I_3^{(p)}(t)$, respectively, denote the first, second, and third terms on the right-hand side of (15). Then the claim in Proposition 3 follows by showing that

(a) $\lim_{t\to\infty} \sup_{p \in [p^*,1]} I_1^{(p)}(t) = 0$,

(b) $\lim_{t\to\infty} \sup_{p \in [p^*,1]} I_2^{(p)}(t) = 0$, and

(c) $\lim_{t\to\infty} \sup_{p \in [p^*,1]} I_3^{(p)}(t) = 0$.

We start by showing (a). By assumption (**A7**),

$$\lim_{t\to\infty} \sup_{p \in [p^*,1]} e^{-2t\alpha_p} \, \text{Var}\,(\phi(t)) \leq \lim_{t\to\infty} e^{-2t\alpha_{p^*}} \, \text{Var}\,(\phi(t)) = 0. \qquad (16)$$

Then (14) and the dominated convergence theorem prove point (a).

Next we show (b). The Cauchy–Schwarz inequality and Proposition 2 imply that

$$\left| \text{Cov}\left(\phi_{\varnothing}(t), \int_0^t m_{t-u}^{(p),\phi} \, e^{-u\alpha_p} \, \Xi_{\varnothing}^{(p)}(du)\right)\right|$$

$$\leq (\,\text{Var}\,(\phi_{\varnothing}(t)))^{1/2} \left(\text{Var}\left(\int_0^t m_{t-u}^{(p),\phi} \, e^{-u\alpha_p} \, \Xi_{\varnothing}^{(p)}(du)\right)\right)^{1/2}$$

$$\leq \left(\text{Var}\,(\phi_{\varnothing}(t)) \, \text{Var}\left(\int_0^t e^{-u\alpha_p} \, \Xi_{\varnothing}^{(p)}(du)\right)\right)^{1/2} \sup_{t\geq 0} \sup_{p \in [p^*,1]} m_t^{(p),\phi}.$$

Since $\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u)$ is dominated by $\mathrm{e}^{-u\alpha_{p^*}}\,\Xi(\mathrm{d}u)$, (**A4**)–(**A5**) and (16) allow us to deduce that

$$\lim_{t\to\infty}\sup_{p\in[p^*,1]}\mathrm{e}^{-t\alpha_p}\left|\mathrm{Cov}\left(\phi_\varnothing(t),\int_0^t m_{t-u}^{(p),\phi}\,\mathrm{e}^{-u\alpha_p}\,\Xi_\varnothing^{(p)}(\mathrm{d}u)\right)\right|=0.$$

Hence an application of the dominated convergence theorem shows (b).

Finally, we prove (c). We show that

$$\lim_{t\to\infty}\sup_{p\in[p^*,1]}g^{(p)}(t-s)=0,\tag{17}$$

which together with an application of the dominated convergence theorem implies (c). Observe that

$$g^{(p)}(t-s)\le g_1^{(p)}(t-s)+g_2^{(p)}(t-s)+g_3^{(p)}(t-s)$$
$$+g_4^{(p)}(t-s)+\mathbb{1}_{\{s\in(t,\infty)\}}(m_\infty^{(p),\phi})^2\,\mathrm{Var}\,(\hat{\Xi}^{(p)}(\alpha_p)),\tag{18}$$

where

$$g_1^{(p)}(t-s)=\mathbb{1}_{\{s\in[0,t]\}}\,\mathrm{Var}\left(\int_0^{t-s}(m_{t-s-u}^{(p),\phi}-m_\infty^{(p),\phi})\,\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u)\right),$$
$$g_2^{(p)}(t-s)=2\mathbb{1}_{\{s\in[0,t]\}}\left|\mathrm{Cov}\left(\int_0^{t-s}(m_{t-s-u}^{(p),\phi}-m_\infty^{(p),\phi})\,\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u),\right.\right.$$
$$\left.\left.m_\infty^{(p),\phi}\int_0^{t-s}\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u)\right)\right|,$$
$$g_3^{(p)}(t-s)=2(m_\infty^{(p),\phi})^2\mathbb{1}_{\{s\in[0,t]\}}\left|\mathrm{Cov}\left(\hat{\Xi}^{(p)}(\alpha_p),\int_{t-s}^\infty\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u)\right)\right|,\quad\text{and}$$
$$g_4^{(p)}(t-s)=\mathbb{1}_{\{s\in[0,t]\}}(m_\infty^{(p),\phi})^2\,\mathrm{Var}\left(\int_{t-s}^\infty\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u)\right).$$

Proposition 2 and (**A5**) imply that

$$\lim_{t\to\infty}\sup_{p\in[p^*,1]}g_4^{(p)}(t-s)=0,\tag{19}$$

and

$$\lim_{t\to\infty}\sup_{p\in[p^*,1]}\mathbb{1}_{\{s\in(t,\infty)\}}(m_\infty^{(p),\phi})^2\,\mathrm{Var}\,(\hat{\Xi}^{(p)}(\alpha_p))=0.\tag{20}$$

Furthermore, Proposition 2 implies that

$$\sup_{t\ge0}\sup_{p\in[p^*,1]}|m_t^{(p),\phi}-m_\infty^{(p),\phi}|<\infty.$$

Recall that $\mathrm{e}^{-s\alpha_p}\,\Xi^{(p)}(\mathrm{d}s)$ is dominated by $\mathrm{e}^{-s\alpha_{p^*}}\,\Xi(\mathrm{d}s)$ and that $\int_0^\infty\mathrm{e}^{-s\alpha_{p^*}}\,\Xi(\mathrm{d}s)<\infty$, by (**A4**). Thus the dominated convergence theorem shows that

$$\lim_{t\to\infty}\sup_{p\in[p^*,1]}\left|\mathbb{1}_{\{s\in[0,t]\}}\int_0^{t-s}(m_{t-s-u}^{(p),\phi}-m_\infty^{(p),\phi})\,\mathrm{e}^{-u\alpha_p}\,\Xi^{(p)}(\mathrm{d}u)\right|=0,$$

almost surely. This implies that

$$\lim_{t\to\infty} \sup_{p\in[p^*,1]} g_1^{(p)}(t-s) = 0, \tag{21}$$

once again using the dominated convergence theorem. Similarly, we can deduce that

$$\lim_{t\to\infty} \sup_{p\in[p^*,1]} g_2^{(p)}(t-s) = 0 \quad \text{and} \quad \lim_{t\to\infty} \sup_{p\in[p^*,1]} g_3^{(p)}(t-s) = 0. \tag{22}$$

Finally, our claim in (17) follows by combining (18), (19), (20), (21), and (22). □

We now have all the ingredients to prove Lemma 2.

Proof of Lemma 2. Note that

$$\mathbb{E}\Big[\sup_{s\geq t} |W_{\varnothing}^{(p),\psi}(s) - W_{\varnothing}^{(p),\psi}(\infty)|^2 \Big]$$

$$\leq 2\mathbb{E}[|W_{\varnothing}^{(p),\psi}(t) - W_{\varnothing}^{(p),\psi}(\infty)|^2] + 2\mathbb{E}\Big[\sup_{s\geq t} |W_{\varnothing}^{(p),\psi}(s) - W_{\varnothing}^{(p),\psi}(t)|^2 \Big] \tag{23}$$

for $t \geq 0$. On the one hand, from properties of square-integrable martingales, we obtain that

$$\mathbb{E}[|W_{\varnothing}^{(p),\psi}(t) - W_{\varnothing}^{(p),\psi}(\infty)|^2] = \mathbb{E}[W_{\varnothing}^{(p),\psi}(\infty)^2] - \mathbb{E}[W_{\varnothing}^{(p),\psi}(t)^2] \tag{24}$$

for $t \geq 0$. On the other hand, by Doob's inequality

$$\mathbb{E}\Big[\sup_{s\geq t} |W_{\varnothing}^{(p),\psi}(s) - W_{\varnothing}^{(p),\psi}(t)|^2 \Big] \leq 4\mathbb{E}[W_{\varnothing}^{(p),\psi}(\infty)^2] - 4\mathbb{E}[W_{\varnothing}^{(p),\psi}(t)^2] \tag{25}$$

for $t \geq 0$. By combining (23), (24), and (25), we deduce that

$$\mathbb{E}\Big[\sup_{s\geq t} |W_{\varnothing}^{(p),\psi}(s) - W_{\varnothing}^{(p),\psi}(\infty)|^2 \Big]$$

$$\leq 10(v_{\infty}^{(p),\psi} - v_t^{(p),\psi}) + 10((m_{\infty}^{(p),\psi})^2 - (m_t^{(p),\psi})^2)$$

$$= 10(v_{\infty}^{(p),\psi} - v_t^{(p),\psi})$$

since $W_{\varnothing}^{(p),\psi}$ is a martingale. Therefore the first statement follows from Proposition 3. We turn our attention to the second claim. Observe that

$$\mathbb{E}[|W_{\varnothing}^{(p),\phi}(t) - m_{\infty}^{(p),\phi} W_{\varnothing}^{(p),\psi}(\infty)|^2]$$

$$\leq 2\mathbb{E}[|W_{\varnothing}^{(p),\phi}(t) - m_{\infty}^{(p),\phi} W_{\varnothing}^{(p),\psi}(t)|^2] + 2(m_{\infty}^{(p),\phi})^2 \mathbb{E}[|W_{\varnothing}^{(p),\psi}(t) - W_{\varnothing}^{(p),\psi}(\infty)|^2], \tag{26}$$

for $t \geq 0$. It follows from the first part that

$$\lim_{t\to\infty} \sup_{p\in[p^*,1]} \mathbb{E}[|W_{\varnothing}^{(p),\psi}(t) - W_{\varnothing}^{(p),\psi}(\infty)|^2] = 0.$$

In order to conclude, it is enough to show that the first term on the right-hand side of (26) tends to 0 uniformly on $p \in [p^*, 1]$ as $t \to \infty$. By Proposition 2, this is equivalent to showing that

$$\lim_{t\to\infty} \sup_{p\in[p^*,1]} \mathrm{Var}\left(W_{\varnothing}^{(p),\phi}(t) - m_{\infty}^{(p),\phi} W_{\varnothing}^{(p),\psi}(t)\right) = 0 \tag{27}$$

since $W_\varnothing^{(p),\phi}(t) + m_\infty^{(p),\phi} W_\varnothing^{(p),\psi}(t) = W_\varnothing^{(p),\phi'}(t)$, where $\phi'(t) = \phi(t) + m_\infty^{(p),\phi}\psi(t)$. But (27) follows from Proposition 3, Remark 3, and the identity

$$\text{Var}\,(W_\varnothing^{(p),\phi}(t) - m_\infty^{(p),\phi} W_\varnothing^{(p),\psi}(t))$$
$$= 2\,\text{Var}\,(W_\varnothing^{(p),\phi}(t)) + 2\,\text{Var}\,(m_\infty^{(p),\phi} W_\varnothing^{(p),\psi}(t)) - \text{Var}\,(W_\varnothing^{(p),\phi}(t) + m_\infty^{(p),\phi} W_\varnothing^{(p),\psi}(t)). \quad \square$$

Finally, we conclude this section with the proof of Lemma 3. The idea of the proof is similar to that of [10, Lemma 3].

Proof of Lemma 3. We first prove that the double limit

$$\lim_{p\to 1, t\to\infty} W_\varnothing^{(p),\phi}(t) \text{ exists in } L_2(\mathbb{P}). \tag{28}$$

Denote the $L_2(\mathbb{P})$-norm by $\|\cdot\|_2$. We claim that

$$\lim_{p\to 1} \|W_\varnothing^{(p),\phi}(t) - W^\phi(t)\|_2 = 0 \quad \text{for } t \geq 0. \tag{29}$$

Recall that at each birth event, independently of all the other individuals, the newborn is a clone of its parent with probability p or a mutant with probability $1 - p$. Then it should be plain from the thinning property of point measure processes that the birth process of the mutant children of the ancestor $\varnothing$, denoted by $\Xi_\varnothing^{(p),\text{m}}$, has intensity measure given $(1 - p)\mu(\mathrm{d}t)$. Furthermore, the latter is independent of the birth process of the clonal children of the ancestor described in Section 1.2. Let $Z^{(p),\text{m}} = (Z^{(p),\text{m}}(t), t \geq 0)$ be the process that counts the number of mutants born up to time $t \geq 0$ and define

$$b_1^{(p)} = \inf\{t \geq 0 \colon Z^{(p),\text{m}}(t) > 0\}$$

to be the first birth time of a mutant. Plainly, $\lim_{p\to 1} b_1^{(p)} = \infty$ in probability, and the probability of the event $\{t \geq b_1^{(p)}\}$ can be made as small as we wish by choosing p sufficiently close to 1. On the one hand, as $Z_\varnothing^{(p),\phi}(t) \leq Z^\phi(t)$, we have

$$\mathbb{E}[|W_\varnothing^{(p),\phi}(t) - W^\phi(t)|^2, t \geq b_1^{(p)}] \leq (e^{2(\alpha-\alpha_p)t} + 1)\mathbb{E}[|W^\phi(t)|^2, t \geq b_1^{(p)}],$$

and the right-hand side goes to 0 as $p \to 1$. On the other hand, on the event $\{t < b_1^{(p)}\}$, we have $Z_\varnothing^{(p),\phi}(t) = Z^\phi(t)$ and hence $W_\varnothing^{(p),\phi}(t) = e^{(\alpha-\alpha_p)t}W^\phi(t)$. This yields

$$\mathbb{E}[|W_\varnothing^{(p),\phi}(t_\varepsilon) - W^\phi(t)|^2, t < b_1^{(p)}] \leq (e^{(\alpha-\alpha_p)t} - 1)\mathbb{E}[|W^\phi(t)|^2],$$

and the right-hand side goes to 0 as $p \to 1$. This establishes the convergence in (29).

Let $\varepsilon > 0$ be arbitrary. By Lemma 2, we can find $t_\varepsilon > 0$ such that

$$\sup_{p\in[p^*,1]} \|W_\varnothing^{(p),\phi}(s) - m_\infty^{(p),\phi} W_\varnothing^{(p),\psi}(\infty)\|_2 \leq \frac{\varepsilon}{6} \quad \text{for all } s \geq t_\varepsilon. \tag{30}$$

By using (29), we obtain that for $s \geq t_\varepsilon$ there is $\delta_\varepsilon > 0$ such that $1 - p < \delta_\varepsilon$ (for $p \in [p^*, 1)$) implies that

$$\|W_\varnothing^{(p),\phi}(s) - W^\phi(s)\|_2 \leq \frac{\varepsilon}{6}. \tag{31}$$

Let us take $s_1, s_2 \geq t_\varepsilon$ and $p_1, p_2 \in [p^*, 1)$ such that $1 - p_1 < \delta_\varepsilon$ and $1 - p_2 < \delta_\varepsilon$. The Minkowski inequality implies that for $s \geq t_\varepsilon$

$$
\begin{aligned}
&\|W_\varnothing^{(p_1),\phi}(s_1) - W_\varnothing^{(p_2),\phi}(s_2)\|_2 \\
&\leq \|W_\varnothing^{(p_1),\phi}(s_1) - m_\infty^{(p_1),\phi} W_\varnothing^{(p_1),\psi}(\infty)\|_2 + \|W_\varnothing^{(p_1),\phi}(s) - m_\infty^{(p_1),\phi} W_\varnothing^{(p_1),\psi}(\infty)\|_2 \\
&\quad + \|W_\varnothing^{(p_1),\phi}(s) - W^\phi(s)\|_2 + \|W_\varnothing^{(p_2),\phi}(s) - W^\phi(s)\|_2 \\
&\quad + \|W_\varnothing^{(p_2),\phi}(s_2) - m_\infty^{(p_2),\phi} W_\varnothing^{(p_2),\psi}(\infty)\|_2 + \|W_\varnothing^{(p_2),\phi}(s) - m_\infty^{(p_2),\phi} W_\varnothing^{(p_2),\psi}(\infty)\|_2.
\end{aligned}
$$

By (30) and (31), we deduce that $\|W_\varnothing^{(p_1),\phi}(s_1) - W_\varnothing^{(p_2),\phi}(s_2)\|_2 \leq \varepsilon$. Thus $W_\varnothing^{(p),\phi}(t)$ is $L_2(\mathbb{P})$-Cauchy. Therefore the claim in (28) follows from the well-known completeness of $L_2(\mathbb{P})$; see [3, Theorem 6.14].

Finally, we show that the claim in Lemma 3 holds. Note that (28) implies that there exists a square-integrable variable $\overline{W}$ such that $\lim_{p \to 1, t \to \infty} W_\varnothing^{(p),\phi}(t) = \overline{W}$ in $L_2(\mathbb{P})$. Thus it is enough to show that $\overline{W} = m_\infty^\phi W^\psi(\infty)$, where ψ is defined in (5) with $p \equiv 1$. In this direction, for $t \geq 0$, recall that (29) shows that $\lim_{p \to 1} W_\varnothing^{(p),\phi}(t) = W^\phi(t)$ in $L_2(\mathbb{P})$. Furthermore, Lemma 2 implies that

$$
\lim_{t \to \infty} \lim_{p \to 1} W_\varnothing^{(p),\phi}(t) = \lim_{t \to \infty} W^\phi(t) = m_\infty^\phi W^\psi(\infty), \quad \text{in } L_2(\mathbb{P}). \tag{32}
$$

On the other hand, (28) implies that for $\varepsilon > 0$ there are $\delta_\varepsilon, t_\varepsilon > 0$ such that for $1 - p < \delta_\varepsilon$ (for $p \in [p^*, 1)$) and $s \geq t_\varepsilon$ we have $\|W_\varnothing^{(p),\phi}(s) - \overline{W}\|_2 \leq \varepsilon$. Hence (29) and an application of the dominated convergence theorem allow us to conclude that

$$
\lim_{p \to 1} \|W_\varnothing^{(p),\phi}(s) - \overline{W}\|_2 = \|W^\phi(s) - \overline{W}\|_2 \leq \varepsilon,
$$

i.e. $\lim_{t \to \infty} \lim_{p \to 1} W_\varnothing^{(p),\phi}(t) = \overline{W}$ in $L_2(\mathbb{P})$, which combined with (32) concludes the proof. $\square$

Acknowledgements

References

[1] ALDOUS, D. (1991). Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Prob.* **1**, 228–266.

[2] ATHREYA, K. B. (2007). Preferential attachment random graphs with general weight function. *Internet Math.* **4**, 401–418.

[3] BARTLE, R. G. (1995). The Elements of Integration and Lebesgue Measure (Wiley Classics *Library*). John Wiley, New York.

[4] BAUR, E. (2016). Percolation on random recursive trees. *Random Structures Algorithms* **48**, 655–680.

[5] BAUR, E. AND BERTOIN, J. (2017). Weak limits for the largest subpopulations in Yule processes with high mutation probabilities. *Adv. Appl. Prob.* **49**, 877–902.

[6] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes* (Cambridge Studies in Advanced Mathematics **102**). Cambridge University Press, Cambridge.

[7] BERTOIN, J. (2013). Almost giant clusters for percolation on large trees with logarithmic heights. *J. Appl. Prob.* **50**, 603–611.

[8] BERTOIN, J. (2014). On the non-Gaussian fluctuations of the giant cluster for percolation on random recursive trees. *Electron. J. Prob.* **19**, 24, 15 pp.

[9] BERTOIN, J. (2014). Sizes of the largest clusters for supercritical percolation on random recursive trees. *Random Structures Algorithms* **44**, 29–44.

[10] BERTOIN, J. AND URIBE BRAVO, G. (2015). Supercritical percolation on large scale-free random trees. *Ann. Appl. Prob.* **25**, 81–103.

[11] BERZUNZA, G. (2015). Yule processes with rare mutation and their applications to percolation on *b*-ary trees. *Electron. J. Prob.* **20**, 43, 23 pp.

[12] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd edn (Cambridge Studies in Advanced Mathematics). Cambridge University Press.

[13] DEVROYE, L. (1993). On the expected height of fringe-balanced trees. *Acta Inform.* **30**, 459–466.

[14] DEVROYE, L. (1999). Universal limit laws for depths in random trees. *SIAM J. Comput.* **28**, 409–432.

[15] DRMOTA, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Springer, New York and Vienna.

[16] DURRETT, R. (2010). *Random Graph Dynamics* (Cambridge Series in Statistical and Probabilistic Mathematics **20**). Cambridge University Press, Cambridge.

[17] GEIGER, J. (1996). Size-biased and conditioned random splitting trees. *Stoch. Process. Appl.* **65**, 187–207.

[18] GEIGER, J. AND KERSTING, G. (1997). Depth-first search of random trees, and Poisson point processes. In *Classical and Modern Branching Processes (Minneapolis, MN, 1994)* (IMA Vol. Math. Appl. **84**), pp. 111–126. Springer, New York.

[19] HOLMGREN, C. AND JANSON, S. (2017). Fringe trees, Crump–Mode–Jagers branching processes and *m*-ary search trees. *Prob. Surv.* **14**, 53–154.

[20] JAGERS, P. (1975). *Branching Processes with Biological Applications* (Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics). Wiley-Interscience, London, New York and Sydney.

[21] JAGERS, P. AND NERMAN, O. (1984). The growth and composition of branching populations. *Adv. Appl. Prob.* **16**, 221–259.

[22] JANSON, S. AND NEININGER, R. (2008). The size of random fragmentation trees. *Prob. Theory Related Fields* **142**, 399–442.

[23] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn (Probability and its Applications (New York)). Springer, New York.

[24] KOLMOGOROFF, A. N. (1941). Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C.R. (Doklady) Acad. Sci. URSS (N.S.)* **31**, 99–101.

[25] LAMBERT, A. (2010). The contour of splitting trees is a Lévy process. *Ann. Prob.* **38**, 348–395.

[26] MAHMOUD, H. M. (1994). A strong law for the height of random binary pyramids. *Ann. Appl. Prob.* **4**, 923–932.

[27] MUNTZ, R. AND UZGALIS, R. (1971). Dynamic storage allocation for binary search trees in a two-level memory. In *Proceedings of 4th Annual Princeton Conference on Information Sciences and Systems*, pp. 345–349.

[28] NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete* **57**, 365–395.

[29] NERMAN, O. AND JAGERS, P. (1984). The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete* **65**, 445–460.

[30] PITMAN, J. (1999). Coalescent random forests. *J. Combinatorial Theory A* **85**, 165–193.

[31] PITMAN, J. (2006). *Combinatorial Stochastic Processes* (Lecture Notes Math. **1875**). Springer, Berlin.

[32] PITTEL, B. (1994). Note on the heights of random recursive trees and random *m*-ary search trees. *Random Structures Algorithms* **5**, 337–347.

[33] RICHARD, M. (2011). Arbres, processus de branchement non markoviens et processus de Lévy. Doctoral thesis, Université Pierre et Marie Curie–Paris VI.

[34] RUDAS, A. AND TÓTH, B. (2009). Random tree growth with branching processes: a survey. In *Handbook of Large-Scale Random Networks* (Bolyai Soc. Math. Stud. **18**), pp. 171–202. Springer, Berlin.

[35] RUDAS, A., TÓTH, B. AND VALKÓ, B. (2007). Random trees and general branching processes. *Random Structures Algorithms* **31**, 186–202.

[36] SZYMAŃSKI, J. (1987). On a nonuniform random recursive tree. In *Random Graphs '85 (Poznań, 1985)* (North-Holland Math. Stud. **144**), pp. 297–306. North-Holland, Amsterdam.

[37] TSIRELSON, B. (2013). From uniform renewal theorem to uniform large and moderate deviations for renewal-reward processes. *Electron. Commun. Prob.* **18**, 52, 13 pp.