On the independence number of some random trees[*]

Svante Janson[†]

Abstract

We show that for many models of random trees, the independence number divided by the size converges almost surely to a constant as the size grows to infinity; the trees that we consider include random recursive trees, binary and m-ary search trees, preferential attachment trees, and others. The limiting constant is computed, analytically or numerically, for several examples. The method is based on Crump–Mode–Jagers branching processes.

Keywords: independence number; random trees; random recursive tree; binary search tree; Crump–Mode–Jagers branching process.
MSC2020 subject classifications: 60C05; 05C05; 05C69.
Submitted to ECP on June 3, 2020, final version accepted on August 24, 2020.
Supersedes arXiv:2003.08712v1.

1 Introduction

The *independence number* i.e., the maximum size of an independent set of nodes, is a quantity that has been studied for various models of random trees (and other random graphs, not considered here). In the present paper we consider rooted trees that can be constructed as family trees of a Crump–Mode–Jagers branching process stopped at a suitable stopping time; this includes, for example, random recursive trees, preferential attachment trees, fragmentation trees, binary search trees and m-ary search trees; see Section 2.1 and [8] for details, and the examples in Sections 4–8 below.

We denote the independence number of T by $I(T)$ and the number of nodes by $|T|$. Our main result, Theorem 3.1, gives a strong law of large numbers for $I(T)$; more precisely, it shows convergence almost surely (a.s.) of $I(T_n)/|T_n|$, the fraction of nodes that belong to a maximum independent set, for a sequence T_n of random trees. The limit ν is a constant depending on the random tree model; the theorem expresses this limit in terms of the solution $p(t)$ of the functional equation (3.3). We show in Sections 4–8 how this equation can be solved and ν found explicitly (at least numerically) in some important examples, viz. random recursive trees, binary search trees, preferential attachment trees, extended binary search tress and m-ary search trees (in particular $m = 3$).

Note that the cases of random recursive trees and binary search trees have been studied before. For random recursive trees, the expectation was found already by Meir and Moon [12, 15]. More recently, both Dadedzi [4] and Fuchs et al. [6] prove (independently, and with different methods) the weak version (i.e., convergence in

---

[†]Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.
  E-mail: svante.janson@math.uu.se

probability) of (3.2) below for random recursive trees and binary search trees, with explicit ν, and also a much stronger central limit theorem. Nevertheless, we think that the present approach is of interest, since it is quite general; moreover, it gives convergence a.s. Furthermore, although we only prove a law of large numbers in the present paper, we hope that future development of our methods will also lead to a central limit theorem.

Similar results have also been proved for other types of random trees. For simply generated trees, see e.g. Meir and Moon [11, 13] and Banderier, Kuba and Panholzer [2]; for uniform unlabelled trees (rooted or unrooted), see Meir and Moon [14].

Remark 1.1. As is well known, for trees, several other quantities are determined by the independence number by linear relations, and our results thus immediately transfer to these quantities. These include, for example:

 (i) The *minimum size of a vertex cover*, i.e., of a vertex set that contains at least one end-point of every edge. This equals $|T| - I(T)$; this is trivial because a set is a vertex cover if and only if its complement is independent.

 (ii) The *matching number*, i.e., the maximum size of a partial matching. This also equals $|T| - I(T)$, see e.g. [5, Theorem 2.1.1].

 (iii) The *nullity*, i.e., the dimension of the kernel of the adjacency matrix, or the multiplicity of the eigenvalue 0 of the adjacency matrix. This equals $2I(T) - |T|$, see e.g. [18, Corollary 5.2.6]. (The results in [4] referred to below are actually stated for the nullity.)

See e.g. [6] for further examples.

2 Preliminaries

We give some definitions and notation, together with some known results that will be used.

If T is a rooted tree, and $v \in T$ (i.e., v is a node in T), then T^v denotes the fringe subtree of T at v, i.e., the subtree consisting of v and all its descendants; T^v is defined as a rooted tree with root v.

$\mathrm{Exp}(\lambda)$ denotes an exponential random variable with rate λ; it thus has mean $1/\lambda$ and density function $\lambda e^{-\lambda x}$, $x > 0$.

2.1 Family trees of branching processes

We follow [8, Section 5], to which we refer for further details. Let $\mathcal{T}_t$ be the family tree of all individuals born up to time $t \geqslant 0$ in a given Crump–Mode–Jagers (CMJ) process, starting at time $t = 0$ with a single individual (the root). Let the children of the root be born at (random) times $(\xi_i)_1^N$, where $0 \leqslant N \leqslant \infty$ and $0 < \xi_1 \leqslant \xi_2 \leqslant \dots$. We regard the (multi)set of birth times as a point process Ξ; formally, Ξ is the random (discrete) measure $\sum_i \delta_{\xi_i}$, where δ_t is the Dirac measure (point mass) at t. Moreover, each individual x has its own copy Ξ_x of Ξ; the processes Ξ_x are i.i.d. (independent and identically distributed). Let σ_x be the time individual x is born. For simplicity we assume that all individuals live forever.

Let Z_t be the number of individuals at time t. In the simplest, and most common, case, we define the stopping time

$$\tau(n) := \inf\{t : Z_t \geqslant n\}, \tag{2.1}$$

the first time the number of individuals is at least n, and $T_n := \mathcal{T}_{\tau(n)}$, the family tree at that time. (By the assumptions below, $\tau(n) < \infty$ a.s.) Thus T_n is a random tree with

$|T_n| \geqslant n$. Typically, the birth times ξ_i are continuous random variables and a.s. no two births are simultaneous, and then $|T_n| = n$.

More generally, we fix a *weight* $\psi(s)$. This is assumed to be a *characteristic*, i.e., a random function $\psi(s) \geqslant 0$ associated to the root and its point process Ξ, and we assume that each individual x is equipped with its own copy $\psi_x(s)$ of ψ; the simplest case is that $\psi_x(s)$ is a deterministic function of the point process Ξ_x. (More generally, ψ_x may also depend on the entire tree of descendants of x, and possibly also on some extra randomness, see [8, in particular Remark 5.10].) We assume $\psi_x \in D[0, \infty)$, and we exclude the trivial case $\psi(t) = 0$ for all $t \geqslant 0$ a.s. The argument $s \geqslant 0$ of $\psi_x(s)$ should be interpreted as the current age of x, which is $t - \sigma_x$ at time t. Let

$$Z_t^\psi := \sum_{x:\sigma_x \leqslant t} \psi_x(t - \sigma_x) \tag{2.2}$$

be the total weight at time $t \geqslant 0$. We then let

$$\tau(n) := \inf\{t : Z_t^\psi \geqslant n\}, \tag{2.3}$$

the first time the total weight is at least n. (We define $\inf \emptyset = \infty$.) Finally, as before, we define $T_n := \mathcal{T}_{\tau(n)}$. Note that the choice $\psi(s) = 1$ gives $Z_t^\psi = Z_t$, and thus the simple definition (2.1) of T_n.

Examples of common random trees that can be constructed as T_n in this way are given in Sections 4–8; see further [8].

Let $\mu := \mathbb{E}\,\Xi$ be the intensity of the point process Ξ. In other words, μ is the (deterministic) measure on $[0, \infty)$ such that, for any Borel set A, $\mu(A)$ is the expected number of children of the root born at times $t \in A$. In particular, with $N \leqslant \infty$ as above the (random) total number of children of the root, $\mu[0, \infty) = \mathbb{E}\,N \leqslant \infty$.

We use the following assumptions throughout the paper:

(A1) $\xi_1 > 0$, i.e., no children are born immediately at their parent's birth. (Equivalently, $\mu\{0\} = 0$.)

(A2) μ is not concentrated on any lattice $h\mathbb{Z}$, $h > 0$. (The results extend to the lattice case with suitable modifications, but we do not know any interesting examples and ignore this case.)

(A3) $N \geqslant 1$ a.s. and $\mathbb{E}\,N > 1$. (Thus, every individual has at least one child, so the process never dies out, and $Z_\infty = \infty$ a.s.)

(A4) There exists a real number α (the *Malthusian parameter*) such that

$$\int_0^\infty e^{-\alpha t} \mu(\mathrm{d}t) = 1. \tag{2.4}$$

(By 2.1, $\alpha > 0$.)

(A5) There exists $\theta < \alpha$ such that

$$\int_0^\infty e^{-\theta t} \mu(\mathrm{d}t) < \infty. \tag{2.5}$$

2.2 Independence numbers

We collect here some simple and well-known properties of independence numbers of (rooted) trees; see e.g. [11].

For a tree T, let $I(T)$ be the independence number of T, i.e., the maximum size of an independent set of nodes. For a rooted tree T, let further $I_1(T)$ be the maximum size of

an independent node set containing the root, and let $I_0(T)$ be the maximum size of an independent node set not containing the root. Thus, with $x \vee y$ denoting $\max\{x, y\}$,

$$I(T) = I_1(T) \vee I_0(T). \tag{2.6}$$

Furthermore, if the children of the root are $v_1, \ldots, v_d$, then it is easily seen that

$$I_0(T) = \sum_{i=1}^{d} I(T^{v_i}), \tag{2.7}$$

$$I_1(T) = 1 + \sum_{i=1}^{d} I_0(T^{v_i}). \tag{2.8}$$

Since $I_0(T) \leqslant I(T)$ by (2.6), it follows that $I_1(T) \leqslant I_0(T) + 1$, and thus

$$I_0(T) \leqslant I(T) \leqslant I_0(T) + 1. \tag{2.9}$$

Define

$$\iota(T) := I(T) - I_0(T) \in \{0, 1\}. \tag{2.10}$$

Then (2.7) yields

$$I(T) = \iota(T) + \sum_{i=1}^{d} I(T^{v_i}), \tag{2.11}$$

which shows that the independence number $I(T)$ is an additive functional on rooted trees with toll function $\iota(T)$.

As is well known, (2.11) is equivalent to

$$I(T) = \sum_{v \in T} \iota(T^v). \tag{2.12}$$

Furthermore, by (2.10), (2.6) and (2.7)–(2.8),

$$\iota(T) = 1 \iff I_1(T) = 1 + I_0(T) \iff \sum_{i=1}^{d} I_0(T^{v_i}) = \sum_{i=1}^{d} I(T^{v_i})$$

$$\iff \iota(T^{v_i}) = 0 \text{ for every child } v_i \text{ of the root.} \tag{2.13}$$

Say that a node $v \in T$ is *essential* if it belongs to every maximum independent set of T^v. This is equivalent to $I_1(T^v) > I_0(T^v)$, and thus to $\iota(T^v) = 1$. In other words,

$$\iota(T^v) = \mathbb{1}\{v \text{ is essential}\}. \tag{2.14}$$

In particular, $\iota(T)$ equals the indicator that the root is essential in T. Note also that, by (2.14) and (2.13),

$$\text{a node is essential if and only if none of its children is.} \tag{2.15}$$

Remark 2.1. By (2.12) and (2.14), the independence number $I(T)$ equals the number of essential nodes in T. Moreover, (2.15) implies that the set of essential nodes is independent, and thus an independent set of maximum size.

3   Main result

We next state our main theorem. Recall that the Laplace functional of the point process Ξ is defined as

$$L_\Xi(f) := \mathbb{E}\, e^{-\int f \, \mathrm{d}\Xi} \tag{3.1}$$

for (measurable) functions $f \geqslant 0$ on $[0, \infty)$.

Theorem 3.1. *Let T_n, $n \geqslant 1$, be random trees that can be defined as stopped family trees of Crump–Mode–Jagers processes as in Section 2.1, for some point process $\Xi = (\xi_i)$ satisfying assumptions 2.1–2.1and some weight ψ. Then, as $n \to \infty$,*

$$\frac{I(T_n)}{|T_n|} \xrightarrow{\text{a.s.}} \nu := \alpha \int_0^\infty e^{-\alpha t} p(t) \, \mathrm{d}t, \tag{3.2}$$

where α is the Malthusian parameter and $p(t)$ is the unique function $[0, \infty) \to (0, 1]$ satisfying

$$p(t) = \mathbb{E} \prod_{i:\xi_i \leqslant t} \big(1 - p(t - \xi_i)\big) = \mathbb{E}\, e^{\int_0^t \log(1 - p(t-s)) \, \mathrm{d}\Xi(s)}$$
$$= L_\Xi\Big(-\log\big(1 - p(t - \cdot)\big)\mathbb{1}\{\cdot \leqslant t\}\Big), \qquad t \geqslant 0. \tag{3.3}$$

Note that the result does not depend on the choice of weigth ψ.

Proof. By (2.12) and (2.14), $I(T_n)/|T_n|$ is the fraction of nodes in T_n that are essential. We apply [8, Theorem 5.14(ii)] to the property that a node is essential. (This theorem is a special case of deep results by Jagers and Nerman [9, 16], see also Aldous [1].) Then [8, (5.23)–(5.24)] yield (3.2), with

$$p(t) := \mathbb{P}\big(\text{the root is essential in } \mathcal{T}_t\big) = \mathbb{P}\big(\iota(\mathcal{T}_t) = 1\big). \tag{3.4}$$

To see (3.3), condition on $\Xi = (\xi_i)_i$, i.e., on the sequence of times that the root gives birth. Then, the children of the root of $\mathcal{T}_t$ are the individuals i born at times $\xi_i \leqslant t$. Each such child has grown a tree $\mathcal{T}_t^i$ that has the same distribution as $\mathcal{T}_{t-\xi_i}$, and thus

$$\mathbb{P}\big(\iota(\mathcal{T}_t^i) = 0 \mid \Xi\big) = 1 - p(t - \xi_i), \qquad \xi_i \leqslant t. \tag{3.5}$$

Furthermore, still conditioned on Ξ, the events in (3.5) for different i are independent, and thus using (2.13),

$$\mathbb{P}\big(\iota(\mathcal{T}_t) = 1 \mid \Xi\big) = \mathbb{P}\big(\iota(\mathcal{T}_t^i) = 0 \text{ for every child } i \text{ of the root} \mid \Xi\big)$$
$$= \prod_{i:\xi_i \leqslant t} \big(1 - p(t - \xi_i)\big). \tag{3.6}$$

Hence, (3.3) follows, using (3.1).

Finally, suppose that $p_1(t)$ is another function $[0, \infty) \to [0, 1]$ that satisfies (3.3), and let $\Delta p(t) := |p(t) - p_1(t)|$. Then,

$$\Delta p(t) = \left| \mathbb{E} \prod_{i:\xi_i \leqslant t} \big(1 - p(t - \xi_i)\big) - \mathbb{E} \prod_{i:\xi_i \leqslant t} \big(1 - p_1(t - \xi_i)\big) \right|$$
$$\leqslant \mathbb{E} \sum_{i:\xi_i \leqslant t} \big|p(t - \xi_i) - p_1(t - \xi_i)\big| = \mathbb{E} \sum_{i:\xi_i \leqslant t} \Delta p(t - \xi_i). \tag{3.7}$$

Fix $\beta > \alpha$, and define $h(t) := \sup_{s \leqslant t} e^{-\beta s} \Delta p(s) \in [0, 1]$. Then, (3.7) yields

$$e^{-\beta t} \Delta p(t) \leqslant e^{-\beta t} \, \mathbb{E} \sum_{i : \xi_i \leqslant t} e^{\beta(t - \xi_i)} h(t) = h(t) \, \mathbb{E} \sum_{i : \xi_i \leqslant t} e^{-\beta \xi_i}$$

$$= h(t) \, \mathbb{E} \int_0^t e^{-\beta x} \, \mathrm{d}\Xi(x) = h(t) \int_0^t e^{-\beta x} \, \mathrm{d}\mu(x). \tag{3.8}$$

Since $h(t)$ is monotone, this implies

$$h(t) = \sup_{s \leqslant t} \big(e^{-\beta s} \Delta p(s) \big) \leqslant h(t) \int_0^t e^{-\beta x} \, \mathrm{d}\mu(x). \tag{3.9}$$

However, by (2.4),

$$\int_0^t e^{-\beta x} \, \mathrm{d}\mu(x) \leqslant \int_0^\infty e^{-\beta x} \, \mathrm{d}\mu(x) < \int_0^\infty e^{-\alpha x} \, \mathrm{d}\mu(x) = 1, \tag{3.10}$$

and thus (3.9) implies $h(t) = 0$ for any $t \geqslant 0$. Thus $p_1(t) = p(t)$, and the solution to (3.3) is unique. $\qquad \square$

Note that $\mathcal{T}_0$ consists of the root only, and thus (3.4) yields the initial condition, also a trivial special case of (3.3),

$$p(0) = 1. \tag{3.11}$$

Remark 3.2. An explanation for the formula (3.2) for ν is that a random fringe tree of T_n converges in distribution to $\mathcal{T}_\tau$, the tree obtained by stopping the branching process at a time $\tau \sim \mathrm{Exp}(\alpha)$ independent of the brancing process; thus ν is the probability that the root of $\mathcal{T}_\tau$ is essential. See further [8].

Remark 3.3. It is sometimes convenient to define $p(t) := 0$ for $t < 0$; then (3.3) may be written

$$p(t) = \mathbb{E} \prod_{i=1}^N \big(1 - p(t - \xi_i) \big), \qquad t \geqslant 0, \tag{3.12}$$

taking the product over all children (born yet or not). We will use this a couple of times, but note that all formulas for $p(t)$ assume $t \geqslant 0$.

4 The random recursive tree

The random recursive tree is an example of a random tree that can be constructed as in Section 2.1, taking Ξ to be a Poisson process with constant intensity 1 on $[0, \infty)$ and the trivial weight $\psi(t) = 1$, see [8, Example 6.1]. Thus Theorem 3.1 applies and shows $I(T_n)/|T_n| \xrightarrow{\text{a.s.}} \nu$ as $n \to \infty$.

To find the limit ν, note first that by the standard formula [10, Theorem 3.9] for the Laplace functional of a rate 1 Poisson process

$$L_\Xi(f) = e^{-\int (1 - e^{-f(s)}) \, \mathrm{d}s}, \tag{4.1}$$

(3.3) yields

$$p(t) = \exp\Big(-\int_0^t p(t - s) \, \mathrm{d}s \Big) = \exp\Big(-\int_0^t p(u) \, \mathrm{d}u \Big). \tag{4.2}$$

This can also be seen directly as follows. The number of children of the root at time t has the Poisson distribution $\mathrm{Po}(t)$. Furthermore, a child born at time $s \leqslant t$ has probability

$p(t-s)$ of being essential at time t, and thus, by the independence properties of the branching process, the children of the root that are essential at time t are born according to a random thinning of the rate 1 Poisson process; this thinning is a Poisson process on $[0,t]$ with intensity $p(t-\cdot)$. In particular, the number of children of the root of $\mathcal{T}_t$ that are essential is $\mathrm{Po}(\lambda(t))$ with $\lambda(t) = \int_0^t p(t-s)\,\mathrm{d}s$. By (2.15), the root is essential if and only if this number is 0, which has probability $e^{-\lambda(t)}$, and (4.2) follows.

Since $p(t) \in [0,1]$, (4.2) implies that $p(t)$ is continuous, and by induction infinitely differentiable. Differentiating (4.2) yields

$$p'(t) = -\exp\Big(-\int_0^t p(u)\,\mathrm{d}u\Big)p(t) = -p(t)^2, \qquad t > 0, \tag{4.3}$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{p(t)} = -\frac{p'(t)}{p(t)^2} = 1. \tag{4.4}$$

Consequently, by the initial condition $p(0) = 1$ (3.11),

$$p(t) = \frac{1}{1+t}. \tag{4.5}$$

The Malthusian parameter $\alpha = 1$, and thus by (3.2) and (4.5)

$$\nu = \int_0^\infty \frac{e^{-t}}{t+1}\,\mathrm{d}t = \int_0^1 \frac{1}{1-\log x}\,\mathrm{d}x = 0.59634736\ldots \tag{4.6}$$

as proved by Meir and Moon [12, 15], Dadedzi [4] and Fuchs et al. [6]; this number is known as the Euler–Gompertz constant.

5 Preferential attachment trees

Consider now a preferential attachment tree, where nodes are added one by one, and each new node chooses a parent at random, with the probability of choosing a node v as the parent is proportional to $\chi d(v) + \rho$, where $d(v)$ is the current outdegree of v, and χ and ρ are given constants. (Here $\rho > 0$ and either $\chi \geqslant 0$ or $\rho/|\chi|$ is an integer $\geqslant 2$. Only the ratio χ/ρ is significant.) This random tree can be constructed by a CMJ process where an individual that already has k children gets the next child with rate $\chi k + \rho$; see [8, Example 6.4]. Again the weight $\psi(t) = 1$. Thus Theorem 3.1 applies and shows $I(T_n)/|T_n| \xrightarrow{\text{a.s.}} \nu$ as $n \to \infty$.

To find $p(t)$ and ν, we use instead of (3.3) the following (closely related) argument. Consider also, for $\lambda > 0$, a modified branching process $\mathfrak{X}_\lambda$, where the starting individual (= the root) is special, and gets children with the rate $\chi k + \lambda$, where k is the current number of children. All other individuals are as before, with rate $\chi k + \rho$. (If $\chi < 0$, we assume that $\lambda/|\chi|$ is an integer; this case will be enough below.) Let $p_\lambda(t)$ be the probability that the root is essential in the family tree of $\mathfrak{X}_\lambda$ at time t. Note that if $\lambda = \rho$, then the modified process equals the original one, and thus $p(t) = p_\rho(t)$.

Consider again the original process, let $t > 0$, and condition on the first child of the root being born at time $s \leqslant t$. The probability that this child is not essential at time t is $1 - p(t-s)$. Furthermore, if we ignore this child and its descendants, the rest of the tree evolves after time s as the modified process $\mathfrak{X}_{\chi+\rho}$. Hence, the probability that no child of the root after the first is essential at time t is $p_{\chi+\rho}(t-s)$. (Note that the property that such a child v is essential depends only on the fringe subtree T^v, and is thus not affected by the deletion of the first child of the root and its offspring.) Consequently, conditioned on the first child being born at time $s \leqslant t$, the probability that the root has

no essential child at time t is $\big(1 - p(t-s)\big)p_{\chi+\rho}(t-s)$. By (2.15), this is the conditional probability that the root is essential at time t, and since the time the first child is born has the distribution $\mathrm{Exp}(\rho)$, we have

$$p(t) = e^{-\rho t} + \rho \int_0^t e^{-\rho s}\big(1 - p(t-s)\big)p_{\chi+\rho}(t-s)\,\mathrm{d}s$$

$$= e^{-\rho t} + \rho e^{-\rho t}\int_0^t e^{\rho u}\big(1 - p(u)\big)p_{\chi+\rho}(u)\,\mathrm{d}u. \tag{5.1}$$

If we have two independent modified processes $\mathfrak{X}_{\lambda_1}$ and $\mathfrak{X}'_{\lambda_2}$, then we may merge them by identifying the two roots. This yields a modified process $\mathfrak{X}_{\lambda_1+\lambda_2}$ with parameter $\lambda = \lambda_1 + \lambda_2$. There are no essential children of the root in the combined process if and only if there are none in both modified processes taken separately; hence, it follows that, for any $t \geqslant 0$ and $\lambda_1, \lambda_2 > 0$,

$$p_{\lambda_1+\lambda_2}(t) = p_{\lambda_1}(t)p_{\lambda_2}(t). \tag{5.2}$$

Fix $t > 0$. Since $0 < p_\lambda(t) \leqslant 1$, (5.2) implies that $\lambda \mapsto p_\lambda(t)$ is decreasing, which in turn implies that (5.2) has the solution $p_\lambda(t) = e^{-C(t)\lambda}$ for some $C(t) \geqslant 0$. Hence,

$$p_\lambda(t) = p_\rho(t)^{\lambda/\rho} = p(t)^{\lambda/\rho}. \tag{5.3}$$

Combining (5.1) and (5.3) yields the functional equation

$$p(t) = e^{-\rho t} + \rho e^{-\rho t}\int_0^t e^{\rho u}\big(1 - p(u)\big)p(u)^{\chi/\rho+1}\,\mathrm{d}u. \tag{5.4}$$

Again, p is infinitely differentiable, and taking the derivative yields

$$p'(t) = -\rho e^{-\rho t} - \rho^2 e^{-\rho t}\int_0^t e^{\rho u}\big(1 - p(u)\big)p(u)^{\chi/\rho+1}\,\mathrm{d}u + \rho\big(1 - p(t)\big)p(t)^{\chi/\rho+1}$$

$$= -\rho p(t) + \rho\big(1 - p(t)\big)p(t)^{\chi/\rho+1}. \tag{5.5}$$

Let $\chi' := \chi/\rho$, and

$$h(x) := x - x^{\chi'+1}(1 - x) = x - x^{\chi'+1} + x^{\chi'+2}. \tag{5.6}$$

Then (5.5) can be written

$$p'(t) = -\rho h(p(t)). \tag{5.7}$$

By (5.6), $h(1) = 1 > 0$ and $h(0) = 0$. (If $\chi < 0$, then $\rho = m|\chi|$ for an integer $m \geqslant 2$, and thus $\chi' = -1/m \in [-\frac{1}{2}, 0)$, so $\chi' + 1 > 0$ also in this case.) Let q be the largest zero of h in $[0,1]$, i.e.,

$$q := \max\{x \in [0,1] : h(x) = 0\}. \tag{5.8}$$

By continuity and $h(0) = 0$, this maximum always exists. Furthermore, if $\chi \geqslant 0$, then (5.6) implies $h(x) > 0$ on $(0,1]$, and thus $q = 0$. On the other hand, if $\chi < 0$, then $h(x) < 0$ for small positive x, and thus $0 < q < 1$.

The function $p(t)$ is continuous on $[0,\infty)$, with $p(0) = 1$. Suppose that $p(t) = q$ for some $t < \infty$, and let t_0 be the smallest such t. Then $p(t) \in [q,1]$ for $t \in [0,t_0]$. However, $h(x)$ is continuously differentiable and thus Lipschitz on $[q,1]$, as is seen by considering the cases $\chi \geqslant 0$ and $\chi < 0$ separately, and thus the differential equation (5.7) has at most one solution for $t \in [0,t_0]$ with $p(t) \in [q,1]$ and $p(t_0) = q$. Since $p(t) = q$ is another

solution of (5.7), this is a contradiction. Hence, $p(t) \neq q$, and thus, by continuity, $p(t) > q$ for all $t \geqslant 0$.

This further implies, by (5.7) again, that $p(t)$ is strictly decreasing on $[0, \infty)$. Hence the limit $p(\infty) := \lim_{t \to \infty} p(t)$ exists. Then (5.7) implies $p'(t) \to -\rho h(p(\infty))$ as $t \to \infty$, and thus $p(\infty) > q$ is impossible; hence,

$$p(t) \to p(\infty) = q, \qquad t \to \infty. \tag{5.9}$$

Consequently, p is a bijection $[0, \infty) \to (q, 1]$.

Let, for $x \in (q, 1]$,

$$\Psi(x) := \int_x^1 \frac{1}{h(y)} \, dy. \tag{5.10}$$

Thus $\Psi(1) = 0$ and $\Psi'(x) = -1/h(x)$. Hence, (5.7) implies by the chain rule

$$\frac{d}{dt} \Psi(p(t)) = \Psi'(p(t)) p'(t) = \frac{-\rho h(p(t))}{-h(p(t))} = \rho, \tag{5.11}$$

and thus

$$\Psi(p(t)) = \rho t, \qquad t \geqslant 0. \tag{5.12}$$

Hence, if $\Psi^{-1} : [0, \infty) \to (q, 1]$ denotes the inverse function, then

$$p(t) = \Psi^{-1}(\rho t). \tag{5.13}$$

It follows from (5.12), letting $t \to \infty$, that the limit $\Psi(q) = \infty$, which also easily can be seen directly from (5.10).

The Malthusian parameter $\alpha = \chi + \rho$, see [8, (6.20)], and thus (3.2) and (5.13) yield

$$\nu = \alpha \int_0^\infty e^{-\alpha t} \Psi^{-1}(\rho t) \, dt = (\chi' + 1) \int_0^\infty e^{-(\chi'+1)s} \Psi^{-1}(s) \, ds. \tag{5.14}$$

The change of variables $s = \Psi(x)$ and an integration by parts yield the formulas

$$\nu = (\chi' + 1) \int_1^q e^{-(\chi'+1)\Psi(x)} x \Psi'(x) \, dx \tag{5.15}$$

$$= (\chi' + 1) \int_q^1 e^{-(\chi'+1)\Psi(x)} \frac{x}{h(x)} \, dx \tag{5.16}$$

$$= 1 - \int_q^1 e^{-(\chi'+1)\Psi(x)} \, dx. \tag{5.17}$$

These integrals can be evaluated numerically.

Example 5.1. Let $\chi = 0$ and $\rho = 1$; this yields the random recursive tree in Section 4. We have $\chi' = 0$, and thus (5.6) yields $h(x) = x^2$. Hence, (5.10) yields $\Psi(x) = 1/x - 1 = (1-x)/x$, and (5.13) yields (4.5) again. Furthermore, (5.16) becomes

$$\nu = \int_0^1 e^{1-1/x} x^{-1} \, dx = \int_1^\infty e^{1-y} y^{-1} \, dy, \tag{5.18}$$

which by a change of variables agrees with (4.6).

Example 5.2. Let $\chi = \rho = 1$; this yields the standard preferential attachment random tree. (This is the same as the plane oriented recursive tree [17]; it is a special case of the preferential attachment graphs [3], [7, Chapter 8].) We have $\chi' = 1$, and thus (5.6) yields

$$h(x) = x - x^2 + x^3. \tag{5.19}$$

We find from (5.10),

$$\Psi(x) = \tfrac{1}{2}\log(x^2 - x + 1) - \log(x) - \frac{1}{\sqrt{3}}\arctan\!\Big(\frac{2x-1}{\sqrt{3}}\Big) + \frac{\pi}{6\sqrt{3}} \qquad (5.20)$$

and thus (5.17) yields, with an integral that magically has an elementary primitive function,

$$
\begin{aligned}
\nu &= 1 - \int_0^1 e^{-2\Psi(x)}\,\mathrm{d}x = 1 - e^{-\frac{\pi}{3\sqrt{3}}} \int_0^1 e^{\frac{2}{\sqrt{3}}\arctan((2x-1)/\sqrt{3})}\frac{x^2}{1 - x + x^2}\,\mathrm{d}x \\
&= 1 - e^{-\frac{\pi}{3\sqrt{3}}}\Big[(x-1)e^{\frac{2}{\sqrt{3}}\arctan((2x-1)/\sqrt{3})}\Big]_0^1 \\
&= 1 - e^{-2\pi/(3\sqrt{3})} = 0.70156394\ldots
\end{aligned}
\qquad (5.21)
$$

Thus, Theorem 3.1 shows that for the standard preferential attachment tree, $I(T_n)/|T_n| \xrightarrow{\text{a.s.}} 1 - e^{-2\pi/(3\sqrt{3})}$ as $n \to \infty$.

6 The binary search tree

The binary search tree is another example where Theorem 3.1 applies. (This is the special case $\chi = -1$, $\rho = 2$ of Section 5, but we prefer to treat it directly.) Now each node gets two children, after waiting times that are independent and $\mathrm{Exp}(1)$. Again, $\psi(t) = 1$.

We proceed to find the limit ν. In this case, (3.3) yields

$$p(t) = \Big(1 - \int_0^t p(t-s)e^{-s}\,\mathrm{d}s\Big)^2 = \Big(1 - e^{-t}\int_0^t e^u p(u)\,\mathrm{d}u\Big)^2, \qquad t \geqslant 0. \qquad (6.1)$$

This can also, perhaps more easily, be seen as follows. As always, if a child is born at time $s \leqslant t$, then the probability that this child is essential at time t is $p(t-s)$. Hence, the probability that the left child of the root is born and is essential at time t is $\int_0^t p(t-s)e^{-s}\,\mathrm{d}s$, and thus the probability that there is no left child that is essential equals $1 - \int_0^t p(t-s)e^{-s}\,\mathrm{d}s$. The same holds for the right child, and since the two children appear and develop independently, (2.15) yields (6.1).

To solve the functional equation (6.1), let $g(t) := p(t)^{1/2}$, so (6.1) may be written

$$g(t) = 1 - e^{-t}\int_0^t e^s g(s)^2\,\mathrm{d}s. \qquad (6.2)$$

It follows, by induction, that g is infinitely differentiable; furthermore, (6.2) yields the differential equation

$$g'(t) = e^{-t}\int_0^t e^s g(s)^2\,\mathrm{d}s - g(t)^2 = 1 - g(t) - g(t)^2. \qquad (6.3)$$

This differential equation is separable and can be written

$$\frac{\mathrm{d}g}{1 - g - g^2} = \mathrm{d}t, \qquad (6.4)$$

which is solved by standard methods as follows.

Let $\gamma_\pm := \frac{-1 \pm \sqrt{5}}{2}$ be the roots of $1 - \gamma - \gamma^2 = 0$. Then

$$\frac{1}{1 - g - g^2} = -\frac{1}{(g - \gamma_+)(g - \gamma_-)} = \frac{1}{\gamma_+ - \gamma_-}\Big(\frac{1}{g - \gamma_-} - \frac{1}{g - \gamma_+}\Big) \qquad (6.5)$$

and thus (6.4) can be integrated to

$$\log(g(t) - \gamma_-) - \log(g(t) - \gamma_+) = (\gamma_+ - \gamma_-)t + C, \qquad (6.6)$$

where $g(0) = 1$ yields $C = \log(1 - \gamma_-) - \log(1 - \gamma_+)$. Hence,

$$\frac{g(t) - \gamma_-}{g(t) - \gamma_+} = \frac{1 - \gamma_-}{1 - \gamma_+} e^{(\gamma_+ - \gamma_-)t} \tag{6.7}$$

and thus

$$g(t) = \frac{\gamma_+(1 - \gamma_-)e^{(\gamma_+ - \gamma_-)t} - \gamma_-(1 - \gamma_+)}{(1 - \gamma_-)e^{(\gamma_+ - \gamma_-)t} - (1 - \gamma_+)}. \tag{6.8}$$

Consequently, $p(t) = g(t)^2$ with $g(t)$ given by (6.8). Note also that we have $\gamma_+ - \gamma_- = \sqrt{5}$ and, with $\phi := \frac{1+\sqrt{5}}{2}$, the golden ratio,

$$\gamma_+ = \phi^{-1}, \qquad\qquad 1 - \gamma_+ = \phi^{-2}, \tag{6.9}$$
$$\gamma_- = -\phi, \qquad\qquad 1 - \gamma_- = \phi^2. \tag{6.10}$$

The Malthusian parameter $\alpha = 1$, and thus (3.2) and (6.8) yield, with $x = e^{-t}$,

$$\begin{aligned}
\nu &= \int_0^\infty e^{-t} g(t)^2 \, \mathrm{d}t = \int_0^\infty e^{-t} \left(\frac{\phi e^{\sqrt{5}\,t} + \phi^{-1}}{\phi^2 e^{\sqrt{5}\,t} - \phi^{-2}} \right)^2 \mathrm{d}t \\
&= \int_0^1 \left(\frac{\phi + \phi^{-1} x^{\sqrt{5}}}{\phi^2 - \phi^{-2} x^{\sqrt{5}}} \right)^2 \mathrm{d}x = \frac{\phi^2}{\sqrt{5}} \int_0^1 \left(\frac{\phi^2 + y}{\phi^4 - y} \right)^2 y^{1/\sqrt{5}-1} \, \mathrm{d}y.
\end{aligned} \tag{6.11}$$

This integral can be evaluated as the sum of a rapidly (geometrically) convergent series by expanding $(\phi^4 - y)^{-2} = \phi^{-8}(1 - \phi^{-4}y)^{-2}$ into a power series, which yields

$$\nu = \sum_{k=0}^\infty (k+1)\phi^{-4k-6} \left(\frac{\phi^4}{k\sqrt{5}+1} + \frac{2\phi^2}{(k+1)\sqrt{5}+1} + \frac{1}{(k+2)\sqrt{5}+1} \right). \tag{6.12}$$

The integrals (6.11) and the sum (6.12) are all easily evaluated numerically, yielding (by Maple)

$$\nu = 0.54287631\ldots \tag{6.13}$$

as found by [4] and [6].

7   Extended binary search trees

An extended binary search tree is a binary search tree where we have added further leaves at all possible places; thus the original nodes (called internal nodes) have all two children each, and the new nodes (called external nodes) have no children. This can be constructed by a CMJ process where each individual gets twins after an $\mathrm{Exp}(1)$ time (and no further children). Note that in the tree $\mathcal{T}_t$, the internal nodes are the ones that have had children, while the others are external nodes.

We may choose to measure the size of an extended binary search tree in three different ways: the total number of nodes, the number of internal nodes, or the number of external nodes. (These are related in simple ways, since in the binary case treated here, the number of external nodes is always $1 +$ the number of internal nodes.) We obtain these three versions as our T_n by choosing different weight functions ψ; $\psi(t) = 1$ as usual gives the total number of vertices, while the number of internal vertices is given by Z_t^ψ with $\psi(t) := \mathbb{1}\{\xi_1 \leqslant t\}$ and the number of external vertices is given by $\psi(t) := \mathbb{1}\{\xi_1 > t\}$. Recall that Theorem 3.1 applies, and gives the same limit, to all three versions.

Since we have $\xi_1 = \xi_2$, (3.3) yields (with $p(t) = 0$ for $t < 0$, see Remark 3.3)

$$p(t) = \mathbb{E}\big[(1 - p(t - \xi_1))^2\big] = e^{-t} + \int_0^t e^{-s}\big(1 - p(t - s)\big)^2\,\mathrm{d}s$$

$$= e^{-t} + e^{-t}\int_0^t e^u\big(1 - p(u)\big)^2\,\mathrm{d}u, \qquad t \geqslant 0. \tag{7.1}$$

This yields the differential equation

$$p'(t) = -p(t) + \big(1 - p(t)\big)^2 = 1 - 3p(t) + p(t)^2. \tag{7.2}$$

Let $q(t) := 1 - p(t)$. Then (7.2) yields

$$q'(t) = -p'(t) = 1 - q(t) - q(t)^2, \tag{7.3}$$

which is the same differential equation as (6.3), although now the initial condition is $q(0) = 1 - p(0) = 0$. The general solution is as in (6.6), and we obtain, cf. (6.7)–(6.8),

$$\frac{q(t) - \gamma_-}{q(t) - \gamma_+} = \frac{\gamma_-}{\gamma_+}e^{(\gamma_+ - \gamma_-)t} \tag{7.4}$$

and, using (6.9)–(6.10),

$$q(t) = \frac{e^{(\gamma_+ - \gamma_-)t} - 1}{\gamma_+ - \gamma_- e^{(\gamma_+ - \gamma_-)t}} = \frac{e^{\sqrt{5}t} - 1}{\phi^{-1} + \phi e^{\sqrt{5}t}} = \frac{\phi e^{\sqrt{5}t} - \phi}{\phi^2 e^{\sqrt{5}t} + 1}. \tag{7.5}$$

Thus, recalling $\phi^2 = \phi + 1$,

$$p(t) = 1 - q(t) = \frac{(\phi^2 - \phi)e^{\sqrt{5}t} + \phi + 1}{\phi^2 e^{\sqrt{5}t} + 1} = \frac{e^{\sqrt{5}t} + \phi^2}{\phi^2 e^{\sqrt{5}t} + 1} \tag{7.6}$$

and, since the Malthusian parameter $\alpha = 1$, using $x = e^{-t}$,

$$\nu = \int_0^\infty p(t)e^{-t}\,\mathrm{d}t = \int_0^1 \frac{1 + \phi^2 x^{\sqrt{5}}}{\phi^2 + x^{\sqrt{5}}}\,\mathrm{d}x = \phi^2 - (\phi^4 - 1)\int_0^1 \frac{\mathrm{d}x}{\phi^2 + x^{\sqrt{5}}}$$

$$= \phi^2 - (3\phi + 1)\sum_{k=0}^\infty (-1)^k \frac{\phi^{-2-2k}}{k\sqrt{5} + 1} = 0.5987899\ldots \tag{7.7}$$

8 m-ary search trees

Consider now an m-ary search tree, for a given $m \geqslant 3$. (The case $m = 2$ was studied in Section 6.) The m-ary search tree T_n generated by n random keys can be constructed by the following CMJ process and weight ψ, see [8, Section 7.2].

Each individual (node) starts by gaining weight; the weight $\psi(t)$ represents the number of keys in the node. It starts with $\psi(0) = 1$, and then increases by 1 after successive independent waiting times $Y_2, \ldots, Y_{m-1}$ with $Y_i \sim \mathrm{Exp}(i)$. At time $S := \sum_{i=2}^{m-1} Y_i$ the weight thus reaches $m - 1$; this marks puberty, and the node becomes fertile and gets m children after further independent waiting times $X_i \sim \mathrm{Exp}(1)$. (Thus, child i is born at $S + X_i$.)

Theorem 3.1 thus applies. To find ν, we condition on S and find that if $0 \leqslant s \leqslant t$, then (with $p(u) = 0$ for $u < 0$, see Remark 3.3)

$$\mathbb{E}\Big(\prod_{i:\xi_i \leqslant t} \big(1 - p(t - \xi_i)\big) \mid S = s \Big) = \mathbb{E}\Big(\prod_{i=1}^m \big(1 - p(t - \xi_i)\big) \mid S = s \Big)$$

$$= \mathbb{E} \prod_{i=1}^m \big(1 - p(t - s - X_i)\big) = \Big(1 - \int_0^{t-s} e^{-x} p(t - s - x)\, \mathrm{d}x \Big)^m$$

$$= \Big(1 - e^{s-t} \int_0^{t-s} e^y p(y)\, \mathrm{d}y \Big)^m. \tag{8.1}$$

Hence, (3.3) or (3.12) yields, for $t \geqslant 0$, if f_S is the density function of S,

$$p(t) = \mathbb{P}(S > t) + \int_0^t \Big(1 - e^{s-t} \int_0^{t-s} e^y p(y)\, \mathrm{d}y \Big)^m f_S(s)\, \mathrm{d}s. \tag{8.2}$$

(We assume $m \geqslant 3$; for $m = 2$ we have $S = 0$ and (8.2) is replaced by (6.1).) We define

$$g(t) := 1 - e^{-t} \int_0^t e^y p(y)\, \mathrm{d}y \tag{8.3}$$

and write (8.2) as

$$p(t) = P(S > t) + \int_0^t g(t - s)^m f_S(s)\, \mathrm{d}s$$

$$= P(S > t) + \int_0^t g(s)^m f_S(t - s)\, \mathrm{d}s. \tag{8.4}$$

For simplicity, we consider in the sequel only the case $m = 3$. Then $S = Y_2 \sim \mathrm{Exp}(2)$, and (8.4) becomes

$$p(t) = e^{-2t} + 2e^{-2t} \int_0^t e^{2s} g(s)^3\, \mathrm{d}s. \tag{8.5}$$

It follows from (8.3) and (8.5) by induction that $g(t)$ and $p(t)$ are infinitely differentiable on $[0, \infty)$, and differentiation yields

$$g'(t) = 1 - g(t) - p(t), \tag{8.6}$$
$$p'(t) = -2p(t) + 2g(t)^3. \tag{8.7}$$

These equations can (as far as we know) only be solved numerically, and then ν can be computed numerically by (3.3), with the Malthusian parameter $\alpha = 1$ [8]. We obtain

$$\nu = \int_0^\infty e^{-t} p(t)\, \mathrm{d}t = 0.58705155\ldots \tag{8.8}$$

(Actually, we consider the system $\{(8.7),\ \nu'(t) = e^{-t} p(t)\}$ with $p(0) = 1$ and $\nu(0) = 0$, and use Maple to find $\nu = \nu(\infty)$.)

Remark 8.1. Extended m-ary search trees may be considered similarly as the case $m = 2$ in Section 7, see [8, Section 7.1]. We leave this case to the reader.

References

[1] David Aldous: Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1** (1991), no. 2, 228–266. MR-1102319

[2] Cyril Banderier, Markus Kuba & Alois Panholzer: Analysis of three graph parameters for random trees. *Random Structures Algorithms* **35** (2009), no. 1, 42–69. MR-2532873

[3] Albert-László Barabási & Réka Albert: Emergence of scaling in random networks. *Science* **286** (1999), no. 5439, 509–512. MR-2091634

[4] Kenneth Dadedzi: *Analysis of Tree Spectra*. Ph.D. dissertation, Stellenbosch University, 2018.

[5] Reinhard Diestel: *Graph Theory*. 5th ed., Springer-Verlag, Heidelberg, 2016. MR-2744811

[6] Michael Fuchs, Cecilia Holmgren, Dieter Mitsche & Ralph Neininger: A note on the independence number, domination number and related parameters of random binary search trees and random recursive trees. arXiv:1909.12767

[7] Remco van der Hofstad: *Random Graphs and Complex Networks. Vol. I*. Cambridge Univ. Press, Cambridge, 2017. MR-3617364

[8] Cecilia Holmgren & Svante Janson: Fringe trees, Crump–Mode–Jagers branching processes and m-ary search trees. *Probability Surveys* **14** (2017), 53–154. MR-3626585

[9] Peter Jagers & Olle Nerman: The growth and composition of branching populations. *Adv. in Appl. Probab.* **16** (1984), no. 2, 221–259. MR-0742953

[10] Günter Last & Mathew Penrose: *Lectures on the Poisson Process*. Cambridge Univ. Press, Cambridge, 2018. MR-3791470

[11] A. Meir & J.W. Moon: The expected node-independence number of random trees. *Indag. Math. (Proceedings)* **35** (1973), 335–341. MR-0345877

[12] A. Meir & J.W. Moon: The expected node-independence number of various types of trees. *Recent Advances in Graph Theory (Proc. Second Czechoslovak Sympos., Prague, 1974)*, pp. 351–363. Academia, Prague, 1975. MR-0401520

[13] A. Meir & J.W. Moon: Packing and covering constants for certain families of trees. I. *J. Graph Theory* **1** (1977), no. 2, 157–174. MR-0463001

[14] A. Meir & J.W. Moon: Packing and covering constants for certain families of trees. II. *Trans. Amer. Math. Soc.* **233** (1977), 167–178. MR-0460161

[15] A. Meir & J.W. Moon: Packing and covering constants for recursive trees. *Theory and Applications of Graphs (Proc. Internat. Conf., Western Mich. Univ., Kalamazoo, Mich., 1976)*, pp. 403–411. Lecture Notes in Math. **642**, Springer, Berlin, 1978. MR-0505717

[16] Olle Nerman & Peter Jagers: The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete* **65** (1984), no. 3, 445–460. MR-0731231

[17] Jerzy Szymański: On a nonuniform random recursive tree. *Annals of Discrete Math.* **33** (1987), 297–306. MR-0930497

[18] Stephan Wagner & Hua Wang: *Introduction to Chemical Graph Theory*. CRC Press, Boca Raton, FL, 2019. MR-3837106