

DEPTH-WEIGHTED TREES

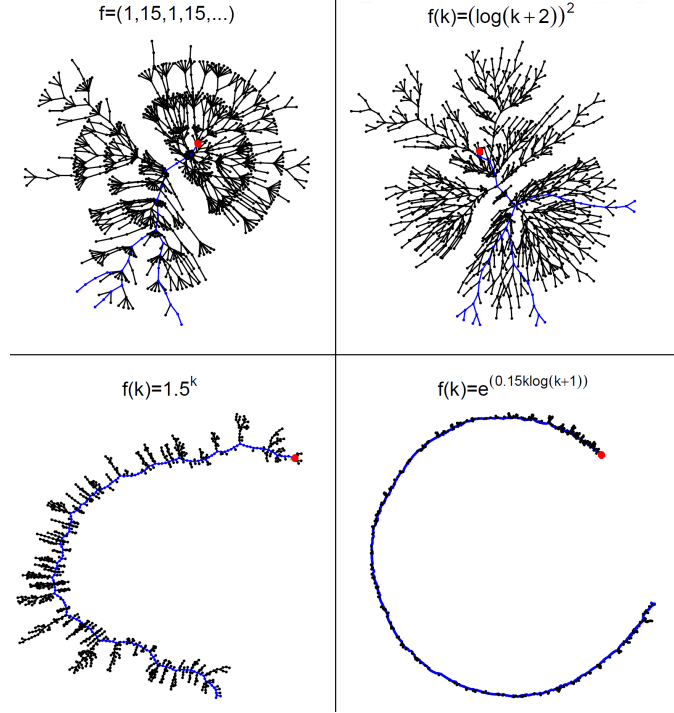


FIGURE 1. Simulations of $DWT(f)$ for different weight functions f . The big red dot denotes the root and the blue paths realise the depth of the tree, that is, the length of the longest path from the root to a leaf.

CONTENTS

1. Introduction	2
2. Comparing depths of depth-weighted trees	2
3. Continuous-time perspectives of the depth-weighted tree	3
3.1. First interpretation: via a continuous-time branching process	4
3.2. Second interpretation: via Cox processes	4
4. Bounded weight functions	5
5. Super-exponential weight functions	9
6. Open problems	14
References	15

1. INTRODUCTION

We study a model of random trees introduced by Leckey, Mitsche, and Wormald [7], called the *depth-weighted random recursive tree* (or *depth-weighted tree*, for short). In this model, vertices are added to the tree one-by-one and a new vertex attaches itself to an existing vertex chosen randomly, where the probability to select a vertex depends on the depth of the vertex in the tree (its graph distance to the root). More formally:

Definition 1.1. Let $f: \mathbb{N}_0 \rightarrow (0, \infty)$ and let $d_T(u)$ denote the depth of vertex u in the tree T . We construct a sequence of trees $(T_n(f))_{n \in \mathbb{N}}$ as follows. Initialise $T_1(f)$ as a single vertex (the root). For $n \in \mathbb{N}$, construct $T_{n+1}(f)$ from $T_n(f)$ by randomly selecting a vertex v_n from $T_n(f)$ and connecting a new vertex to v_n by an edge. Here, conditionally on $T_n(f)$, v_n is distributed as

$$\mathbb{P}(v_n = u | T_n(f)) = \frac{f(d_{T_n(f)}(u))}{\sum_{v \in T_n(f)} f(d_{T_n(f)}(v))} \quad \text{for } u \in T_n(f).$$

A natural question to ask, is how structural properties of the tree $T_n(f)$ depend on the weight function f . In particular, we will discuss the behaviour of $d(T_n(f))$, the *depth* of the tree, defined as

$$d(T_n(f)) := \max_{v \in T_n} d_{T_n(f)}(v).$$

2. COMPARING DEPTHS OF DEPTH-WEIGHTED TREES

When $f \equiv 1$ is the constant function, the depth-weighted tree is equivalent to the random recursive tree model, where at each step the vertex to connect to is selected uniformly at random from the tree constructed thus far. The following result provides a simple comparison of depth-weighted trees with increasing/decreasing weight functions to the random recursive tree, for which we know that $\mathbb{E}[d(T_n(1))] = e \log n - \frac{3}{2} \log \log n + \mathcal{O}(1)$ and $d(T_n(1)) = e \log n - \frac{3}{2} \log \log n + \mathcal{O}_{\mathbb{P}}(1)$.

Proposition 2.1 (Proposition 2.7 in [7]). *If $f: \mathbb{N}_0 \rightarrow (0, \infty)$ is an increasing (resp. decreasing) function, then $d(T_n(1)) \preceq d(T_n(f))$ (resp. $d(T_n(f)) \preceq d(T_n(1))$).*

Proof. We prove the claim when f is increasing, the decreasing case follows in a similar way. We prove a stronger result, namely that we can couple the construction of $T_1(f), \dots, T_n(f)$ and $T_1(1), \dots, T_n(1)$ in such a way that each vertex v in $T_n(f)$ is at least as deep as in $T_n(1)$. Let $N_j(\ell, f)$ and $N_j(\ell, 1)$ denote the number of vertices at depth ℓ in $T_j(f)$ and $T_j(1)$ for $\ell \in \mathbb{N}_0$ and $j \in \mathbb{N}$, respectively. We thus construct a coupling such that

$$\sum_{i=\ell}^{\infty} N_j(i, f) \geq \sum_{i=\ell}^{\infty} N_j(i, 1) \quad \text{for all } \ell \in \mathbb{N}_0 \text{ and } j \in \mathbb{N}. \quad (2.1)$$

From this, it directly follows that $d(T_n(f)) \geq d(T_n(1))$.

It is clear that $T_1(f) = T_1(1)$ and $T_2(f) = T_2(1)$, so that (2.1) holds for $j = 1, 2$ and all ℓ . Now, suppose that (2.1) is satisfied for some $j \in \mathbb{N}$ and all ℓ . We partition the interval $(0, 1)$ into sub-intervals in two ways. We set, for $k \in \mathbb{N}_0$,

$$I_{k,j}(f) := \left[\frac{\sum_{\ell=0}^{k-1} f(\ell) N_j(\ell, f)}{\sum_{i=0}^{\infty} f(i) N_j(i, f)}, \frac{\sum_{\ell=0}^k f(\ell) N_j(\ell, f)}{\sum_{i=0}^{\infty} f(i) N_j(i, f)} \right),$$

and

$$I_{k,j}(1) := \left[\frac{\sum_{\ell=0}^{k-1} N_j(\ell, 1)}{\sum_{i=0}^{\infty} N_j(i, 1)}, \frac{\sum_{\ell=0}^k N_j(\ell, 1)}{\sum_{i=0}^{\infty} N_j(i, 1)} \right).$$

Sample a uniform $(0, 1)$ random variable U_j independent of everything else. To construct $T_{j+1}(f)$ and $T_{j+1}(1)$, attach vertex $j+1$ to a vertex at depth $k \in \mathbb{N}_0$ in $T_j(f)$ if $U_j \in I_{k,j}(f)$ and to

a vertex at depth $k \in \mathbb{N}_0$ in $T_j(1)$ if $U_j \in I_{k,j}(f)$. We claim that (2.1) combined with f being increasing yields that

$$\bigcup_{\ell=0}^k I_{\ell,j}(f) \subseteq \bigcup_{\ell=0}^k I_{\ell,j}(1) \quad \text{for any } k \in \mathbb{N}_0. \quad (2.2)$$

This directly implies $U_j \in I_{k_1,j}(f)$ and $U_j \in I_{k_2,j}(f)$ for some $k_1 \geq k_2$, so that vertex $j+1$ is at least as deep in $T_{j+1}(f)$ as it is in $T_{j+1}(1)$, so that (2.1) holds for $j+1$ and all $\ell \in \mathbb{N}_0$ as well. To prove the claim, we write

$$\bigcup_{\ell=0}^k I_{\ell,j}(f) = \left[0, \frac{\sum_{\ell=0}^k f(\ell)N_j(\ell, f)}{\sum_{\ell=0}^k f(\ell)N_j(\ell, f) + \sum_{\ell=k+1}^{\infty} f(\ell)N_j(\ell, f)} \right),$$

and notice that, since $f(\ell) \geq f(k+1)$ for all $\ell \geq k+1$ and by (2.1),

$$\begin{aligned} \frac{\sum_{\ell=0}^k f(\ell)N_j(\ell, f)}{\sum_{\ell=0}^k f(\ell)N_j(\ell, f) + \sum_{\ell=k+1}^{\infty} f(\ell)N_j(\ell, f)} &\leq \frac{\sum_{\ell=0}^k f(\ell)N_j(\ell, f)}{\sum_{\ell=0}^k f(\ell)N_j(\ell, f) + f(k+1) \sum_{\ell=k+1}^{\infty} N_j(\ell, f)} \\ &\leq \frac{\sum_{\ell=0}^k (f(\ell)/f(k+1))N_j(\ell, f)}{\sum_{\ell=0}^k (f(\ell)/f(k+1))N_j(\ell, f) + \sum_{\ell=k+1}^{\infty} N_j(\ell, 1)} \\ &\leq \frac{\sum_{\ell=0}^k N_j(\ell, f)}{\sum_{\ell=0}^k N_j(\ell, f) + \sum_{\ell=k+1}^{\infty} N_j(\ell, 1)} \\ &\leq \frac{\sum_{\ell=0}^k N_j(\ell, 1)}{\sum_{\ell=0}^k N_j(\ell, 1) + \sum_{\ell=k+1}^{\infty} N_j(\ell, 1)}. \end{aligned}$$

so that we arrive at (2.2). $\square$

Proposition 2.1 allows us to compare depth-weighted trees with increasing/decreasing weight functions to the random recursive tree (i.e. with a constant weight function). A natural question is whether there exist conditions under which we can compare other depth-weighted trees. Intuitively, if a weight function f increases/decreases faster than another weight function g , then the depth of $T_n(f)$ should stochastically dominate that of $T_n(g)$. This is reflected in the following conjecture.

Conjecture 2.2 ([7]). *Let $f, g: \mathbb{N}_0 \rightarrow (0, \infty)$ be two weight functions such that*

$$\frac{f(k+1)}{f(k)} \geq \frac{g(k+1)}{g(k)} \quad \text{for all } k \in \mathbb{N}_0.$$

Then, $d(T_n(f)) \succeq d(T_n(g))$ for all $n \in \mathbb{N}$.

Remark 2.3. (i) Note that the condition $f(k) \geq g(k)$ for all $k \in \mathbb{N}_0$ is *not* sufficient for $d(T_n(f)) \succeq d(T_n(g))$ to hold. Indeed, set $f(k) = g(k)$ for all $k \in \mathbb{N}$, but $f(0) > g(0)$. As the weight of the root in $T_n(f)$ is slightly larger compared to $T_n(g)$, should be $d(T_n(f))$ stochastically dominated by $d(T_n(g))$.

(ii) Proving Conjecture 2.2 with the same approach as in the proof of Proposition 2.1, i.e. by coupling $T_n(f)$ and $T_n(g)$ in such a manner that each vertex in $T_n(f)$ is at least as deep as in $T_n(g)$ does not work, as one can find counterexamples. $\blacktriangleleft$

3. CONTINUOUS-TIME PERSPECTIVES OF THE DEPTH-WEIGHTED TREE

In this section we describe two continuous-time interpretations of the DWT model. These are equivalent to the discrete-time description, but come with certain analytical advantages.

3.1. First interpretation: via a continuous-time branching process. This continuous-time version of the model appears in the original paper of Leckey, Mitsche and Wormald [7]. It has the advantage that the discrete model can be completely recovered from it (which is not the case with the second interpretation presented in Section 3.2).

A central object in this construction is the *Ulam-Harris tree* with vertex set

$$\text{UH} := \{\emptyset\} \cup \left(\bigcup_{n \in \mathbb{N}} \mathbb{N}^n \right).$$

This tree grows away from a root $\emptyset$ and the children of each vertex $v \in \text{UH}$ are $v1, v2, \dots$. Moreover, we denote by $d_{\text{UH}}: \text{UH} \rightarrow \mathbb{N}_0$ the depth function for the Ulam-Harris tree; namely, $d_{\text{UH}}(\emptyset) = 0$ and, for every vertex $v = v_1 v_2 \dots v_k$ with $k, v_1, \dots, v_k \in \mathbb{N}$, $d_{\text{UH}}(v) = k$. We also define a *ray* to be an infinite sequence of positive integers and denote by $\mathbb{N}^\infty$ the set of rays.

To every $v \in \text{UH}$, we then assign an independent Poisson process ξ^v of rate $f(d_{\text{UH}}(v))$ where f is the weight function associated with the depth-weighted tree. The points $\sigma_1^v, \sigma_2^v, \dots$ in the Poisson process ξ^v correspond to the difference between the birth-times of the children $v1, v2, \dots$ of v and the birth-time of v itself. We can thus recursively construct the *absolute* birth times $(\mathcal{B}(v))_{v \in \text{UH}}$ as

$$\mathcal{B}(\emptyset) := 0 \quad \text{and} \quad \mathcal{B}(vj) := \mathcal{B}(v) + \sigma_j^v \quad \text{for every } v \in \text{UH} \text{ and } j \in \mathbb{N}.$$

Finally, we construct the continuous-time branching process $(\mathcal{T}(t))_{t \geq 0}$ by defining

$$\mathcal{T}(t) := \{v \in \text{UH} : \mathcal{B}(v) \leq t\} \quad \text{for all } t \geq 0.$$

In particular, $d(\mathcal{T}(t)) := \max_{v \in \mathcal{T}(t)} d_{\text{UH}}(v)$ denotes the depth of the branching process $\mathcal{T}$ at time t . To justify the correspondence between the continuous-time branching process $\mathcal{T}$ and the depth-weighted tree process, we define the stopping times $(\tau_n)_{n \in \mathbb{N}}$ as

$$\tau_n := \inf\{t \geq 0 : |\mathcal{T}(t)| = n\} \quad \text{for } n \in \mathbb{N}.$$

We also introduce $N_r(t) := \{v \in \mathcal{T}(t) : d_{\text{UH}}(v) = r\}$ as the number of individuals at depth r at time t for $r \in \mathbb{N}_0$ and $t \geq 0$.

$$\tau_{j,r} := \inf\{t \geq 0 : |\{v \in \mathcal{T}(t) : d_{\text{UH}}(v) = r\}| = j\} \quad \text{for } j, r \in \mathbb{N}_0.$$

We observe (as follows from [7, Lemma 3.1]) that, by the memoryless property of the exponential distribution and the strong Markov property,

$$\{T_n : n \in \mathbb{N}\} \stackrel{d}{=} \{\mathcal{T}(\tau_n) : n \in \mathbb{N}\}.$$

3.2. Second interpretation: via Cox processes. Sometimes the exact structure of the tree T_n is irrelevant and only the number of vertices at different depths matters. The continuous-time construction from 3.1 carries more information than needed in such a case, which can complicate the analysis. To this end, we provide an interpretation using Cox processes.

Fix a measure space $(S, \mathcal{F}, \Lambda)$. A *Poisson point process* N with intensity measure Λ is a family of random variables $\{N(A) : A \in \mathcal{F}\}$ with values in $\mathbb{N}_0$ such that the following properties hold:

- for each $A \in \mathcal{F}$, $N(A)$ is a Poisson random variable with parameter $\Lambda(A)$, and
- for each family (A_i) of pairwise disjoint measurable sets, $(N(A_i))$ are jointly independent random variables.

Furthermore, if the intensity measure Λ varies according to a stochastic process on the set of measures on $(S, \mathcal{F})$, N is instead called a *Cox process*.

We now introduce the following continuous-time interpretation of the DWT model, for which $S = \mathbb{R}_+$.

- First, we define a Poisson point process $N_1(t)$ with intensity $\Lambda_1(dt) = f(0)dt$ (counting the vertices on depth 1).
- Recursively, for every $r \geq 2$, we define a Cox process $N_r(t)$ with intensity $\Lambda_r(dt) = f(r-1)N_{r-1}(t)dt$ (counting the vertices on depth r).

For every $r \in \mathbb{N}$, we call Λ_r the *directing measure* of N_r and additionally denote $N_0(t) = 1$ for all $t \geq 0$ to account for the root at depth zero. Further, for every $t \geq 0$, we define

$$N(t) := \sum_{r=0}^{\infty} N_r(t) = 1 + \sum_{r=1}^{\infty} N_r(t),$$

interpreted as the total number of individuals at time t . In relation to the branching process $\mathcal{T}(t)$ defined in Section 3.1, the $N_r(t)$ denote the generation sizes and we have $N(t) = |\mathcal{T}(t)|$, as well as

$$\tau_n = \inf\{t \geq 0 : N(t) = n\} \quad \text{and} \quad \tau_{j,r} = \inf\{t \geq 0 : N_r(t) = j\}.$$

The following result formalises the connection with the original model.

Lemma 3.1. *For every weight function f , the distribution of the sequences $(N_r(\tau_n))_{r \in \mathbb{N}}$ and $(|\{v \in T_n : d_{T_n(f)}(v) = r\}|)_{r \in \mathbb{N}}$ coincide.*

Proof. We proceed by recursively constructing a coupling of the two sequences that identifies them almost surely. The coupling of the roots is immediate. Suppose that, for some $n \geq 2$, a coupling with the above description has been constructed so that $N_r(\tau_{n-1}) = |\{v \in T_{n-1} : d(v) = r\}|$ for every $r \in \mathbb{N}$, and define $s_r = s_r(n) := N_r(\tau_{n-1}) + 1$. Then, for every $r \in \mathbb{N}$, $(\tau_{s_r, r} - \tau_{n-1})_{r \in \mathbb{N}}$ is a family of independent exponential random variables with rates $f(r-1)N_{r-1}(\tau_{n-1})$, respectively. Thus, since $N_r(\tau_n) = s_r$ holds only for the index r of the minimum among $(\tau_{s_r, r} - \tau_{n-1})_{r \in \mathbb{N}}$,

$$\mathbb{P}(N_r(\tau_n) = s_r \mid N_r(\tau_{n-1})_{r \in \mathbb{N}_0}) = \frac{f(r-1)N_{r-1}(\tau_{n-1})}{\sum_{i=0}^{\infty} f(i)N_i(\tau_{n-1})}.$$

However, this is precisely the probability that the n -th vertex attaches to a vertex on depth $r-1$ in the discrete model, which concludes the coupling until time τ_n and the proof. $\square$

4. BOUNDED WEIGHT FUNCTIONS

Suppose that the weight function f is such that $0 < \inf_{k \in \mathbb{N}_0} f(k) \leq \sup_{k \in \mathbb{N}_0} f(k) < \infty$. One can think of the depth-weighted tree $T_n(f)$ as a perturbation of the random recursive tree $T_n(1)$. We show that its depth grows at a logarithmic order.

Theorem 4.1 (Bounded weight functions, [8]). *For a weight function f such that $0 < \inf_{k \in \mathbb{N}_0} f(k) \leq \sup_{k \in \mathbb{N}_0} f(k) < \infty$, there exist constants $c = c(f), C = C(f) > 0$, such that almost surely,*

$$c \leq \liminf_{n \rightarrow \infty} \frac{d(T_n(f))}{\log n} \leq \limsup_{n \rightarrow \infty} \frac{d(T_n(f))}{\log n} \leq C.$$

We use the continuous-time branching process interpretation in Section 3.1 to prove Theorem 4.1 and start with the observation that $d(T_n(f)) \stackrel{d}{=} d(\mathcal{T}(\tau_n))$ and

$$d(\mathcal{T}(\tau_n)) = \sup\{k \in \mathbb{N}_0 : \tau_{1,k} \leq \tau_n\}. \quad (4.1)$$

Our aim is thus to find lower and upper bounds for $\tau_{1,k}$ and τ_n when k and n are large. To this end, we consider the Yule process $(\mathcal{Y}_\lambda(t))_{t \geq 0}$ of rate $\lambda > 0$, which is a continuous-time branching process where, for each $v \in \text{UH}$, the Poisson process ξ^v has rate λ . With

$$I = I(f) := \inf_{k \in \mathbb{N}_0} f(k) \quad \text{and} \quad S = S(f) := \sup_{k \in \mathbb{N}_0} f(k),$$

we can couple our branching process $\mathcal{T}$ to two Yule processes $\mathcal{Y}_I$ and $\mathcal{Y}_S$, such that

$$\mathcal{Y}_I(t) \subseteq \mathcal{T}(t) \subseteq \mathcal{Y}_S(t) \quad \text{for all } t \geq 0.$$

If we let $\tau_n^\lambda, \tau_{1,k}^\lambda$ be the corresponding stopping times for the branching process $\mathcal{Y}_\lambda$, the coupling thus yields that

$$\tau_{1,k}^S \leq \tau_{1,k} \leq \tau_{1,k}^I \quad \text{and} \quad \tau_n^S \leq \tau_n \leq \tau_n^I \quad \text{for all } k, n \in \mathbb{N}_0. \quad (4.2)$$

We first analyse the stopping times τ_n^λ for $\lambda > 0$.

Proposition 4.2 (Growth rate of Yule processes, [9]). *Fix $\lambda > 0$. Then, $M_t := e^{-\lambda t} |\mathcal{Y}_\lambda(t)|$ is a martingale with respect to the filtration $\mathcal{F}_t := \sigma(\mathcal{Y}_\lambda(s); s \leq t)$. As a result, $\tau_n^\lambda - \frac{1}{\lambda} \log n$ converges almost surely.*

The almost sure convergence of the size of the branching process $\mathcal{Y}_\lambda(t)$ when rescaled, is true for a much larger class of branching processes (including Bienaymé–Galton–Watson branching processes) known as Crump–Mode–Jagers branching processes. See e.g. the early papers of Jagers and Nerman [9, 3, 4] or the classical book on branching processes by Athreya and Ney [1].

Proof. We claim that $\mathbb{E}[\mathcal{Y}_\lambda(t)] = e^{\lambda t}$, so that $\mathbb{E}[M_t] = 1$ for all $t \geq 0$. Fix $0 \leq s \leq t$. By the branching property, conditionally on $\mathcal{F}_s$ we can think of $|\mathcal{Y}(t)|$ as a sum of $|\mathcal{Y}(s)|$ many independent copies $|\mathcal{Y}_\lambda^{(i)}(t-s)|$ of $|\mathcal{Y}_\lambda(t-s)|$ that are independent of $\mathcal{F}_s$. As a result,

$$\mathbb{E}[M_t | \mathcal{F}_s] = e^{-\lambda t} \mathbb{E} \left[\sum_{i=1}^{|\mathcal{Y}_\lambda(s)|} |\mathcal{Y}_\lambda^{(i)}(t-s)| \middle| \mathcal{F}_s \right] = e^{-\lambda t} |\mathcal{Y}_\lambda(s)| \mathbb{E}[|\mathcal{Y}_\lambda(t-s)|] = e^{-\lambda s} |\mathcal{Y}_\lambda(s)|,$$

as desired, where the last step follows from the claim. As M_t is a non-negative martingale, it follows that there exists a random variable W such that $M_t \xrightarrow{\text{a.s.}} W$. By the definition of τ_n^λ , we have $|\mathcal{Y}(\tau_n^\lambda)| = n$. Moreover, since $\mathbb{E}[Y_t] < \infty$ and thus $Y_t < \infty$ almost surely for any $t \geq 0$, we conclude that $\tau_n^\lambda \rightarrow \infty$ almost surely with n . Hence,

$$e^{-\lambda \tau_n^\lambda} |\mathcal{Y}(\tau_n^\lambda)| = e^{\log n - \lambda \tau_n^\lambda}$$

converges almost surely, so that the second result follows.

To prove the claim, we consider the time at which the root produces its first child, so that we can write

$$\mathbb{E}[\mathcal{Y}_\lambda(t)] = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda s} 2 \mathbb{E}[\mathcal{Y}_\lambda(t-s)] ds = e^{-\lambda t} + \int_0^t \lambda e^{-\lambda(t-s)} 2 \mathbb{E}[\mathcal{Y}_\lambda(s)] ds.$$

If we set $g(t) := e^{\lambda t} \mathbb{E}[\mathcal{Y}_\lambda(t)]$, we obtain the functional equation

$$g(t) = 1 + 2\lambda \int_0^t g(s) ds.$$

It is readily checked that $g(t) = e^{2\lambda t}$ solves the equation, from which $\mathbb{E}[\mathcal{Y}(t)] = e^{\lambda t}$ follows. $\square$

We then consider $\tau_{1,k}^\lambda$.

Theorem 4.3 (First-birth time of Yule processes, [2]). *Fix $\lambda > 0$. Then,*

$$\frac{\tau_{1,k}^\lambda}{k} \xrightarrow{\text{a.s.}} \frac{1}{\lambda e}.$$

This theorem is an interesting result in its own right and, again, holds true for a more general class of CMJ branching processes (with a different constant as the limit, see the work of Biggins [2]). We shall prove a lower bound only, and briefly sketch the argument for a matching upper bound.

To prove a lower bound for $\tau_{1,k}^\lambda$, let $N_k^\lambda(t)$ denote the number of individuals in $\mathcal{Y}_\lambda(t)$ at depth k . We notice that

$$\{\tau_{1,k}^\lambda < t\} = \{N_k^\lambda(t) \geq 1\}.$$

Deriving the mean of $N_k^\lambda(t)$ thus allows us to bound $\tau_{1,k}$ from below, as follows from the next lemma.

Lemma 4.4. *For every $k \in \mathbb{N}_0$ and $t \geq 0$,*

$$\mathbb{E}[N_k(t)] = \frac{t^k}{k!} \prod_{i=0}^{k-1} f(i). \quad (4.3)$$

In particular, when $f(i) = \lambda > 0$ for all $i \in \mathbb{N}_0$,

$$\mathbb{E}[N_k^\lambda(t)] = \frac{(\lambda t)^k}{k!}. \quad (4.4)$$

Proof. We prove the result by induction over k . The result holds for $k = 0$ as $N_0(t) = 1$ for all $t \geq 0$. Assume (4.4) for some $k \in \mathbb{N}_0$. Then,

$$\mathbb{E}[N_{k+1}(t)] = \mathbb{E}[\mathbb{E}[N_{k+1}(t) \mid (N_k(s))_{s \in [0, t]}]] = \mathbb{E}\left[\int_0^t f(k)N_k(s) \, ds\right].$$

After switching expectation and integration by Fubini's theorem, the induction hypothesis gives

$$\mathbb{E}[N_{k+1}(t)] = \int_0^t f(k)\mathbb{E}[N_k(s)] \, ds = \frac{1}{k!} \prod_{i=0}^k f(i) \int_0^t s^k \, ds = \frac{t^{k+1}}{(k+1)!} \prod_{i=0}^k f(i),$$

as desired. $\square$

With Lemma 4.4, we directly obtain that for any $x < 1/(\lambda e)$,

$$\mathbb{P}(\tau_{1,k}^\lambda < xk) = \mathbb{P}(N_k^\lambda(xk) \geq 1) \leq \mathbb{E}[N_k^\lambda(xk)] = \frac{(\lambda xk)^k}{k!} = \frac{1 + o(1)}{\sqrt{2\pi k}} (\lambda e x)^k,$$

where the final step uses Sterling's approximation $k! = (1 + o(1))\sqrt{2\pi k}(k/e)^k$. Since $\lambda e x < 1$, the Borel-Cantelli lemma implies that almost surely,

$$\liminf_{k \rightarrow \infty} \frac{\tau_{1,k}^\lambda}{k} \geq \frac{1}{\lambda e}.$$

Sketch of the upper bound proof. To obtain a matching upper bound, we cannot directly use a second moment method, since $\mathbb{E}[N_k^\lambda(t)]^2 / \mathbb{E}[N_k^\lambda(t)^2]$ is bounded away from 0 and 1 for all large enough k when $t = ck$ for some constant $c > 1/(\lambda e)$. However, by the Paley-Zygmund inequality,

$$\mathbb{P}(\tau_{1,k}^\lambda \leq t) = \mathbb{P}(N_k^\lambda(t) \geq 1) \geq \frac{\mathbb{E}[N_k^\lambda(t)]^2}{\mathbb{E}[N_k^\lambda(t)^2]} > 0. \quad (4.5)$$

Now, we take $\varepsilon > 0$ small such that $(1 - \varepsilon)c > 1/(\lambda e)$. Then, let

$$\mathcal{E}_t := \{N_1(\varepsilon t) \geq \frac{\lambda \varepsilon}{2} t\}$$

denote the event that the root has produced at least $\frac{\lambda \varepsilon}{2} t$ many children by time εt . Since $N_1(\varepsilon t) \sim \text{Poi}(\lambda \varepsilon t)$, this event occurs with probability at least $1 - e^{-a_1 t}$ for some small constant $a_1 > 0$, so that

$$\mathbb{P}(\tau_{1,k}^\lambda > t) \leq \mathbb{P}(\{\tau_{1,k}^\lambda > t\} \cap \mathcal{E}_t) + \mathbb{P}(\mathcal{E}_t^c) \leq \mathbb{P}(\{\tau_{1,k}^\lambda > t\} \cap \mathcal{E}_t) + e^{-at}.$$

Each of the children born by time εt grows an independent Yule process up to time t , so each subtree grows at least $(1 - \varepsilon)t$ time. The event $\mathcal{E}_t \cap \{\tau_{1,k}^\lambda > t\}$ then implies that each of the subtrees rooted at the (at least) $\frac{\lambda \varepsilon}{2} t$ many children of the root does not reach depth $k - 1$ by time $(1 - \varepsilon)t$. So,

$$\mathbb{P}(\{\tau_{1,k}^\lambda > t\} \cap \mathcal{E}_t) \leq \mathbb{P}(\tau_{1,k-1}^\lambda > (1 - \varepsilon)t)^{\lambda \varepsilon t / 2} \leq \mathbb{P}(\tau_{1,k-1}^\lambda > (1 - \varepsilon)c(k - 1))^{\lambda \varepsilon t / 2}$$

By (4.5), using that $(1 - \varepsilon)t = (1 - \varepsilon)c \frac{k}{k-1}(k - 1)$ and $(1 - \varepsilon)c \frac{k}{k-1} > (1 - \varepsilon)c > 1/(\lambda e)$, we obtain an upper bound $e^{-a_2 t}$ for some constant $a_2 > 0$. Hence,

$$\mathbb{P}(\tau_{1,k}^\lambda > ck) \leq e^{-a_1 t} + e^{-a_2 t} = e^{-a_1 ck} + e^{-a_2 ck},$$

so that the Borel-Cantelli lemma yields the desired result.

To show that the second moment of $N_k^\lambda(t)$ and the first moment squared are of the same order, we determine the variance. We have

$$\text{Var}(N_k^\lambda(t)) = \mathbb{E}[N_k^\lambda(t)] \sum_{\ell=0}^{k-1} \frac{(\lambda t)^\ell k!}{(k + \ell)!} \binom{2\ell}{\ell} = \sum_{\ell=0}^{k-1} \frac{(\lambda t)^{k+\ell}}{(k + \ell)!} \binom{2\ell}{\ell}. \quad (4.6)$$

To arrive at this equality, we use that, for each $j \in \mathbb{N}$, the point process N_j is a Cox process with directing random measure Λ_j , with density

$$\Lambda_j(B) := \int_B f(j-1)N_{j-1}(s) ds, \quad B \subseteq [0, \infty).$$

The measure Λ_j has intensity measure ν_{j-1} with density

$$\nu_{j-1}(ds) := \frac{s^{j-1}}{(j-1)!} \prod_{i=0}^{j-1} f(i) ds.$$

We note that N_j and Λ_j have the *same* intensity measure. Indeed, as the law of N_j , conditionally on Λ_j , is equal to that of a Poisson point process with intensity measure Λ_j , it follows that $\mathbb{E}[N_j(B) | \Lambda_j] = \Lambda_j(B)$, and thus $\mathbb{E}[N_j(B)] = \mathbb{E}[\Lambda_j(B)] = \nu_{j-1}(B)$ for any $B \subseteq [0, \infty)$ and $j \in \mathbb{N}$. Continuing this recursion, we can for each $j \in \mathbb{N}$ define the measures $\xi_\ell := N_\ell \prod_{i=\ell}^{j-1} f(i)$ for $\ell \in \{0, \dots, j\}$. Then, ξ_ℓ is a Cox process with directing random measure $\xi_{\ell-1}$ for each $\ell \in [j]$, and ξ_ℓ has intensity measure ν_ℓ , given by

$$\nu_\ell(ds) := \frac{s^\ell}{\ell!} \prod_{i=0}^{j-1} f(i), \quad \ell \in \{0, \dots, j\}.$$

We can then apply [6, Proposition 13.6] to conclude that, for any $t > 0$ and with $g : [0, \infty) \rightarrow [0, \infty)$, $g(x) := \mathbb{1}_{\{x \in [0, t]\}}$,

$$\text{Var}(N_j(t)) = \text{Var}(N_j(g)) = \nu_{j-1}(g^2) + \text{Var}(\Lambda_j(g)) = \mathbb{E}[N_j(t)] + \text{Var}\left(\int_0^t f(j-1)N_{j-1}(s) ds\right).$$

We observe that

$$\int_0^t f(j-1)N_{j-1}(s) ds = \int_0^t f(j-1)(t-s)N_{j-1}(ds) = f(j-1)N_{j-1}(h_1),$$

where we define the functions $h_i : \mathbb{R} \rightarrow \mathbb{R}$ by $h_i(s) := \mathbb{1}_{\{s \in [0, t]\}}(t-s)^i/i!$ for any $i \in \mathbb{N}$. As such, we obtain

$$\begin{aligned} \text{Var}(f(j-1)N_{j-1}(h_1)) &= f(j-1)^2 \nu_{j-2}(h_1^2) + f(j-1)^2 \text{Var}\left(\int_0^t f(j-2)(t-s)N_{j-2}(s) ds\right) \\ &= f(j-1)^2 \int_0^t (t-s)^2 \frac{s^{j-2}}{(j-2)!} \prod_{i=0}^{j-2} f(i) ds \\ &\quad + f(j-1)^2 \text{Var}\left(\int_0^t f(j-2) \frac{(t-s)^2}{2} N_{j-2}(ds)\right) \\ &= \frac{2f(j-1)t}{j+1} \mathbb{E}[N_j(t)] + (f(j-1)f(j-2))^2 \text{Var}(N_{j-2}(h_2)). \end{aligned}$$

Continuing this recursion, we arrive at

$$\text{Var}(N_j(t)) = \mathbb{E}[N_j(t)] \sum_{i=0}^{j-1} t^i \frac{(2i)!}{(i!)^2} \prod_{\ell=1}^i \frac{f(j-\ell)}{j+\ell}.$$

For the case $f(i) = \lambda$ for all $i \in \mathbb{N}_0$, we thus obtain (4.6).

We complete this section by proving Theorem 4.1.

Proof of Theorem 4.1. By (4.1) and (4.2) we have for any $x > 0$ and $k_n := \lceil x \log n \rceil$,

$$\{d(\mathcal{T}(\tau_n)) \geq k_n\} = \{\sup\{k \in \mathbb{N}_0 : \tau_{1,k} \leq \tau_n\} \geq k_n\} \subseteq \{\sup\{k \in \mathbb{N}_0 : \tau_{1,k}^S \leq \tau_n^I\} \geq k_n\} \subseteq \{\tau_{1,k_n}^S \leq \tau_n^I\}.$$

Since, by Proposition 4.2 and Theorem 4.3,

$$\frac{\tau_{1,k_n}^S}{k_n} \xrightarrow{\text{a.s.}} \frac{1}{Se} \quad \text{and} \quad \frac{\tau_n^I}{k_n} = \left(\frac{1}{x} + o(1)\right) \frac{\tau_n^I}{\log n} \xrightarrow{\text{a.s.}} \frac{1}{xI},$$

it follows that the event $\{d(\mathcal{T}(\tau_n)) \geq k_n\}$ occurs only finitely often almost surely when $x > I/(Se)$. In a similar manner, we can prove a lower bound for $d(\mathcal{T}(\tau_n))$. $\square$

Convergence of the depth.

Under additional assumptions on f , beyond that it is bounded, it is possible to prove more precise results on the growth rate on the depth. We will not prove these here, but just state them for the interested reader.

Theorem 4.5 ([5]). *Let $f: \mathbb{N}_0 \rightarrow (0, \infty)$ be a weight function such that $0 < \inf_{k \in \mathbb{N}_0} f(k) \leq \sup_{k \in \mathbb{N}_0} f(k) < \infty$ and*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \log(f(i)) = \delta, \quad (4.7)$$

for some $\delta \in \mathbb{R}$ (without loss of generality, we may set $\delta = 0$). Then,

$$\frac{d(T_n(f))}{\log n} \xrightarrow{\text{a.s.}} e.$$

Remark 4.6. The condition in (4.7) says something about the arithmetic mean of f , and is used to control the expected generation sizes. Indeed, under this condition, Lemma 4.4 yields that $\mathbb{E}[N_k(t)] \approx (e^\delta t)^k / k!$ for k large, so that $\mathbb{E}[N(t)] \approx \exp(e^\delta t)$.

Natural examples of weight functions f that satisfy (4.7) are converging and periodic f . $\blacktriangleleft$

It is also possible to construct bounded weight functions that do not yield trees whose depth is asymptotically $e \log n$.

Proposition 4.7 ([5]). *There exists a weight function $f: \mathbb{N}_0 \rightarrow \{1, 3\}$, such that, almost surely,*

$$\liminf_{n \rightarrow \infty} \frac{d(T_n(f))}{\log n} \leq \frac{5}{2}, \quad \limsup_{n \rightarrow \infty} \frac{d(T_n(f))}{\log n} \geq e.$$

5. SUPER-EXPONENTIAL WEIGHT FUNCTIONS

In this section we study weight functions f that grow super-exponentially. That is, assume that

$$\lim_{k \rightarrow \infty} \frac{f(k+1)}{f(k)} = \infty, \quad (5.1)$$

and define

$$I_n := \sum_{k=0}^n \frac{f(k)}{f(k+1)}.$$

We then have the following result.

Theorem 5.1. *Suppose the weight function f satisfies (5.1).*

- (a) *If I_n converges, there exists a random variable K such that $n - d(T_n(f)) \xrightarrow{\text{a.s.}} K$.*
- (b) *If I_n diverges, then*

$$\frac{n - d(T_n(f))}{I_n} \xrightarrow{\mathbb{P}} 1.$$

We prove the case when I_n converges in these notes.

As $\sum_{i=0}^{\infty} f(i)^{-1}$ is finite, the branching process $\mathcal{T}$ explodes (grows infinitely large in finite time). This is relatively easy to see, by letting T denote the time until the infinite ray $(\emptyset, 1, 11, 111, 1111, \dots)$ is constructed. As each individual on the ray (apart from the root) is the first child of the previous individual on the ray, it follows that

$$T = \sum_{i=0}^{\infty} E_i,$$

where $E_i \sim \text{Exp}(f(0))$. As a result, $\mathbb{E}[T] = \sum_{i=0}^{\infty} f(i)^{-1}$, which is finite. It thus follows that the branching process is infinitely large at the almost surely finite time T . As a result, the random variable

$$\tau_{\infty} := \lim_{n \rightarrow \infty} \tau_n$$

is finite almost surely (since $\tau_{\infty} \leq T$ almost surely). Moreover, we also know that there exists a unique infinite ray $\gamma \in \mathbb{N}^{\infty}$ in $\mathcal{T}(\tau_{\infty})$ almost surely, see Proposition 5.6.

Intuitively, the theorem follows from the fact that the probability that a vertex at depth $j \in \mathbb{N}_0$ produces more than one child before the process explodes is of the order $f(j)/f(j+1)$. The condition that I_n converges, combined with the Borel-Cantelli lemma, implies that only finitely many vertices along the infinite ray γ produce more than one child. To make this intuition more precise, we introduce the following auxiliary continuous-time branching process (CTBP). To aid notation, we let $p(v)$ denote the parent of $v \in \text{UH} \setminus \{\emptyset\}$.

Definition 5.2 (CTBP with dormant non-first-born individuals). Let $(\tilde{\mathcal{T}}_{\emptyset}(t))_{t \geq 0}$ be a CTBP rooted at $\emptyset$, where an individual $v \in \text{UH}$ gives birth at rate $f(d_{\text{UH}}(v))$ when $v = \emptyset$ or $v = p(v)1$ (i.e. v is the first child of $p(v)$), and does not give birth to children, otherwise. We call vertices that do not give birth to children *dormant*, and vertices that do give birth to children *active*.

In general, for $u \in \text{UH}$, we let $(\tilde{\mathcal{T}}_u(t))_{t \geq 0}$ be the CTBP rooted at u . That is, vertices $v \in \{uw : w \in \text{UH}\}$ give birth at rate $f(d_{\text{UH}}(v))$ when $v = u$ or $v = p(v)1$ and are dormant otherwise. $\blacktriangleleft$

Whilst the processes $\tilde{\mathcal{T}}_u$ for $u \in \text{UH}$ are different from the CTBP $\mathcal{T}$ that describes the depth-weighted tree evolution, we use these processes to describe the infinite branching process $\mathcal{T}(\tau_{\infty})$ at the time the process has exploded.

It is clear that the process $\tilde{\mathcal{T}}_{\emptyset}$ produces a single infinite ray $\gamma = \gamma_{\emptyset} := (\emptyset, 1, 11, 111, \dots)$, as all other individuals that are not on this infinite ray are dormant and thus do not produce children. Moreover, the ray γ is produced in finite time almost surely, say in time $T(\gamma)$. Similarly, $\tilde{\mathcal{T}}_u$ produces a single infinite ray $\gamma_u := (u, u1, u11, \dots)$ in finite time $T(\gamma_u)$, almost surely. The following proposition determines the number of dormant children in $\tilde{\mathcal{T}}_u$ at time $T(\gamma_u)$.

Proposition 5.3 (Expected number of dormant children). *Let the weight function f satisfy (5.1). Fix $u \in \text{UH}$ and consider the CTBP $\tilde{\mathcal{T}}_u$. Fix $u \in \text{UH}$ and let $(N_w)_{w \in \gamma_u}$ denote the number of dormant children w produces before time $T(\gamma_u)$. If I_n converges,*

$$\mathbb{E} \left[\sum_{w \in \gamma_u} N_w \right] = (1 + o(1)) \sum_{j=d_{\text{UH}}(u)}^{\infty} \frac{f(j)}{f(j+1)} = (1 + o(1)) \frac{f(d_{\text{UH}}(u))}{f(d_{\text{UH}}(u) + 1)}, \quad (5.2)$$

where the $o(1)$ term tends to zero as $d_{\text{UH}}(u)$ tends to infinity. If I_n diverges,

$$\mathbb{E} \left[\sum_{w \in \gamma_u} N_w \right] = \infty. \quad (5.3)$$

Proof. The birth process ξ^w associated with $w \in \gamma_u$ is a Poisson process of rate $f(d_{\text{UH}}(w))$. Let $(E_i)_{i \in \mathbb{N}_0}$ be a sequence of independent exponential random variables, where E_i has rate $f(i)$, that are also independent of ξ^w . Since the ray γ_u can be created only after w has produced its first child, it follows that N_w is distributed as a Poisson random variable with mean $f(d_{\text{UH}}(w)) \sum_{i=d_{\text{UH}}(w)+1}^{\infty} E_i$, conditionally on $(E_i)_{i \geq d_{\text{UH}}(w)+1}$. Hence,

$$\mathbb{E} \left[\sum_{w \in \gamma_u} N_w \right] = \sum_{w \in \gamma_u} \mathbb{E} \left[f(d_{\text{UH}}(w)) \sum_{i=d_{\text{UH}}(w)+1}^{\infty} E_i \right] = \sum_{j=d_{\text{UH}}(u)}^{\infty} f(j) \sum_{i=j+1}^{\infty} \frac{1}{f(i)}.$$

By omitting all but the first term in the inner sum (and the outer sum), we arrive at the desired lower bound in the first (and second) step in (5.2). When I_n diverges, this also directly

implies (5.3). For a matching upper bound in (5.2), we write

$$\sum_{i=j+1}^{\infty} \frac{1}{f(i)} = \frac{1 + \varepsilon_j}{f(j+1)},$$

where,

$$\varepsilon_j := \sum_{i=j+2}^{\infty} \frac{f(j+1)}{f(i)} = \sum_{i=j+2}^{\infty} \prod_{k=j+1}^{i-1} \frac{f(k)}{f(k+1)}$$

By (5.1), it follows that for any $\delta \in (0, 1)$ there exists $J = J(\delta) \in \mathbb{N}$ such that for all $j \geq J$,

$$\varepsilon_j \leq \sum_{i=j+2}^{\infty} \delta^{i-j-1} = \frac{\delta}{1 - \delta}.$$

Hence, for $u \in \text{UH}$ such that $d_{\text{UH}}(u) \geq J(\delta)$,

$$\mathbb{E} \left[\sum_{w \in \gamma_u} N_w \right] \leq \frac{1}{1 - \delta} \sum_{j=d_{\text{UH}}(u)}^{\infty} \frac{f(j)}{f(j+1)}.$$

With the same approach, we can bound the sum from above by $(1 - \delta)^{-1} f(d_{\text{UH}}(u)) / f(d_{\text{UH}}(u) + 1)$. As δ is arbitrary, this yields the desired upper bound. $\square$

We then prove the first part of Theorem 5.1.

Proof of Theorem 5.1 (a). The aim is to show that the continuous-time branching process $\mathcal{T}$ almost surely satisfies that, eventually, only individuals along a single infinite path are born. To this end, we use the following steps:

- (i) We construct an infinite tree $\widehat{\mathcal{T}}$ based on the processes $(\widetilde{\mathcal{T}}_u)_{u \in \text{UH}}$. This tree $\widehat{\mathcal{T}}$ consists of at least one infinite ray (and possibly infinitely many infinite rays).
- (ii) We define an auxiliary discrete branching process $(B_n)_{n \in \mathbb{N}}$, where each individual in the process corresponds to an infinite ray of $\widehat{\mathcal{T}}$.
- (iii) We prove that the discrete branching process $(B_n)_{n \in \mathbb{N}}$ is subcritical for all generations $n \geq N$, for some $N \in \mathbb{N}$.
- (iv) We argue that $\mathcal{T}(\tau_{\infty})$ is a subset of $\widehat{\mathcal{T}}$ and that $\mathcal{T}(\tau_{\infty})$ contains a single infinite ray, from which the desired result follows.

(i). We construct $\widehat{\mathcal{T}}$ recursively by constructing a sequence of trees $(\widehat{\mathcal{T}}_j)_{j \in \mathbb{N}}$ and setting $\widehat{\mathcal{T}} = \bigcup_{j \in \mathbb{N}} \widehat{\mathcal{T}}_j$. We recall that $\gamma_u := (u, u1, u11, \dots)$ for $u \in \text{UH}$, and that $T(\gamma_u)$ denotes the (almost surely finite) time at which the process $\widetilde{\mathcal{T}}_u$ has created γ_u (where we omit the subscript if $u = \emptyset$). Further, we introduce the independent random variables $(E_j^{(u)})_{u \in \text{UH}, j \in \mathbb{N}}$, where $E_j^{(u)}$ is an exponential random variable with rate $f(d_{\text{UH}}(u))$.

We initialise the construction by setting $\widehat{\mathcal{T}}_1 := \widetilde{\mathcal{T}}_{\emptyset}(T(\gamma))$, where an active individual $v \in \gamma$ gives birth according to the exponential clocks $(E_j^{(v)})_{j \in \mathbb{N}}$. That is, $\widehat{\mathcal{T}}_1$ consists of the infinite ray $\gamma = (\emptyset, 1, 11, \dots)$ and dormant vertices, which are non-first-born children of $\emptyset, 1, 11, \dots$. We then construct $\widehat{\mathcal{T}}_2$ from $\widehat{\mathcal{T}}_1$ as follows: For each dormant $u \in \widehat{\mathcal{T}}_1$, start the process $\widetilde{\mathcal{T}}_u$, and run it up to time $T(\gamma_u)$, that is, until the infinite ray γ_u is created. Here, active vertices $v \in \gamma_u$ give birth according to the exponential clocks $(E_j^{(v)})_{j \in \mathbb{N}}$. We add each such process $\widetilde{\mathcal{T}}_u$ to $\widehat{\mathcal{T}}_1$ to construct $\widehat{\mathcal{T}}_2$. More precisely, we set

$$\widehat{\mathcal{T}}_2 := \widehat{\mathcal{T}}_1 \cup \bigcup_{\substack{u \in \widetilde{\mathcal{T}}_{\emptyset}(T(\gamma)) \\ u \text{ dormant}}} \widetilde{\mathcal{T}}_u(T(\gamma_u)).$$

To create $\widehat{\mathcal{T}}_3$, we take all dormant vertices u' born in the process of creating $\widehat{\mathcal{T}}_2$, and repeat the same procedure. That is, activate u' , run $\widetilde{\mathcal{T}}_{u'}$ until time $T(\gamma_{u'})$, where each $v \in \gamma_{u'}$ gives birth according to the exponential clocks $(E_j^{(v)})_{j \in \mathbb{N}}$, and attach $\widetilde{\mathcal{T}}_{u'}(T(\gamma_{u'}))$ to $\widehat{\mathcal{T}}_2$. This recursive construction is

visualised in Figure 2. The construction either continues forever or until, at some step $j \in \mathbb{N}$, no dormant vertices are created, in which case we set $\hat{\mathcal{T}}_k = \hat{\mathcal{T}}_j$ for all $k > j$. Then $\hat{\mathcal{T}} = \cup_{j \in \mathbb{N}} \hat{\mathcal{T}}_j$.

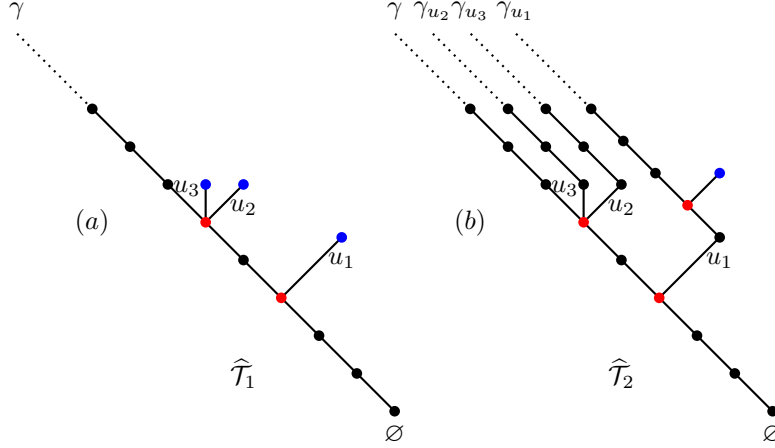


FIGURE 2. (a) The process $\tilde{\mathcal{T}}_\emptyset$ that creates the initial tree $\hat{\mathcal{T}}_1$. The red vertices have produced dormant children before γ is created, which results in three *dormant* individuals, u_1 , u_2 , and u_3 , in blue.

(b) Each of the three dormant individuals is activated and processes $\tilde{\mathcal{T}}_{u_i}$ are started, for each $i = 1, 2, 3$, to create $\hat{\mathcal{T}}_2$. One individual in red in the process $\tilde{\mathcal{T}}_{u_1}$ produces a dormant child before the infinite ray γ_{u_1} is created. This results in one dormant individual, in blue.

(ii) We define the following discrete branching process $(B_n)_{n \in \mathbb{N}}$, where B_n is a finite subset of UH. We set $B_1 = \{\emptyset\}$. Then, $\emptyset$ gives birth to

$$\sum_{w \in \gamma} N_w$$

many children, and let B_1 be the set of dormant children born in the process $\tilde{\mathcal{T}}_\emptyset$. That is, an individual $u \in B_2$ is a dormant vertex in $\tilde{\mathcal{T}}(T(\gamma))$. This process continues analogously for further generations. That is, for $n \geq 2$, each individual u in B_n gives birth to

$$\sum_{w \in \gamma_u} N_w$$

many children in generation $n+1$, each of which corresponds to a unique infinite ray created in the construction of $\hat{\mathcal{T}}_{n+1}$. The branching process either survives forever, or dies out at some generation j , in which case we set B_k as the empty set for all $k > j$. Observe that, by the construction of the $(B_n)_{n \in \mathbb{N}}$, the discrete branching process dying out at generation j is equivalent to no dormant vertices being created in the construction of $\hat{\mathcal{T}}_j$.

(iii) We prove that $(B_n)_{n \in \mathbb{N}}$ is eventually subcritical. By Proposition 5.3, there exists $C > 0$ such that for any $u \in \text{UH}$,

$$\mathbb{E} \left[\sum_{w \in \gamma_u} N_w \right] \leq C \frac{f(d_{\text{UH}}(u))}{f(d_{\text{UH}}(u) + 1)}.$$

We now make the following crucial observation: individuals $u \in B_n$ satisfy that $d_{\text{UH}}(u) \geq n - 1$. As a result, since I_n converges and thus $f(k)/f(k+1)$ tends to zero with k , it follows that there exists $N \in \mathbb{N}$ such that $Cf(n-1)/f(n) < 1$ for all $n \geq N$. As a result, the branching process $(B_n)_{n \in \mathbb{N}}$ is subcritical from generation N onwards.

(iv) Since $(B_n)_{n \in \mathbb{N}}$ is eventually subcritical, it follows that this branching process dies out almost surely. Since each individual in $(B_n)_{n \in \mathbb{N}}$ corresponds to an infinite ray in $\hat{\mathcal{T}}$, it follows that $\hat{\mathcal{T}}$ contains a finite number of infinite rays almost surely. This is visualised in part (c) of Figure 3.

To conclude, we couple the process $(\mathcal{T}(t))_{t \leq \tau_\infty}$ with the tree $\hat{\mathcal{T}}$, where we recall that τ_∞ is the almost surely finite time when $\mathcal{T}$ has grown infinitely large. Let us construct the CTBP $\mathcal{T}$ by letting individual v born in $\mathcal{T}$ give birth to offspring according to the exponential clocks $(E_j^{(v)})_{j \in \mathbb{N}}$. It then directly follows from the construction of $\hat{\mathcal{T}}$ that $\mathcal{T}(\tau_\infty) \subseteq \hat{\mathcal{T}}$. Indeed, any vertex v born in $\mathcal{T}(\tau_\infty)$ must also be born in some process $(\tilde{\mathcal{T}}_u(t))_{t \leq T(\gamma_u)}$ that was started to construct the tree $\hat{\mathcal{T}}$. Since, by Proposition 5.6, the tree $\mathcal{T}(\tau_\infty)$ contains a unique infinite ray almost surely, it follows that this is one of the finite number of infinite rays of $\hat{\mathcal{T}}$. Only finite parts of the other infinite rays of $\hat{\mathcal{T}}$ are contained in $\mathcal{T}(\tau_\infty)$. This is visualised in part (d) of Figure 3. It immediately implies that, in the process $(\mathcal{T}(t))_{t \leq \tau_\infty}$, there exists some (random) time τ_L , at which time the process contains L individuals (L is also random), after which only vertices on a unique infinite path are born, which is the unique infinite path that is part of $\mathcal{T}(\tau_\infty)$.

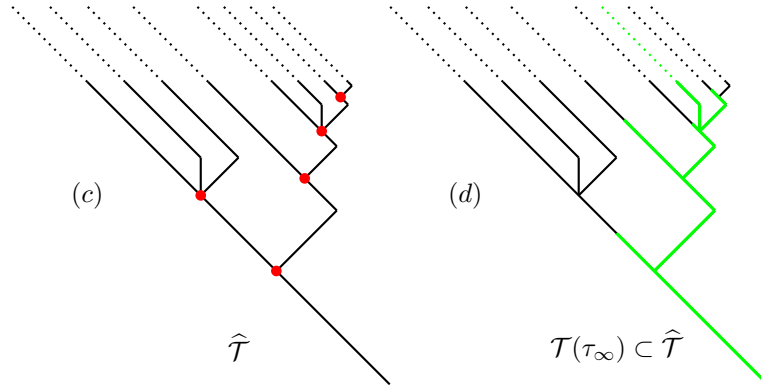


FIGURE 3. (c) The limiting tree $\hat{\mathcal{T}}$. It contains a finite number of infinite rays and a finite number of individuals with more than one child, in red. (d) The infinite tree $\mathcal{T}(\tau_\infty)$ in green as a strict subset of $\hat{\mathcal{T}}$. $\mathcal{T}(\tau_\infty)$ contains a unique infinite ray, which is one of the infinite rays of $\hat{\mathcal{T}}$.

To conclude, we translate this back to the discrete sequence of trees $(T_n)_{n \in \mathbb{N}}$. There exists a random step L , after which new vertices connect only to a unique path that grows infinitely large almost surely. With D the depth of $T_L = \mathcal{T}(\tau_L)$, it follows that $d(T_n) = n - (L - D)$ for all $n \geq L$, almost surely, so that the desired result follows with $K := L - D$ and which concludes the proof. $\square$

Idea behind case (b) in Theorem 5.1.

When I_n does not converge, we can still determine by Proposition 5.3 that

$$\mathbb{E} \left[\sum_{w \in \gamma_\emptyset} \mathbb{1}_{\{d_{\text{UH}}(w) \leq n\}} N_w \right] \approx \sum_{j=0}^n \frac{f(j)}{f(j+1)} = I_n,$$

so that along a path of length n (consisting of firstborn children), one expects to see approximately I_n many additional children. To extend this when we considering all individuals that are not part of the unique infinite path in $\mathcal{T}(\tau_\infty)$ is rather technical and not particularly insightful, and therefore we do not include it here.

Uniqueness of infinite ray.

Lemma 5.4 (Explosion time has no atoms). *Suppose that the weight function f satisfies $\sum_{k=0}^{\infty} f(k)^{-1} < \infty$. Then, the distribution of τ_∞ has no atoms, i.e. $\mathbb{P}(\tau_\infty = t) = 0$ for all $t \geq 0$.*

Remark 5.5. Via an equivalent proof, it follows that the distribution of τ_∞^v has no atoms, for any $v \in \text{UH}$.

Proof. We argue by contradiction. Suppose that the distribution of τ_∞ has an atom, i.e. there exists $t \geq 0$ such that $\mathbb{P}(\tau_\infty = t) > 0$. We let $(E_k)_{k \in \mathbb{N}}$ be a sequence of i.i.d. exponential random variables with rate $f(0)$, to write

$$0 < \mathbb{P}(\tau_\infty = t) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} \left\{ \sum_{k=1}^i E_k + \tau_\infty^{(i)} = t \right\}\right) \leq \sum_{i=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^i E_k + \tau_\infty^{(1)} = t\right),$$

where we use a union bound and that $\tau_\infty^{(i)} \stackrel{d}{=} \tau_\infty^{(1)}$ for any $i \in \mathbb{N}$. Now, the distribution of $\tau_\infty^{(1)}$ has a countable set of atoms (any finite measure has a countable set of atoms), but the distribution of $\sum_{k=1}^i E_k$ has no atoms, for any $i \in \mathbb{N}$. It follows that each probability on the right-hand side equals zero, so that the sum equals zero, leading to a contradiction. $\square$

We use this result to prove the following result.

Proposition 5.6 (Uniqueness of infinite rays in explosive depth-weighted branching process). *Suppose that the weight function f satisfies $\sum_{k=0}^{\infty} f(k)^{-1} < \infty$. Then, the infinite tree $\mathcal{T}(\tau_\infty)$ contains a unique infinite ray.*

Proof. For $v \in \text{UH}$, let $(E_k)_{k \in \mathbb{N}}$ be a sequence of i.i.d. exponential random variables with rate $f(d(v))$. We observe that

$$\tau_\infty^{(v)} = \inf_{i \in \mathbb{N}} \left(\sum_{\ell=1}^i E_\ell + \tau_\infty^{(vi)} \right), \quad \text{or, equivalently,} \quad \mathcal{B}(v) + \tau_\infty^{(v)} = \inf_{i \in \mathbb{N}} (\mathcal{B}(vi) + \tau_\infty^{(vi)}),$$

and that the infimum is attained by at least one integer i (that is, by at least one child vi). We aim to prove that this child is, in fact, almost surely unique. To this end, we write

$$\mathbb{P}\left(\bigcup_{1 \leq i < j < \infty} \{\mathcal{B}(vi) + \tau_\infty^{(vi)} = \mathcal{B}(vj) + \tau_\infty^{(vj)}\}\right) \leq \sum_{1 \leq i < j < \infty} \mathbb{P}\left(\tau_\infty^{(vi)} = \sum_{\ell=i+1}^j E_\ell + \tau_\infty^{(vj)}\right).$$

Note that $\mathcal{B}(vi) + \tau_\infty^{(vi)}$ and $\mathcal{B}(vj) + \tau_\infty^{(vj)}$ are finite almost surely (and hence so are the random variables on the right-hand side), so that the equality is non-trivial. Furthermore, the random variables on the right-hand side are all independent. Then, as $\tau_\infty^{(vi)} \stackrel{d}{=} \tau_\infty^{(vj)}$ have a distribution without atoms by Lemma 5.4, and a sum of i.i.d. exponential random variables has no atoms, each probability on the right-hand side equals zero. As a result,

$$\mathbb{P}(\exists! i \in \mathbb{N} : \mathcal{B}(v) + \tau_\infty^{(v)} = \mathcal{B}(vi) + \tau_\infty^{(vi)}) = 1.$$

Each $v \in \text{UH}$ thus has a unique child whose subtree explodes before the subtrees rooted at all other children of v . Hence, we can construct an infinite path $\gamma = (\emptyset = u_0, u_1, u_2, \dots)$, where u_i is the unique child of u_{i-1} whose subtree explodes before those of the other children of u_i , for each $i \in \mathbb{N}$. Any other path $\gamma' = (\emptyset = u'_0, u'_1, u'_2, \dots)$ is born strictly after γ . As each $v \in \text{UH}$ produces children at a finite rate and $\tau_\infty < \infty$ almost surely, each individual in $\mathcal{T}(\tau_\infty)$ has finitely many children. By Königs lemma, the infinite and locally finite tree $\mathcal{T}(\tau_\infty)$ thus contains an infinite path, almost surely, so that $\mathcal{T}(\tau_\infty)$ contains a unique infinite path, which is γ , almost surely. This concludes the proof. $\square$

6. OPEN PROBLEMS

In these lecture notes we have seen some results on the behaviour of the depth of depth-weighted trees, focussing on the bounded and super-exponential regimes. There are several other results that can be found in [7, 8, 5], but, most importantly, many open problems left to solve! Here are a few that I think are interesting to think about.

Open Problem 6.1 (Polynomial weight functions and super-logarithmic depth). *If f is bounded, the depth of T_n grows logarithmically. In fact, a result in [7] shows that the depth of T_n is of logarithmic order when f grows at most polynomially. Two questions are: Suppose f grows polynomially with exponent $\gamma > 0$, i.e.*

$$\lim_{k \rightarrow \infty} \frac{\log f(k)}{\log k} = \gamma.$$

What is the behaviour of the depth of T_n ? Is it true that $d(T_n) = (1 + \gamma)e \log n + o(\log n)$?

Furthermore, is it true that for any f that grows faster than polynomial, i.e.

$$\lim_{k \rightarrow \infty} \frac{\log f(k)}{\log k} = \infty,$$

we have $d(T_n)/\log n \xrightarrow{\text{a.s.}} \infty$?

Open Problem 6.2 (Polynomial depth). *For a fixed $\beta \in (0, 1)$, for which weight functions f do we observe that $d(T_n) = n^{\beta+o(1)}$? Is it $f(k) = \exp(ck/\log(k+2))$ for some $c > 0$, as suggested in [8]?*

Open Problem 6.3 (Vanishing weight functions). *There are currently no results, beyond Proposition 2.1, regarding weight functions f that are decreasing and tend to 0. Related to Problem 6.1, does $d(T_n)$ still grow at a logarithmic rate when $f(k) = k^{-\gamma}$ for some $\gamma > 0$? It should also be possible to get the depth to grow as slowly as one would want by letting f tend to zero sufficiently quickly. Can the growth of $d(T_n)$ be characterised for such f ?*

The above open problems, as these entire lecture notes, consider themselves only with the depth of the tree. One can, of course, study many other properties and/or questions, too:

Open Problem 6.4 (Degree distribution and local limits). *What is the limiting degree distribution $(p_k)_{k \in \mathbb{N}}$ of T_n ? For f such that T_n grows like a random recursive tree, we expect the degree distribution to also match that of the random recursive tree, i.e. $p_k = (1/2)^k$. What can be said about p_k when f grows quickly?*

Beyond the degree distribution, one can also ask about local convergence. What does the neighbourhood of a typical vertex look like?

Open Problem 6.5 (Root finding, with thanks to Guillem Perarnau). *Suppose that you construct the tree T_n and remove the labels, so that you do not know in which order the vertices arrived. For a fixed $\varepsilon \in (0, 1)$, select a subset of the vertices of size $K(\varepsilon)$, such that the root is contained in this set with probability $1 - \varepsilon$. This problem is known as root finding (in random trees), and tight bound on $K(\varepsilon)$ are known in the case of the random recursive tree, among other models. How big does $K(\varepsilon)$ need to be for depth-weighted trees, and how does this depend on the weight function f ?*

REFERENCES

- [1] K. B. Athreya and P. E. Ney. *Branching processes*. Springer Science & Business Media, 2012.
- [2] J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Applied Probability*, 8(3):446–459, 1976.
- [3] P. Jagers and O. Nerman. The growth and composition of branching populations. *Advances in applied probability*, 16(2):221–259, 1984.
- [4] P. Jagers and O. Nerman. Limit theorems for sums determined by branching and other exponentially growing processes. *Stochastic processes and their applications*, 17(1):47–71, 1984.
- [5] E. Kammerer. Depth profile of depth-weighted trees with bounded weights. *arXiv preprint arXiv:2607.05366*, 2026.

- [6] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7. Cambridge University Press, 2017.
- [7] K. Leckey, D. Mitsche, and N. Wormald. The height of depth-weighted random recursive trees. *Random Structures & Algorithms*, 56(3):851–866, 2020.
- [8] L. Lichev, A. Linker, B. Lodewijks, and D. Mitsche. On the depth of depth-weighted trees. *arXiv preprint arXiv:2602.15709*, 2026.
- [9] O. Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete*, 57(3):365–395, 1981.