STeady state clusters and the Ráth-óth forest fire model

Edward Crane

Abstract. We introduce a random finite rooted tree $C$ which we call the steady state cluster. Its law is characterized by a recursive description: $C$ is a singleton with probability $1/2$ and otherwise is obtained by joining by an edge the roots of two independent trees $C'$ and $C''$, each having the law of $C$, then re-rooting the resulting tree at a uniform random vertex.

We construct and study a stationary regenerative stochastic process $C(t)$ on the space of finite rooted trees, which we call the steady state cluster growth process. It is characterized by a simple fixed-point property. Its stationary distribution is the law of the steady state cluster $C$. We conjecture that $C(t)$ is the local limit of the evolution of the cluster of a tagged vertex in the stationary state of the mean field forest fire model of Ráth and Tóth. In particular we conjecture that the size-biased cluster distribution in the stationary state of the Ráth-Tóth model converges to the law of the steady state cluster as the size of the model tends to infinity.

We couple the steady state cluster growth process to a growth process of genealogical trees, whose stationary measure is the law of the critical binary Galton-Watson tree. We show how to compute many statistics of the steady state cluster growth process. We describe its explosions in terms of a Lévy subordinator, using a state-dependent time change.

We give an alternative description of the steady state cluster as a multitype Galton-Watson tree with a continuum of types. We exhibit the steady state cluster conditioned on its size as a random weighted spanning tree of the complete graph equipped random edge weights with a simple explicit joint distribution. We describe the dynamics of the steady state cluster growth process in terms of coupled multitype Galton-Watson trees. The time-reversal of the steady state cluster growth process is realised as the component of a ‘uniform’ vertex in a logging process of a critical multitype Galton-Watson tree conditioned to be infinite.

Finally we construct a stationary forest fire model on the infinite rooted tree $\mathbb{Z}^*$ with the property that the evolution of the cluster of the root is a version of the steady state cluster growth process. This infinite forest fire model is similar in spirit to Aldous’ frozen percolation model on the rooted infinite binary tree. We conjecture that it is the local weak limit of the stationary Ráth-Tóth model.

1. Introduction

In this paper we introduce and study a random finite tree $C$ which we call the steady state cluster. Its law is characterized by the following simple recursive distributional equation (RDE). With probability $1/2$, $C$ is a singleton. Otherwise,
$\mathcal{C}$ consists of two independent samples from the law of $\mathcal{C}$, joined by an edge that connects two vertices chosen independently uniformly at random from their vertex sets. We will see that this RDE has a unique solution, which is a critical random tree: it is almost surely finite but its expected size is infinite. The steady state cluster is intimately related to the critical binary Galton-Watson tree, which satisfies a similar RDE.

A substantial part of the paper is devoted to an alternative description of the steady state cluster as a critical multitype Galton-Watson tree. When conditioned on its size it may also be seen as a random weighted spanning tree of the complete graph equipped with certain random edge weights. A key to the connection between these descriptions is to consider the stationary distribution of the local dynamics of a forest fire model. In particular we will show that the steady state cluster is the distribution of the connected component of the root vertex in a certain stationary forest fire model on the infinite tree $\mathbb{Z}^*$. 

Our motivation for studying the steady state cluster was to approximate the distribution of clusters in the stationary state of the mean field forest fire model of Ráth and Tóth [19]. This model is a homogeneous càdlàg Markov process on the space of graphs on the finite vertex set $\{1, \ldots, n\}$. We denote it by MFFF($n$). The generator has two terms. The first term is the generator of the dynamical Erdős-Rényi random graph process. Each possible edge that is not already present appears at rate $1/n$, independently. The second term is a Poisson rain of lightning strikes on the vertices. Each vertex is struck by lightning at times of a Poisson process of rate $\lambda n$. The lightning strike processes of the vertices are independent and also independent of the edge arrival process. When lightning strikes a vertex, all the edges in the connected cluster of that vertex are instantaneously removed: we say that the cluster is burned. The vertices survive the fire and continue to acquire edges following the dynamical Erdős-Rényi transition rates. Edges that have been burned may appear again at later times.

Ráth and Tóth studied the limit as $n \to \infty$ of the empirical size-biased cluster size distribution in MFFF($n$), considered as a stochastic process. They concentrated on the asymptotic regime in which the total lightning rate $n\lambda n$ tends to infinity but the lightning rate per site $\lambda_n$ tends to zero as $n \to \infty$. For suitable initial conditions, they showed that in this regime the limit of the cluster size distribution exists as a process as $n \to \infty$. The limit does not depend on the lightning rate except through the above conditions. The limit process is deterministic and exhibits the phenomenon of self-organized criticality (SOC), in the following sense. There is a positive gelation time, before which the cluster size distribution has an exponential tail and the fires have a negligible global effect. At the gelation time the cluster size distribution has a polynomially decaying tail, as it does at criticality in the Erdős-Rényi model. But after the gelation time, the system remains critical: the cluster size distribution continues to have a polynomial tail and finite cluster sizes account for all the mass. In particular there is no giant component.

In [10] we described an integer-valued continuous-time process that is the limit of the evolution of the size of the cluster of a tagged vertex in MFFF($n$), as $n \to \infty$. This local limit is an explosive continuous time Markov branching process in an environment that varies deterministically in time. It is regenerative, returning instantaneously to state 1 at each explosion time. Its distribution at each fixed time agrees with the size-biased global cluster size distribution that defines the
environment. This environment is the global limit of the size-biased cluster size distribution in MFFF$(n)$ and was shown in [19] to be the unique solution of an infinite system of coupled ODEs subject to a conservativity condition, which they called the critical forest fire equations. This system has a unique fixed point, or steady state. In this paper we aim to understand the graph structure of the clusters in MFFF$(n)$ for large $n$ when the system is close to this steady state.

Our main object of study, the steady state cluster growth process, is constructed in [2]. The construction does not rely on any results about MFFF$(n)$ but is motivated by the local limit described in [10]. We start by defining a regenerative cluster growth process of finite rooted trees growing in a constant and deterministic environment of rooted trees. The environment is described by a law $W$ on the set of isomorphism classes of rooted finite trees. The cluster grows by coalescing with trees sampled from $W$, at rate proportional to its own size, and if it explodes then it instantaneously returns to the singleton state. We show that there is a unique law $W_0$ with the property that the stationary distribution of the cluster growth process in the environment $W_0$ is equal to $W_0$. This turns out to be the law of the steady state cluster, as described above via an RDE. The steady state cluster growth process is defined to be the cluster growth process in the environment $W_0$. The size distribution corresponding to $W_0$ is the fixed point $(w_k)_{k=1}^\infty$ of the critical forest fire equations.

In [2] we also describe an enriched process in which each cluster is decorated with a genealogical tree, which is a rooted binary tree whose leaves are identified with the vertices of the cluster. The motivation from thinking about the local limit of MFFF$(n)$ as $n \to \infty$ is that the genealogical tree records information about the coalescences which formed the cluster. The genealogical tree of the steady state cluster also satisfies a fixed point equation, which identifies it as the critical binary Galton-Watson tree. We give a combinatorial description of the joint distribution of the steady state cluster and its genealogical tree. As a consequence we prove re-root invariance of the steady state cluster, and show that the steady state cluster is the unique solution of the RDE given in the first paragraph.

We expect that the steady state cluster growth process arises as the local limit of the stationary states of the mean field forest fire model in the SOC regime, but we are not yet able to prove this. MFFF$(n)$ is aperiodic and irreducible, so it has a unique stationary law $P_n^{stat}$ for each model size $n$. We expect that the steady state cluster growth process should be the local weak limit of $P_n^{stat}$ as $n \to \infty$. However it is not currently known whether the empirical cluster size distribution under $P_n^{stat}$ converges to a deterministic limit as $n \to \infty$. The convergence of solutions of the critical forest fire equations as $t \to \infty$ to the fixed point $(w_k)_{k=1}^\infty$ is also not known; as far as we know it may be the case that there are random non-constant solutions of the critical forest fire equations that have stationary law. Note that the fixed point $(w_k)_{k=1}^\infty$ does not satisfy the third moment condition in the main theorem of Ráth and Tóth [19], so we cannot apply their results to a sequence of initial conditions that converge to the fixed point. Nevertheless, we conjecture that the stationary states do converge to the fixed point solution of the critical forest fire equations in the following sense. Let $v_k^n(t)$ be the random proportion of vertices at time $t$ that belong to clusters of size $k$, under the stationary law $P_n^{stat}$. 


Conjecture 1. For every $\epsilon > 0$ and $T > 0$, 

\[
\mathbb{P}_{\text{stat}}^n\left(\sup_{t \in [0,T]} \sup_{k \geq 1} |v^n_k(t) - w_k| > \epsilon\right) \to 0 \quad \text{as } n \to \infty.
\]

In §3 we derive some basic properties of the steady state cluster growth process. We compute the distribution of the excess time to explosion for a cluster given its size. We obtain closed-form generating functions for various joint distributions of the cluster size, the remaining survival time of the cluster, the age of the root vertex and the degree of the root vertex. We find the joint distribution of the cluster size and the number of jumps in cluster size seen by the root vertex since its last fire. We find a scaling limit for the number of jumps needed to exceed a given large size. We examine the nature of the explosions, proving a limit theorem that describes a time-change of the size process in terms of an exponential functional of a Lévy process. This result is analogous to the behaviour of the tagged fragment in a self-similar fragmentation process.

In §4 we give an alternative description of the law of the steady state cluster, as a multitype Galton-Watson tree. The type of a vertex in this tree corresponds to the age of a vertex in MFFF($n$). The age is defined to be the length of the interval during which a vertex has survived unburned either since time 0 or since its most recent fire. In MFFF($n$), if we condition on the ages of all the vertices, then the set of edges present is the following inhomogeneous random graph. The probability that an edge is present between two vertices of ages $a$ and $b$ is $1 - \exp(-(a \wedge b)/n)$ and the states of all possible edges are independent. We expect that the empirical age distribution in the stationary state $\mathbb{P}_{\text{stat}}^n$ should converge to a certain deterministic distribution $\pi$. We therefore expect that the exploration process of the cluster of a tagged vertex in the stationary state of the mean field forest fire model when $n$ is large should be approximated by a critical multitype Galton-Watson tree. For the purposes of the present paper this argument, due to Balázs Ráth and Dominic Yeo, is just motivation. We define a multitype Galton-Watson tree $H$ without reference to MFFF($n$), and show using the characterizing RDE that $H$ has the law of the steady state cluster. We also define a stationary process $H(t)$, marginally distributed like $H$, and show that it is a version of the steady state cluster growth process. We describe $H(t)$ conditioned on its next explosion time as another multitype Galton-Watson tree size-biased by its total progeny. We give an alternative structural description of this distribution in terms of trees hanging from a finite spine. We derive as a limit the distribution of the fires in the steady state cluster growth process. The fires are distributed as a random one-ended infinite rooted tree $\hat{H}^{(0)}$ that is built up from an infinite spine, along which the vertex ages comprise a Markov chain, with a copy of $H$ conditioned on its root age hanging off each spinal vertex. This random infinite tree is also the local limit of the steady state cluster conditioned to have size $k$, as $k \to \infty$.

The culmination of the present paper, in §5, is the realization of the steady state cluster growth process as the cluster of the root vertex in a stationary forest fire model on an infinite tree. An informal description of this model is quite simple but the rigorous construction is complicated. The difficulty is similar to that encountered by Aldous in the construction of his frozen percolation model on the infinite rooted binary tree [2]. We construct an essentially stationary forest fire model on the infinite rooted plane tree $\mathbb{Z}^*$. This has the property that the cluster
of the root is a version of the steady-state cluster growth process. The future edge arrivals are described by Aldous’ Poisson-weighted infinite tree. The model can be viewed as a candidate for the local weak limit of the Ráth-Tóth model in its steady state, as a process. It satisfies an important necessary condition to be a local weak limit identified by Aldous and Steele [4], namely that it is involution-invariant, or equivalently unimodular. In particular the steady state cluster itself is a re-root invariant random tree and the distribution of the fires as rooted one-ended infinite trees is unimodular. An interesting and probably difficult question is whether the infinite forest fire model is measurable on the sigma algebra generated by its edge arrival times. This is the endogeny problem, which has not yet been resolved in the simpler case of Aldous’ frozen percolation model on the infinite rooted binary tree [2].

In a companion paper, we study the scaling limit of the steady state cluster conditioned to have \( k \) vertices, as \( k \to \infty \). We treat the cluster as a measured metric space equipped with the graph metric and the normalized counting measure on the vertices. We show there that when the metric is rescaled so that each edge has length \( 3/(2\sqrt{2}k) \), the limit exists and is the Brownian continuum random tree (BCRT). The convergence is in distribution with respect to the Gromov-Prokhorov topology.

1.1. Acknowledgements. The author would like to acknowledge helpful conversations on the material in this paper and the companion paper on the scaling limit of the steady state cluster with many people including Louigi Addario-Berry, Omer Angel, Nicolas Curien, Nic Freeman, Christina Goldschmidt, James Martin, Grégory Miermont, Bálint Réth, Oliver Riordan, Bélint Tóth and Dominic Yeo. The author would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for its hospitality during the 2015 programme on Random Geometry, where some of this research was carried out.

2. The geometric cluster growth process

2.1. Growing graphs in a constant environment. Fix a probability distribution \( \mathcal{W} \) on the set Graph of isomorphism classes of rooted unlabelled finite simple undirected graphs. For a graph \( G \) with root vertex \( r \) we will write \( [\{G, r\}] \) for the isomorphism class of \((G, r)\) as a rooted graph. We will treat the distribution \( \mathcal{W} \) as an environment in which to grow a random time-dependent rooted graph \( C^W_1(t) \). The subscript 1 indicates that the process is a graph of size 1 at time 0. To be precise, \( C^W_1(\cdot) \) is a continuous-time Graph-valued Markov process, defined on a random time interval \([0, t_\infty)\), where the distribution of \( t_\infty \) depends on \( \mathcal{W} \). When \( \mathbb{E}(t_\infty) < \infty \) we can subsequently construct a stationary process \( C^W(\cdot) \).

We begin by constructing the process \( C^W_1(\cdot) \). Let \( (C_j, r_j)_{j=1}^\infty \) be an i.i.d. sequence of rooted graphs such that the isomorphism class \([\{C_1, r_1\}]\) of \((C_1, r_1)\) has law \( \mathcal{W} \). Let \( \gamma_1, \gamma_2, \ldots \) be an i.i.d. sequence of exponential random variables with mean 1, independent of the sequence \((C_j, r_j)\). Let \( t_0 = 0 \) and let \( C^W_1(0) \) be a graph consisting of a single vertex \( \rho \), the root. Now we define \( t_i \) and \( C^W_1(t_i) \) inductively for \( i \in \mathbb{N} \). Given \( t_i \) and \( C^W_1(t_i) \), we let

\[
t_{i+1} = t_i + \frac{\gamma_i}{|C^W_1(t_i)|}.
\]
Define $C^W_1(t) = C^W_1(t_i)$ for all $t \in [t_i, t_{i+1})$. Construct $C^W_1(t_{i+1})$ by joining $C_i$ to $C^W_1(t_i)$ by adding an edge from the root $r_i$ to a random vertex $x_i$, chosen uniformly from the vertices of $C^W_1(t_i)$, independently of the sequences $(\gamma_i)$ and $(G_i, r_i)$ and of all $x_j$ for $j < i$. Let $\rho$ be the root of $C^W_1(t_{i+1})$. We denote this joining construction as follows:

$$C^W_1(t_{i+1}) = C^W_1(t_i) \bigcup_{(x_i, r_i)} C_i.$$ 

Let $t_\infty = \lim_{i \to \infty} t_i$. The process $C^W_1(\cdot)$ is defined only on $[0, t_\infty)$. For times $s < t$, $C^W_1(s)$ is naturally a subgraph of $C^W_1(t)$. If $t_\infty < \infty$ then we define $C^W_1(t_\infty)$ to be the nested union of $C^W_1(t)$ over all $t < t_\infty$. Note that this is an infinite rooted graph.

The size process $|C^W_1(t)|$ is a time-homogeneous continuous time Markov branching process. There is a well-known necessary and sufficient condition for explosivity of such a process (see [13, Chapter 5, Thm 9.1]). Let

$$W(z) = \sum_{k=1}^{\infty} \mathbb{P}(|G_1| = k)$$

be the probability generating function for the sizes of graphs distributed according to the law $W$. If the integral $\int_{0}^{1} \frac{dz}{1-W(z)}$ diverges then $t_\infty = \infty$ a.s. Otherwise $t_\infty$ has an exponential tail, so $E(t_\infty) < \infty$.

Remark. The size process $|C^W_1(t)|$ can also be thought of as a time-change of a compound Poisson process whose jump distribution is the marginal size distribution corresponding to $W$. We will exploit this connection, along with another time-change, in [3].

2.2. Construction of a stationary process. Suppose that $E(t_\infty) < \infty$. Then there exists a unique stationary regenerative Markov process $C^W(t)$ for $t \in \mathbb{R}$, whose state space and evolution are identical to those of $C^W_1$ except that at each explosion time the process instantaneously returns to the singleton graph. The sequence of explosion times is a stationary renewal process. We construct $C^W(\cdot)$ by concatenating independent instances of $C^W_1(\cdot)$, using a standard procedure called stationarization. In order to start the process $C^W(t)$ in its stationary state at $t = 0$ we first take a sample of $C^W_1(\cdot)$ that is size-biased in proportion to $t_\infty$. This is possible exactly when $E(t_\infty) < \infty$. Let $s$ be a $U([0, 1])$ random variable independent of $C^W_1(\cdot)$ and define $C(t) = C^W(t + st_\infty)$ for $t \in [-st_\infty, (1 - s)t_\infty)$. Finally, concatenate two independent i.i.d. sequences of instances of $C^W_1(t)$ before and after the interval $[-s, t_\infty + s)$, translated in time so that their domains of definition form a partition of $\mathbb{R}$.

2.3. The unique fixed point.

Lemma 2. There exists a unique probability distribution $W_0$ for which $C^W(t)$ almost surely explodes and the distribution of the isomorphism class of $C^W(0)$ is precisely $W_0$. For $(G, r)$ distributed according to $W_0$, the distribution of $|G|$ is given by

$$\mathbb{P}(|G| = k) = \frac{2}{k} \binom{2k - 2}{k - 1} 4^{-k}.$$ 

$W_0$ is supported on the set $\mathcal{T}_*$ of isomorphism classes of rooted trees.
Proof. Let \(W\) be a law on \(\text{Graph}_*\), and let \((G, r)\) be a random finite rooted graph such that the isomorphism class \([(G, r)]\) is distributed according to \(W\). Write \(w_i = W(|G| = i)\). First we will show that there is a unique size distribution \((w_i)_{i=1}^\infty\) such that \(C^W_1(t)\) almost surely explodes, \(\mathbb{E}(t_\infty) < \infty\) and the resulting stationary distribution of \(|C^W(0)|\) equals \((w_i)_{i=1}^\infty\).

Define
\[
p_r := \mathbb{P}(\exists t \in (t_0, t_\infty) : |C^W_1(t)| = r) .
\]
We have \(p_1 = 1\). Considering the size of the last jump taken to reach size \(r\) we have, for \(r \geq 2\),
\[
p_r = \sum_{i=1}^{r-1} p_{r-i} w_i .
\]
Define the ordinary generating functions
\[
P(z) = \sum_{i=1}^\infty p_i z^i \quad \text{and} \quad W(z) = \sum_{i=1}^\infty w_i z^i .
\]
Both converge on \(|z| < 1\) and \(2\) gives, for \(|z| < 1\),
\[
P(z) = z + P(z)W(z) .
\]
Conditioned on \(\{|C^W(t)| : t \in (t_0, t_\infty)\}\), the holding times form a sequence of independent exponential random variables, with the mean holding time at size \(k\) being \(1/k\). Therefore
\[
\mathbb{E}(t_\infty) = \sum_{r=1}^\infty \frac{1}{r} p_r = \lim_{z \uparrow 1} \int_0^z \frac{P(s)}{s} \, ds = \int_0^1 \frac{1}{1 - W(z)} \, dz
\]
Suppose that the condition \(\mathbb{E}(t_\infty) < \infty\) holds. Then the stationary process \(C^W(t)\) can be constructed as described above so that it is defined for all \(t \in \mathbb{R}\) and is right-continuous. Let \(\pi\) denote the stationary size distribution \(\pi_k := \mathbb{P}(|C^W(0)| = k)\).

Then for \(k \geq 2\) we have
\[
k \pi_k = \sum_{i=1}^{k-1} (k-i) \pi_{k-i} w_i .
\]
Define \(\Pi(z) = \sum_{i=1}^\infty \pi_i z^i\) for \(|z| < 1\). Then \(3\) implies
\[
z \Pi(z) - z \pi_1 = z \Pi'(z) W(z) ,
\]
for \(|z| < 1\), and \(\Pi(0) = 0\). Thus
\[
\Pi(z) = \pi_1 \int_0^z \frac{1}{1 - W(s)} \, ds .
\]
Since \(\lim_{z \uparrow 1} \Pi(z) = 1\), and \(\int_0^1 (1 - W(s))^{-1} \, ds < \infty\), we find \(\pi_1 = \mathbb{E}(t_\infty)^{-1}\), and then \(3\) determines \(\pi\) inductively. We have shown that for each distribution \(W\) for which \(\mathbb{E}(t_\infty) < \infty\) there exists a unique stationary distribution \((\pi_i)\) for \(|C^W(0)|\).

The fixed-point condition on the size distributions is that \(\pi_i = w_i\) for all \(i\), which leads to
\[
W'(z) = \frac{w_1}{1 - W(z)} .
\]
Using $W(0) = 0$ we integrate to get \( W(z) - W(z)^2/2 = zw_1 \). The condition \( \lim_{z \to 1} W(z) = 1 \) fixes \( w_1 = 1/2 \). Hence

\[
W(z)^2 - 2W(z) + z = 0,
\]

which has the unique analytic solution \( W(z) = 1 - \sqrt{1 - z} \) on the unit disc satisfying the condition \( W(0) = 0 \). For this solution of (4) the condition \( E(\infty) < \infty \) is indeed satisfied, since

\[
E(\infty) = \int_0^1 \frac{1}{1 - (1 - \sqrt{1 - z})} \, dz = 2 = \frac{1}{\pi} = \frac{1}{w_1}.
\]

We recover the fixed point size distribution \( (w_k)_{k=1}^{\infty} \) from the Taylor series:

\[
1 - \sqrt{1 - z} = \sum_{k=1}^{\infty} w_k z^k = \sum_{k=1}^{\infty} \frac{2k - 2}{k} 4^{-k} z^k.
\]

To complete the proof of the lemma, we have to show that there is a unique probability distribution \( W_0 \) on \( \text{Graph}_* \) such that if the environment is \( W_0 \) then the distribution of \( C_{W_0}(0) \) is also \( W_0 \). We have already shown that for such a \( W_0 \) the corresponding size distribution \( (w_i) \) must be given by the generating function \( W(z) = 1 - \sqrt{1 - z} \). In particular the probability that \( C_{W_0}(0) \) is a singleton must be \( 1/2 \). It is easy to see that any such distribution \( W_0 \) must be supported on the set \( \mathcal{T}_* \) of isomorphism classes of rooted trees, by considering a non-tree of minimal size in the support of \( W_0 \): there is no way for such a cluster to be formed by joining two trees.

Considering now an arbitrary law \( W \) on \( \mathcal{T}_* \), we next explain how for any rooted tree \( (T, r) \in \mathcal{T}_* \) with \( |T| \geq 2 \) vertices the probability that \( C^W(0) \) is isomorphic to \( (T, r) \) can be computed from the corresponding probabilities for isomorphism classes of rooted trees with fewer than \( |T| \) vertices. The rate at which the stationary process \( C^W(\cdot) \) leaves the state \( (T, r) \) is simply \( |T| \, P(C^W(0) \cong (T, r)) \), and this must equal the rate at which it enters the state \( (T, r) \). By summing rates of possible jumps to \( (T, r) \), we obtain

\[
(5) \quad |T| \, P(C^W(0) \cong (T, r)) = \sum_{[(A, a)] \in \mathcal{T}_*} \sum_{[(B, b)] \in \mathcal{T}_*} P(C^W(0) \cong (A, a)) \, W([(B, b)]) \left| \{v \in A : (A_{\{v,b\}} B, a) \cong (T, r)\} \right| \frac{|T|}{|A|}.
\]

There are only finitely many non-zero terms in this double sum, and they occur in cases where \( |A| < |T| \). Therefore all of these probabilities may be computed inductively with respect to \( |T| \).

Now we can solve the equation

\[
(6) \quad P(C^{W_0}(0) \cong (T, r)) = W_0([T, r])
\]

for all finite rooted trees \( (T, r) \) simultaneously as follows. Assign \( W_0 \)-mass \( 1/2 \) to the class of the singleton graph. Let \( \mathcal{T}_n \) be the set of isomorphism classes of rooted trees with exactly \( n \) vertices. Then we make the following inductive definition,
where the induction is with respect to $k = |T|$. 

$$W_0([T, r]) = \frac{1}{k} \sum_{i=1}^{k-1} \sum_{(A, a) \in T_i} \sum_{(B, b) \in T_{k-i}} W_0([A, a]) W_0([B, b]) \frac{|\{v \in A : (A, a) \sqcup (v, b) \sim (T, r)\}|}{|A|}.$$ 

Define 

$$\tilde{w}_k = \sum_{(T, r) : |T| = k} W_0([T, r]).$$ 

Then we have $\tilde{w}_1 = 1/2$ and for $k \geq 2$ we sum (7) over isomorphism classes of $(T, r)$ with $|T| = k$ to find 

$$\tilde{w}_k = \sum_{i=1}^{k-1} i \tilde{w}_i \tilde{w}_{k-i}.$$ 

As we have seen, the unique solution of these equations is $\tilde{w}_k = 2_k \left(\frac{2k-2}{k-1}\right) 4^{-k}$, and as this sequence sums to 1, the inductive definition (7) defines a probability distribution on $T_\ast$ as required. The rate at which $C W_0$ enters the singleton state is $1/E(t_\infty) = 1/2$. Equating this with the rate at which $C W_0$ leaves the singleton state, we find $P(|C W_0| = 1) = 1/2$, so (6) is satisfied for the singleton tree. Taking (5) and (7) together, one can show by induction that (6) is satisfied for all finite rooted trees.

2.4. The steady state cluster: definitions and notation.

**Definition 1.** We call the stationary process $C W_0(t)$ the steady state cluster growth process, and from now on we denote it simply as $C(t)$. We call the random rooted tree $C$ whose law is $W_0$ the steady state cluster. We denote by $C_1(t)$ the process $C W_0(t)$ that starts as a singleton and is defined on the random time interval $[0, t_\infty)$. We denote by $\rho$ the root vertex of either of these rooted tree valued processes.

Let $(\theta_i)_{i=-\infty}^{\infty}$ be the doubly infinite sequence of explosion times of the stationary process $C(t)$, where almost surely 

$$\cdots < \theta_{-2} < \theta_{-1} < \theta_0 < 0 < \theta_1 < \theta_2 < \cdots.$$ 

We say that the age of the root $\rho$ at time $t$ is 

$$a_\rho(t) := \min\{t - \theta_i : t \geq \theta_i}\).$$ 

Note that $t_\infty$ has the law of $\theta_1$ conditioned on the event $|C(0)| = 1$, and $C_1$ has the law of the restriction of $C(\cdot)$ to the time interval $[0, \theta_1)$ conditioned on $|C(0)| = 1$. Extending this notation we write $C_\ell(t)$ for the process that has the law of the restriction of $C$ to the time interval $[0, \theta_1)$ conditioned on the event $|C(0)| = \ell$. Conditional on $|C(0)|$ the distribution of $\theta_1$ is independent of the structure of $C(0)$ as a rooted tree. However, conditional on $|C(0)|$ the structure of $C(0)$ is not independent of $\theta_0$.

For the rest of the paper, $(w_k)_{k=1}^\infty$ will denote the sequence 

$$w_k := \frac{2}{k} \left(\frac{2k-2}{k-1}\right) 4^{-k}$$
and \( W(z) := 1 - \sqrt{1 - z} \) will denote its ordinary generating function. From Stirling’s approximation we find
\[
w_k \sim \frac{k^{-3/2}}{2\sqrt{\pi}} \text{ as } k \to \infty.
\]
The law of \( \mathcal{C}(\cdot) \) restricted to the time interval \([0, \theta_1)\) is the mixture of the laws of \( \mathcal{C}_\ell(\cdot) \) for \( \ell = 1, 2, 3, \ldots \) with weights \( w_\ell \).

Figure 1 shows the mass that the law of \( \mathcal{C} \) assigns to each of the rooted trees with up to five vertices. Observe that for each unrooted tree the possible rooted versions of the tree occur in proportions consistent with choosing the root vertex uniformly at random from the vertices of the tree. This property is called re-root invariance and we will prove it in the next section.

2.5. The genealogical tree of the steady state cluster. In this section we enrich the steady state cluster growth process by equipping it with a genealogical tree. Recall that the steady state cluster is a candidate for the limit in distribution of the cluster of a tagged vertex in the stationary state of the mean field forest fire model on \( n \) vertices, as \( n \to \infty \). For a cluster \( C \) in the stationary state of the mean field forest fire model, consider the collection of subsets of the vertex set of \( C \) which at some time during the formation of \( C \) formed a cluster. Call these subclusters. The genealogical tree of the cluster is a rooted binary tree \( G \) whose vertices correspond to the subclusters of \( C \). The root vertex \( r \) of \( G \) corresponds to \( C \) itself. Every singleton subset of the vertex set of \( C \) is a subcluster and corresponds to a leaf of the binary tree \( G \). Every non-singleton subcluster \( S \) corresponds to a vertex of \( G \) that has two child vertices, which correspond to the subclusters that were joined together by an edge arrival to form \( S \). In this section we construct a candidate for the limit in distribution of the pair \( (G, C) \) as \( n \to \infty \).

Instead of rooted binary trees it will be technically more convenient to use rooted plane binary trees. Let \( \{0, 1\}_* \) denote the infinite binary tree whose vertices are the finite strings over the alphabet \( \{0, 1\} \). For us, a rooted plane binary tree is a connected subgraph of \( \{0, 1\}_* \) that contains the root \( \emptyset \) and in which every vertex \( v \neq \emptyset \) has degree 1 or 3. That is, each vertex has either 0 or 2 children. The vertices of \( \{0, 1\}_* \) have a lexicographic order, which we will think of as left-to-right ordering. In addition, we will label each vertex of the rooted plane binary tree with a non-negative real number, its spent time. This label is intended to correspond to the age of the corresponding subcluster at the time of its coalescence with another subcluster, or, in the case of the root, the length of time since the whole cluster was formed.

In section 2.4 we considered an environment of rooted graphs. We will now consider a richer environment of pairs \((G, C)\), where \( G \) is a finite rooted plane binary tree with non-negative vertex labels and \( C \) is a rooted tree whose vertex set is the set of leaves of \( G \). We will refer to \( C \) as a cluster and \( G \) as the genealogical tree of the cluster. We write \( \overline{G} \) for the rooted plane binary tree obtained from \( G \) by forgetting the spent time labels. To relate this to our description of genealogical trees in \( \text{MFFF}(n) \), think of each vertex in \( G \) corresponding to the set of leaves above it, which once formed a cluster. We will denote the root of \( C \) by \( \rho \) and the root of \( G \) by \( r \). Let \( T_{\text{gen}} \) be the space of such pairs \((G, C)\), with the obvious topology and Borel \( \sigma \)-algebra. Given an environment that is a law \( \mathcal{V} \) on \( T_{\text{gen}} \), we will now
Figure 1. The probabilities of clusters of up to 5 vertices under $W_0$, the law of $C$.

$n = 1$  $n = 2$  $n = 3$

\[ \begin{array}{ccc} 1 & 1 & 1 \\ 2 & 8 & 48 \\ \end{array} \]

$n = 4$: total probability $w_4 = 5/128$

\[ \begin{array}{cc} 11 & 11 \\ 768 & 768 \\ \end{array} \]

$n = 5$: total probability $w_5 = 7/256$

\[ \begin{array}{cccc} 19 & 19 & 19 & 1 \\ 3840 & 3840 & 7680 & 960 \\ \end{array} \]

\[ \begin{array}{c} \frac{7}{2560} \frac{7}{2560} \frac{7}{1280} \frac{7}{2560} \end{array} \]
Figure 2. An example of a jump of \((G_i^V(t), C_i^V(t))\)

\[
\begin{align*}
C_1(t_i^-) &+ C_i \\
G_i &
\end{align*}
\]

\[
\begin{align*}
\mathcal{G}_1(t_i^-) &
\end{align*}
\]

The rooted plane tree \(G_i\) is constructed as follows. Introduce a new root vertex \(r_i\) with two children, and spent time label 0. If \(b_i = 0\) then the left child of \(r_i\) is identified with the root of \(G_i^V(t_{i-1})\), while the right child is identified with the root of \(G_i\). If \(b_i = 1\) the right child of \(r_i\) is identified with the root of \(G_i^V(t_{i-1})\), and the left child is identified with the root of \(G_i\). Note that we are considering \(G_i^V(t_i)\) as a finite subtree of the infinite rooted plane binary tree \(\{0,1\}_\times\). To construct the tree \(C_i^V(t_i)\), take the trees \(C_i^V(t_{i-1})\) and \(C_i\) and join them by introducing an edge from the root of \(C_i\) to a vertex of \(C_i^V(t_{i-1})\) chosen uniformly at random, independently of all other variables. The root of \(C_i^V(t_i)\) is the vertex that was the root of \(C_i^V(t_{i-1})\).

Notice that the sequence of times \((t_i)\) has a law that only depends on the common distribution of the size increments \(|C_i|\). Likewise the growth process of rooted trees \(C_i^V(t)\) on its own depends only on the marginal distribution \(W\) of \(C_1\) as a rooted tree. The process \(C_i^V(t)\) is therefore a version of the process \(C_i^W(\cdot)\) constructed in §2.1 as is implied by the notation.
We let $t_\infty = \lim_{t \to \infty} t_i$. If $\mathbb{E}(t_\infty) < \infty$ then we construct a corresponding stationary process $(G^V(t), C^V(t))$ by size-biased sampling and concatenation, exactly as we did to construct $C^W(t)$ in §2.2.

**Proposition 3.** There is a unique law $\mathcal{V}_0$ on $\mathcal{T}_\text{gen}$ such that $\mathbb{E}(t_\infty) < \infty$ and the law of $(G^{V_0}(0), C^{V_0}(0))$ is equal to $\mathcal{V}_0$. For $(G, C) \sim \mathcal{V}_0$, the genealogical tree $G$ has the law of the critical binary Galton-Watson tree with vertices labelled by spent rates of entry to state $g$, probability 1 at each jump the previous value of the genealogical tree becomes the left subtree with exits state $b$. Hence by stationarity we obtain $\mathcal{V}_0(G = g) = \mathcal{V}_0(G = g_{\text{left}}) \mathcal{V}_0(G = g_{\text{right}})/2$. Moreover, by comparing rates of entry to state $g$, we see that conditional on $G = g$, the probability that the root $\rho$ belongs to the subtree $g_{\text{left}}$ is $a/(a+b)$, and it follows by induction that conditional on $G$, the root $\rho$ is uniformly distributed among the leaves of $G$. The complete description of $C$ now follows easily by induction over the size of $G$.

Finally we must establish that there is a unique conditional law of the spent time labels of $G$ given $\overline{G}$ for which the law of the spent time labels of $G^{V_0}(0)$ given $\overline{C}^W(0)$ coincides with the environment. The law of the spent time label of the root given $\overline{G}$ and all the other spent time labels only depends on the number of leaves $k$ of $G$. It is a product of a $U([0,1])$ random variable with a size-biased sample of
an exponential holding time with mean $1/k$. This is simply an exponential random variable with mean $1/k$, and the claim now follows by induction over the size of $G$.

\[\square\]

**Corollary 4.** The distribution of $(C, \rho)$ is re-root invariant.

Re-root invariance is a property that we expected, given that we believe the process $C_t$ is the local weak limit of the stationary states of MFFF($n$), in which the vertices are exchangeable. But since we did not construct the steady state cluster growth process as a local weak limit of an exchangeable model, re-root invariance required proof.

**Corollary 5.** The law $W_0$ of $C$ is the unique solution of the following RDE: $C$ is a singleton with probability $1/2$ and otherwise is obtained by taking two independent rooted trees $C'$ and $C''$ distributed according to $W_0$ and joining their roots by an edge, then selecting the root of $C$ uniformly at random among the vertices resulting tree.

\[\text{Proof.}\] Let $(\mathcal{G}, C)$ be the pair distributed as $V_0$. To see that the law of $C$ satisfies the RDE, look at the root $r$ of $\mathcal{G}$. It has no children with probability $1/2$, in which case $\mathcal{G}$ and $C$ are singletons. Otherwise $r$ has two child subtrees $C_{\text{left}}$ and $C_{\text{right}}$ corresponding to an edge $e$ in $C$. The edge $e$ connects the subclusters $C_{\text{left}}$ and $C_{\text{right}}$ corresponding to the left and right child subtrees of $r$ by joining vertices $v_{\text{left}}$ and $v_{\text{right}}$ chosen independently uniformly at random from $C_{\text{left}}$ and $C_{\text{right}}$. By re-root invariance, we may take $v_{\text{left}}$ to be the root of $C_{\text{left}}$ etc. Then the pairs $(G_{\text{left}}, C_{\text{left}})$ and $(G_{\text{right}}, C_{\text{right}})$ are independent and each distributed according to $V_0$.

For uniqueness, note that for any solution $W$ of the RDE, a sample may be taken from $W$ by recursively applying the RDE. This recursion is indexed by a critical binary Galton-Watson tree. Since this is almost surely finite, it easily follows that $W = W_0$.

\[\square\]

The edge described in the RDE corresponds to the root of the genealogical tree. One can understand the RDE in the context of MFFF($n$) by considering the youngest edge in a size-biased cluster.

**Definition 2.** Let $(\mathcal{G}_1(t), C_1(t))$ and $(\mathcal{G}(t), C(t))$ denote the processes described above in the environment $V_0$, and write $(\mathcal{G}, C)$ for the random variable whose law is $V_0$. In particular $C$ has the law of the steady-state cluster which we previously called $C$. For each vertex $v$ in the cluster $C$, the age of $v$, denoted $a(v)$, is defined to be the sum of the spent times of the vertices along the unique path in $\mathcal{G}$ from $v$ to $r$. Likewise for each edge $e$ of $C$, corresponding to a vertex $w$ of $\mathcal{G}$, the age of $e$, denoted $a(e)$, is defined to be the sum of the spent times of the vertices along the unique path in $\mathcal{G}$ from $w$ to $r$. (The endpoints are included in these summations.)

**Corollary 6.** Conditional on $|C| = k$, the sum of the ages of the vertices in $C$ is a $\Gamma(2k - 1, 1)$ random variable and the age of the youngest edge is an exponential random variable with mean $1/k$.

\[\text{Proof.}\] The spent time labels of the vertices of $\mathcal{G}(0)$ are independent exponential random variables. For each vertex $s$ of $\mathcal{G}(0)$, that corresponds to a subcluster of size $m$, the spent time label of $s$ contributes to $m$ terms in the sum, and is an exponential random variable of mean $1/m$. Thus the sum of ages in the cluster is
the sum of $2k - 1$ independent exponential random variables of mean 1. The age of the youngest edge is the spent time of the root, which is exponential with mean $1/k$ given $|\mathcal{C}| = k$. □

3. Properties of the steady state cluster growth process

3.1. The distribution of time to explosion. Let $t_\infty$ be the explosion time of the process $\mathcal{C}_1(\cdot)$, and write $c_t := |\mathcal{C}_1(t)|$, defined for $0 \leq t < t_\infty$. Let $p_r$ be the probability that there exists $t \in (0, t_\infty)$ such that $c_t = r$.

**Lemma 7.** $p_r = \left(\frac{2r-2}{r-1}\right)4^{1-r}$ for each $r \geq 1$.

**Proof.** We have $p_1 = 1$. For $r \geq 2$, the probability that $c_t$ ever jumps from $r - k$ to $r$ is $p_{r-k}w_k$ and these jumps are mutually exclusive, so

$$p_r = \sum_{k=1}^{r-1} p_{r-k}w_k.$$  

Let $P(z) = \sum_{r=1}^{\infty} p_r z^r$. Then $P(z) = z + P(z)W(z)$, so $P(z) = z(1-z)^{-1/2}$ and the result follows. □

By summing the expected holding time at size $r$ we recover

(8) $$E(t_\infty) = \sum_{r=1}^{\infty} \frac{1}{r} \left(\frac{2r-2}{r-1}\right)4^{1-r} = 2.$$  

A similar partition by the last jump to reach $r$ can be applied to compute the probability $p_{r,k}$ that $c_t = r$ for some $s < t < t_\infty$, conditional on $t_\infty > s$ and $c_s = k$. For $r \geq k$ it is simply $p_{r-k+1}$. Therefore

$$E(t_\infty - s \mid t_\infty > s \text{ and } c_s = k) = \sum_{r=k}^{\infty} \frac{1}{r} p_{r-k+1} = \int_0^1 \frac{e^{2z} - 1}{\sqrt{1-e^{2z}}} dz = \frac{4^k}{k(2k)} \sim \frac{\pi}{k} \text{ as } k \to \infty.$$  

Therefore have the following simple expression for the martingale $E(t_\infty \mid \mathcal{F}_t)$, where $\mathcal{F}_t$ is the filtration generated by the process $\mathcal{C}_1(\cdot)$.

$$E(t_\infty \mid \mathcal{F}_t) = \begin{cases} t + \frac{4^k}{k(2k)}, & \text{if } t_\infty > t \text{ and } c_t = k, \\ t_\infty, & \text{if } t_\infty \leq t. \end{cases}$$

Note that the expected amount of time that $\mathcal{C}_1$ spends at size greater than $K$ between time 0 and its explosion time $t_\infty$ is given by the tail of the series in equation (8), which is $O(K^{-1/2})$.

We now compute the distribution of $\theta_1$, the time of the first explosion after time 0 of the stationary process $\mathcal{C}(t)$.

**Lemma 8.** For $x \geq 0$ and $k \geq 1$,

$$P(\theta_1 < x \mid |\mathcal{C}(0)| = k) = 1 - (\cosh(\frac{x}{2}))^{-2k},$$

in particular

$$P(t_\infty < x) = \tanh^2\left(\frac{x}{2}\right),$$

and

$$P(\theta_1 < x) = \tanh\left(\frac{x}{2}\right).$$
Consequently
\[ E(\theta_1) = 2 \log 2 \]

and
\[ E(\theta_1 \mid |C(0)| = k) = \frac{4^k}{k^{(2k)}}. \]

**Proof.** Let
\[ F(x) := \mathbb{P}(\theta_1 < x) \]
and for \( k \geq 1 \) let
\[ F_k(x) := \mathbb{P}(\theta_1 < x \mid |C(0)| = k). \]
In particular \( F_1(x) = \mathbb{P}(t_\infty < x) \). We will relate \( F \) and \( F_1 \) in two ways:

1. \[ F(x) = \sqrt{F_1(x)} , \]
2. \[ 2F'(x) = 1 - F_1(x). \]

To prove (10), suppose \( C(0) \) has \( k \) vertices. Then \( \theta_1 \) is the minimum of \( k \) independent random variables, one for each vertex, each with cumulative distribution \( F_1 \). Hence
\[ F_k(x) = 1 - (1 - F_1(t))^k, \]
and
\[
F(x) = \sum_{k=1}^{\infty} w_k \left( 1 - (1 - F_1(x))^k \right) = \left( \sum_{k=1}^{\infty} w_k \right) - \sum_{k=1}^{\infty} w_k (1 - F_1(x))^k = 1 - \left( 1 - \sqrt{1 - (1 - F_1(x))} \right) = \sqrt{F_1(x)}. \]

To prove (10), note that \( \Theta \) is a the jump process of a stationary renewal process whose inter-arrival times have the law of \( t_\infty \). Therefore \( \theta_1 \) has the law of \( U.S \) where \( S \) is a size-biased sample from the law of \( t_\infty \) and \( U \) is a \( U([0,1]) \) random variable independent of \( S \). Hence
\[ F(x) = \mathbb{P}(\theta_1 < x) = \int_0^x s dF_1(s) + \int_x^\infty \frac{x}{s} dF_1(s). \]
The denominator is \( \mathbb{E}(t_\infty) = 2 \). Differentiating with respect to \( x \) we obtain
\[ 2F'(x) = xF_1'(x) - xF_1'(x) + \int_x^\infty F_1'(s) ds = 1 - F_1(x). \]
Combining (9) and (10) we obtain the differential equation
\[ 2F'(x) = 1 - F(x)^2, \quad F(0) = 0. \]
This has the unique solution \( F(x) = \tanh(x/2) \). Hence \( F_1(x) = \tanh^2(x/2) \) and
\[ F_k(x) = 1 - (1 - F_1(x))^k = 1 - \cosh(x/2)^{-2k}. \]
Finally we integrate to obtain
\[ e_k := \mathbb{E}(\theta_1 \mid |C(0)| = k) = \int_0^\infty \mathbb{P}(\theta_1 > x \mid |C(0)| = k) dx = \frac{4^k}{k^{(2k)}}. \]
\[
\mathbb{E}(\theta_1 1_{|C(0)| = k}) = c_k w_k = \frac{1}{k(2k - 1)}.
\]
and
\[
\mathbb{E}(\theta_1) = \sum_{k=1}^{\infty} e_k w_k = \int_0^{\infty} 1 - F(x) = 2 \log 2.
\]

Straightforward Bayesian calculations, summations and integrations, and Stirling’s approximation now yield the following; we omit the proofs.

**Corollary 9.**

\[
\mathbb{E}(|C(0)| 1(\theta_1 > t)) = \frac{1}{\sinh t}.
\]
\[
\mathbb{P}(|C(0)| = k | \theta_1 = t) = 2k w_k \text{sech}^{2k-2} \left( \frac{t}{2} \right) \tan h \left( \frac{t}{2} \right).
\]
\[
\mathbb{E}(|C(0)| | \theta_1 = t) = 1 + \frac{1}{2 \sinh^2 t} = \frac{\cosh t}{\cosh t - 1}.
\]
\[
\mathbb{E}((|C(0)|^{-1} | \theta_1 = t) = 1 - e^{-t}.
\]
\[
\mathbb{E}(1/\sinh(\theta_1/2) | |C(0)| = k) = k(2k - 1)w_k \pi.
\]

We have the asymptotics

\[
\mathbb{E}(\theta_1 | |C(0)| = k) \sim \frac{\sqrt{\pi / k}}{k}
\]
\[
\mathbb{E}(\theta_1^{-1} | |C(0)| = k) \sim \frac{1}{2} \sqrt{\pi k} \quad \text{as } k \to \infty,
\]
\[
\mathbb{E}(|C(0)| | \theta_1 = t) \sim 2t^{-2}
\]
\[
\mathbb{E}(|C(0)|^{-1} | \theta_1 = t) \sim t \quad \text{as } t \to 0.
\]

The conditional distribution of \((\theta_1/2)^2 |C(0)|\) given \(C(0) = k\) converges to the law of a standard exponential random variable as \(k \to \infty:\)

\[
\mathbb{P}((\theta_1/2)^2k > x | |C(0)| = k) = \text{sech}^{2k} \left( \sqrt{x/k} \right) \to e^{-x} \quad \text{as } k \to \infty.
\]

The conditional distribution of \(\theta_1^2 |C(0)|/2\) given \(\theta_1 = t\) converges to the law of the square of a standard normal random variable as \(t \to 0\).

**3.2. Counting jumps.** Define \(n(t)\) to be the number of jumps of the stationary process \(C(\cdot)\) in the time interval \((\theta_t, t]\) where \(\theta_t \leq t < \theta_{t+1}\).

**Lemma 10.**

\[
(11) \quad \mathbb{E} \left( z^{\left| C(0) \right| n(0)} \right) = \frac{1}{x} \left( W(z) + \frac{1-x}{x} \log(1 - xW(z)) \right).
\]

**Proof.** The pair \((|C(t)|, n(t))\) is an \(\mathbb{N} \times \mathbb{N}\)-valued stationary Markov process. It enters the state \((1, 0)\) at rate 1/2. Equating rates of entry and exit from all the states we obtain

\[
\frac{\partial}{\partial z} \mathbb{E} \left( z^{\left| C(0) \right| n(0)} \right) + \frac{z}{2} = 0,
\]

with the boundary value \(W(x)\) along \(x = 1\). The stated function is the unique solution. ■
Corollary 11. 
\[ \mathbb{E}(n(0) \mid |C(0)| = k) = \frac{1}{2ku_k} - 1 \sim \frac{\sqrt{k}}{2\sqrt{\pi}} \quad \text{as} \quad k \to \infty. \]

The stationary distribution of \( n(t) \) is the Yule-Simon distribution with parameter 1:
\[ \mathbb{P}(n(0) = n) = \frac{1}{(n+1)(n+2)} \quad \text{for} \quad n \geq 0. \]

Proof. Take the partial derivative of the expression (11) with respect to \( x \) and evaluate at \( x = 1 \), to get
\[ \mathbb{E}(z^{\mid C(0) \mid} n(0)) = -W(z) - \frac{1}{2} \log(1 - z). \]

Now extract coefficients and divide by \( \mathbb{P}(\mid C(0) \mid = k) \) to obtain the first equation. For the distribution of \( n(0) \) we evaluate the expression (11) at \( z = 1 \) and extract coefficients. \( \blacksquare \)

Lemma 12. Let \( J_n \) be the number of jumps taken by \( |C_1(\cdot)| \) to exceed \( n \). Then for each \( \alpha > 0 \) we have
\[ \mathbb{P}(J_n \leq \alpha\sqrt{n}) \to \text{erf}(\alpha) \quad \text{as} \quad n \to \infty. \]

Proof. Let \( X_1, X_2, \ldots \) be the sizes of the successive jumps of \( |C_1(\cdot)| \). Then \( J_n > \alpha \sqrt{n} \) if and only if \( X_1 + \cdots + X_{\lfloor \alpha \sqrt{n} \rfloor} < n \). Now consider a simple symmetric random walk \( Z_n \) on \( \mathbb{Z} \) with \( Z_0 = 0 \). For each \( k \geq 0 \) let \( T_k = \min\{n \geq 0 : Z_n = -k \} \) be the first hitting time of \(-k\). Then \( (T_1 - T_0, T_2 - T_1, T_3 - T_2, \ldots) \) is identical in distribution to \( (2X_1 - 1, 2X_2 - 1, 2X_3 - 1, \ldots) \). Therefore
\[ \mathbb{P}(J_n > m) = \mathbb{P}(X_1 + \cdots + X_m < n) = \mathbb{P}(T_m \leq 2n - m - 1) = 2 \mathbb{P}(Z_{2n-m-1} \leq -m), \]
by the reflection principle. Setting \( m = \lfloor \alpha \sqrt{n} \rfloor \), the normal approximation to the binomial distribution yields the result. \( \blacksquare \)

3.3. Fluctuations on the way to explosion. Figure 3 is a plot of \( \log |C_1(t)| \) against \(- \log(t_\infty - t)\) for three simulated sample paths of \( |C_1(\cdot)| \). We will explain the features of this picture by proving a scaling limit theorem which approximates both \( \log(t_\infty - t) \) and \( \log |C_1(\cdot)| \) by functional of a certain Lévy subordinator. We begin by defining a random time change \( \tau(t) \) for the process \( C_1(\cdot) \), which is closely related to \( \log(t_\infty - t) \).
\[ \tau(t) := \int_0^t \sqrt{|C_1(s)|} \, ds. \]

Lemma 13. Almost surely \( \tau(t) \to \infty \) as \( t \nearrow t_\infty \).

Proof. As in the proof of Lemma 12 let \( X_1, X_2, \ldots \) be the random sequence of jumps of \( |C_1(\cdot)| \). This is an i.i.d. sequence with \( \mathbb{P}(X_1 = k) = w_k \). Let \( S_k = 1 + \sum_{i=1}^k X_i \) for \( k \geq 0 \). Then conditional on \( (X_1, X_2, \ldots) \), the limit as \( t \to t_\infty \) of \( \tau(t) \) is the sum of a sequence of independent exponential random variables with means \( 1/\sqrt{S_k} \) for \( k = 0, 1, 2, \ldots \). So it suffices to show that \( \sum_{k=0}^{\infty} 1/\sqrt{S_k} \) is almost surely infinite. Let \( E_n \) be the event that \( |C_1(\cdot)| \) hits \( n \) and then takes at least \( \sqrt{n} \) distinct values in the interval \([n + 1, 2n]\). From Lemma 12 we see that there is a constant \( c > 0 \) such that for all \( n \geq 1 \), \( \mathbb{P}(E_n \mid |C_1(\cdot)| \text{ hits } n) > c \). Conditional on \( |C_1(\cdot)| \) hitting \( n \), the event \( E_n \) is independent of the jumps prior to hitting \( n \). When
Figure 3. Three simulated explosions: the $x$-axis is $-\log (t_\infty - t)$ and the $y$-axis is $\log (|C_1(t)|)$.

$E_n$ occurs, the contribution to $\sum 1/\sqrt{S_k}$ from the passage through the interval $[n + 1, 2n]$ is at least $\sqrt{n}/\sqrt{2n}$. It is now straightforward to show that almost surely $E_n$ occurs for all $n$ in some random infinite sequence $n_1, n_2, \ldots$ such that $n_{i+1} > 2n_i$, and hence $\sum_{k=0}^{\infty} 1/\sqrt{S_k}$ is almost surely infinite. ■

**Remark.** The jump distribution is in the normal domain of attraction of the totally skewed stable law $S_{1/2}(1/4, 1, 0)$ of index $1/2$, also called a Lévy distribution. Thus $S_k/k^2$ has this law as its limiting distribution as $k \to \infty$, and the almost sure divergence of $\sum_{k=0}^{\infty} 1/\sqrt{S_k}$ may instead be proved by comparing the sequence $S_k$ to a stable subordinator of index $1/2$ and applying the law of the iterated logarithm for large times (see for example [6, Ch. 3, Thm. 14]).

Define $Z_\tau$ for every $\tau \in [0, \infty)$ by $Z_\tau(t) = \log |C_1(t)|$. According to Lemma 13 this almost surely defines $Z_\tau$ for all $\tau \in [0, \infty)$. Let $Y_\tau$ be the pure-jump subordinator given by the jump measure

$$\Pi(ds) = \frac{e^s ds}{2\sqrt{\pi}(e^s - 1)^{3/2}} 1(s > 0).$$

$Y_\tau$ is the Lamperti transformation of the standard stable subordinator $S_t$ of index $1/2$. Lamperti’s transformation is a correspondence between self-similar Markov processes and Lévy processes, which in particular relates continuous state Markov branching processes with no killing to pure-jump subordinators. See for example [15, Chap. 10]. A calculation shows that $Y_\tau$ has Laplace exponent

$$\Phi(\lambda) = \frac{\Gamma(\lambda + \frac{1}{2})}{2\Gamma(\lambda)}.$$  \hfill (12)

**Proposition 14.** $Z_\tau$ and $Y_\tau$ may be coupled so that $Z_\tau - Y_\tau$ converges a.s. to a random finite value. The time remaining until explosion may be expressed as
The remaining jumps of $Y$ are by increments smaller than $\log(1 + \frac{\alpha}{\pi e^x})$, and these are not coupled to any jumps of $Z$. We make two claims:

1. There is a constant $c > 0$ such that for every possible simultaneous jump of $Y$ by $\alpha$ and of $Z$ by $\log(1 + ne^{-Z})$ we have

$$|\alpha - \log(1 + ne^{-Z})| \leq \frac{c}{e^{Z}}.$$ 

2. Conditional on the path of $Z$, the expectation of the sum of the jumps of $Y$ that were not coupled to jumps of $Z$ is almost surely finite.

These two claims together imply that the total variation of $Z - Y$ is almost surely finite. The rest of the lemma follows easily.

Proof of claim 1: We have

$$\sum_{i=n}^{\infty} w_i = \left(\frac{2n-2}{n-1}\right)^{4-n} = \frac{1}{\sqrt{\pi n}}(1 + O(1/n)) \quad \text{as} \quad n \to \infty.$$ 

Given simultaneous jumps of $Y$ by $\alpha$ and of $Z$ by $\log(1 + ne^{-Z})$ we have

$$\frac{1}{\sqrt{\pi n}} e^{-b/n} \leq \frac{1}{\sqrt{\pi (e^{\alpha} - 1)}} e^{Z} \leq \frac{1}{\sqrt{\pi n}} e^{b/n},$$

for some constant $b > 0$. In particular, $n \geq e^{-2b/n}(e^{\alpha} - 1)e^{Z}$ so

$$\frac{1 - e^{-\alpha}}{n} \leq e^{2b/n} e^{-\alpha} e^{Z} = O(e^{-Z}).$$
Rearranging (14) we find
\[ \log \left( 1 + \frac{n}{e^{\tau}} \right) = \alpha + \log(1 + O(1/n)(1 - e^{-\alpha})) = \alpha + O \left( e^{-Z_\tau} \right), \]
as required.

Proof of claim 2: Conditional on \( Z_\tau \), the component of the drift of \( Y_\tau \) coming from jumps of size less than \( u := \log \left( 1 + \frac{1}{\pi e^{Z_\tau}} \right) \) is
\[ \int_0^u \frac{s e^s \, ds}{2\sqrt{\pi}(e^s - 1)^{3/2}} = -\log \left( 1 + \frac{1}{\pi e^{Z_\tau}} \right) e^{Z_\tau/2} + \frac{2}{\sqrt{\pi}} \tan^{-1} \left( \frac{1}{\sqrt{\pi} e^{Z_\tau}} \right). \]

As \( Z_\tau \to \infty \), this is asymptotic to \( \frac{1}{\pi} e^{-Z_\tau/2} \). Given that \( Z_\tau \) takes the value \( \log k \) at some time, the expected increment of \( Y \) due to the uncoupled jumps during the interval that \( Z \) spends at \( \log k \) is therefore asymptotic to \( 1/(\pi k) \) as \( k \to \infty \). It follows that the expected total increment from the uncoupled jumps of \( Y_\tau \) is almost surely finite.

Corollary 15. Let \( 1 < \lambda < \infty \). Almost surely there are infinitely many random jump times \( t \) such that \( |C_1(t)| > \lambda |C_1(t^-)| \).

Remark. Proposition 14 is strikingly similar to the description of the process describing the size of a tagged fragment in a self-similar fragmentation process (see Bertoin [7, Theorem 3.2 and Cor. 3.1]). For a binary fragmentation with dislocation measure
\[ \sqrt{\frac{2}{\pi}} x^{-3/2}(1-x)^{-3/2} \, dx \text{ on } \left[ \frac{1}{2}, 1 \right], \]
the process \( \exp(-Y_\tau) \) describes the size of the fragment containing a tagged particle. See also Haas, [12, §4.3.1-2] for a similar result for a fragmentation process which arises from a randomly cutting a random tree. Later in §4.10 we will see that one can view the time-reversal of \( C_1(\cdot) \) as a fragmentation process of a certain random infinite tree.

3.4. Conditioning the size process on the next explosion time.

Lemma 16. The process \( \{|C(s)|, s \in [0,t)\} \) conditioned on \( \theta_1 > t \), and the process \( \{|C_1(s)|, s \in [0,t)\} \) conditioned on \( t_\infty > t \) are time-inhomogeneous Markov jump processes whose jump rate from state \( k \) to state \( k + j \) at time \( s \) is
\[ k w_j \left( \text{sech}^2 \left( \frac{t-s}{2} \right) \right)^j. \]
The process \( \{|C(s)|, s \in [0,t)\} \) conditioned on \( \theta_1 = t \), and the process \( c_s = \{|C_1(s)|, s \in [0,t)\} \) conditioned on \( t_\infty = t \) are time-inhomogeneous Markov jump processes whose jump rate from state \( k \) to state \( k + j \) at time \( s \) is
\[ (k + j) w_j \left( \text{sech}^2 \left( \frac{t-s}{2} \right) \right)^j. \]

Proof. Immediate from Lemma 8. ■

Lemma 16 implies that the drift of \( \{|C(s)| \) conditioned on \( \theta_1 > t \) is
\[ \frac{|C(s)|}{\sinh(t-s)}, \]
and likewise for \( \{|C_1(s)| \) conditioned on \( t_\infty > t \). Solving the resulting differential equation we can find the expected size of both processes conditioned on surviving to time \( t \).
Corollary 17. For $0 \leq s < t$,

\begin{align}
\mathbb{E}(|\mathcal{C}(s)| t^\infty > t) &= \operatorname{tanh} \left( \frac{t}{2} \right) \operatorname{coth} \left( \frac{t-s}{2} \right), \quad (15) \\
\mathbb{E}(|\mathcal{C}(s)| | \theta_1 > t) &= \frac{\operatorname{coth} \left( \frac{t-s}{2} \right)}{1 + e^{-t}}. \quad (16)
\end{align}

Note that these expectations have a simple pole at $s = t$ even though the conditioned processes are almost surely finite at time $t$. Similarly, the drift of $\mathcal{C}(s)$ conditioned on $\theta_1 = t$ is

\[ \frac{1 + |\mathcal{C}(s)|}{\sinh(t-s)} + \frac{1}{2 \sinh(t-s) \sinh^2 \left( \frac{t-s}{2} \right)} . \]

Corollary 18. For $0 \leq s < t$,

\begin{align}
\mathbb{E}(|\mathcal{C}(s)| | t^\infty = t) &= -1 + \operatorname{coth} \left( \frac{t-s}{2} \right) \left( \operatorname{coth}(t-s) + 2 \tanh \frac{t}{2} - \operatorname{coth} t \right) \quad (17) \\
\mathbb{E}(|\mathcal{C}(s)| | \theta_1 = t) &= -1 + \operatorname{coth} \left( \frac{t-s}{2} \right) \left( \operatorname{coth}(t-s) + \tanh \frac{t}{2} \right). \quad (18)
\end{align}

These expectations have a double pole at $t = s$.

3.5. The distribution of $|\mathcal{C}(s)|$ conditioned on $t^\infty$.

Lemma 19.

\begin{align}
\mathbb{E} \left( z^{\mathcal{C}(t)} \big| t^\infty > t \right) &= \frac{z}{(1 + \tanh(\frac{t}{2}) \sqrt{1 - z})^2}. \quad (19)
\end{align}

Proof. Define

\[ F(z,t) := \mathbb{E} \left( z^{\mathcal{C}(t)} \big| t^\infty > t \right) = \sum_{k=1}^{\infty} z^k \mathbb{P}(|\mathcal{C}(t)| = k \text{ and } t^\infty > t) . \]

We consider the rates of entries and exits of $\mathcal{C}_1(\cdot)$ to and from the set of rooted trees of size $k$. Notice that the entry term is a convolution of the partial derivative $F_z$ with $W(z)$. This gives the PDE

\[ \frac{\partial F(z,t)}{\partial t} = -z \sqrt{1-z} \frac{\partial F(z,t)}{\partial z} . \]

The substitution $z = \operatorname{sech}^2(w/2)$, $G(w,t) = F(\operatorname{sech}^2(w/2), t)$ converts this into the simple PDE

\[ \frac{\partial G(w,t)}{\partial t} = \frac{\partial G(w,t)}{\partial w} . \]

This implies that $G(w,t)$ is constant along the lines $w + t = c$. The boundary conditions are

\[ F(1,t) = \mathbb{P}(t^\infty > t) = \operatorname{sech}^2 \left( \frac{t}{2} \right), \quad F(z,0) = z . \]

Converting these into boundary conditions for $G$, we find $G(0,t) = \operatorname{sech}^2 \left( \frac{t}{2} \right)$ and $G(w,0) = \operatorname{sech}^2(w/2)$, so $G(w,t) = \operatorname{sech}^2 \left( \frac{w+t}{2} \right)$ and

\[ F(z,t) = \frac{z}{(\cosh(\frac{t}{2}) + \sqrt{1-z \sinh^2(\frac{t}{2})})^2} . \]
Applying a standard result in singularity analysis (see [11, Cor. 2]) we deduce that for \( t > 0 \) we have
\[
P(t_\infty > t \text{ and } |C_1(t)| = k) \sim \frac{\text{sech}^2 \left( \frac{t}{2} \right) \tanh \left( \frac{t}{2} \right)}{\sqrt{\pi} k^{3/2}} \quad \text{as } k \to \infty,
\]
and
\[
P(|C_1(t)| = k \mid t_\infty > t) \sim \frac{\tanh \left( \frac{t}{2} \right)}{\sqrt{\pi} k^{3/2}} \quad \text{as } k \to \infty.
\]
Moreover, it follows from the singularity analysis that these asymptotics hold locally uniformly in \( t > 0 \).

We now compute the distribution of \(|C_1(s)|\) conditioned on \( t_\infty = t \), for \( 0 \leq s < t \). This coincides with the distribution of \(|C(0)|\) given both \( \theta_0 = -s \) and \( \theta_1 = t - s \). The density of \( t_\infty \) given both \( t_\infty > s \) and \( C_1(\cdot)_{[0,s]} \) depends only on \(|C_1(s)|\), and from Lemma 8 we can compute the density ratio
\[
P(|C_1(s)| = k \mid t_\infty = t) = \frac{P(t_\infty > s \text{ and } |C_1(s)| = k)}{P(t_\infty > s)} \quad \text{for } 0 \leq s < t.
\]
Hence the probability generating function of \( C_1(s) \) conditioned on \( t_\infty = t \) is
\[
E(z |C_1(s)| \mid t_\infty = t) = \frac{\tanh \left( \frac{t-s}{2} \right)}{\text{sech}^2 \left( \frac{t}{2} \right) \tanh \left( \frac{t}{2} \right)} \left( w \frac{\partial}{\partial w} F(w, s) \bigg|_{w = z \text{ sech}^2 \left( \frac{t-s}{2} \right)} \right).
\]
By differentiating the latter expression with respect to \( z \) it is possible to recover our earlier equation (17) for \( E(|C_1(s)| \mid t_\infty = t) \).

From (20) and (21) we obtain that for \( 0 < s < t \), as \( k \to \infty \),
\[
P(|C_1(s)| = k \mid t_\infty = t) \sim \frac{\text{sech}^2 \left( \frac{t-s}{2} \right) \tanh \left( \frac{t}{2} \right) \text{sech}^2 \left( \frac{t-s}{2} \right) \tanh \left( \frac{t}{2} \right)}{\sqrt{\pi} k^{1/2} \text{sech}^2 \left( \frac{t}{2} \right) \tanh \left( \frac{t}{2} \right)}.
\]
This asymptotic holds locally uniformly in \( s \) and uniformly in \( t - s \). Similarly, for \( 0 \leq s \leq t \),
\[
P(|C_1(s)| = k \mid t_\infty > t) = \frac{P(t_\infty > s \text{ and } |C_1(s)| = k)}{\text{sech}^2 \left( \frac{t-s}{2} \right)} \quad \text{as } k \to \infty.
\]
and hence
\[
E \left( z |C_1(s)| \mid t_\infty > t \right) = \frac{F(z \text{ sech}^2 \left( \frac{t-s}{2} \right), s)}{\text{sech}^2 \left( \frac{t}{2} \right)}.
\]
From (20) and (23) we obtain
\[
P(|C_1(s)| = k \mid t_\infty > t) \sim \frac{\text{sech}^2 \left( \frac{t-s}{2} \right) \tanh \left( \frac{t}{2} \right) \text{sech}^2 \left( \frac{t-s}{2} \right) \tanh \left( \frac{t}{2} \right)}{\sqrt{\pi} k^{3/2} \text{sech}^2 \left( \frac{t}{2} \right)} \quad \text{as } k \to \infty.
\]
Notice also that the conditional distribution of \(|C_1(s)|\) given \( t_\infty = t \) is the size-biased version of the conditional distribution of \(|C_1(s)|\) given \( t_\infty > t \). Later, in section 4.8 we will describe explicitly the conditional distribution of \( C_1(s) \) as a rooted tree given \( t_\infty > t \) or given \( t_\infty = t \).
3.6. The distribution of $|C(s)|$ conditioned on $\theta_1$.

**Lemma 20.** For $t \geq 0$,

\[
E\left(z^{\left|C(t)\right|} \mid \theta_1 > t\right) = \frac{1 - \sqrt{1 - z}}{1 + \tanh(\frac{t}{2})\sqrt{1 - z}}.
\]

**Proof.** Conditioning on $|C(0)| = \ell$ instead of $|C(0)| = 1$, we have

\[
F_\ell(z, t) := \sum_{k=1}^{\infty} z^k P(|C(t)| = k \text{ and } \theta_1 > t \mid |C(0)| = \ell) = \frac{z^\ell}{(\cosh(\frac{t}{2}) + \sqrt{1 - z} \sinh(\frac{t}{2}))^{2\ell}}.
\]

Now we can compute the generating function analogous to $F(z, t)$ for the stationary process $C(t)$:

\[
H(z, t) := E\left(z^{\left|C(t)\right|} \mathbf{1}(\theta_1 > t)\right) = \sum_{k=1}^{\infty} z^k P(\theta_1 > t \text{ and } |C(t)| = k)
\]

\[
= \sum_{r=1}^{\infty} \sum_{k=1}^{\infty} P(|C(0)| = r)P(t_\infty > t \mid |C(t)| = k \mid |C(0)| = r) z^k
\]

\[
= \sum_{r=1}^{\infty} P(|C(0)| = r) \prod_{i=1}^{r} \sum_{k_i=1}^{\infty} z^{k_i} P(t_\infty > t \text{ and } c_i = k_i)
\]

\[
= W(F(z, t)).
\]

Continuing with the substitution $z = \text{sech}^2(\frac{w}{2})$, we have

\[
W(F(z, t)) = 1 - \sqrt{1 - \text{sech}^2\left(\frac{w + t}{2}\right)} = 1 - \tanh\left(\frac{w + t}{2}\right)
\]

Using the addition formula for tanh, and noting that $\tanh(\frac{w}{2}) = \sqrt{1 - z}$, we find

\[
H(z, t) = W(F(z, t)) = \frac{(1 - \sqrt{1 - z})(1 - \tanh(\frac{t}{2}))}{1 + \tanh(\frac{t}{2})\sqrt{1 - z}}.
\]

\[\blacksquare\]

Similar calculations to those in the previous section yield

**Corollary 21.** For $t \geq 0$,

\[
\mathbb{P}(|C(t)| = k \mid \theta_1 > t) \sim \frac{1 + \tanh(\frac{t}{2})}{2\sqrt{\pi} k^{3/2}} \text{ as } k \to \infty.
\]

For $0 \leq s \leq t$,

\[
\mathbb{P}(|C(s)| = k \mid \theta_1 > t) \sim \frac{\text{sech}^2(\frac{s}{2}) \text{sech}^{2k}(\frac{t-s}{2})}{(1 - \tanh(\frac{s}{2})) \cdot 2\sqrt{\pi} k^{3/2}} \text{ as } k \to \infty.
\]

and

\[
\mathbb{P}(|C(s)| = k \mid \theta_1 = t) \sim \frac{\text{sech}^2(\frac{s}{2}) \text{sech}^{2k}(\frac{t-s}{2}) \tanh(\frac{t-s}{2})}{\text{sech}^2(\frac{t}{2}) \cdot \sqrt{\pi} k^{1/2}} \text{ as } k \to \infty.
\]
Lemma 22.
\[ \mathbb{E}(z | \mathcal{C}(0) | \theta_0 = -t) = \mathbb{E}(z | \mathcal{C}(0) | \theta_0 \leq -t \text{ and } |\mathcal{C}(-t)| = 1) = \mathbb{E}(z | \mathcal{C}(t) | t_{\infty} > t). \]

Proof. The first equality holds because \( C(\cdot) \) is a Markov process. For the second equality, note that conditional on \( \theta_0 = -t \), the process \( C(\cdot) \) restricted to the random interval \([-t, \theta_1)\) has the same law as \( C_1(\cdot + t) \) conditioned on \( t_{\infty} > t \).

An interesting consequence of Lemmas 19, 20 and 22 is the limiting distribution of the cluster size given that the age of the root is large:
\[
\lim_{t \to \infty} \mathbb{P}(|\mathcal{C}(0)| = k | \theta_0 = -t) = \lim_{t \to \infty} \mathbb{P}(|\mathcal{C}(0)| = k | \theta_0 < -t) = 2w_{k+1}.
\]

4. The steady state cluster as a multitype Galton-Watson tree

Definition 3. We define a multitype Galton-Watson tree \( H \) whose type space is \([0, \infty)\). The root \( \rho \) of \( H \) is a vertex whose type is distributed according the measure \( \pi \), the steady state age distribution, with density \( \frac{1}{2} \text{sech}^2(\frac{x}{2}) \) on \([0, \infty)\).

The multiset of types of the offspring of an individual of type \( s \) is a Poisson random measure (PRM) of intensity \( \lambda_s(dx) = (x \wedge s)\pi(dx) \). We consider \( H \) as a rooted tree \( H = (V, \rho, E) \) equipped with a function \( a : V \to [0, \infty) \) giving the types of the vertices. We will refer to \( a(v) \) as the age of \( v \). We also label each edge \( e = (v, w) \) of \( H \) by a random age \( a(e) \) which, conditional on the ages \( a(v), a(w) \) of the endpoints, is uniformly distributed in the interval \([0, a(v) \wedge a(w)]\). The edge ages are conditionally mutually independent given the tree and vertex ages.

Let us explain the meaning of the PRM in this definition. We consider the offspring of each vertex as an ordered sequence (not necessarily sorted by age). Then the density for the offspring sequence of a vertex of type \( s \) to be \((x_1, x_2, \ldots, x_k) \in [0, \infty)^k\) is
\[
\frac{1}{k!} e^{-|\lambda_s|} \prod_{i=1}^{k} d\lambda_s(x_i),
\]
where \(|\lambda_s|\) denotes the total mass of \( \lambda_s \). This density has support \([0, \infty)^k\). If we treat the offspring of a given vertex as a set rather than a sequence, then accounting for all the possible orderings, the density for the offspring set to be \( \{y_1, \ldots, y_k\} \), where \( y_1 < y_2 < \cdots < y_k \), is
\[
(25) \quad e^{-|\lambda_s|} \prod_{i=1}^{k} d\lambda_s(y_i).
\]

The total mass of the measure \( \lambda_s \) is
\[
|\lambda_s| = \int_0^\infty (s \wedge x) \frac{1}{2} \text{sech}^2(x/2) \, dx = 2 \log(1 + \tanh(\frac{x}{2})).
\]

Thus the number of offspring of a vertex of type \( s \) has the Poisson distribution with mean \( 2 \log(1 + \tanh(\frac{x}{2})) \), and the factor \( \exp(-|\lambda_s|) \) appearing in the density \((25)\) is \((1 + \tanh(\frac{x}{2}))^{-2} \).


4.1. The distribution of the total progeny of $H$.

**Lemma 23.** The total progeny $|H|$ has distribution
\[ P(|H| = k) = w_k. \]

In particular $H$ is critical: it is almost surely finite but its expected size is infinite.

Conditioning on the age of the root, we have for $|z| < 1$
\[ (26) \quad \mathbb{E}(z^{|H|} | a(\rho) = x) = \frac{z}{(1 + \tanh(x/2)\sqrt{1 - z})^2}. \]

**Proof.** Let $\tilde{h}(x, z)$ denote the function on the right-hand side of (26), which could have been guessed from Lemma 22. For $|z| < 1$, let $\tilde{h}(x, z) = \mathbb{E}(z^{|H|} | a(\rho) = x)$. Then for each value of $x \in [0, \infty)$ the function $\tilde{h}(x, \cdot)$ is analytic on the open unit disc. By conditioning on the offspring at the first generation we obtain the nonlinear recurrence relation
\[ (27) \quad h(x, z) = z \exp\left( -\int_0^\infty (1 - h(s, z))(x \wedge s) d\pi(s) \right). \]

To show that $\tilde{h}$ is a solution of (27) is a routine exercise in integration: split the integral at $x$, substitute $u = \tanh(x/2)$ and integrate by parts.

Next, we will show that $\tilde{h}$ is the unique solution of (27) when $z$ is real, positive, and sufficiently small. Then since $h(x, z)$ and $\tilde{h}(x, z)$ are analytic in $z$ for each fixed $x$, we deduce that $h = \tilde{h}$ for all $z$ in the disc $|z| < 1$. Suppose $0 \leq z < 1$. Then we have a priori that $0 < h(x, z) \leq z < 1$, so that
\[ \int_0^\infty (1 - h(s, z))(x \wedge s) d\pi(s) > 0. \]

The same statements hold with $\tilde{h}$ in place of $h$. Hence for every $x \in [0, \infty]$ we have
\[ |h(x, z) - \tilde{h}(x, z)| \leq z \int_0^\infty |h(s, z) - \tilde{h}(s, z)|(x \wedge s) d\pi(s) \]
\[ \leq z \|h(\cdot, z) - \tilde{h}(\cdot, z)\|_\infty \int_0^\infty s d\pi(s) \]
\[ = 2z \log 2 \|h(\cdot, z) - \tilde{h}(\cdot, z)\|_\infty. \]

In particular for $0 < z < 1/(2 \log 2)$ we must have $h(x, z) = \tilde{h}(x, z)$, as required.

Now that we have (26), we compute
\[ \mathbb{E}(z^{|H|}) = \int h(x, z) d\pi(x) = 1 - \sqrt{1 - z}. \]

The integration here is straightforward using the substitution $u = \tanh(x/2)$. Hence $P(|H| = k) = w_k$, as claimed. ■

4.2. The joint distribution of root age, root degree and total progeny.

**Lemma 24.**
\[ (28) \quad \mathbb{E}(z^{|H|} \delta(\rho) | a(\rho) = x) = \frac{z}{(1 + \tanh(x/2)\sqrt{1 - z})^{2s}} \left( \frac{2}{1 + e^{-x}} \right)^{2(s-1)}. \]
Proof. $z^{\lvert H \rvert - 1} s^{\deg(\rho)}$ is the product of $s z^{\lvert H_i \rvert}$ over the offspring of the root, where $H_i$ is the subtree rooted at the $i^{\text{th}}$ child of $\rho$. Since the types of the offspring of the root are a sample of the PRM with intensity $(x \wedge y) d\pi(y)$, we have

$$E(z^{\lvert H \rvert} s^{\deg(\rho)} | a(\rho) = x) = z \exp \left( \int_0^\infty (s h(y, z) - 1)(x \wedge y) d\pi(y) \right),$$

where

$$h(y, z) = E(z^{\lvert H \rvert} | a(\rho) = y) = z \left( 1 + \tanh(y/2) \sqrt{1 - z} \right)^{-2}$$

was computed in the proof of Lemma 23. Using the recurrence (27) that is satisfied by $h$, and computing

$$\int_0^\infty (x \wedge y) d\pi(y) = x - 2 \log \cosh\left( \frac{x}{2} \right) = 2 \log \left( 1 + \tanh\left( \frac{x}{2} \right) \right),$$

we obtain the result. ■

Integrating equation (28) against $d\pi(x)$, we find

(29) $E(z^{\lvert H \rvert} s^{\deg(\rho)}) = \frac{2}{2s - 1} \left( \frac{1 + \sqrt{1 - z}}{2} \right)^{2 - 2s} - \left( \frac{1 + \sqrt{1 - z}}{2} \right).$

Setting $s = 1$ in (28) we recover Lemma 22. Setting $z = 1$ in (28) we recover the fact (which follows from the description of the offspring by a PRM) that conditional on $a(\rho) = x$, the degree of the root has a Poisson distribution with mean $2 \log \left( 1 + \tanh\left( \frac{x}{2} \right) \right) = |\lambda_x|$. Thus the degree of the root of the steady state cluster has a compound Poisson distribution, where the mean of the random Poisson distribution is $2 \log(1 + U)$ where $U \sim U([0, 1])$. Once we have shown that $H$ and $C$ have the same law, it will also be possible to prove (29) directly by using the dynamics of $C(t)$ to write down a recurrence relation for $w_{k,i} = \mathbb{P}(\lvert C \rvert = k, \text{ and } \deg(\rho) = i)$, namely for $k \geq 2$ and $i \geq 1$,

$$0 = -kw_{k,i} + \sum_{j=1}^{k-1} w_{j,i} w_{k-j}(j - 1) + \sum_{j=1}^{k-1} w_{j,i-1} w_{k-j},$$

and converting this into a PDE for $E(z^{\lvert C \rvert} s^{\deg(\rho)})$ whose unique solution is the right-hand side of (29).

4.3. The density of the distribution of $H$. $H$ is a random finite rooted tree with age function $a : V(H) \rightarrow [0, \infty)$. Since the intensity measures $\lambda_x$ are non-atomic, the types in a tree are almost surely distinct. The density for $H$ to be any particular age-decorated rooted finite tree $(T, \rho)$ labelled with vertex ages $a$ is

(30) $\prod_{e = (v, w) \in E(T)} (a(v) \wedge a(w)) \prod_{v \in V(T)} e^{-\lambda_a([0, \infty))} d\pi(a(v))$

$$= \prod_{(v, w) \in E(T)} (a(v) \wedge a(w)) \prod_{v \in V(T)} \frac{e^{-a(v)}}{2} d(a(v)).$$
Including the edge age variables as well we get the joint density for all the age labels of a given rooted labelled tree \((T, \rho)\):

\[
(31) \prod_{e=(v,w) \in E(T)} da(e) 1(a(e) \in [0, a(v) \wedge a(w)]) \prod_{v \in V(T)} e^{-a(v)} 2da(a(v)).
\]

Notice that this simple expression does not depend on the choice of \(\rho\), which is to say that the distribution of \(H\) is re-root invariant.

Note: here we are considering as isomorphic any two trees that are related by a graph isomorphism that preserves the root and preserves ages.

4.4. The marginal distribution of the set of vertex ages in \(H\).

**Lemma 25.** The density for the sorted sequence of ages of the vertices in \(H\) on the chamber \(A_k := \{(a_1, \ldots, a_k) \in \mathbb{R}^k | 0 < a_1 < a_2 < \cdots < a_k\}\) is

\[
(32) 2^{-k} \left( \prod_{j=1}^{k-1} (k-j+1)a_j + \sum_{m=1}^{j-1} a_m \right) \prod_{i=1}^{k} e^{-a_i} da_i.
\]

The sum of the ages is a \(\Gamma(2k-1, 1)\) random variable. Conditioned on the vertex ages, \(H\) is a random weighted spanning tree of the complete graph with edge weights \(w_{ij} = a_i \wedge a_j\).

**Proof.** There are \(n\) ways to assign the root to one of the ages, so the density with respect to the reference measure \(2^{-k} \prod_{i=1}^{k} e^{-a_i} da_i\) on \(A_k\) is \(n\) times the sum over spanning trees \(T\) of the complete graph on the set \(\{1, \ldots, k\}\) of the weight \(\prod_{(i,j) \in E(T)} a_{i \wedge j}\). We can evaluate this sum using Kirchhoff’s matrix-tree theorem: up to sign it reduces to the determinant of the \((k-1)\) by \((k-1)\) matrix \(A\) defined by

\[
A_{ij} = \begin{cases} a_{i \wedge j}, & \text{if } i \neq j, \\ (i-k)a_i - \sum_{m=1}^{i-1} a_m & \text{if } i = j. \end{cases}
\]

If we add columns 2, \ldots, \(k-1\) of \(A\) to column 1 and then add column 1 onto each of columns 2, \ldots, \(k-1\), we obtain a lower triangular matrix with the same determinant, whose \(j^{th}\) entry on the diagonal is \(-a_1\) if \(j = 1\), and \((j-k-1)a_i - \sum_{m=1}^{j-1} a_i\) if \(j > 1\). The sum of weights of spanning trees is the absolute value of the product of the diagonal entries, yielding \((32)\).

Set \(e_1 = ka_1\) and \(e_i = (k + 1 - i)(a_i - a_{i-1})\) for \(i > 1\). In terms of the variables \(e_i\), the density \((32)\) is given on \([0, \infty)^k\) by

\[
\frac{1}{2^k k!} \prod_{j=1}^{k-1} \left( \sum_{m=1}^{j} e_m \right) \prod_{i=1}^{k} e^{-e_i} de_i.
\]

Make the further change of variable \(u_i = \sum_{m=1}^{i} e_m\), for \(i = 1, \ldots, k\), to express this density as

\[
\frac{e^{-u_k} du_k}{2^k k!} \prod_{i=1}^{k-1} u_i du_i 1(0 < u_1 < u_2 < \cdots < u_k).
\]

It is now easy to integrate out the variables \(u_1, \ldots, u_{k-1}\), getting the marginal density for \(u_k\):

\[
\frac{1}{2^{2k-1}k!(k-1)!} \prod_{i=1}^{k} u_i^{2(k-1)} e^{-u_k} du_k.
\]
This is $w_k$ times the density of a $\Gamma(2k - 1, 1)$ random variable; this gives a different proof of the distribution of the total progeny $|H|$. Given $u_k$, the random variables $u_1, \ldots, u_{k-1}$ are the sorted sequence of $k - 1$ independent size-biased $U[0, u_k]$ random variables. Note that
\[ u_k = \sum_{m=1}^{k} e_m = ka_k + \sum_{m=2}^{k} (k + 1 - m)(a_m - a_{m-1}) = \sum_{m=1}^{k} a_m, \]
which is the sum of ages of vertices. So we have shown that conditional on $|H| = k$, the sum of vertex ages is a $\Gamma(2k - 1, 1)$ random variable. ■

**Remark.** The proof showed how to sample the vertex ages efficiently conditioned on $|H| = k$. Since there are efficient algorithms for sampling random weighted spanning trees, such as Wilson’s algorithm, one can efficiently sample $H$ conditioned on its size. However, we will see that $H$ and $C$ are identically distributed, so it would be more efficient still to sample a critical binary Galton-Watson tree $G$ conditioned to have $k$ leaves using Rémy’s algorithm, label its vertices with independent $\text{Exp}(1)$ random variables, and set $H$ as its genealogical tree, and deduce the vertex and edge ages of $C$ from the spent times of $G$.

**4.5. The marginal joint density of edge ages.**

**Lemma 26.** The density for $H$ with the vertex ages forgotten to be isomorphic to a particular finite rooted tree $(T, \rho)$ with edges labelled by ages $(a(e) : e \in E(T))$ is
\[ \prod_{v \in V(T)} \frac{1}{2} \exp(- \max\{a(e) : e \text{ incident on } v\}). \] (33)

**Proof.** Integrate out the vertex ages from (31). ■

**4.6. $H$ and $C$ are identically distributed.**

**Theorem 27.** Let $T$ be a finite unrooted tree with a marked oriented edge $e_0 = (v_1, v_2)$ (oriented from $v_1$ to $v_2$). Let $T_1$ and $T_2$ be the two trees obtained from $T$ by cutting the edge $e_0$, rooted at $v_1$ and $v_2$ respectively. Let $p_1 = \Pr(H \cong (T_1, v_1))$ be the probability that $H$ is isomorphic to $(T_1, v_1)$ as a rooted tree, and similarly $p_2 = \Pr(H \cong (T_2, v_2))$. Orient the youngest edge $e$ in $H$ uniformly at random. Then the probability that $(H, \overrightarrow{e})$ is isomorphic to $(T, \overrightarrow{e_0})$, as unrooted trees with a marked oriented edge, is $p_1p_2$.

Hence $H$ satisfies the same RDE as $C$ and consequently $H$ and $C$ have the same law as rooted trees. In fact they have the same law as rooted trees with vertices and edges labelled by ages.

**Proof.** Take two independent copies $H_1$ and $H_2$ of $H$, with roots $\rho_1$ and $\rho_2$. Let $Y$ be an $\text{Exp}(1)$ random variable independent of $H_1$ and $H_2$ and set $X = Y/(|H_1| + |H_2|)$. Make an age-labelled tree $\tilde{H}$ by adding $X$ to all the vertex and edge labels in both $H_1$ and $H_2$ and joining the roots $\rho_1$ and $\rho_2$ by an edge labelled with age $X$. Choose a root $\tilde{\rho}$ uniformly at random from the vertices of $\tilde{H}$. This produces an age-labelled rooted tree with the youngest edge oriented (from $\rho_1$ to $\rho_2$). Conditional on $(\tilde{H}, \tilde{\rho})$, the orientation of the youngest edge is uniform, by the symmetry of the construction. We claim $(\tilde{H}, \tilde{\rho})$ is distributed as $(H, \rho)$ conditioned on $|H| \geq 2$. Because of re-root invariance of $|H|$ conditioned on its size, this reduces to checking...
that their densities agree as unrooted age-labelled trees with oriented youngest edge. Let us say that an age labelling $a : V(T) \cup E(T) \rightarrow [0, \infty)$ is legal if all the ages are in $[0, \infty)$ and for every vertex the age of each of its incident edges is less than its own age. Then the density for either $(H, \tilde{T})$ or $(\tilde{H}, (\rho_1, \rho_2))$ to be isomorphic to $(T, (v_1, v_2))$ with age labels $a : V(T) \cup E(T) \rightarrow [0, \infty)$, where $|V(T)| = k$, is

$$
\frac{k}{2^k} \prod_{e \in E(T)} da(e) \prod_{v \in V(T)} e^{-a(v)+a(v_1,v_2)} da(v) \mathbf{1}(a \text{ is legal})
$$

$$
= \frac{k e^{-ke(v_1,v_2)}}{2^k} \prod_{i=1}^2 \prod_{e \in E(T_i)} da(e) \prod_{v \in V(T_i)} e^{-a(v)} da(v) \mathbf{1}((a - a(v_1,v_2))|_{T_i} \text{ is legal}).
$$

In the first expression the factor $k$ comes from the choice of $k$ roots for $H$, all equally likely. In the second expression, $k \exp(-ke(v_1,v_2))$ corresponds to the density of $X$ given $|H_1| + |H_2| = k$.

We saw already that $P(|H| = 1) = \frac{1}{2}$. Conditional on $|H| \geq 2$, there is almost surely a unique youngest edge in $H$, and the calculation above shows that uniformly orienting and then cutting this edge yields two independent copies of $H$, age-shifted by an exponential random variable with mean $1/|H|$. It follows that $H$ satisfies the RDE that characterizes $C$. Therefore $H$ and $C$ have the same law as rooted trees.

In section 2.3 we equipped the vertices of the genealogical tree $G$ of $C$ with spent time labels, allowing us to define the age of each vertex and each edge of $C$ in Definition 2. Using the first part of the theorem, we can now couple $H$ and $(G, C)$ so that the vertex ages in $H$ correspond to the vertex ages in $C$. To obtain a sample of the genealogical tree starting from $H$ proceed as follows. If $H$ is a singleton with age $x$ then so is $G$, and we are done. Otherwise, we find the youngest edge of $H$, (which has age $x$, say); we orient it uniformly at random and cut it to obtain subtrees $H_1$ and $H_2$. Give the root $r$ of $G$ spent time $x$ and let its left and right child subtrees be obtained from $H_1$ and $H_2$ respectively by the same procedure, recursively. Since $x$ has distribution $\text{Exp}(k)$ given $|H| = k$, and $kx$ is independent of $H_1$ and $H_2$, we find that $G$ is a critical binary Galton-Watson tree with independent spent time labels that are exponentially distributed with the correct means. Each edge of $H$ corresponds to a vertex of $G$ and joins a pair of leaves of $G$ chosen uniformly and independently from the subtrees of $G$ above that vertex; these choices are independent for different edges of $H$. Keeping track of the age-shifting, we see that the edge and vertex ages as defined in Definition 2 agree with the original edge and vertex labels of $H$. 

4.7. Dynamical version of the multitype Galton-Watson tree. We have exhibited a measure-preserving map from pairs $(G, C)$ distributed according to $\mathcal{V}_0$ to age-labelled multitype Galton-Watson trees distributed according to the law of $H$. We will apply this map to the stationary process $(G(t), C(t))$ to obtain a stationary process $H_t$, whose invariant measure is the law of $H$.

We first consider the non-stationary process $H_1(t)$ for $t \in [0, t_\infty)$, obtained by applying the bijection to the process $(G_1(t), C_1(t))$. The process $H_1(t)$ takes values in the space of finite rooted trees with legal age labellings. It is increasing as a function of $t$ both in the sense of inclusion and in the sense that the labels all increase at rate 1. $H_1(0)$ consists of a single vertex labelled with age 0. The process $H_1(t)$ evolves as a continuous-time Markov process. Let $H_1, H_2, \ldots$ be
the i. i. d. samples of the multitype Galton-Watson tree $H$ obtained by applying the bijection to the pairs $(G_i, C_i)$ in the construction of section 2.5. Recall that $\gamma_1, \gamma_2, \ldots$ are independent exponential random variables with mean 1, and that they define a sequence of random times by $t_0 = 0$, and for $i \geq 1$,

$$t_i = t_{i-1} + \frac{\gamma_i}{|C_1(t_{i-1})|} = t_{i-1} + \frac{\gamma_i}{\sum_{j=1}^{i-1} |H_i|}.$$  

At time $t_i$, $H_1(t)$ jumps by the addition of an edge that joins a uniform random vertex $v_i$ of $H_1(t_{i-1})$ to the root vertex of $H_i$. The vertex $v_i$ is chosen independently of the sequences $H_i$, $\xi_i$ and $v_1, \ldots, v_{i-1}$. Recall that $t_\infty := \lim_{i \to \infty} t_i$.

The stationary process $H(t)$ is now obtained by concatenating instances of $H_1$, with the instance containing time 0 size-biased by its lifetime, exactly as we did to create $C(t)$ from $C_1(t)$. This means that the age of the root of $H(t)$ follows the steady state age distribution $\pi$, with density $(1/2) \theta e^{2x/2}$ on $[0, \infty)$.

Note that the entire description of the process $H(t)$ could be given in terms of the law of $H$ alone, without reference to $(G, C)$. But it would not then be straightforward to show that the invariant measure of $H(t)$ is the law of $H$.

**Lemma 28.** For each $t \geq 0$, the law of $H_1(t)$ conditioned on $t_\infty > t$ is that of the multitype Galton-Watson tree $H$ conditioned on $a(\rho) = t$.

**Proof.** Recall that $P(t_\infty > t) = \theta^2(\frac{t}{2})$ and $P(\theta_1 > t) = 1 - \tanh(\frac{t}{2})$. The law of $H$ conditioned on $a(\rho) = t$ is the law of $H(0)$ conditioned on $\theta_0 = -t$. The joint density of $(\theta_0, \theta_1)$ is

$$\frac{1}{2} \tanh(\frac{2\theta_1 - \theta_0}{2}) \sech^2(\frac{2\theta_1 - \theta_0}{2}) 1(\theta_0 < 0 < \theta_1).$$

Indeed, given $\theta_1 - \theta_0 = \ell$, $\theta_0$ is uniformly distributed on $[-\ell, 0]$, where it has density $1/\ell$; this exactly cancels the size-biasing of the inter-explosion interval containing 0. The law of $H(0)$ conditioned on $\theta_0 = -t$ and $\theta_1 = s$ is the law of $H_1(t)$ conditioned on $t_\infty = s + t$. Now integrate over $s$.  

4.8. The steady state cluster conditioned on its explosion time.

**Definition 4.** For $x \geq 0$, let $H^{(x)}$ be the following multitype Galton-Watson tree with types (ages) in $[0, \infty)$. The root $\rho^{(x)}$ has a random age $a(\rho^{(x)})$, where

$$P(a(\rho^{(x)}) > y) = \frac{1 - \tanh(\frac{x+y}{2})}{1 - \tanh(\frac{x}{2})}.$$  

The offspring of any vertex of age $b$ are described by a PRM whose intensity is the measure $\lambda^{(x)}_b$ defined by the density $(b \wedge a) \frac{1}{2} \sech^2(\frac{a+b}{2}) da$. The edges to the offspring vertices are labelled by random ages that are independent conditional on the offspring ages, where the edge to a vertex of age $a$ has the uniform distribution over the interval $[0, b \wedge a]$.

In particular $H^{(0)}$ has the law of $H$. Note that $a(\rho^{(x)})$ is stochastically decreasing in $x$. Moreover, the intensity of the offspring measure of a vertex of age $b$ in $H^{(x)}$ is increasing in $b$ and decreasing in $x$. Thus the offspring of the root $\rho^{(x)}$ is stochastically decreasing in $x$: for $x < y$ it is possible to couple $H^{(x)}$ and $H^{(y)}$ so that the offspring of $\rho^{(y)}$ form a subset of the offspring of $\rho^{(x)}$ with the same ages. It is therefore possible to couple $H^{(x)}$ and $H^{(y)}$ so that $H^{(y)}$ is isomorphic as a rooted tree to a subtree of $H^{(x)}$. If we condition $H^{(x)}$ and $H^{(y)}$ both to have root
age \( s \), then they may be coupled so that \( H^{(y)} \) is isomorphic to a subtree of \( H^{(x)} \) as rooted trees with identical age labels on all common edges and vertices.

**Lemma 29.** For \( 0 \leq s \leq t \), the law of \( H(s) \) conditioned on \( \theta_1 > t \) is the law of \( H^{(t-s)} \) conditioned to have root age \( a(\rho) > s \). In particular, the law of \( H(0) \) conditioned on \( \theta_1 > t \) is the law of \( H^{(t)} \).

**Proof.** By stationarity, sampling \( H(t) \) conditioned on \( \theta_1 > t \) is the same as sampling \( H \) conditioned on \( a(\rho) > t \). We can then run the clock backwards from time \( t \) to time \( s \). This means that we delete every edge whose age at time \( t \) is less than \( t - s \), and keep only the connected component of the root. In particular we delete all vertices whose age at time \( t - s \) is less than \( t \), but possibly some older vertices as well. Consider a vertex that survives this pruning procedure and has age \( b \) in \( H(s) \). It is a vertex of age \( b + t - s \) in \( H(t) \). Its offspring (vertex, edge) pairs in \( H(t) \) are described by a PRM of intensity

\[
1(0 < a(e) < (b + t - s) \wedge a(v)) \, da(e) \cdot \frac{1}{2} \text{sech}^2\left(\frac{a(v)}{2}\right) \, da(v).
\]

Keeping only the offspring whose edges have age at least \( t - s \) at time \( t \), and winding the clock back to time \( s \), we get a PRM of intensity

\[
1(0 < a(e) < b \wedge a(v)) \, da(e) \cdot \frac{1}{2} \text{sech}^2\left(\frac{a(v) + t - s}{2}\right) \, da(v).
\]

The age of the root at time \( s \) is \( X - t + s \) where \( X \) is distributed according to the steady state age distribution \( \pi \) conditioned on \( X > t \). This is the distribution of \( a(\rho^{(t-s)}) \) conditioned to be at least \( s \).

One could also prove Lemma 29 by computing the density of \( H(t) \) explicitly, similarly to equation (30), and observing that it is the tilt of the density of \( H \) by the Bayes factor

\[
(\text{sech}^2\left(\frac{1}{2}\right)) |V(T)|/(1 - \tanh\left(\frac{1}{2}\right)) = \frac{\mathbb{P}(\theta_1 > t \mid H(0) = T)}{\mathbb{P}(\theta_1 > t)}.
\]

The next lemma describes the dynamics of the process \( H_1(\cdot) \) conditioned on the event \( t_\infty > t \).

**Lemma 30.** For \( 0 < s < t \) the law of \( H_1(s) \) conditioned on \( t_\infty > t \) is the law of \( H(0) \) conditioned on \( \theta_0 = -s \) and \( \theta_1 > t - s \), which is the law of \( H^{(t-s)} \) conditioned to have root age \( s \). Keeping \( t \) fixed and letting \( s \) increase, conditional on \( t_\infty > t \) and \( H_1(s) \), new offspring arrive at each vertex of \( H_1(s) \) independently as a Poisson rain of intensity \( \frac{1}{2} \text{sech}^2\left(\frac{a(\rho^{(t-s)})}{2}\right) \, da \) on the type space \( [0, \infty) \). Each new offspring vertex of age \( a \) comes with a subtree that has the law of the Galton-Watson tree \( H^{(t-s)} \) conditioned to have root age \( a \).

**Proof.** The first sentence is proved in the same way as Lemma 28. The rest follows from the proof of Lemma 29 by considering how the pruned subtrees are added as we wind the clock forward again.

**Remark.** Note that the arrivals process at each vertex depends only on \( t \), and not on the structure or age-labelling of \( H_1(s) \), as we would expect from thinking about the process \( C_1(s) \) conditioned on \( t_\infty > t \).
In Lemma 30 we found the law of $H_1(s)$ conditioned on $t_\infty > t$. We now examine the law of $H_1(s)$ conditioned on $t_\infty = t$, for $0 \leq s < t$. We obtain this by taking the limit as $\epsilon \to 0$ of the law of $H_1(s)$ conditioned on $t_\infty \in [t, t + \epsilon)$. From (35) we have

$$
P(|C_1(s)| = k \mid t_\infty = t) = \frac{k \sech^2 \frac{t-s}{2} \tanh \frac{t-s}{2} \sech^2 \frac{s}{2}}{\sech^2 \frac{t}{2} \tanh \frac{s}{2}} P(|C_1(s)| = k \mid t_\infty > s).
$$

Conditioning on $t_\infty = t$ tilts the density of $C_1(s)$ by a factor that depends only on $|C_1(s)|$. We saw in Lemma 28 that the distribution of $C_1(s)$ given $t_\infty > s$ is the law of $H$ conditioned to have $a(\rho) = s$.

We compute

$$
\lambda^{(t)}(s) = a - 2 \log \cosh \frac{a+s}{2} + 2 \log \cosh \frac{s}{2}.
$$

The marginal density for $H^{(t-s)}$ conditioned to have root age $s$ to be a given tree $(T, \rho)$ labelled with vertex ages $a : V(T) \to [0, \infty)$ with $a(\rho) = s$ is

$$
\prod_{(v, w) \in E(T)} (a(v) \wedge a(w)) \prod_{v \in V(T)} e^{-\lambda^{(t-s)}(v)} \prod_{v \in V(T) \setminus \{\rho\}} \frac{1}{2} \sech^2 \left( \frac{a(v)+t-s}{2} \right) da(v)
$$

$$
= \frac{2 \cosh^2 \left( \frac{t}{2} \right)}{\cosh^{2|E(T)|} \left( \frac{t-s}{2} \right)} \prod_{(v, w) \in E(T)} (a(v) \wedge a(w)) \prod_{v \in V(T)} e^{-a(v)} \prod_{v \in V(T) \setminus \{\rho\}} da(v).
$$

This is the tilt of the density of $H$ conditioned on $a(\rho) = s$ by the factor $\sech^{2|E(T)|} \left( \frac{t-s}{2} \right) \cosh^2 \left( \frac{t}{2} \right) \sech^2 \left( \frac{s}{2} \right)$, which depends on $(T, \rho, a)$ only through the factor $\sech^{2|E(T)|} \left( \frac{t-s}{2} \right)$. Comparing (34) and (35) we obtain

**Lemma 31.** For $0 \leq s < t$, the law of $H_1(s)$ conditioned on $t_\infty = t$ is obtained by size-biasing the law of $H^{(t-s)}$ conditioned to have root age $s$. The size-biasing here means that we tilt the law by the total progeny $|H^{(t-s)}|$. The normalizing factor for this tilt is given by

$$
\mathbb{E} \left( |H^{(t-s)}| \mid a(\rho) = s \right) = \frac{\tanh \frac{s}{2}}{\tanh \frac{t}{2}}.
$$

**Lemma 32.** For $0 \leq s < t$ the law of $H(s)$ conditioned on $\theta_1 = t$ is obtained by size-biasing the law of $H^{(s)}$ conditioned on $a(\rho) > s$, i.e. tilting that distribution by the total progeny.

**Proof.** The Bayesian calculations above show that for $0 \leq s < t$ we have

$$
P(|C(s)| = k \mid \theta_1 = t) = k \tanh \left( \frac{t-s}{2} \right) \left( 1 + e^{-t} \right).
$$

Thus the law of $H(0)$ conditioned on $\theta_1 = t$ is obtained by size-biasing the law of $H(0)$ conditioned on $\theta_1 > t$, which was shown in Lemma 29 to be the law of $H(t)$ conditioned on $a(\rho) > s$. ■
4.9. Spinal representation of the size-biased $\hat{H}^{(x)}$. We define another multitype Galton-Watson tree $\hat{H}^{(x)}$. The root $\hat{\rho}^{(x)}$ has age $a(\hat{\rho}^{(x)})$ where

$$
\mathbb{P}(a(\hat{\rho}^{(x)}) > y) = \frac{1 - \tanh^2(\frac{x+y}{2})}{1 - \tanh^2(\frac{x}{2})} = \frac{\cosh^2(\frac{x}{2})}{\cosh^2(\frac{x+y}{2})}.
$$

The density of $a(\hat{\rho}^{(x)})$ is $\operatorname{sech}^2(\frac{a+x}{2}) \tanh(\frac{a+x}{2}) \cosh(\frac{x}{2}) da$ on $[0, \infty)$. The vertices of $\hat{H}^{(x)}$ fall into two classes: spinal and non-spinal. The root $\hat{\rho}^{(x)}$ is a spinal vertex. The offspring measure of a non-spinal vertex of age $b$ is a PRM of intensity $\lambda_b^{(x)}$ defined by the density $(b \wedge a)^{\frac{1}{2}} \operatorname{sech}^2(\frac{a+x}{2}) da$, (just as for the tree $H^{(x)}$). All offspring of a non-spinal vertex are non-spinal. For a spinal vertex of age $b$, the offspring is the union of a set of non-spinal vertices given by a PRM with intensity measure $\lambda_b^{(x)}$ with either zero or one spinal vertices. The spinal offspring are independent of the non-spinal offspring. There is no spinal child with probability $\tanh(\frac{x}{2}) / \tanh(\frac{b+x}{2})$. If there is a spinal child, its age is distributed according to the measure $\lambda_b^{(x)}$ which is the measure $\lambda_b^{(x)}$ tilted by the factor $\tanh(\frac{a+x}{2})$. Explicitly, $\lambda_b^{(x)}$ is the probability measure with density

$$
\frac{(b \wedge a) \operatorname{sech}^2(\frac{a+x}{2}) \tanh(\frac{a+x}{2})}{2(\tanh(\frac{b+x}{2}) - \tanh(\frac{x}{2}))} da.
$$

Thus the density for a spinal vertex of age $b$ to have a spinal child of age $a$ is

$$
\frac{(b \wedge a) \operatorname{sech}^2(\frac{a+x}{2}) \tanh(\frac{a+x}{2})}{2 \tanh(\frac{b+x}{2})} da.
$$

**Lemma 33.** Let $x > 0$. After forgetting the spine (but keeping the root), the tree $\hat{H}^{(x)}$ has the law of $H^{(x)}$ tilted in proportion to the total progeny, as in Lemmas 31 and 32. The same holds when we condition both distributions on the root age. Moreover, conditioned on $H^{(x)}$ being isomorphic to $(T, \rho, a)$, the spine is the shortest path from the root to a uniform random vertex of $T$. Conditioned on $H^{(x)}$ being isomorphic as an age-labelled unrooted tree to $(T, a)$, the spine is the oriented path between two independent uniform random vertices of $T$.

**Proof.** For a spinal vertex of age $b$, the density for the offspring to consist of $k \geq 0$ non-spinal vertices with ages $a_1 < \cdots < a_k$ and no spinal vertex is

$$
\frac{\tanh(\frac{x}{2})}{\tanh(\frac{b+x}{2})} e^{-|\lambda_b^{(x)}|} \prod_{i=1}^{k} (b \wedge a_i) \frac{1}{2} \operatorname{sech}^2(\frac{a_i+x}{2}).
$$

For a spinal vertex of age $b$, the density for the offspring to consist of $k \geq 0$ non-spinal vertices with ages $a_1 < \cdots < a_k$ and one spinal vertex of age $b'$ is

$$
(b \wedge b') \frac{1}{2} \operatorname{sech}^2(\frac{b'+x}{2}) \frac{\tanh(\frac{b'+x}{2})}{\tanh(\frac{b+x}{2})} e^{-|\lambda_b^{(x)}|} \prod_{i=1}^{k} (b \wedge a_i) \frac{1}{2} \operatorname{sech}^2(\frac{a_i+x}{2}).
$$

Recall that $e^{-|\lambda_b^{(x)}|} = e^{-b} \operatorname{cosh}^2(\frac{b+x}{2}) \operatorname{sech}^2(\frac{x}{2})$.

Let $(T, \rho, a)$ be a rooted tree with vertices labelled by distinct ages. There are $|T|$ ways to choose a spine $S \subseteq V(T)$, being the shortest path from the root to a marked vertex $v_0$. For any choice of spine $S$, the density for $\hat{H}^{(x)}$ to be $(T, \rho, a)$
with spine $S$ is
\[ 2 \text{sech}^2|T| - \frac{2}{2} \tanh(\frac{x}{2}) \prod_{v \in V(T)} \frac{e^{-a(v)}}{2} \prod_{(v,w) \in E(T)} (a(v) \wedge a(w)). \]

Note that this density does not involve the choice of root or the choice of $v_0$.

In computing this density, for each vertex $v \in S$ there were cancelling factors of $\tanh(\frac{a(v) + x}{2})$ in the numerator and denominator. For comparison, the density for $\hat{H}^{(x)}$ to be $(T, \rho, a)$ is
\[ \frac{\text{sech}^2|T| - \frac{2}{2}}{1 - \tanh(\frac{x}{2})} \prod_{v \in V(T)} \frac{e^{-a(v)}}{2} \prod_{(v,w) \in E(T)} (a(v) \wedge a(w)). \]

Since there are $|T|$ choices of spine, after forgetting the spine the density of $\hat{H}^{(x)}$ is the tilt of the density of $H^{(x)}$ by the factor $|T| 2 \sinh(\frac{x}{2}) e^{-x/2}$. In light of Lemma 29 and Lemma 32, this agrees with equation (16) which tells us that $1/E(|H(0)|: \theta_1 > x) = 2 \sinh(\frac{x}{2}) e^{-x/2}$.

Repeating this calculation with the root conditioned to have age $s$ we find that $\hat{H}^{(x)}$ conditioned to have root age $s$ has the size-biased distribution of $H^{(x)}$ conditioned to have root age $s$. ■

The offspring measure of a spinal vertex of age $b$ in $\hat{H}^{(x)}$ is the tilt of the PRM with intensity $\lambda_0^{(x)}$ by the factor $\left(\tanh(\frac{x}{2}) + \sum_{i=1}^k \tanh(\frac{a_i + x}{2})\right)$, where $a_1 < \cdots < a_k$ are all the offspring ages. The vertex with age $a_i$ is the spinal child with probability $\tanh(\frac{a_i + x}{2})/(\tanh(\frac{x}{2}) + \sum_{i=1}^k \tanh(\frac{a_i + x}{2}))$. There is no spinal child with probability $\tanh(\frac{x}{2})/(\tanh(\frac{x}{2}) + \sum_{i=1}^k \tanh(\frac{a_i + x}{2}))$.

4.10. Dynamics of the steady state cluster conditioned on its explosion time. In Lemma 16 we described the dynamics of the size processes $|C(\cdot)|$ and $|C_1(\cdot)|$ conditioned on $\theta_1$ and $t_\infty$, respectively. In 4.1 we described the dynamics of $C_1(\cdot)$ conditioned on $t_\infty > t$ and the dynamics of $C(\cdot)$ conditioned on $\theta_1 > t$. We used this to understand the distribution of the rooted trees $C(s)$ and $C_1(s)$ conditioned on their next explosion time as size-biased multitype Galton-Watson trees. In 4.9 we gave an explicit spinal representation of these laws. Using the spinal representation we can now easily describe the dynamics of $C(\cdot)$ and $C_1(\cdot)$ conditioned on their next explosion time being $t$. The processes have the same generator on the interval $[0, t]$; the difference is in the initial law at time 0. In this section we will only consider $C(\cdot)$ conditioned on $\theta_1 = t$.

We construct a monotone coupling of $\hat{H}^{(x)}$ conditioned on $a(\hat{p}^{(x)} > (t - x)$ over all $x \in [0, t]$. We start by noting that the tree $\hat{H}^{(0)}$ makes sense. The age of the root $\hat{p}^{(0)}$ has density $\text{sech}^2(\frac{x}{2}) \tanh(\frac{x}{2}) \, da$, so it has the same distribution as $t_\infty$. In $\hat{H}^{(0)}$ the law of the offspring distribution of a spinal vertex of age $b$ is simply the usual (untitled) PRM with intensity $\lambda_b$ together with an extra, independent, point (the spinal offspring), which is distributed as the tilt of $\lambda_b$ by $\tanh(\frac{x}{2})$. So $\hat{H}^{(0)}$ has an infinite spine starting at the root, and the ages of the vertices up the spine form a Markov chain. We will study this Markov chain in 4.11 where we will see that its invariant measure has cumulative distribution function $\tanh^3(\frac{x}{2})$. This means that the root age is out of equilibrium for the spinal Markov chain. Attached to the spine are independent multitype Galton-Watson trees, one rooted at each vertex of
the spine. The tree rooted at spinal vertex \( v \) has the law of \( H \) conditioned on its root age being \( a(v) \). Almost surely, all of these countably many attached trees are finite, so \( \hat{H}(0) \) is an infinite tree with one end. Each edge in the tree can be assigned a random age. As usual, we can assign edge ages independently, so that for an edge \( e \) joining vertices \( v \) and \( w \) the age \( a(e) \) is uniformly distributed in \( [0, a(v) \wedge a(w)] \).

Let \( T_t \) be a random infinite tree with root \( \rho_t \), age labelling \( a_t \) and spine \( S_t \), distributed according to the law of \( \hat{H}(0) \) conditioned on \( a_t(\rho_t) > t \). That is, the density of the root age is

\[
cosh^2 \left( \frac{a}{2} \right) \text{sech}^2 \left( \frac{a}{2} \right) \tanh \left( \frac{a}{2} \right) 1(a > t) da .
\]

We now wind the clock backwards from time \( t \). As we rewind the clock we decrease all ages at rate 1 and any edge whose age reaches 0 is deleted. We keep only the connected component of the root, and call this tree \( T_s \). We denote by \( S_s \) the part of the spine \( S_t \) that is contained in the subtree \( T_s \), and by \( a_s \) the age-labelling of the tree at time \( s \). Thus we get an increasing family of rooted trees \( (T_s, \rho_s, a_s, S_s) \), each decorated with a spine and with vertex and edge ages.

**Lemma 34.** For each \( s \in [0, t] \) the law of \( (T_s, \rho_s, a_s, S_s) \) is the law of \( \hat{H}^{(t-s)} \).

Now we can run the clock forwards from time \( 0 \) and observe that \( (T_s, \rho_s, a_s, S_s) \) is a Markov process with respect to the filtration that it generates. At time \( 0 \) it is a sample of \( \hat{H}^{(t)} \), consisting of a rooted tree \( (T, \rho) \) equipped with age labels on the vertices and edges and a spine from the root to a marked vertex \( v_0(0) \). Its dynamics are as follows:

- The age of each existing vertex and each existing edge increases at rate 1.
- For each vertex \( v \) independently, at rate \( 1 - \tanh \left( \frac{a(v)}{2} \right) \) a new edge \( e \) arrives joining \( v \) to the root \( w \) of an independent sample of \( \mathcal{H}^{(t-s)} \). The age of \( e \) is 0. The new edge does not form part of the spine.
- In addition, at the vertex \( v_0(s) \) that is at the top of the spine at time \( s \) there is an extra arrivals process, independent of the previous one: at rate \( \cosech \left( \frac{a(s)}{2} \right) \) a new edge \( e \) arrives joining \( v_0(s) \) to the root of an independent sample of \( \mathcal{H}^{(s)} \). The spine after the arrival is the concatenation of the previous spine, \( e \) and the spine of the new subtree.

A beautiful property of this Markov process is that if we watch the process \( (T_s, \rho_s, a_s) \) from time \( 0 \) up to some time \( s < t \) but we are not shown the spine, then we cannot tell where the spine is, in the following sense: conditioned on \( (T_s, \rho_t, a_s)|_{0,s} \), the spine at time \( s \) is the path from the root to a uniform random vertex of \( T_s \). Only once we reach time \( t \), when the tree almost surely becomes infinite, are we certain where the spine is.

**Theorem 35.** The law of \( C(\cdot) \) on \( [0, t] \) conditioned on \( \theta_1 = t \) is the law of the process \( (T_t, \rho_t, a_t, S_t) \) with the spine forgotten, when it is started at time 0 as a sample of \( \mathcal{H}^{(t)} \). Similarly the law of \( C_1(\cdot) \) on \( [0, t] \) conditioned on \( t_\infty = t \) is the law of the process \( (T_t, \rho_t, a_t, S_t) \) with the spine forgotten, when it is started at time 0 as a singleton of age 0. In particular, \( C_1(t_\infty) \) has the law of \( \hat{H}(0) \).

We now see that the time-reversal of the steady-state cluster growth process can be described entirely in terms of a bi-infinite sequence of independent copies of the random infinite tree \( \hat{H}(0) \). Each copy contains the information necessary to describe one inter-explosion period of \( C(\cdot) \), whose lifetime is given by the age
of the root. As usual, the inter-explosion period containing time 0 must be size-biased. If we are given a sample of $\hat{H}(0)$ with edge ages forgotten, then running time backwards gives an interesting fragmentation or logging process in which there are immediately infinitely many fragments, all finite. In this logging process the edges are cut independently but not all at the same rate: an edge whose end vertices have ages $a$ and $a'$ is cut at rate $1/(a \wedge a')$. The vertex ages all decrease at rate 1, so each edge is almost surely cut strictly before either of its endpoints reaches age 0 and disappears.

Remark. The infinite rooted tree $C_1(t_\infty)$ is a unimodular random tree. Unimodularity is an analog of re-root invariance that applies to possibly infinite random networks. See for example Aldous and Lyons [3] and Benjamini, Lyons and Schramm [5] for the definition and some important properties of unimodularity. $\hat{H}(0)$ is unimodular because it is the Benjamini-Schramm limit, or local weak limit, of $H_x$ as $x \downarrow 0$, and $H_x$ is re-root invariant for $x > 0$. In particular, the law of $\hat{H}(0)$ is invariant under the re-rooting (and consequent change of spine) that is induced by any continuous-time random walk of the root whose jump rates are determined by (possibly random) edge conductances: this property is called involution invariance and is known from [3] to be equivalent to unimodularity. Note that although in $H_x$ the spine joins two independent uniform random vertices, we cannot make sense of choosing two independent uniform vertices in an infinite random tree. The length of the spine of $\hat{H}(x)$ diverges in probability as $x \downarrow 0$.

From Lemma 33, we can deduce the Benjamini-Schramm limit of the steady-state cluster conditioned on $|C| = k$, as $k \to \infty$. We conjecture that this is also the Benjamini-Schramm limit as $n \to \infty$ of the distribution of the fires seen in the long run by a tagged vertex in MFFF($n$). Depending on the lightning rate $\lambda(n)$, the typical fires may in fact have cycles, but the length of the shortest cycle tends to infinity in probability as $n \to \infty$, so the cycles are not captured by the Benjamini-Schramm limit.

**Theorem 36.** The Benjamini-Schramm (or local weak) limit of $C$ conditioned on $|C| = k$ as $k \to \infty$ exists and has the law of $\hat{H}(0)$.

**Proof.** Fix a radius $r \in \mathbb{N}$ and consider the ball $B(\rho,r)$ of radius $r$ about the root vertex in the graph distance. We have to show that restriction of $C$ conditioned on $|C(0)| = k$ to $B(\rho,r)$ converges in distribution as $k \to \infty$ to the restriction of $\hat{H}(0)$ to $B(\rho,r)$. Note that $\hat{H}(x)$ has the law of $C(0)$ conditioned on $\theta_1 = x$, and that it converges in the local weak sense to $\hat{H}(0)$ as $x \downarrow 0$.

To compare these two conditionings, we consider $C(0)$ conditioned on both $|C(0)| = k$ and $\theta_1 = x$, where $x \to 0$ and $k \to \infty$. (One could in fact take $x = k^{-1/2}$, so that $k$ is a typical value for $|\hat{H}(x)|$, but this will not be necessary for our argument). We can perform this double conditioning in either order.

On one hand, once we have conditioned on $|C(0)| = k$, further conditioning on $\theta = x$ does not affect the distribution of $C(0)$ at all, so the doubly-conditioned measure is the law of $C$ conditioned on $|C| = k$.

On the other hand, we claim that conditioning the tree $\hat{H}(x)$ to have exactly $k$ vertices hardly affects the distribution of the ball $B(\rho,r)$ in $\hat{H}(x)$, in the strong sense of total variation distance. To make this precise, for each $r \in \mathbb{N}$ let $U_r$ be the subtree of $\hat{H}(x)$ consisting of all vertices whose most recent spinal ancestor is
at distance less than \( r \) from \( \rho \), let \( s_r \) be the \( r^{th} \) vertex up the spine, if it exists, let \( a_r \) be the age of \( s_r \) and let \( V_r \) be the subtree rooted at \( s_r \). Then \( H^{(x)}(z) = U_r \cup V_r \) and \( B(\rho, r) \subseteq U_r \cup \{s_r\} \).

Now fix \( r \) and let \( \epsilon > 0 \). Choose \( a_{\min} > 0, M < \infty \) and \( x_0 > 0 \) such that for \( x < x_0 \) we have

\[
\mathbb{P}(s_r \text{ exists, } a_r > a_{\min}, \text{ and } |U_r| < M) > 1 - \epsilon.
\]

We have \( |\hat{H}^{(x)}| = |U_r| + |V_r| \). Conditional on the event that \( s_r \) exists and \( a_r = a \), \( U_r \) and \( V_r \) are independent, and \( |V_r| \) has the distribution of \( |H^{(x)}| \) conditioned to have root age \( a \). Further conditioning on \(|U_r| + |V_r| = k\) affects the distribution of \( U_r \) through a tilt that depends only on \(|U_r|\) and \( a_r\):

\[
\frac{\text{density}(a_r = a, U_r = U) |\{U_r| + |V_r| = k\}}{\text{density}(a_r = a, U_r = U)} = \frac{\mathbb{P}(\hat{H}^{(x)} = k | a_r = a, U_r = U)}{\mathbb{P}(\hat{H}^{(x)} = k)} = \frac{\mathbb{P}(\hat{H}^{(x)} = k - |U| | a(\rho) = a)}{\mathbb{P}(\hat{H}^{(x)} = k)}.
\]

Since \( \hat{H}^{(x)} \) has the distribution of \( \mathcal{C}(0) \) conditioned on \( \theta_1 = x \), we have from Corollary 9 that

\[
\mathbb{P}(\hat{H}^{(x)} = k) = 2k \sech^{2k-2} \frac{x}{2} \tanh \frac{a}{2} w_k \sim \frac{\sech^{2k-2} \frac{x}{2} \tanh \frac{a}{2}}{\sqrt{\pi k}} \text{ as } k \to \infty,
\]

uniformly in \( x \). Likewise equation (22) gives

\[
\mathbb{P}(\hat{H}^{(x)} = k | a(\rho) = a) \sim \frac{\sech^{2k} \frac{x}{2} \tanh \frac{a}{2} \sech^2 \frac{a+x}{2} \tanh \frac{a+x}{2}}{\sqrt{\pi k \sech^2 \frac{a+x}{2} \tanh \frac{a+x}{2}}} \text{ as } k \to \infty,
\]

uniformly in \( x > 0 \) and locally uniformly in \( a > 0 \). Hence the density ratio computed above converges to 1 as \( x \to 0 \) and \( k \to \infty \), locally uniformly in \( a \) and \( |U| \).

We have shown that for any \( \epsilon > 0 \), outside a set of measure at most \( \epsilon \), conditioning on \( |\hat{H}^{(x)}| = k \) tilts the joint distribution of \( U_r \) and \( a_r \) by a factor that tends uniformly to 1 as \( x \to 0 \) and \( k \to \infty \). Hence the total variation distance between the distributions of \( B(\rho, r) \) in \( H^{(x)} \) before and after conditioning on \( |\hat{H}^{(x)}| = k \) tends to 0 as \( x \to 0 \) and \( k \to \infty \) together. Since \( H^{(x)} \) converges to \( \hat{H}^{(0)} \) in the Benjamini-Schramm sense as \( x \to 0 \), this implies that the distribution of \( \mathcal{C} \) conditioned on \(|\mathcal{C}| = k \) also converges in the Benjamini-Schramm sense to \( \hat{H}^{(0)} \). ☐

**Remark.** Much is known about size-biased Galton-Watson trees. See Janson’s comprehensive survey [16]. Kesten made sense of size-biasing a critical (single-type) Galton-Watson tree, where the total progeny is almost surely finite but has infinite expectation. The size-biased tree is an infinite rooted tree with one end, described by an infinite spine of vertices whose offspring have the size-biased version of the offspring distribution. See for instance Lyons and Peres [17] Ch. 12. In the special case of the critical binary Galton-Watson tree, the corresponding Kesten tree consists of an infinite spine, each vertex along the spine other than the root has exactly one non-spinal child, which is identified with the root of an independent critical binary Galton-Watson tree. This is a description of \( \mathcal{G}_1(t^-) \), the genealogical tree of \( \mathcal{C}_1 \) at its explosion time.

A multi-type Kesten tree was introduced by Kurtz, Lyons, Pemantle and Peres [14]. Recently Pénisson [18], Stephenson [20] and Abraham et al [1] have studied
the local weak convergence of critical multi-type Galton-Watson trees with finitely many types conditioned to have large total progeny to the corresponding multi-type Kesten tree. Theorem 36 is similar to their results, though not a consequence since $H$ has a continuum of types. Our result is reasonably precise in that the conditioning is on $|H| = k$, rather than on $|H| > k$. In [1] the authors condition on the multi-set of types, which we have not done here.

4.11. Diagonalization of the transfer operator for the multitype Galton-Watson process associated to $H$. In this section it will be convenient to work with a compact space of types, so we will make a change of variable, taking the type of a vertex to be $\tanh(a/2)$ where $a$ is its age. The result of this transformation is that the space of types is $[0, 1]$ and the distribution of the type of the root of $C$ is the uniform distribution on $[0, 1]$.

Now the $k$th generation of the multitype Galton-Watson process is described by a purely atomic measure on the type $[0, 1]$, say $\mu_k$. In this section we study the operator $T$ that sends $\mu_k$ to the expected value of $\mu_k + 1$ given $\mu_k$. Starting at generation 0 with a measure $\mu_0 = \delta_x$, where $x$ is a random type with law $U([0, 1])$, we have

$$E(\mu_k \mid a(\rho) = 2 \arctanh(x)) = T^k(\delta_x).$$

It is convenient to study the corresponding transfer operator $T^*$ acting on $C([0, 1])$, defined by

$$\int T^* f(s)d\mu(s) = \int f(t)dT(\mu)(t)$$

Explicitly, we have

$$T^* f(s) = E \left( \int f(t)d\mu_1(t) \mid \mu_0 = \delta_s \right) = \int_0^1 2 \arctanh(s \wedge t)f(t)dt.$$

Lemma 37. $T^*$ extends to a compact self-adjoint operator on $L^2([0, 1])$, a contraction with simple spectrum $\{1/(n(2n - 1)) : n \in 1, 2, 3, \ldots\}$.

Proof. Note that $T^*$ extends to a bounded operator on $L^2([0, 1])$ defined by the same formula. Indeed by Cauchy-Schwarz the operator norm is bounded by the Hilbert-Schmidt norm, which is finite:

$$\int_0^1 |T^* f(s)|^2ds = \int_0^1 \left( \int_0^1 2 \arctanh(s \wedge t)f(t)dt \right)^2 ds \leq \int_0^1 \left( \int_0^1 4 \arctanh(s \wedge t)^2dt \right) \left( \int_0^1 |f(t)|^2dt \right) ds = \left( \frac{2\pi^2}{3} - 8 \log 2 \right) \|f\|_2^2.$$

Since $T^*$ is bounded, self-adjointness follows from the symmetry of the kernel. Compactness follows from the finiteness of the Hilbert-Schmidt norm.

Next we observe that the operator $T^*$ is diagonalized by the Legendre polynomials of odd degree. From self-adjointness we know that the eigenvectors are orthogonal in $L^2([0, 1])$. It is simple to check that for $n \geq 1$ odd, $T^*(x^n)$ is an odd polynomial of degree $n$ with leading coefficient $1/(n(2n - 1))$. This can be done by
integration by parts and an induction. It follows that the eigenvectors are the odd Legendre polynomials $P_{2n-1}(x)$ for $n = 1, 2, 3, \ldots$, with eigenvalues $1/(n(2n-1))$. The odd Legendre polynomials form an orthogonal basis for $L^2([0,1])$, so we are done. ■

Note that the (unique) invariant probability distribution $\mu^*$ of $T$ is given by the density $2x$. The invariant distribution $\mu^*$ of $T$ corresponds to the age distribution with cdf $\tanh(a/2)^2$, which is is the distribution of $t_\infty$. For any choice of $\mu_0$, $\mathbb{E}(\mu_k)$ converges at exponential rate in the total variation norm to $\left(\frac{3}{4} \int_0^1 x \, d\mu_0(x)\right)^n$. In fact this convergence is uniform over probability measures $\mu_0$, because the Legendre polynomials normalized by $P_n(1) = 1$ are uniformly bounded by 1 in the sup norm on $[-1,1]$. In this normalization we have

$$\int_0^1 P_{2i-1}(x) \, dx = (-1)^{i+1} w_i \quad \text{and} \quad \int_0^1 P_{2i-1}(x) P_{2j-1}(x) \, dx = \delta_{ij}/(4i-1).$$

Thus when $\mu_0$ is the Lebesgue measure on $[0,1]$, the distribution of the type of the root of $C$, by expanding the density of $\mu_0$ in the odd Legendre polynomials we obtain the following expression for the density of $\mathbb{E}(\mu_k)$:

$$d\mathbb{E}(\mu_k)/d\mu_0(x) = \sum_{i=1}^{\infty} (4i-1)(-1)^{i+1} w_i P_{2i-1}(x)/(i(2i-1))^k.$$

Hence the expected number of vertices at distance exactly $k$ from the root in $C$ is given by

$$\mathbb{E}(\mu_k) = \sum_{i=1}^{\infty} \frac{(4i-1)w_i^2}{i(2i-1)^k},$$

from which we find that $\frac{3}{4} < \mathbb{E}(\mu_k) \leq \frac{3}{4} + \frac{1}{6} 6^{-k}$ for every $k \geq 0$.

The transfer operator for the spinal Markov chain of $\hat{H}(0)$ is very closely related to $T^*$. In fact it is the conjugate of $T^*$ by the map $f \mapsto x f$, meaning that it is given by

$$f \mapsto x T^*(f/x).$$

Hence the eigenvectors for $T^*$ are the polynomials $x P_{2n-1}(x)$, with corresponding eigenvalues $1/(n(2n-1))$ for $n = 1, 2, 3, \ldots$. The first of these gives us the invariant distribution for the spinal Markov chain, which in terms of the transformed type space $[0,1]$ has the density $3x^2$.

5. Construction of an infinite forest fire model

Although the connection between the steady state cluster process and the Ráth-Tóth mean field forest fire model is currently only heuristic, it is possible to view the steady state cluster growth process as the process of the cluster of a tagged vertex in an infinite forest fire model, which we will construct in this section. It will have the property that a cluster burns at the moment it becomes infinite; there is no need for a lightning process. The model is defined on an infinite tree. One can think of the fires as being ignited at the ideal boundary of this tree.

The Ráth-Tóth model is exchangeable, but that is too much to ask for in a countably infinite forest fire model. If such a model were exchangeable then the probability that any particular pair of vertices become joined by an edge during any fixed time interval must be zero, since the expected number of edges arriving
at any vertex in that time interval is finite. But then since there are only countably
many possible edges, almost surely no edges would arrive at all! Instead of being
exchangeable, our infinite forest fire model can be thought of as a candidate for
a local weak limit of $\text{MFFF}(n)$ in its stationary state, as $n \to \infty$. As Aldous
and Steele discuss in [4], local weak limits of exchangeable models are involution
invariant, or unimodular. We have already seen that $\hat{H}$ has this property, and
the infinite forest fire model that we construct here will also possess it.

We will construct our model on a labelled infinite tree $Z$, whose vertices are
labelled by the set $Z^*$ of finite strings of integers. This labelling by strings is a
convenient scaffolding for the construction, which we will eventually erase in order
to obtain a stationary process. If $v \in Z^*$ and $n \in \mathbb{Z}$ then we denote by $v.n$ the string
obtained by appending $n$ to the string $v$. The edges of $Z$ are the pairs $(v, v.n)$ for
$v \in Z^*$ and $n \in \mathbb{Z}$. The tree $Z$ is rooted at the empty string, which we denote $\emptyset$.
We denote by $|v|$ the length of the string, so that $|v|$ is the height of $v$ above the
root $\emptyset$. We denote the subtree of $Z$ rooted at $v$ by $v. Z$. This consists of all vertices
whose strings have $v$ as a prefix. The induced subgraph $\mathcal{N}$ of $Z$ on the vertex set
$\mathbb{N}^*$ is called the Harris-Ulam-Neveu tree.

Our forest fire model will be a càdlàg process $\text{FF}_t : t \in [0, \infty)$. A state of
$\text{FF}_t$ consists of a spanning subgraph of $Z$, (necessarily a forest), together with an
assignment of an age $a(e) \in \mathbb{R}$ to every edge $e$, and an age $a(v) \in [0, \infty)$ to every
vertex $v$. We let $\text{Live}_t$ denote the spanning forest consisting of all the edges of
$\text{FF}_t$ that have non-negative age. We call these the live edges. We let $\text{Future}_t$
denote the spanning forest consisting of all the edges of $\text{FF}_t$ that have negative age.
We call these the future edges. We insist that for every vertex $v$, the set
of $n \in \mathbb{Z}$ such that $(v, v.n)$ belongs to $\text{FF}_t$ must be an interval of $Z$ of the form
$Z_{\geq m} = \{n \in \mathbb{Z} | n \geq m\}$ for some $m$. Thus as a rooted plane tree $\text{FF}_t$ is always
isomorphic to $\mathcal{N}$. For each vertex $v$ the sequence of ages of the edges to its children
must be strictly decreasing: if $m < n$ and $(v, v.m)$ is an edge of $\text{FF}_t$ then $(v, v.n)$ is
also an edge of $\text{FF}_t$ and $a((v, v.m)) > a((v, v.n))$. For every $t \geq 0$, conditional on
the subgraph $\text{Live}_t$, for each vertex $v \in Z^*$ the sequence of the ages of child edges
$(v, v.n)$ in $\text{Future}_t$ forms a Poisson point process of unit intensity on $(-\infty, 0)$, and
these processes are independent. At time 0, the graph $\text{Future}_0$ is defined to be $\mathcal{N}$.
The reader may recognise that $\text{Future}_0$ is isomorphic to Aldous’ Poisson-weighted
infinite tree, where the edge weights play the rôle of negative ages.

Next, we specify how to sample from the distribution of $\text{Live}_0$. Sample inde-
dependent copies $H_v$ of the multitype Galton-Watson tree $H$, one for each of the
countably many vertices $v$ of $Z$. We will define inductively an embedding $\phi$ of the
disjoint union of some of these trees to make a spanning forest of $Z$. Suppose the
components in $\text{Live}_0$ of all the vertices up to height $h$ have been specified; when
we begin, $h = -1$. Then for each vertex $v$ at height $h + 1$ that does not belong to
any of the components already specified, we extend $\phi$ by mapping the root of $H_v$
to $v$ and then for each vertex $x \in H_v$ with children $w_1, \ldots, w_k$ ordered so that the
age in $H_v$ satisfy $a((x, w_1)) < \cdots < a((x, w_k))$, we define $\phi(w_i) = \phi(x),(-i)$
for $i = 1, \ldots, k$.

Informally, the process $\text{FF}_t$ evolves as follows. The age of each edge of $\text{FF}_t$
and the age of each vertex increases at rate 1. Let $e$ be any edge of $Z$. At some
random time $\tau(e)$ such that $a_{\tau(e)}(e) \geq 0$, the component of $\text{Live}_t$ that contains $e$
may explode. When this happens, all edges in this component are instantaneously
deleted from $\mathbf{FF}_i$, and all vertices in the component have their ages reset to 0. The vertices are not deleted. We refer to such a deletion event as a fire.

Warning: we have not yet defined a process! The previous paragraph apparently describes a deterministic evolution - so all the randomness should be in the sampling of the initial state. However, it is not at all clear that the fire times $\tau(v)$ can all be defined as functions that are measurable with respect to the $\sigma$-algebra generated by $\mathbf{FF}_0$. If it were true that the initial state almost surely determines the fire times uniquely, then the process would be called endogenous, and we would already have specified a stochastic process (though some work would be required to prove this). However, we are not currently able to resolve this endogeny problem, so instead we will give a rigorous construction of a process that does satisfy the description above but is defined on a larger sigma algebra than the one generated by $\mathbf{FF}_0$.

Although it is clear that any process $\mathbf{FF}_i$ meeting the above informal description is not stationary, we will construct such a process with the property that it becomes stationary once the labelling of the vertices by $\mathbb{Z}^+$ is forgotten, and only the rooted plane graph structure and the age labelling are retained. After forgetting the labelling by strings, the process $\text{Live}_i$ will be a candidate for the local weak limit of the Ráth-Tóth forest fire model $\mathbf{MFFF}(n)$ in its stationary state, as $n \to \infty$.

5.1. Rigorous forward construction of a version of $\mathbf{FF}_i$. In order to construct the process meeting the above description of $\mathbf{FF}_i$, we will use Kolmogorov’s consistency theorem. For each $h \in \mathbb{N}$ we will construct a forest fire process $\mathbf{FF}_i^h$ on the truncation $Z_h$ of $Z$, which is the induced subgraph on strings of length at most $h$. We will show that for $h' < h$, the restriction of $\mathbf{FF}_i^h$ to $Z_{h'}$ is identical in law to $\mathbf{FF}_i^{h'}$. It will follow that there exists a process $\mathbf{FF}_i^\infty$, taking values in the age-decorated spanning forests of $Z$, whose restriction to $Z_h$ is identical in law to $\mathbf{FF}_i^h$ for all $h$. Finally, we will check that $\mathbf{FF}_i^\infty$ has the law of $\mathbf{FF}_i^0$ and that the evolution of $\mathbf{FF}_i^\infty$ agrees with the informal description given above of the evolution of $\mathbf{FF}_i$.

To construct $\mathbf{FF}_i^h$, we define $\mathbf{FF}_0^h$ to be the restriction of $\text{Live}_0 \cup \text{Future}_0$ (as defined above) to the truncated tree $Z_h$. We will denote the age labelling in $\mathbf{FF}_i^h$ by $a$. Thus $\mathbf{FF}_0^h$ is a spanning forest of $Z_h$ whose vertices and edges are labelled with ages. Moreover, the initial states $\mathbf{FF}_0^h$ for different $h$ are coupled together, related by restriction. In addition, independently of $\mathbf{FF}_0^h$ we equip each leaf $v$ of $Z_h$ with a Poisson point process $\mathcal{P}_0(v)$ on $[0, \infty)^2$ with intensity at $(t, y)$ given by

$$1(y < t + a_0(v)) \frac{1}{2} \text{sech}^2 \frac{y}{2} dt dy.$$  

Conditional on the ages of the leaves, the Poisson point processes $\mathcal{P}_0(v)$ are independent. A point $(t, y)$ in $\mathcal{P}_0(v)$ represents the potential ignition of a fire at the leaf $v$ at time $t$. It will also be possible for the points to be deleted before they ignite a fire. For the moment the reader may think of the point process $\mathcal{P}_0(v)$ simply as a device to encode a random ignition process at each leaf whose rate depends on the age of the leaf as part of the initial data. Later we will see that the $y$ co-ordinate plays the role of the arrival time of an edge from $v$ to a child of $v$.

Given these initial data, the evolution of $\mathbf{FF}_i^h$ is deterministic. Each edge age $a_e(t)$ and each vertex age $a_i(v)$ increases at rate 1. As above we partition the edges of $\mathbf{FF}_i^h$ into $\text{Live}_i^h$ and $\text{Future}_i^h$, according to the sign of $a_e(t)$. Thus, edges move from $\text{Future}_i^h$ to $\text{Live}_i^h$ when their age passes 0.
For each leaf $v$ of $Z_h$, the first co-ordinate of each point of each Poisson point process $\mathcal{P}_t(v)$ decreases at rate 1. When any of these points reaches the boundary so that it is a point $(0, y) \in \mathcal{P}_t(v)$, the vertex $v$ is ignited. This means the following:

- $(0, y)$ is deleted so that it is not present in $\mathcal{P}_t(v)$ (recall we want to define a càdlàg process),
- the edges in the connected component of $v$ in $\text{Live}^h_t$ are deleted from $\mathcal{F}^h_t$ to obtain $\mathcal{F}^h_t$,
- the vertex ages of all the vertices in that component (including $v$ itself) are reset to be 0 at time $t$,
- for each vertex $w$ at height $h$ in that component, the Poisson point process $\mathcal{P}_t(w)$ is obtained by removing all points $(t, y)$ such that $y \geq t$ from $\mathcal{P}_t(w)$. 

Now the burning times of the edges are well-defined measurable functions of the initial data. Indeed, even if we suppressed the ignitions, no infinite connected component could form in $\text{Live}^h_t$, since this would require some vertex to have infinitely many live edges at some finite time, which almost surely does not occur. Therefore the event that any given edge $e$ burns before a given time $t$ almost surely depends on only finitely many of the ignition processes $\mathcal{P}_t(v)$, and hence on finitely many edge arrival times and potential ignitions. It follows that it is measurable. Almost surely each edge does eventually burn in $\mathcal{F}^h_t$.

For the consistency argument, we have to prove that for each $h \geq 1$ the restriction of $\mathcal{FF}^h_t$ to $Z_{h-1}$ is identically distributed to $\mathcal{FF}^{h-1}_t$. For each edge $e = (v, v, n)$ in $\mathcal{FF}^h_t$ where $|v| = h - 1$, we define the arrival time $\alpha(e)$ of $e$ to be $-\alpha_0(e)$, and the a priori burning time $\Theta(e)$ of $e = (v, v, n)$ to be

$$\Theta(e) = \min\{t : (0, y) \in \mathcal{P}_t(v, n), t \geq \alpha(e)\}.$$ 

This is the unique time at which an ignition at the leaf $v, n$ could possibly cause a fire in which the edge $e$ is burned. However, it is also possible for the edge $e$ to be burned earlier than this time, if $v$ is burned in a fire ignited at a different child leaf of $v$ at some time in $[\alpha(e), \Theta(e))$. Either way, $e$ does not belong to $\mathcal{FF}^h_t$ for any time $t > \Theta(e)$. We define $\mathcal{P}_t(v)$ to be the point process

$$\mathcal{P}_t(v) = \{(\Theta(e) - t, \Theta(e) - \alpha(e)) : e = (v, v, n) \in \mathcal{FF}^h_t, |v| = h - 1\}.$$ 

The first co-ordinate of each points in $\mathcal{P}_0(v)$ represents the waiting time until a time at which the vertex $v$ could potentially burn due to an ignition at one of its children, given complete information about the arrival times of the edges $(v, v, n)$ and the processes $\mathcal{P}_0(v, n)$ of potential ignitions at the children. The second co-ordinate in each point enables us to enforce the constraint on the age of $v$ at each of these potential burning times that comes from the fact that a live edge incident on $v$ could not have survived through the time at which $v$ previously burned. Nevertheless, each of the points in $\mathcal{P}(v)$ still only represents a potential burning time, because $v$ and the live edges to its children could be burned earlier by a fire transmitted to $v$ from the parent of $v$.

After conditioning on the ages of the vertices at height $h - 1$ in $\mathcal{FF}^h_0$ (which are mutually dependent), the processes $\mathcal{P}_0(v) : |v| = h - 1$ are conditionally independent, as they are functions of disjoint sets of conditionally independent variables. We also need to compute the distribution of $\mathcal{P}_t(v)$ to complete the consistency argument.
Lemma 38. $\tilde{P}_0(v)$ is a Poisson point process on $[0,\infty)^2$ with intensity

$$1(y < a_0(v) + t) \frac{1}{2} \text{sech}^2 \frac{y}{2} \, dt \, dy.$$ 

Proof. Consider an edge $e = (v, v.n)$ that belongs to $\mathbf{FF}_0^h$, where $v$ is at height $h - 1$. There are two cases to consider: $e \in \mathbf{Live}_0^h$ and $e \in \mathbf{Future}_0^h$.

Suppose $e \in \mathbf{Future}_0^h$. We compute the distribution of $a_{\alpha(e)}(v.n)$. Before the edge $e$ becomes live at time $\alpha(e) = -a_0(e)$ (which is positive), the burning times of $v.n$ follow a stationary renewal process with interarrival times distributed like $t_\infty$. This is because the initial age $a_0(v.n)$ is the root age of $H_{v.n}$, which has the distribution of $\theta_1$, and this is the stationary distribution of the spent time of the renewal process. The hazard function of the renewal process when the spent time is $a$ is tanh $\frac{a}{2}$, which is $\int_0^a \frac{1}{2} \text{sech}^2 \frac{y}{2} \, dy$, so the renewals are correctly described by the actual ignitions implied by $\mathcal{P}(v.n)$. Hence conditional on $\alpha(e)$, the age $a_{\alpha(e)}(v.n)$ is also distributed like $\theta_1$.

Claim: conditional on all the initial data apart from $\mathcal{P}(v.n)$, $\Theta(e) - \alpha(e)$ is distributed like $\theta_1$. We have

$$\mathbb{P}(\Theta(e) > \alpha + s \mid \alpha(e) = \alpha, a_{\alpha}(v.n) = a) = \mathbb{P}(\mathcal{P}(v.n) \cap \Delta_{\alpha,s} = \emptyset),$$

where

$$\Delta_{\alpha,s} = \{(t, y) : \alpha \leq t < \alpha + s, y < t - \alpha - a\}.$$

This probability is

$$\exp \left(- \int_0^s \int_0^{u+a} \frac{1}{2} \text{sech}^2 \frac{y}{2} \, dy \, du \right) = \exp \left(- \int_0^s \tanh \frac{u+a}{2} \, du \right) = \text{sech}^2 \frac{a+s}{2} \cosh^2 \frac{s}{2}.$$

Here we used the fact that $a_t(v.n) \leq t + a_0(v.n)$ to see that the indicator function $1(y < t+a_0(v.n))$ in the density of $\mathcal{P}(v.n)$ is identically 1 on $\Delta_{\alpha,s}$. Now we convolve with the distribution of $a_{\alpha(e)}(v.n)$ to obtain the distribution of $\Theta(e)$:

$$\mathbb{P}(\Theta(e) > \alpha(e) + s) = \int_0^\infty \text{sech}^2 \frac{a+s}{2} \cosh^2 \frac{s}{2} \cdot \frac{1}{2} \text{sech}^2 \frac{a}{2} \, da = 1 - \tanh \frac{s}{2}.$$

Hence $\Theta(e) - \alpha(e) = \Theta(e) + a_0(e)$ is distributed like $\theta_1$, as claimed.

We must also consider the edges from $v$ to its children that are already present in $\mathbf{Live}_0^h$. We condition on $\mathbf{FF}_0^{h-1}$, and consider a leaf $v$ of $\mathcal{Z}_{h-1}$, with initial age $a_0(v)$. The ages of the offspring of $v$ in the spanning forest $\mathbf{Live}_0^h$ and the ages of the corresponding edges are described by a PRM of pairs $(a_v, a_e) \in [0,\infty)^2$ with density

$$1(0 \leq a_e \leq a_0(v) \land a_v) \cdot \frac{1}{2} \text{sech}^2 \frac{a_v}{2} \, da_v \, da_e.$$

Each offspring of $v$ in $\mathbf{Live}_0^h$ is a leaf $v.n$ of $\mathcal{Z}_h$ with its own Poisson process $\mathcal{P}(v.n)$ of potential ignitions. Conditional on all the initial data apart from $\mathcal{P}(v.n)$, the a priori burning time of the edge $e = (v, v.n)$ is distributed as follows:

$$\mathbb{P}(\Theta(e) > s \mid e \in \mathbf{Live}_0^h, a_0(v.n) = a) = \mathbb{P}(\mathcal{P}(v.n) \cap \Delta_{a_0,s} = \emptyset),$$

which we computed above to be $\text{sech}^2 \frac{a+s}{2} \cosh^2 \frac{s}{2}$. Differentiating this with respect to $s$ gives the conditional density of $\Theta(e)$. Integrating out the leaf vertex age, we
find that the pairs \((a_0(e), \Theta(e))\) for edges \(e\) of the form \((v, v, n)\) in \(\text{Live}_0^v\) are the atoms of a PRM whose intensity is

\[
\left(\int_0^\infty \frac{1}{2} \sech^2 \frac{a_0 + \Theta_e}{2} \tanh \frac{a_0 + \Theta_e}{2} 1(0 \leq a_e \leq a_e \wedge a_0(v)) \, da_e \right) \, d\Theta_e.
\]

This density simplifies to

\[1(0 \leq a_e < a_0(v)) \cdot \frac{1}{2} \sech^2 \frac{a_0 + \Theta_e}{2} \, da_e \, d\Theta_e.
\]

Combining the \((a_0(e), \Theta(e))\) pairs for the edges \(e\) from \(v\) to height \(h\) in \(\text{Live}_0^v\) and \(\text{Future}_0^v\), we find that they form a Poisson point process on \(\mathbb{R}^2\) with intensity

\[1(a_e < a_0(v), a_e + \Theta_e > 0, \Theta_e > 0) \cdot \frac{1}{2} \sech^2 \left(\frac{a_0 + \Theta_e}{2}\right) \, da_e \, d\Theta_e.
\]

Changing variables to \((\Theta_e, \Theta_e + a_e)\) we find that \(\tilde{P}(v)\) is a PRM with intensity

\[1(y < a_0(v) + t) \cdot \frac{1}{2} \sech^2 \left(\frac{y}{2}\right) \, dt \, dy,
\]

as required. \(\blacksquare\)

Lemma 38 shows that the laws of \(\text{FF}_t^h\) are compatible as \(h\) varies. In particular we obtain a sample of \(\text{FF}_t^{h-1}\) from a sample of \(\text{FF}_t^h\) by restricting to \(Z_{h-1}\) and letting the process \(\tilde{P}(v)\) serve as \(P(v)\). The construction causes the (deterministic) evolutions to stay coupled in the sense that the restriction of \(\text{FF}_t^h\) to \(Z_{h-1}\) is \(\text{FF}_t^{h-1}\) for all \(t \geq 0\). Applying Kolmogorov’s consistency theorem, we find that there exists an essentially unique random process \(\text{FF}_t^\infty\) on the whole tree \(\mathcal{Z}\) whose truncation to \(Z_h\) has the law of the process \(\text{FF}_h^h\) for every \(h\). In \(\text{FF}_t^\infty\) there is no explicit ignition process analogous to the processes \(P(v)\), but we have not shown that the evolution of \(\text{FF}_t^\infty\) is deterministic, i.e. that \(\text{FF}_t^\infty\) is a measurable function of \(\text{FF}_0^\infty\).

### 5.2. Steady state cluster growth in \(\text{FF}_t^\infty\).

**Theorem 39.** The connected component of the root in \(\text{FF}_t^\infty\) is stationary up to order-preserving relabelling of the vertices by strings. The live cluster of the root in \(\text{FF}_t^\infty\) is a version of the stationary multitype Galton-Watson tree process \(\mathcal{H}(t)\).

By the connected component of the root in \(\text{FF}_t^\infty\), we mean the connected component using both live and future edges, but not burned edges. By stationary up to order-preserving relabelling we mean that if we forget the labelling of the vertices by strings but retain the rooted plane tree structure of \(\mathcal{Z}\) and the labelling of vertices and edges by ages, then we obtain a stationary process.

**Proof.** \(\text{FF}_t^\infty\) is always isomorphic as a rooted plane tree to \(\mathcal{N}\). The restriction of \(\text{Live}_0^\infty\) to \(\text{FF}_t^\infty\) is a spanning forest whose connected components are finite at all times. The restriction of \(\text{FF}_t^\infty\) to the connected component of the root is a forest fire process in which when an edge burns it is deleted along with the entire subtree of \(\text{FF}_t^\infty\) that it cuts off from the root. For stationarity, we have to show that the restriction of \(\text{Live}_0^\infty\) to the connected component of the root in \(\text{FF}_t^\infty\) is stationary. This suffices because after dropping the string-labelling the process \(\text{Future}_t\) is stationary and for each fixed time \(t\), \(\text{Future}_t\) is independent of \((\text{Live}_s)_{s \in [0, t]}\).

For each \(v \in \mathcal{Z}\) denote by \(\mathcal{L}_t(v)\) the cluster of \(v\) in \(\text{Live}_t\). Denote by \(\text{pa}(v)\) the parent of \(v\) in \(\mathcal{Z}\). From the construction (and in particular the uniqueness part of Kolmogorov’s consistency theorem), we see that for each \(v \in \mathcal{Z}\) other than the root, the conditional distribution of \(\mathcal{L}_t(v)\) given \(\text{pa}(\text{pa}(v), v)) > t\) is the distribution of \(\mathcal{L}_t(\emptyset)\). Hence \(\mathcal{L}_t(\emptyset)\) is a geometric cluster growth process in a (possibly dynamic) environment of clusters identically distributed to \(\mathcal{L}_t(\emptyset)\). We already know from
that the steady state cluster growth process is a steady state solution of this problem, and we are starting at time 0 with \( \mathcal{L}_0(\emptyset) \) distributed like the steady state cluster. But the question of whether this recursive property of \( \mathcal{L}_t(\emptyset) \) forces it to be stationary is a chicken-and-egg problem which we have not solved.

Instead we will prove only that \( \mathbf{FF}_t^\infty \) is stationary, by showing inductively that for each \( h \geq 0 \), the live cluster of the root in \( \mathbf{FF}_t^h \) is a stationary process, having the distribution of the truncation of \( \mathcal{H}(t) \) at height \( h \). This will also imply the second statement of the theorem, by the uniqueness part of Kolmogorov’s consistency theorem. Denote the live cluster of \( v \) in \( \mathbf{FF}_t^h \) by \( \mathcal{L}_t^h(v) \). By construction, for \( \emptyset \neq v \in \mathcal{Z}_h \) the conditional distribution of \( \mathcal{L}_t^h(v) \) given that \( \alpha(\text{pa}(v),v) > t \) is equal to the distribution of \( \mathcal{L}_t^{h-|v|}(\emptyset) \).

To begin the induction, \( \mathbf{FF}_t^0 \) consists of the root vertex, labelled by an age \( a_t(\emptyset) \) and equipped with an ignition process \( \mathcal{P}_t(\emptyset) \). As remarked in the construction, the resulting ignitions are the arrival times of a stationary renewal process whose interarrival time is distributed like \( t_\infty \) and \( a_t(\emptyset) \) records the spent time since the last ignition; the conditional distribution of \( \mathcal{P}_t(\emptyset) \) given \( a_t(\emptyset) \) is a Poisson point process that is conditionally independent of the renewal process before the last ignition. Hence \( \mathcal{L}_t^0(\emptyset) \) is stationary with the distribution of the truncation to height 0 of \( \mathcal{H}(t) \), as required.

The induction step is simple: consider the dynamics of the truncation \( \mathcal{H}^h(t) \) of \( \mathcal{H}(t) \) at height \( h \), which is stationary by its construction. At time 0 it is distributed like the truncation \( H^h \) of \( H \) at height \( h \). At rate 1 an edge arrives joining the root to a new child, whose subtree is distributed like \( H^{h-1} \) and subsequently evolves like \( \mathcal{H}^{h-1}(t) \). The first of these child subtrees to burn determines the burning time of the root, at which time \( \mathcal{H}^h(t) \) becomes a singleton, and then continues to evolve in the same way. Likewise \( \mathcal{L}_t^0(\emptyset) \) is distributed like \( H^h \) and subsequently at rate 1 an edge arrives joining the root to a new child, whose subtree is distributed like \( \mathcal{L}_t^{h-1} \); the first of these child subtrees to burn determines the next burning time of the root, at which time \( \mathcal{L}_t^h(\emptyset) \) becomes a singleton, and then continues to evolve in the same way. The induction hypothesis ensures that these descriptions agree, and we are done.

5.3. A stationary forest fire model on \( \mathcal{Z} \). Because of the relatively simple way in which we sampled \( \textbf{Live}_0 \), the whole infinite forest fire model \( \mathbf{FF}_t^\infty \) is not a stationary process. Even if we drop the labelling of the vertices by strings and think of \( \mathcal{Z} \) as an infinite rooted plane tree, (meaning that we remember the ordering of the children of each vertex), we do not obtain a stationary process. Indeed, in \( \mathbf{FF}_0^\infty \) almost surely all the vertices have distinct ages. However, in \( \mathbf{FF}_t^\infty \) at any positive time there have already been fires, and infinitely many of the vertices involved in any one of those fires will have survived unburned afterwards until time \( t \); so there will be infinite sets of vertices sharing the same age.

For a richer stationary forest fire model in which vertices are not deleted at burning times, we have to modify the construction of the initial state to fill in the past. We will only sketch the construction and will not give a proof of stationarity.

In the following description, all samplings are independent unless otherwise stated. The goal is to construct a rooted plane tree isomorphic to \( \mathcal{Z} \) by gluing countably infinitely many independent copies of \( \hat{H}^{(0)} \) representing the fires. We glue by identifying vertices. We begin by constructing the past and future fires of the root. To do this, we sample a bi-infinite i.i.d. sequence of one-ended infinite
rooted plane trees \( \text{Fire}_i(\emptyset) : i \in \mathbb{Z} \), each distributed like \( \hat{H}(0) \), except that \( \text{Fire}_0(\emptyset) \) is size-biased in proportion to the age of the root. We then sample the initial root age \( a_0(\emptyset) \) by sampling from \( U([0, a]) \) where \( a \) is the age of the root in \( \text{Fire}_0(\emptyset) \). This determines the time of the first fire at the root after time 0: it is \( \theta_1(\emptyset) := a - a_0(\emptyset) \). The time of the last fire of the root before time 0 is \( \theta_0(\emptyset) := -a_0(\emptyset) \). Each tree \( \text{Fire}_i(\emptyset) \) has its own root age, and these ages allow us to construct the complete sequence of times \( \theta_i(\emptyset) : i \in \mathbb{Z} \) at which the root burns, so that \( \theta_{i+1}(\emptyset) - \theta_i(\emptyset) \) is the age of the root in \( \text{Fire}_i(\emptyset) \). The sequence \( (\theta_i(\emptyset))_{i \in \mathbb{Z}} \) has the distribution of a stationary renewal process extending infinitely into the past and the future. We glue all of the rooted plane trees \( \text{Fire}_i(\emptyset) \) together, in order, at their roots to create an infinite rooted plane tree, whose root we label with the empty string \( \emptyset \). The root has edges with the order type of the integers and every other vertex has only finite degree. Each edge \( e \) in \( \text{Fire}_i(\emptyset) \) is equipped with an age and hence acquires an arrival time \( \alpha(e) \) by subtracting this age from \( \theta_{i+1}(\emptyset) \). Therefore we can label the children of the root by integers, in order of the arrival times of their edges to the root. We do this so that the first child whose edge to the root has non-negative arrival time has label 0.

We have now constructed all of the edges incident on \( \emptyset \) and all of the vertices at height 0 or 1. We now continue the construction inductively with respect to height. Suppose we have constructed all of the edges and vertices up to height \( h \). For each vertex \( v \) at height \( h \) we already have a unique fire in which the vertex \( v \) was constructed: we let \( \text{Fire}_0(v) \) be this fire. This provides a pair of consecutive fire times of \( v \): the burning time of \( \text{Fire}_0(v) \) and the preceding burning time of \( v \). We sample a bi-infinite i.i.d. sequence of fires \( \text{Fire}_i(v) : i \in \mathbb{Z} \setminus \{0\} \), each distributed like \( \hat{H}(0) \). We identify the root of each \( \text{Fire}_i(v) \), \( (i \neq 0) \), with \( v \). The root ages give the inter-fire times of \( v \) in the past and the future, and therefore yield the complete sequence of burning times of \( v \). Now we have constructed all the children \( v.n \) of \( v \), ordered by the time of arrival of the edge \( (v, v.n) \), with the children labelled so that the edge to \( v.0 \) is the first one to arrive after time 0.

Once the inductive construction is finished, we have constructed a copy of \( Z \) in which every vertex has a bi-infinite sequence of burning times distributed like a stationary renewal process with interarrival times distributed like \( t_\infty \). Every edge \( e \) has an arrival time \( \alpha(e) \) and a burning time \( \theta(e) \). The edge set of \( Z \) is partitioned into countably infinitely many fires; each of these is a one-ended infinite rooted plane tree consisting of edges all of which have the same burning time.

At any time \( t \) we can partition the edge set into three spanning forests:

- **Future** \( _{t} \), the spanning forest consisting of all edges \( e \) such that \( \alpha(e) > t \),
- **Live** \( _{t} \), consisting of all edges \( e \) such that \( \alpha(e) \leq t < \theta(e) \), and
- **Burned** \( _{t} \), consisting of all edges \( e \) such that \( \theta(e) \leq e \).

At each time \( t \) we can label each vertex with an age,

\[
a_0(v) := t - \sup((-\infty, t] \cap \{\theta_i(v) : i \in \mathbb{Z}\})
\]

Likewise we can label each edge of \( \text{Live}_t \) with an age,

\[
a_0(e) := t - \alpha(e)
\]

When we forget the labelling of vertices by strings, but keep the rooted plane tree structure of \( Z \), the spanning forest \( \text{Live}_t \) labelled with vertex and edge ages constitutes a stationary process. This follows easily from the fact that we started with a stationary renewal process of fires at the root.
Also by construction the connected component of the root $\emptyset$ in $\text{Live}_t$ is a version of the dynamic multi-type Galton-Watson tree $\mathcal{H}(t)$, in other words a version of the steady state cluster growth process.

What is the dynamics of the stochastic process $\text{Live}_t$? At each vertex $v$, the arrival times of the edges to the children of $v$ form a Poisson process of rate 1; these arrival processes are independent at distinct vertices, and the arrivals are independent of $\text{Live}_t$. Each edge burns just once, at the time when its connected component in $\text{Live}_t$ explodes.

By considering only the connected component of $\emptyset$ in $\text{Live}_t \cup \text{Future}_t$, we obtain a stationary forest fire process in which both edges and vertices are deleted when they burn.

**References**

[1] Abraham, R., Delmas, J.-F., and Guo, H., *Critical Multi-Type Galton-Watson trees conditioned to be large*, to appear in J. of Th. Probab. (2018), [arXiv:1511.01721](https://arxiv.org/abs/1511.01721).

[2] Aldous, D., *The percolation process on a tree where infinite clusters are frozen*, Math. Proc. Cam. Phil. Soc., Vol. 128 (2000), no. 3, 465–477.

[3] Aldous, D. and Lyons, R., *Processes on unimodular random networks*, Electron. J. Probab., 12 (2007), 1454–1508.

[4] Aldous, D. and Steele, J. M., *The objective method: probabilistic combinatorial optimization and local weak convergence*, in *Probability on Discrete Structures*, Encyclopedia of Mathematical Sciences 110, (H. Kesten, ed.), 1–72, Springer (Berlin). MR2023650.

[5] Benjamini, I., Lyons, R. and Schramm, O., *Unimodular random trees*, Ergodic Theory and Dynamical Systems 35 (2015) no. 2, 359–373.

[6] Bertoin, J., *Lévy Processes*, Cambridge University Press (1996).

[7] Bertoin, J., *Random fragmentation and coagulation processes*, Cambridge University Press (2006).

[8] Bertoin, J. and Yor, M., *Exponential functionals of Lévy processes*, Probability Surveys 2 (2005), 191–212.

[9] Carmona, P., Petit, F. and Yor, M., *Exponential functionals of Lévy processes, in Lévy processes: theory and applications*, Barndorff-Nielsen, O.E. et al, eds., Birkaüser (2001), 41–55.

[10] Crane, E., Freeman, N., and Tóth, B., *Cluster growth in the dynamical Erdős-Rényi process with forest fires*, Electron. J. Probab. 20 (2015), no. 101, 33p, [arXiv:1405.5044](https://arxiv.org/abs/1405.5044).

[11] Flajolet, P., and Odlyzko, A., *Singularity analysis of generating functions*, SIAM J. Discrete Math. 3 (1990), no. 2, 216–240.

[12] Haas, B., *Scaling limits of Markov-branching trees and applications*, [arXiv:1605.07873](https://arxiv.org/abs/1605.07873).

[13] Harris, T., *The Theory of Branching Processes*, Grundlehren der Math. Wiss., Springer (1963).

[14] Kurtz, T., Lyons, R., Pemantle, R. and Peres, Y., *A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes, in Classical and modern branching processes (Minneapolis, 1994)*, vol. 84 of IMA Vol. Math. Appl., Springer (1997), 181–185.

[15] Kyprianou, A., *Introductionary lectures on fluctuations of Lévy process with applications*, Springer Universitext (2006).

[16] Janson, S., *Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation*, Probability Surveys Vol. 9 (2012), 103–252.

[17] Lyons, R. and Peres, Y., *Probability on trees and networks*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press (2016).

[18] Pénisson, S., *Beyond the $Q$-process: various ways of conditioning the multitype Galton-Watson process*, ALEA, Lat. Am. J. Probab. Math. Stat., Vol. 13 no. 1, 223–237.

[19] Ráth, B., and Tóth, B., *Erdős-Rényi random graphs + forest fires = self-organized criticality*, Electron. J. Probab. 14 (2009), no. 45, 1290–1327.

[20] Stephenson, R., *Local convergence of large critical multi-type Galton-Watson trees and applications to random maps*, arXiv 1412.6911
HEILBRONN INSTITUTE FOR MATHEMATICAL RESEARCH, UNIVERSITY OF BRISTOL, BS8 1TW, UNITED KINGDOM

E-mail address: edward.crane@bristol.ac.uk