

Regenerative tree growth: structural results and convergence*

Jim Pitman[†]

Douglas Rizzolo[†]

Matthias Winkel[‡]

September 30, 2013

Abstract

We introduce regenerative tree growth processes as consistent families of random trees with n labelled leaves, $n \geq 1$, with a regenerative property at branch points. This framework includes growth processes for exchangeably labelled Markov branching trees, as well as non-exchangeable models such as the alpha-theta model, the alpha-gamma model and all restricted exchangeable models previously studied. Our main structural result is a representation of the growth rule by a σ -finite dislocation measure κ on the set of partitions of \mathbb{N} extending Bertoin's notion of exchangeable dislocation measures from the setting of homogeneous fragmentations. We use this representation to establish necessary and sufficient conditions on the growth rule under which we can apply results by Haas and Miermont for unlabelled and not necessarily consistent trees to establish self-similar random trees and residual mass processes as scaling limits. While previous studies exploited some form of exchangeability, our scaling limit results here only require a regularity condition on the convergence of asymptotic frequencies under κ , in addition to a regular variation condition.

AMS 2000 subject classifications: 60J80.

Keywords: regenerative composition, Markov branching model, fragmentation, self-similar tree, continuum random tree, \mathbb{R} -tree, weighted \mathbb{R} -tree, recursive random tree

1 Introduction to regenerative tree growth processes

For each $n \geq 1$, denote by \mathbb{T}_n the set of rooted leaf-labelled combinatorial trees with no degree-2 vertices and $n + 1$ degree-1 vertices, one of which is called the *root*, the others *leaves*. We distinguish the leaves by labels $1, \dots, n$. Vertices of degree 3 or higher are called *branch points*. Consider a family \mathcal{T}_n , $n \geq 1$, of random trees in \mathbb{T}_n , $n \geq 1$. For $n \geq 2$, we refer to the vertex adjacent to the root as the *first branch point*. It induces the *first split*, a random partition $\Pi_n = (\Pi_{n,1}, \dots, \Pi_{n,K_n})$ of the label set $[n] := \{1, \dots, n\}$ into the label sets of the *subtrees above the branch point*, the connected components of the tree with the first branch point removed. Here, we put the *blocks* $\Pi_{n,i}$ of Π_n in the order of their least elements. For illustration, we write

$$\mathbb{T}_1 = \left\{ \begin{array}{c} \textcircled{1} \\ | \\ \text{---} \end{array} \right\}, \quad \mathbb{T}_2 = \left\{ \begin{array}{c} \textcircled{1} \textcircled{2} \\ \diagup \quad \diagdown \\ \text{---} \end{array} \right\}, \quad \mathbb{T}_3 = \left\{ \begin{array}{c} \textcircled{1} \textcircled{2} \textcircled{3} \\ \diagup \quad \diagdown \quad \diagdown \\ \text{---} \end{array}, \begin{array}{c} \textcircled{1} \textcircled{3} \textcircled{2} \\ \diagup \quad \diagdown \quad \diagdown \\ \text{---} \end{array}, \begin{array}{c} \textcircled{1} \textcircled{2} \textcircled{3} \\ \diagup \quad \diagdown \quad \diagup \\ \text{---} \end{array}, \begin{array}{c} \textcircled{1} \textcircled{3} \textcircled{2} \\ \diagup \quad \diagdown \quad \diagup \\ \text{---} \end{array} \right\}, \quad \text{etc.,}$$

where we have ordered subtrees by their least labels to uniquely choose plane tree representatives.

We suppose that the family $(\mathcal{T}_n, n \geq 1)$ is *consistent* in the sense that removal of leaf $n + 1$ (and the resulting degree-2 vertex, if any) from \mathcal{T}_{n+1} yields \mathcal{T}_n . Reversing this removal gives a tree growth step from n to $n + 1$. A consistent family $(\mathcal{T}_n, n \geq 1)$ is called a *tree growth process*. For $B \subseteq [n]$, let \mathbb{T}_B be the set of trees with $\#B$ leaves labelled by B , so that $\mathbb{T}_{[n]} = \mathbb{T}_n$. Let $\mathcal{T}_{n,B} \in \mathbb{T}_B$ be the *reduced subtree* of \mathcal{T}_n spanned by the root of \mathcal{T}_n and leaves in B , and let $\tilde{\mathcal{T}}_{n,B} \in \mathbb{T}_{[\#B]}$ be the image of $\mathcal{T}_{n,B}$ after *relabelling of leaves* by the increasing bijection from B to $[\#B]$.

*This research was supported by the National Science Foundation Awards 0405779 and 0806118, DR's work also in part by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1106400 and in part by NSF DMS-1204840.

[†]University of California at Berkeley; email pitman@stat.berkeley.edu, drizzolo@math.berkeley.edu

[‡]University of Oxford; email winkel@stats.ox.ac.uk

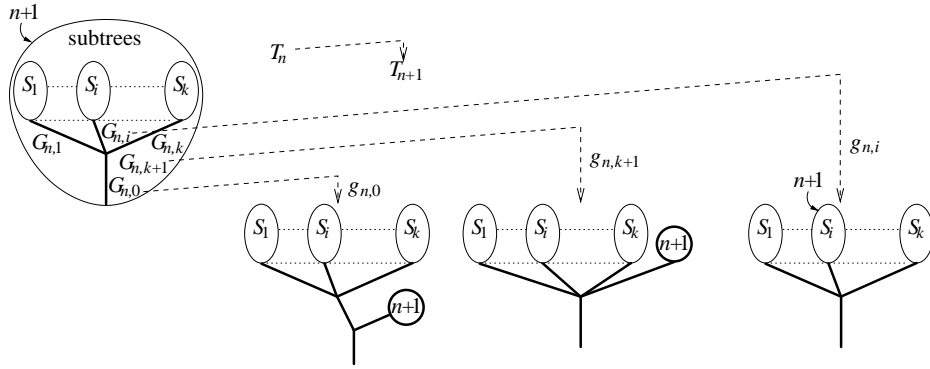


Figure 1: Illustration of the regenerative tree growth step

Definition 1 We call a tree growth process $(\mathcal{T}_n, n \geq 1)$ *regenerative* if for each $n \geq 2$, conditionally given that the first split of \mathcal{T}_n is $\Pi_n = (B_1, \dots, B_k)$, the relabelled subtrees $\tilde{\mathcal{T}}_{n,B_i}$, $1 \leq i \leq k$, above the first branch point are independent copies of $\mathcal{T}_{\#B_i}$.

While this property is well-known for many tree growth processes, the goals of the present paper are to provide general structural results and to study implications for continuum tree asymptotics in this general framework. In the terminology of [6], the trees in a regenerative tree growth process as defined here are “consistent labelled Markov branching trees”. The *exchangeable* case, where the distribution of \mathcal{T}_n is invariant under all permutations of labels, was initiated by Aldous [3], who posed the problem of providing a Kingman-type representation in this case. Bertoin’s [4] theory of homogeneous fragmentations solved that problem as explained in [17]. Then [17, 18] studied tree growth processes associated with fragmentation processes. Natural non-exchangeable tree growth processes were described in terms of simple growth rules that admit regenerative descriptions based on the first split and its subtrees, see particularly [5, 10, 28], as reviewed in Examples 16 and 17 below. We remark, however, that not all natural models fall into our current framework. For example, if \mathcal{T}_n is uniform on \mathbb{T}_n then $(\mathcal{T}_n, n \geq 1)$ is not a regenerative tree growth process because it is not consistent (see [27] for weak limits). Haulk and Pitman [19] give de Finetti representations for exchangeable tree growth processes that are not necessarily regenerative.

An important consequence of Definition 1 is that all regenerative tree growth processes admit descriptions in terms of a growth rule (cf. Figure 1).

Proposition 2 *In the tree growth step from n to $n+1$ for $n \geq 2$, there are the following disjoint events, $G_{n,i}$ for $i = 0, \dots, K_n + 1$, where $K_n \geq 2$ is the number of blocks of the first split of \mathcal{T}_n :*

- $G_{n,0}$: leaf $n+1$ is attached to a new branch point between the root and the first branch point of \mathcal{T}_n ;
- $G_{n,i}$, $1 \leq i \leq K_n$: label $n+1$ is inserted into the i th block of the first split;
- G_{n,K_n+1} : leaf $n+1$ is attached to the first branch point, as singleton block of the first split.

A tree growth process $(\mathcal{T}_n, n \geq 1)$ is regenerative if and only if $\mathbb{P}(G_{n,0} | \mathcal{T}_n) = \mathbb{P}(G_{n,0})$ does not depend on \mathcal{T}_n and $\mathbb{P}(G_{n,i} | \mathcal{T}_n) = \mathbb{P}(G_{n,i} | \Pi_n)$, $1 \leq i \leq K_n + 1$, only depends on the partition Π_n of the first split. In the event $G_{n,i}$, $1 \leq i \leq K_n$, label $n+1$ is inserted into the i th subtree of \mathcal{T}_n of size $\#\Pi_{n,i}$ following the same rule, up to relabelling by the increasing bijection from $\Pi_{n,i} \cup \{n+1\}$ to $[\#\Pi_{n,i} + 1]$.

See Appendix A for a proof of this proposition.

We denote by \mathcal{P}_n the set of partitions $\pi = (B_1, \dots, B_k)$ of $[n]$, with blocks B_i ordered by their least elements. We use the notation $g_n(\pi, i) = \mathbb{P}(G_{n,i} | \Pi_n = \pi)$, $0 \leq i \leq k + 1$, for $\pi \neq 1_{[n]} := ([n])$, $n \geq 2$, and write $g_n(0) = g_n(\pi, 0)$, since we require that this quantity is independent of $\pi \in \mathcal{P}_n \setminus \{1_{[n]}\}$.

One of our main results is that regenerative tree growth rules are (almost) in one-to-one correspondence with σ -finite measures on \mathcal{P} , the set of partitions of $\mathbb{N} = \{1, 2, 3, \dots\}$. Before

stating this, let us introduce the notation $\mathcal{P}^\pi = \{\Gamma \in \mathcal{P} : \Gamma^{[n]} = \pi\}$ where $\Gamma^{[n]} \in \mathcal{P}_n$ is the partition whose blocks are the non-empty blocks of $(\Gamma_i \cap [n], i \geq 1)$. Often we will abuse this notation and for a partition $\pi = (B_1, \dots, B_k)$ write $\mathcal{P}^{B_1, \dots, B_k}$ instead of \mathcal{P}^π . The most common occurrence of this will be the use of $\mathcal{P}^{[n]}$ instead of $\mathcal{P}^{1^{[n]}}$. We equip \mathcal{P} with the σ -algebra generated by $\{\mathcal{P}^\pi, \pi \in \mathcal{P}_n, n \geq 1\}$, which is also the Borel σ -algebra generated by the metric $d(\Gamma, \hat{\Gamma}) = \exp(-\inf\{n \geq 1 : \Gamma^{[n]} \neq \hat{\Gamma}^{[n]}\})$.

Theorem 3 (i) *Let $(g_n, n \geq 2)$ be a regenerative growth rule such that $g_j(0) < 1$ for all $j \geq 2$. Then there exists a unique σ -finite measure κ on \mathcal{P} with $\kappa(\{1_{\mathbb{N}}\}) = 0$ and $\kappa(\mathcal{P}^{\{1\}, \{2\}}) = 1$ such that*

$$g_n(0) = 1 - \frac{\lambda_n}{\lambda_{n+1}}, \quad g_n(\pi, i) = \frac{\lambda_n}{\lambda_{n+1}} \frac{\kappa(\mathcal{P}^{B_1, \dots, B_{i-1}, B_i \cup \{n+1\}, B_{i+1}, \dots, B_k})}{\kappa(\mathcal{P}^{B_1, \dots, B_k})}, \quad i \in [k+1], \quad (1)$$

where $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]})$ and $\pi = (B_1, \dots, B_k)$. Moreover, in this case for $\pi \in \mathcal{P}_n \setminus \{1_{[n]}\}$ we have the splitting rule $p_n(\pi) := \mathbb{P}(\Pi_n = \pi) = \lambda_n^{-1} \kappa(\mathcal{P}^\pi)$.

(ii) *If κ is any measure on \mathcal{P} such that $\kappa(\{1_{\mathbb{N}}\}) = 0$, $0 < \kappa(\mathcal{P}^{\{1\}, \{2\}}) < \infty$, and $\kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) < \infty$ for all $n \geq 2$, then $(g_n, n \geq 2)$ defined by (1) is a regenerative growth rule such that $g_j(0) < 1$ for all $j \geq 2$. In particular, there is a regenerative tree growth process associated with κ .*

We remark that part (ii) of this theorem shows how the relation between $(g_n, n \geq 2)$ and κ fails to be one-to-one. That is, if κ produces a regenerative growth rule $(g_n, n \geq 2)$ by (1), then any constant multiple of κ produces the same growth rule $(g_n, n \geq 2)$ by (1). If κ as in part (ii) and $(g_n, n \geq 2)$ are related by (1), we call κ a *dislocation measure* for $(g_n, n \geq 2)$.

Many of the asymptotic properties of a regenerative tree growth process can be obtained by analysing the asymptotic properties of the associated measure κ . In fact, the two most important considerations turn out to be the growth rate of λ_n and the regularity of the convergence of asymptotic frequencies under κ . Let us expand on the second point. For $\Gamma \in \mathcal{P}$ and $n \geq 1$, consider the decreasing rearrangement $|\Gamma^{[n]}|^\downarrow = (|\Gamma^{[n]}|_i^\downarrow, i \geq 1)$ of relative frequencies $(|\Gamma_i^{[n]}|, i \geq 1)$, where $|\Gamma_i^{[n]}| = \#\Gamma_i^{[n]}/n$. If the limit as $n \rightarrow \infty$ of $|\Gamma^{[n]}|_i^\downarrow$ or of $|\Gamma_i^{[n]}|$ exists, this is denoted by $|\Gamma|_i^\downarrow$ and $|\Gamma_i|$, respectively, and we say that *an asymptotic frequency exists* for that part. If $|\Gamma_i|$ exists for all i we say Γ has *asymptotic frequencies* while if $|\Gamma|_i^\downarrow$ exists for all i we say Γ has *asymptotic ranked frequencies*. Moreover, if the asymptotic (ranked) frequencies exist and sum to 1 κ -a.e., we say they are *proper*. If Γ has asymptotic ranked frequencies then $|\Gamma|^\downarrow = (|\Gamma|_i^\downarrow, i \geq 1)$ naturally lives in the space

$$\mathcal{S}^\downarrow = \left\{ (s_1, s_2, \dots) : s_1 \geq s_2 \geq \dots \geq 0 \text{ and } \sum_{i \geq 1} s_i \leq 1 \right\}. \quad (2)$$

We will equip \mathcal{S}^\downarrow with the topology of pointwise convergence (which is also the topology of ℓ_p convergence for any $p > 1$). We also introduce $\mathcal{S}_1^\downarrow = \{\mathbf{s} \in \mathcal{S}^\downarrow : \sum_{i \geq 1} s_i = 1\}$.

We can then prove the following theorem, the background for which will be fully developed later.

Theorem 4 *Let $(\mathcal{T}_n, n \geq 1)$ be a regenerative tree growth process associated with a dislocation measure κ . Assume that κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies in $\mathcal{S}_1^\downarrow \setminus \{(1, 0, \dots)\}$, define ν to be the push-forward of κ under $\Gamma \mapsto |\Gamma|^\downarrow$ and suppose $\int_{\mathcal{S}^\downarrow} (1 - s_1) \nu(d\mathbf{s}) < \infty$ and $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$ for some slowly varying function ℓ and $\gamma > 0$. If*

$$\lim_{n \rightarrow \infty} \int_{\mathcal{P}} \left(|\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow \right) \kappa(d\Gamma) = 0, \quad (3)$$

then $\frac{\mathcal{T}_n^\circ}{n^\gamma \ell(n)} \rightarrow \mathcal{T}_{\gamma, \nu}$ in distribution, as $n \rightarrow \infty$, in the rooted Gromov-Hausdorff-Prokhorov (GHP) sense, where $\mathcal{T}_{\gamma, \nu}$ is a self-similar fragmentation tree with characteristics (γ, ν) and \mathcal{T}_n° is the tree obtained from \mathcal{T}_n by delabelling the leaves, considered as a metric measure space with the graph metric and the uniform probability measure on the leaves.

We remark that when considering \mathcal{T}_n° purely as a tree we treat it as an element of the set \mathbb{T}_n° of rooted unlabelled trees with n leaves and no degree-2 vertices.

This theorem provides conditions for the existence of a scaling limit of \mathcal{T}_n° , where the label structure of \mathcal{T}_n has been forgotten. However, the leaf labels are an integral part of the tree growth processes under consideration here, so it is natural to ask what happens to the labels. Ideally, one would like a notion of labelled continuum trees to serve as scaling limits of regenerative tree growth processes, just as there is a notion of ordered continuum trees that serve as scaling limits of ordered Galton-Watson trees [2]. However, the appropriate notion is elusive, so we content ourselves with studying the leaf $\{1\}$ and the structure of the path from the root to this leaf. We obtain several results relating the convergence of the residual mass process of the leaf $\{1\}$ to the existence of a scaling limit of the whole tree. Here, by the residual mass process of $\{1\}$ we mean the Markov chain in $m \geq 0$ starting from $X_0^{(n)} = n$, decreasing to $X_1^{(n)} = \#\Pi_{n,1}$ and further according to successive splits of the block containing $\{1\}$ until $M_n = \inf\{m \geq 0 : X_m^{(n)} = 1\}$, when label 1 becomes a singleton. We set $X_m^{(n)} = 0$, $m > M_n$. The limiting processes are decreasing self-similar Markov processes in $[0, \infty)$, which Lamperti [20] represented in terms of subordinators ξ , as

$$X_t = \exp\left(-\xi_{\tau_\xi(t)}\right), \quad \text{where} \quad \tau_\xi(t) = \inf\left\{u \geq 0 : \int_0^u \exp(-\gamma \xi_s) ds > t\right\}. \quad (4)$$

Theorem 5 *Let $(\mathcal{T}_n, n \geq 1)$ be a regenerative tree growth process with dislocation measure κ and $X^{(n)}$ the residual mass process of $\{1\}$ in \mathcal{T}_n . Assume that the first block Γ_1 of κ -a.e. $\Gamma \in \mathcal{P}$ has an asymptotic frequency $|\Gamma_1| \in (0, 1)$ and define Λ as push-forward of κ under $\Gamma \mapsto -\log(|\Gamma_1|)$. Suppose $\int_{(0, \infty)} (1 - e^{-x}) \Lambda(dx) < \infty$ and $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$ for some slowly varying function ℓ and $\gamma > 0$. If*

$$\lim_{n \rightarrow \infty} \int_{\mathcal{P}} \left(|\Gamma_1^{[n]}| - |\Gamma_1|\right) \kappa(d\Gamma) = 0 \quad (5)$$

then $X_{\lfloor \lambda_n t \rfloor}^{(n)} / \lambda_n \rightarrow X_t$ in distribution, as $n \rightarrow \infty$, in the Skorohod sense as functions of $t \geq 0$, where X is a self-similar Markov process and $\mathbb{E}(e^{-s\xi_r}) = \exp(-r \int_{(0, \infty)} (1 - e^{-sy}) \Lambda(dy))$ in Lamperti's representation (4). Moreover, letting A_n be the absorption time of $X^{(n)}$ at 0, the above convergence in distribution holds jointly with the convergence of A_n / λ_n to the absorption time of X at 0.

If, in addition, κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies then $\frac{\mathcal{T}_n^\circ}{n^\gamma \ell(n)} \rightarrow \mathcal{T}_{\gamma, \nu}$ in distribution, as $n \rightarrow \infty$, in the rooted GHP sense, as in Theorem 4.

Assuming that κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies, the remaining conditions of Theorem 5 are stronger than those of Theorem 4. In particular, note that

$$\int_{S^\downarrow} (1 - s_1) \nu(ds) \leq \int_{(0, \infty)} (1 - e^{-x}) \Lambda(dx)$$

and that (5) implies (3) (see the proof of Theorem 28). In Example 19 we construct a regenerative tree growth process where the conditions of Theorem 4 are satisfied, but the conditions of Theorem 5 are not. We note again that leaf $\{1\}$ is generally not typical (i.e. uniformly random) and a heuristic interpretation of the last part of Theorem 5 is that the natural conditions that imply the convergence of the residual mass process of leaf $\{1\}$ are strong enough to imply that the residual mass process of a typical leaf converges as well.

In the other direction, there is a natural strengthening of the hypotheses of Theorem 4 that implies that the conclusions of Theorem 5 are satisfied.

Corollary 6 *If, in addition to the hypotheses of Theorem 4, including (3), we assume that*

$$\kappa(|\Gamma_1| \neq |\Gamma_1^\downarrow|) < \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \int_{\{|\Gamma_1| = |\Gamma_1^\downarrow|\}} \left(|\Gamma_1^{[n]}|^\downarrow - |\Gamma_1^{[n]}|\right) \kappa(d\Gamma) = 0$$

then, with the notation of Theorem 5, $X_{\lfloor \lambda_n t \rfloor}^{(n)}/n \rightarrow X_t$ in distribution, as $n \rightarrow \infty$, in the Skorohod sense as functions of $t \geq 0$ and this convergence in distribution happens jointly with the convergence of $\lambda_n^{-1} A_n$ to the absorption time of X at 0.

When κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies, Theorem 5, combined with previous results about the residual mass process of a typical leaf, provides a description of how leaf $\{1\}$ differs from a typical leaf. To see this, let $(U_k^{(n)}, k \geq 0)$ be the residual mass process of a leaf picked uniformly at random from \mathcal{T}_n for $n \geq 1$. Under the assumptions of Theorem 4, Lemma 28 in [16] implies that $U_{\lfloor \lambda_n t \rfloor}^{(n)}/n \rightarrow U_t = \exp(-\zeta_{\tau_\zeta(t)})$ where $(\zeta_t, t \geq 0)$ is a subordinator with

$$\mathbb{E}(e^{-s\zeta_r}) = \exp \left(-r \int_{\mathcal{P}} \left(1 - \sum_{i \geq 1} (|\Gamma|_i^\downarrow)^{s+1} \right) \kappa(d\Gamma) \right).$$

It is easy to check that this agrees with the expression for $\mathbb{E}(e^{-s\zeta_r})$ in Theorem 5 when κ is exchangeable (see Example 12), but these two expressions may differ in general. Thus we have identified the scaling limit of the residual mass process of $\{1\}$ and scaling limit of the residual mass process of a uniform leaf in terms of subordinators whose Laplace exponents we know explicitly in terms of κ . This provides insight into the difference between what the tree looks like from $\{1\}$'s perspective versus that of a typical leaf.

The remainder of this paper is organized as follows. In Section 2 we give a detailed analysis of the laws of \mathcal{T}_n . The proof of Theorem 3 can be found here, along with a number of other structural results. Section 3 is devoted to examples. One of the nice aspects of the theory presented in this paper is that it gives a coherent framework for many particular models that have been studied previously in the literature. As a result, we are able to give simplified proofs of a number of previously known results about these models. Moreover, our framework makes it easy to specify regenerative growth processes with desired asymptotic properties and this allows us to construct examples illustrating what can go wrong if some of the hypotheses of our theorems are left out. Section 4 provides the necessary background to understand the precise meaning of our statements about scaling limits. We define the limit objects $\mathcal{T}_{\gamma, \nu}$, the GHP topology, and provide the main results from the literature on which our present theorems are built. In Section 5 and 6 respectively, we provide the proofs of Theorems 4 and 5 based on general convergence criteria by Haas and Miermont [15, 16] for (not necessarily consistent) Markov branching models and non-increasing Markov chains. Actually, our results are stronger but a bit more technical than Theorems 4 and 5, so we will prove results that have these theorems as obvious consequences. Section 7 gives some pointers at further problems and related work.

2 Laws of regenerative growth processes

2.1 Explicit formulas in terms of the growth rule

The regenerative nature of the growth processes conditioned on the partition at the first split shows that much of the analysis of these processes can be reduced to analyzing the laws of the partition at the first split of \mathcal{T}_n , i.e. the *splitting rule* p_n , $n \geq 2$. We first find the splitting rule in terms of $(g_n, n \geq 2)$ and then obtain a formula for the law of \mathcal{T}_n . From the growth rule, we have for all $\pi = (B_1, \dots, B_k) \in \mathcal{P}_n \setminus \{1_{[n]}\}$,

$$\begin{aligned} p_2(\{1\}, \{2\}) &= 1, & p_{n+1}([n], \{n+1\}) &= g_n(0), & n &\geq 2, \\ p_{n+1}(B_1, \dots, B_{i-1}, B_i \cup \{n+1\}, B_{i+1}, \dots, B_k) &= p_n(B_1, \dots, B_k) g_n(\pi, i), & i &\in [k+1]. \end{aligned} \quad (6)$$

Using the natural convention $g_1(0) = 1$, the solution to these equations can be written as

$$p_n(\pi) = p_n(B_1, \dots, B_k) = g_{\min B_2 - 1}(0) \prod_{j=\min B_2}^{n-1} g_j(\pi^{[j]}, I_j), \quad \text{where } I_j = i \text{ if } j+1 \in B_i, \quad (7)$$

and $\pi^{[j]}$ is the vector of non-empty $B_i \cap [j]$. The RHS of this formula is the probability of successively creating a new first branch point when $\min B_2$ is added and inserting all higher labels such that the resulting partition at the first split is π . By the regenerative property of \mathcal{T}_n , we can write tree probabilities as a product over branch points; for a tree $T \in \mathbb{T}_n$, we identify each vertex with the set B of labels in the subtrees above this vertex, write $\pi(B)$ for the partition of the split at B , and $\tilde{\pi}(B)$ for the partition of $[\#B]$ obtained when relabelling $\pi(B)$ by the increasing bijection from B to $[\#B]$:

$$\mathbb{P}(\mathcal{T}_n = T) = \prod_{B \in T: \#B \geq 2} p_{\#B}(\tilde{\pi}(B)) = \prod_{B \in T: \#B \geq 2} \left(g_{\min \tilde{\pi}(B)_2 - 1}(0) \prod_{j=\min \tilde{\pi}(B)_2}^{\#B-1} g_j(\tilde{\pi}(B)^{[j]}, I_j(B)) \right), \quad (8)$$

where $I_j(B) = i$ if $j+1 \in \tilde{\pi}(B)_i$, and where $\tilde{\pi}(B) = (\tilde{\pi}(B)_1, \dots, \tilde{\pi}(B)_{k(B)})$.

The residual mass process of the leaf $\{1\}$ can be described in terms of p_n . Recall that the residual mass process is a Markov chain in $m \geq 1$ starting from $X_0^{(n)} = n$, decreasing to $X_1^{(n)} = \#\Pi_{n,1}$ and further according to successive splits until $M_n = \inf\{m \geq 0: X_m^{(n)} = 1\}$, when label 1 becomes a singleton. We represent this Markov chain as a composition of n

$$\mathcal{C}_n = (C_0^{(n)}, \dots, C_{M_n}^{(n)}) = (X_0^{(n)} - X_1^{(n)}, X_1^{(n)} - X_2^{(n)}, \dots, X_{M_n-1}^{(n)} - X_{M_n}^{(n)}, X_{M_n}^{(n)}).$$

Proposition 7 *In a regenerative tree growth process, the family $(\mathcal{C}_n, n \geq 1)$ of compositions is regenerative in the sense that conditionally given $C_0^{(n)} = j$, the composition $(C_1^{(n)}, \dots, C_{M_n}^{(n)})$ of $n-j$ has the same distribution as \mathcal{C}_{n-j} . The entries of the transition probability matrix are*

$$\mathbb{P}(C_0^{(n)} = n-j) = \mathbb{P}(X_1^{(n)} = j) = \sum_{\pi=(B_1, \dots, B_k) \in \mathcal{P}_n: \#B_1=j} p_n(\pi), \quad 1 \leq j \leq n-1.$$

This is a straightforward consequence of Definition 1. We stress that we have consistency in the sense that \mathcal{C}_n can be obtained from \mathcal{C}_{n+1} by reducing one part of \mathcal{C}_{n+1} by 1 (the one corresponding to label $n+1$ in \mathcal{T}_{n+1}), but $(\mathcal{C}_n, n \geq 1)$ is not *sampling consistent* in the sense of [11] as this part is not a size-biased pick from \mathcal{C}_{n+1} , in general. In special cases, versions of this proposition are in the literature; in the exchangeable (sampling consistent) case, it is implicit in Bertoin's [4] study of tagged particles and explicit in [18].

2.2 The dislocation measure

Recall the notation \mathcal{P}_n for the set of partitions $\pi = (B_1, \dots, B_k)$ of $[n] = \{1, \dots, n\}$ with blocks indexed in increasing order of their least elements and the notation \mathcal{P} for the set of all partitions $\Gamma = (\Gamma_i, i \geq 1)$ of \mathbb{N} , with blocks ordered by their least element and $\Gamma_i = \emptyset$ if there are fewer than i blocks. Theorem 3 relates growth rules g_n and splitting rules p_n on $\mathcal{P}_n \setminus \{1_{[n]}\}$ to dislocation measures κ on \mathcal{P} .

Taking our cues from the exchangeable case, cf. [4], one thing we want from our dislocation measures is to be able to use them to embed regenerative tree growth processes in continuous time, making the trees the genealogical trees of continuous-time fragmentation processes. This κ is to provide rates

$$\lambda_n = \kappa(\{\Gamma \in \mathcal{P}: \Gamma^{[n]} \neq 1_{[n]}\}) = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]})$$

for the first split of $[n]$, $n \geq 2$, which allow us to consistently embed the evolution of blocks in \mathcal{T}_n , $n \geq 1$, into continuous time (see Theorem 10). Observe that the rate λ_n of the first split of $[n]$ can then be thinned by the event that this split also splits $[n-1]$, an event with probability $1 - g_{n-1}(0)$, where $(g_n, n \geq 2)$ is the growth rule of the regenerative tree growth process, so that we need

$$\lambda_n(1 - g_{n-1}(0)) = \lambda_{n-1}, \quad n \geq 3, \quad \text{and hence } \lambda_n = \lambda_2 \prod_{j=2}^{n-1} \frac{1}{1 - g_j(0)}, \quad \text{if } g_j(0) \neq 1, j \geq 2. \quad (9)$$

Note that $g_j(0) = 1$ for any $j \geq 2$ means that all insertions in a subtree with j leaves are made below the first split; if scaling limits of $(\mathcal{T}_n, n \geq 1)$ exist at all, such subtrees with j leaves will collapse in the scaling as $n \rightarrow \infty$. We will exclude such behaviour in the sequel and make the

Assumption (A) $g_j(0) < 1$ for all $j \geq 2$.

Proposition 8 *Consider a regenerative tree growth rule $(g_n, n \geq 2)$ satisfying Assumption (A) with splitting rule $(p_n, n \geq 2)$ given by (7), and let $\lambda_2 > 0$ be arbitrary. With $\lambda_n, n \geq 3$, defined by (9), define*

$$\kappa(\mathcal{P}^\pi) = \lambda_n p_n(\pi), \quad \pi \in \mathcal{P}_n \setminus \{1_{[n]}\}, n \geq 2; \quad \kappa(\{1_{\mathbb{N}}\}) = 0. \quad (10)$$

Then κ extends uniquely to a measure on \mathcal{P} .

Proof. This is essentially the same as the analogous result for exchangeable fragmentations, cf. Bertoin's argument [4, Proposition 3.2]. We have defined κ on $\{\mathcal{P}^\pi, \pi \in \bigcup_n (\mathcal{P}_n \setminus \{1_{[n]}\}) \cup \{1_{\mathbb{N}}\}\}$ and it clearly extends to a countably additive measure on the ring generated by these sets. Carathéodory's Extension Theorem provides the unique extension to the σ -ring these sets generate in \mathcal{P} . It is then straightforward to check that this σ -ring is a σ -algebra and, in fact, is the Borel σ -algebra on \mathcal{P} . \square

Note that we can condition κ on splitting $[n]$ and write p_n as

$$p_n(\pi) = \kappa(\mathcal{P}^\pi) / \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}), \quad \pi \in \mathcal{P}_n \setminus \{1_{[n]}\}, n \geq 2. \quad (11)$$

Proof of Theorem 3. Theorem 3 is a direct consequence of (6) and Proposition 8. \square

Let us discuss how Bertoin's [4] notion of a \mathcal{P} -valued homogeneous fragmentation process finds a natural extension where his exchangeable dislocation measure is replaced by a dislocation measure in the sense defined above.

Definition 9 A \mathcal{P} -valued process $\Pi = (\Pi(t), t \geq 0)$ is called *refining* if for all $s < t$ and all blocks $\Pi_j(t)$ of $\Pi(t)$, there is a block $\Pi_i(s)$ of $\Pi(s)$ that contains $\Pi_j(t)$. For a refining process Π , we define *genealogical trees* $\mathcal{T}_n \in \mathbb{T}_n, n \geq 1$, using the representation above (8): \mathcal{T}_n has as branch points and leaves all blocks $\Pi_i^{[n]}(t), i \geq 1, t \geq 0$, visited by the restriction $\Pi^{[n]}$ of Π to $[n]$.

Every regenerative tree growth process can be represented by a nice refining \mathcal{P} -valued process:

Theorem 10 *For each dislocation measure κ as defined after Theorem 3, there exists a \mathcal{P} -valued Feller process $\Pi = (\Pi(t), t \geq 0)$ such that the genealogical trees \mathcal{T}_n of the restrictions $\Pi^{[n]}$ of Π to $[n], n \geq 1$, form a regenerative tree growth process associated with dislocation measure κ .*

Proof. We will use κ in a Poissonian construction based on independent \mathcal{P} -valued Poisson point processes $(\Xi^{(i)}(t), t \geq 0), i \geq 1$, with intensity measure κ . Roughly, we construct Π with $\Pi(0) = 1_{\mathbb{N}}$ such that for all i and t the partition $\Xi^{(i)}(t)$ fragments the i th block $\Pi_i(t)$ of $\Pi(t)$ into the image $\tilde{\Xi}^{(i)}(t)$ of $\Xi^{(i)}(t)$ under the increasing bijection from \mathbb{N} , or $[\#\Pi_i(t)]$, to $\Pi_i(t)$.

More precisely, we build consistent \mathcal{P}_n -valued continuous-time Markov chains $(\Pi^{[n]}(t), t \geq 0), n \geq 1$, with jump times $S^{[n]}(k) \geq 0$ and states $M^{[n]}(k) = (M_1^{[n]}(k), \dots, M_{K^{[n]}(k)}^{[n]}(k)) \in \mathcal{P}_n$: we set $\Pi^{[n]}(t) = M^{[n]}(k), S^{[n]}(k) \leq t < S^{[n]}(k+1), k \geq 0$, where $S^{[n]}(0) = 0, M^{[n]}(0) = 1_{[n]}$,

$$S^{[n]}(k+1) = \inf \left\{ t > S^{[n]}(k) : [\#\Pi_i^{[n]}(k)] \notin \Xi_1^{(i)}(t) \text{ for some } 1 \leq i \leq K^{[n]}(k) \right\}$$

and, if $S^{[n]}(k+1) < \infty$, let $M^{[n]}(k+1)$ be the partition obtained from $M^{[n]}(k)$ by replacing the i th block by the blocks of $\tilde{\Xi}^{(i)}(S^{[n]}(k+1))$, the image of $\Xi^{(i)}(S^{[n]}(k+1)) \cap [\#\Pi_i^{[n]}(k)]$ under the increasing bijection from $[\#\Pi_i^{[n]}(k)]$ to $\Pi_i^{[n]}(k)$. Note that $S^{[n]}(k+1) = \infty$ if and only if $M^{[n]}(k) = 0_{[n]} := (\{1\}, \dots, \{n\})$, as we require $\lambda_2 = \kappa(\mathcal{P}^{\{1\}, \{2\}}) > 0$ for all dislocation measures.

Since Π is uniquely determined by $(\Pi^{[n]}, n \geq 1)$, standard properties of Poisson point processes, and of the space \mathcal{P} complete the proof. \square

By Theorem 3, a growth rule $(g_n, n \geq 2)$ determines a measure κ only up to a multiplicative factor $\lambda_2 > 0$. This is reflected in the fragmentation processes Π of Theorem 10 in the fact that the genealogical trees \mathcal{T}_n , $n \geq 1$, are unaffected by (linear) time changes of Π .

From the consistency of $(\mathcal{T}_n, n \geq 1)$, it is clear that there is a unique branch point of \mathcal{T}_n , where 1 and 2 are separated into different blocks. Moreover, as n varies, the partitions at this branch point define a partition of some random subset of \mathbb{N} , whose distribution when relabelled by the increasing bijection is described by the splitting rule conditioned on partitions that restrict to $(\{1\}, \{2\})$, hence by $\kappa(\cdot | \mathcal{P}^{\{1\}, \{2\}})$. In the Poissonian construction, this partition after relabelling is $\Xi^{(1)}(S^{[2]}(1))$. More generally, while there may be Poisson points that do not induce branch points of \mathcal{T}_n , $n \geq 1$, e.g. when κ is finite or when κ can produce blocks of finite size, those points $\Xi^{(i)}(S^{[n]}(k))$ used in the Poissonian construction describe the partition at a branch point of \mathcal{T}_m for all $m \geq n$. The partition at every branch point, separating labels j and ℓ say, has a distribution that is absolutely continuous with respect to κ .

The Poissonian construction formulated here differs from Bertoin's [4, Section 3.1.3] in the relabelling by increasing bijections: Bertoin uses $\Pi_i(t) \cap \Xi^{(i)}(t)$ instead of $\Xi^{(i)}(t)$. In the exchangeable case this yields the same processes, in distribution. A notable consequence is that under assumptions that ensure that there are always infinitely many blocks and that they are all infinite, we can recover the $(\Xi^{(i)}, i \geq 1)$ from Π in our setting. It is now possible to explore some more of Bertoin's fragmentation theory [4, Chapter 3] in our extended generality, notably erosion effects, stopping lines and extended branching properties, and, under conditions that ensure the existence of asymptotic frequencies, also self-similar partition-valued fragmentation processes. More generally, it would be interesting to characterise Markov processes (with a suitable branching property) whose genealogical trees are regenerative.

2.3 Unlabelled Markov branching trees

Our scaling limit results take advantage of recent progress on scaling limits of rooted unlabelled Markov branching trees, which we now introduce. For $n \geq 1$, let \mathbb{T}_n° be the image of \mathbb{T}_n under the map that delabels the leaves of a tree. Define

$$\mathcal{P}_n^\circ = \left\{ (n_1, \dots, n_p) \in \bigcup_{k \geq 1} \mathbb{N}^k : n \geq n_1 \geq n_2 \geq \dots \geq n_p \text{ and } \sum_{i=1}^p n_i = n, p \geq 1 \right\}.$$

Let $(p_n^\circ, n \geq 2)$ be a sequence such that for each n , p_n° is a probability function on $\mathcal{P}_n^\circ \setminus \{(n)\}$. A sequence $(\mathcal{T}_n^\circ, n \geq 1)$ of random variables such that $\mathcal{T}_n^\circ \in \mathbb{T}_n^\circ$ is called a *Markov branching model* based on $(p_n^\circ, n \geq 2)$ if for each $n \geq 2$, the law of \mathcal{T}_n° is the same as the law of the tree $\widehat{\mathcal{T}}$ constructed as follows: choose (N_1, \dots, N_p) according to p_n° ; conditionally given that $(N_1, \dots, N_p) = (n_1, \dots, n_p)$, let $(\widehat{\mathcal{T}}_1, \dots, \widehat{\mathcal{T}}_p)$ be a vector of independent trees such that $\widehat{\mathcal{T}}_i$ is distributed as $\mathcal{T}_{n_i}^\circ$, $1 \leq i \leq p$; the tree $\widehat{\mathcal{T}}$ is then obtained by identifying the roots of $\widehat{\mathcal{T}}_1, \dots, \widehat{\mathcal{T}}_p$ as a single vertex and attaching a new root to this vertex.

The following proposition is an immediate consequence of these definitions.

Proposition 11 *If $(\mathcal{T}_n, n \geq 1)$ is a regenerative tree growth process with associated dislocation measure κ and $(\mathcal{T}_n^\circ, n \geq 1)$ is the sequence of trees such that \mathcal{T}_n° is obtained from \mathcal{T}_n by delabelling the leaves, then $(\mathcal{T}_n^\circ, n \geq 1)$ is a Markov branching model based on the functions*

$$p_n^\circ(n_1, \dots, n_k) = \sum_{\pi \in \mathcal{P}_n : (\#\pi)^\downarrow = (n_1, \dots, n_k)} p_n(\pi) = \sum_{\pi \in \mathcal{P}_n : (\#\pi)^\downarrow = (n_1, \dots, n_k)} \frac{\kappa(\mathcal{P}^\pi)}{\lambda_n}, \quad \lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}),$$

where we write $(\#\pi)^\downarrow$ for the decreasing rearrangement of the block sizes of π .

3 Examples

An important motivation for our results is that they allow a unified treatment of previously studied models. In this section we discuss these models, recall or construct their dislocation measures

and demonstrate how our Theorems 4 and 5 apply. We also develop some further examples that explore the conditions (3) and (5) that appear in Theorems 4 and 5. Before proceeding with the examples, we introduce paintbox partitions, which are a recurring theme in the construction of dislocation measures. For $\mathbf{s} \in \mathcal{S}^\downarrow$, where \mathcal{S}^\downarrow is defined in (2), we define *Kingman's paintbox* $\kappa_{\mathbf{s}}$ as the distribution of the random partition Π of \mathbb{N} where $i, j \in \mathbb{N}$ are in the same block if $i = j$ or $R_i = R_j \geq 1$, where the R_i , $i \in \mathbb{N}$, are independent random variables with $\mathbb{P}(R_i = k) = s_k$, $k \geq 0$, and where $s_0 = 1 - \sum_{i \geq 1} s_i$. Note that the Strong Law of Large Numbers implies that $\kappa_{\mathbf{s}}$ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies $|\Gamma|^\downarrow = \mathbf{s}$.

Example 12 (Exchangeable models [4, 17]) Bertoin classified all *exchangeable* dislocation measures, i.e. measures that are invariant under the action of permutations of \mathbb{N} on \mathcal{P} , giving an integral representation

$$\kappa = \sum_{j \geq 1} c \delta_{\varepsilon(j)} + \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(\cdot) \nu(d\mathbf{s}),$$

where $c \geq 0$, $\varepsilon(j)$ is the partition with blocks $\{j\}$ and $\mathbb{N} \setminus \{j\}$, and ν is a measure on \mathcal{S}^\downarrow with

$$\nu(\{(1, 0, 0, \dots)\}) = 0 \quad \text{and} \quad \int_{\mathcal{S}^\downarrow} (1 - s_1) \nu(d\mathbf{s}) < \infty. \quad (12)$$

Then ν is the push-forward of κ under $\Gamma \mapsto |\Gamma|^\downarrow$, restricted to $\mathcal{S}^\downarrow \setminus \{(1, 0, \dots)\}$.

The splitting rules $(p_n, n \geq 2)$ associated with Bertoin's exchangeable dislocation measures κ via (10) give rise to the consistent exchangeably labelled Markov branching trees of [17]. For exchangeable $\kappa = \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(\cdot) \nu(d\mathbf{s})$, it was demonstrated in [16] that if $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$ for some $\gamma > 0$ and some slowly varying function ℓ , then (3) also holds. It is also easy to verify the condition of Theorem 5 in this case. \square

Example 12 includes an important subclass of models previously studied in their own right, whose dislocation measures are of Poisson-Dirichlet type. For this example, we can also calculate the growth rule explicitly up to a sequence of normalisation constants:

Example 13 (Poisson-Dirichlet model [18, 22]) According to [22], the only consistent exchangeable model with splitting rule of the Gibbs form

$$p_n(\pi) = \frac{a_k}{c_n} \prod_{i=1}^k w_{\#B_i}, \quad \pi = (B_1, \dots, B_k) \in \mathcal{P}_n \setminus \{1_{[n]}\}, \quad \text{for some } w_j \geq 0, a_k \geq 0, c_n > 0,$$

is given by a two-parameter family. Most relevant for us are $0 < \alpha < 1$ and $\theta \geq -2\alpha$ with $w_j = \Gamma(j - \alpha)/\Gamma(1 - \alpha)$, $j \geq 1$, and $a_k = \alpha^{k-2} \Gamma(k + \theta/\alpha)/\Gamma(2 + \theta/\alpha)$, $k \geq 2$, with normalisation constants $c_n = c_{\alpha, \theta}(n)$ satisfying $c_{\alpha, \theta}(2) = 1$ and $c_{\alpha, \theta}(n+1) = (n + \theta)c_{\alpha, \theta}(n) + \Gamma(n - \alpha)/\Gamma(1 - \alpha)$, $n \geq 2$. Case $\alpha = 0$ is a limiting case. These yield growth rules for $\pi = (B_1, \dots, B_k)$ of the form

$$\begin{aligned} g_n(0) &= p_{n+1}([n], \{n+1\}) = \frac{\Gamma(n - \alpha)}{\Gamma(1 - \alpha)c_{\alpha, \theta}(n+1)}, \\ g_n(\pi, i) &= \frac{p_{n+1}(B_1, \dots, B_{i-1}, B_i \cup \{n+1\}, B_{i+1}, \dots, B_k)}{p_n(B_1, \dots, B_k)} = \frac{(\#B_i - \alpha)c_{\alpha, \theta}(n)}{c_{\alpha, \theta}(n+1)}, \quad i \in [k], \\ g_n(\pi, k+1) &= \frac{(k\alpha + \theta)c_{\alpha, \theta}(n)}{c_{\alpha, \theta}(n+1)}. \end{aligned}$$

Dislocation measures $\kappa_{\alpha, \theta}^{\text{PD}^*}$ are exchangeable with, for $\theta > -2\alpha$, $\alpha \in (0, 1)$, coefficient $c = 0$ and

$$\nu(d\mathbf{s}) = \text{PD}_{\alpha, \theta}^*(d\mathbf{s}) = \mathbb{E}(\sigma_1^\theta; \sigma_1^{-1} \Delta \sigma_{[0,1]} \in d\mathbf{s}) = \int_{(0, \infty)} x^\theta \mathbb{P}(x^{-1} \Delta \sigma_{[0,1]} \in d\mathbf{s} | \sigma_1 = x) \mathbb{P}(\sigma_1 \in dx),$$

where $(\sigma_t, t \geq 0)$ is a stable subordinator with Laplace transform $\mathbb{E}(e^{-\lambda \sigma_t}) = e^{-t\lambda^\alpha}$ and $\Delta \sigma_{[0,1]}$ the decreasing rearrangement of its jumps $\Delta \sigma_t = \sigma_t - \sigma_{t-}$, $t \in [0, 1]$, see [18, 22] for details. \square

There are simpler growth rules, which in general lead to models that are not fully exchangeable. Before we present these simpler growth rules, we mention a large class of models that retain a weak form of exchangeability and for which scaling limits have been obtained.

Example 14 (Restricted exchangeable models [6]) Let us define *restricted exchangeable* dislocation measures by their integral representation, referring to [6] for a full discussion:

$$\kappa = c_1 \delta_{\varepsilon(1)} + \sum_{j \geq 1} \left(c_j \delta_{\varepsilon(j+1)} + k_j \delta_{\omega[j]} + \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^{[j], \{j+1\}}) \nu_j(d\mathbf{s}) \right)$$

where $c_j \geq 0$, $k_j \geq 0$, $\omega[j] = ([j], \{j+1\}, \{j+2\}, \dots)$, and ν_j is a measure on \mathcal{S}^\downarrow satisfying

$$\nu_j(\{(1, 0, \dots), (0, 0, \dots)\}) = 0 \quad \text{and} \quad \int_{\mathcal{S}^\downarrow} \left(s_0 1_{\{j=1\}} + \sum_{i \geq 1} s_i^j (1 - s_i) \right) \nu_j(d\mathbf{s}) < \infty, \quad j \geq 1.$$

This includes all exchangeable dislocation measures, for $c_j = c$, $k_j = \nu(\{(0, \dots)\})$, $\nu_j = \nu - k_j \delta_{(0, \dots)}$. The splitting rules $(p_n, n \geq 2)$ associated with restricted exchangeable dislocation measures κ give rise to the consistent restricted exchangeable labelled Markov branching trees of [6].

Consider the case where $c_j = k_j = 0$ and $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$ for some slowly varying function ℓ . The push-forward of κ under $\Gamma \mapsto |\Gamma|^\downarrow$ is given by

$$\nu(ds) = \sum_{j \geq 1} \left(\sum_{i \geq 1} s_i^j (1 - s_i) \right) \nu_j(ds).$$

Suppose that for each $j \geq 1$, ν_j has its support in $\mathcal{S}_1^\downarrow \setminus \{(1, 0, \dots)\}$ and that $\int_{\mathcal{S}^\downarrow} (1 - s_1) \nu(ds) < \infty$. Assuming further that $\nu_j = \nu_m$ for all $j \geq m$ for some $m \geq 1$, as in [6, Theorem 7] where scaling limits were established for convergence in probability, we deduce that (3) holds for

$$\kappa = \kappa_{\nu_m} - \kappa_{\nu_m}(\cdot \cap (\mathcal{P} \setminus \mathcal{P}^{[m]})) + \sum_{j=1}^{m-1} \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^{[j], \{j+1\}}) \nu_j(d\mathbf{s}),$$

from the exchangeable case and by dominated convergence, because on the RHS only the measure $\kappa_{\nu_m} = \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(\cdot) \nu_m(d\mathbf{s})$ is infinite. \square

One of the early families of regenerative tree growth processes to be studied was Ford's alpha-model. It has also been a main driver for much of the literature on scaling limits of Markov branching trees, both for general models and for further models with special structure.

Example 15 (Ford's alpha-model [10]) This family is parametrized by $\alpha \in [0, 1]$ as follows. For each edge e of \mathcal{T}_n , give e weight α if both of its vertices are internal and weight $1 - \alpha$ if one of its vertices is a leaf. Choose an edge with probability proportional to its weight and attach $n+1$ to a new branch point between the two vertices of the selected edge. From this description it is easy to check that $(\mathcal{T}_n, n \geq 1)$ is a family of binary trees that forms a regenerative tree growth process. Moreover, for $\pi = (B_1, B_2)$ we have

$$g_n(\pi, 0) = \frac{\alpha}{n - \alpha} \quad \text{and} \quad g_n(\pi, i) = \frac{\#B_i - \alpha}{n - \alpha}, \quad i \in \{1, 2\}.$$

This model was introduced in [10] as a model on cladograms that interpolates between the Yule model ($\alpha = 0$), the uniform model ($\alpha = 1/2$), and the comb ($\alpha = 1$).

The alpha-model is a restricted exchangeable model of binary trees that admits (at least) two natural extensions. The alpha-gamma model, which is restricted exchangeable but not binary, and the alpha-theta model, which is binary but not, in general, restricted exchangeable. The details of these models are our next two examples.

Example 16 (Alpha-gamma model [5]) For $0 \leq \gamma \leq \alpha \leq 1$ and $\pi = (B_1, \dots, B_k) \in \mathcal{P}_n$, let

$$g_n(\pi, 0) = \frac{\gamma}{n - \alpha}, \quad g_n(\pi, i) = \frac{\#B_i - \alpha}{n - \alpha}, \quad i \in [k], \quad g_n(\pi, k+1) = \frac{(k-1)\alpha - \gamma}{n - \alpha}.$$

The κ -measures are restricted exchangeable with

$$c_j = k_j = 0, \quad \text{and} \quad \nu_1(ds) = (1 - \alpha)PD_{\alpha, -\alpha-\gamma}^*(ds), \quad \nu_j(ds) = \gamma PD_{\alpha, -\alpha-\gamma}^*(ds), \quad j \geq 2,$$

if $0 < \gamma < \alpha < 1$, see [6]. The convergence results of Example 14 include this as a special case. \square

Example 17 (Alpha-theta model [28]) For $0 \leq \alpha \leq 1$, $\theta \geq 0$ and $\pi = (B_1, B_2) \in \mathcal{P}_n$, let

$$g_n(\pi, 0) = \frac{\alpha}{n - 1 + \theta}, \quad g_n(\pi, 1) = \frac{\#B_1 - 1 + \theta}{n - 1 + \theta}, \quad g_n(\pi, 2) = \frac{\#B_2 - \alpha}{n - 1 + \theta}, \quad g_n(\pi, k) = 0, \quad k \geq 3.$$

This model is not restricted exchangeable except in the case $\theta = 1 - \alpha$, where the model reduces to Ford's alpha model. Moreover, the dislocation measure for the alpha-theta model has not previously appeared in the literature. To describe it, we introduce an *ordered paintbox* $\tilde{\kappa}_{(u, 1-u)}$, $0 < u < 1$, as the distribution of $\Pi = (\{i \geq 1: R_i = 1\}, \{i \geq 1: R_i = 2\})$ where $R_1 = 1$ and the R_i , $i \geq 2$, are independent random variables with $\mathbb{P}(R_i = 1) = u = 1 - \mathbb{P}(R_i = 2)$.

For $0 < \alpha < 1$ and $\theta > 0$, the κ -measure of the alpha-theta model is now given by

$$\kappa = \alpha \tilde{\kappa}_{\theta, -\alpha}^{\text{beta}}(\cdot \cap \mathcal{P}^{[2]}) + \theta \tilde{\kappa}_{\theta, -\alpha}^{\text{beta}}(\cdot \cap \mathcal{P}^{\{1\}, \{2\}}), \quad \text{where} \quad \tilde{\kappa}_{\theta, -\alpha}^{\text{beta}} = \int_0^1 \tilde{\kappa}_{(u, 1-u)}(\cdot) u^{\theta-1} (1-u)^{-\alpha-1} du.$$

To see this, note that from (7)

$$p_n(B_1, B_2) = (\alpha 1_{\{2 \in B_1\}} + \theta 1_{\{2 \in B_2\}}) \frac{\Gamma(\#B_1 - 1 + \theta) \Gamma(\#B_2 - \alpha)}{\Gamma(n - 1 + \theta) \Gamma(1 - \alpha)},$$

for $(B_1, B_2) \in \mathcal{P}_n$ with $\#B_1 \geq 1$ and $\#B_2 \geq 1$, $n \geq 2$. The result now follows from (11) and the fact that $\tilde{\kappa}_{\theta, -\alpha}^{\text{beta}}(\mathcal{P}^{B_1, B_2}) = \int_0^1 u^{\theta-1+\#B_1-1} (1-u)^{-\alpha-1+\#B_2} du$ is a beta integral.

Scaling limits for these trees were established in [16] using criteria that are directly equivalent to what appears in our paper as Condition (i) in Theorem 26 below. However, we can now give a shorter argument. We have (similar to [16] for exchangeable paintboxes)

$$\int_{\mathcal{P}} \left(1 - |\Gamma^{[n]}|_1^{\downarrow}\right) \tilde{\kappa}_{(u, 1-u)}(d\Gamma) \leq \int_{\mathcal{P}} \left(1 - |\Gamma_1^{[n]}|\right) \tilde{\kappa}_{(u, 1-u)}(d\Gamma) = (1-u) \left(1 - \frac{1}{n}\right)$$

But then (3) follows for $\tilde{\kappa}_{\theta, -\alpha}^{\text{beta}}$ and for κ , which is bounded by a multiple of $\tilde{\kappa}_{\theta, -\alpha}^{\text{beta}}$, by dominated convergence. Thus Theorem 4 applies. \square

For the alpha-gamma and alpha-theta models, the regenerative property was shown in [28, Proposition 11] and [6, Proposition 8], respectively. They both contain as special case for $\alpha = 1/2$ and, respectively, $\theta = 1/2$ and $\gamma = 1/2$, the exchangeable uniform model on binary trees, related to Aldous's Brownian Continuum Random Tree [1]. Ford's binary alpha-model [10] is also included in both examples.

Aldous's binary beta model [3] is included in the alpha-theta model for $\theta = -2\alpha$. Both the Poisson-Dirichlet model and the alpha-gamma model contain as special cases for $\alpha \in [1/2, 1)$ and, respectively, $\theta = -1$ and $\gamma = 1 - \alpha$, the exchangeable model related to the stable Continuum Random Tree [7, 21, 23].

All the examples we have given so far satisfy the hypotheses of Theorem 4. In fact, in these examples both tree convergence and residual mass process convergence were previously known to hold. The exchangeable case is [17, Proposition 7], the particle labelled 1 in the restricted exchangeable case is [6, Proposition 28] and particle labelled 1 in the alpha-theta model is [28, Proposition 6(iv)]. Let us provide some very different examples that show what can go wrong. We first give an example where tree convergence fails.

Example 18 Some of the most elementary non-trivial dislocation measures are of the form

$$\kappa = \sum_{j \geq 2} (\lambda_j - \lambda_{j-1}) \delta_{\Gamma(j)} \quad \text{for some } \Gamma(j) \in \mathcal{P}^{[j-1], \{j\}}, j \geq 2.$$

To ensure $\lambda_n \sim n^\gamma$ for some $\gamma \in (0, 1)$, let $\lambda_n - \lambda_{n-1} = \gamma n^{\gamma-1}$. For simplicity, we take $\Gamma(j)$ binary with asymptotic frequencies $(x^{(j)}, 1 - x^{(j)})$, where $x^{(j)} = 1 - 1/j$. This implies

$$\int_{S^\downarrow} (1 - s_1) \nu(ds) = \int_{\mathcal{P}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma) = \sum_{j \geq 2} (1 - x^{(j)}) (\lambda_j - \lambda_{j-1}) = \gamma \sum_{j \geq 2} j^{\gamma-2} < \infty,$$

with ν as push-forward of κ . Consider a γ -self-similar tree $\mathcal{T}_{\gamma, \nu}$ with dislocation measure ν . We explore two examples illustrating the validity/violation of (3), which now reads

$$\int_{\mathcal{P}} \left(|\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow \right) \kappa(d\Gamma) = \sum_{j \geq 2} \left(|\Gamma(j)^{[n]}|_1^\downarrow - x^{(j)} \right) (\lambda_j - \lambda_{j-1}) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

- (a) For $j \geq 2$ and $x^{(j)} = 1 - 1/j \in (0, 1)$, we construct $\Gamma(j)$ as a sequence $(\Gamma(j)^{[n]}, n \geq j)$, starting from $\Gamma(j)^{[j]} = ([j-1], \{j\})$, and using Step $A_{x^{(j)}}$ inductively for $n \geq j$, where for $0 \leq x \leq 1$:

- Step A_x : Given $\Gamma^{[n]}$, if $|\Gamma_1^{[n]}| > x$, set $\Gamma_1^{[n+1]} = \Gamma_1^{[n]}$, otherwise set $\Gamma_1^{[n+1]} = \Gamma_1^{[n]} \cup \{n+1\}$.

The purpose of Step A_x is to change the relative frequency towards x . For $x = x^{(j)}$ and $\Gamma = \Gamma(j)$, we get $|\Gamma(j)_1| = |\Gamma(j)_1^\downarrow| = x^{(j)}$ and $||\Gamma(j)_1^{[n]}| - x^{(j)}| \leq 1 - x^{(j)}$ for all $n \geq 1$, $j \geq 1$, equality for $j \geq n$ and strict inequality for $j < n$, since $1/n < 1/j = 1 - x^{(j)}$. By the Dominated Convergence Theorem, (3) is satisfied.

- (b) For (3) to fail, first let $\Gamma(j)_1^{[n]}$ approach frequency $1/2$, applying Step $A_{1/2}$ for $n < a_j$, so that $|\Gamma(j)_1^{[2j]}| = 1/2$ and $|\Gamma(j)_1^{[n]}| \approx 1/2$ for $n \in [2j, a_j]$. Choose (a_j) increasing with $2 \geq \sum_{i \geq 2: n \in [2i, a_i]} (\lambda_i - \lambda_{i-1}) > 1$ and apply Step $A_{x^{(j)}}$ for $n \geq a_j$. Then we will have $|\Gamma(j)_1| = x^{(j)}$ for all $j \geq 2$, but for all n sufficiently large,

$$\sum_{j \geq 2} \left(|\Gamma(j)_1^{[n]}| - x^{(j)} \right) (\lambda_j - \lambda_{j-1}) \leq -\frac{1}{3} \sum_{j \geq 2: n \in [2j, a_j]} (\lambda_j - \lambda_{j-1}) < -\frac{1}{3} < 0.$$

Intuitively, the approximating trees have too many even branchpoints splitting into two equal-sized subtrees making trees wide and small in height, while the proposed limiting distribution produces uneven branch points leading to thin and high trees with higher probability. Gromov-Hausdorff convergence fails, if total heights do not converge [8]. \square

In our next example, we show that the hypotheses of Theorem 5 are strictly stronger than the hypotheses of Theorem 4.

Example 19 In the general setting of Example 18, consider $\Gamma(j)_1^{[n]}$ that first approaches (the wrong!) frequency $1 - x^{(j)}$, applying Step $A_{1-x^{(j)}}$ for $n < a_j$, so that $|\Gamma(j)_1^{[n]}| \approx 1 - x^{(j)}$ for $n \in [j/(1 - x^{(j)}), a_j]$. Then we apply Step $A_{x^{(j)}}$ for $n \geq a_j$ to achieve $|\Gamma(j)_1| = x^{(j)}$. We call these partitions “evil”. If we did this for all $j \geq 2$, too many partitions would have intermediate frequencies around $1/2$ when restricted to $[n]$ and tree convergence may fail. Note that while at $1 - x^{(j)}$, the block not containing 1 has frequency $x^{(j)}$ and is the larger block size that appears in the tree convergence criterion, while frequency $1 - x^{(j)}$ is relevant for the residual mass process.

To control the influence of partitions at intermediate frequencies, we also consider “good” partitions from Example 18(a). The following strategy gives the right mix of “good” and “evil”:

1. For $j = 2$ and $j = 3$, start with two evil partitions $\Gamma(2)$ and $\Gamma(3)$, with $|\Gamma(3)_1^{[\ell]}| \approx 1 - x^{(3)}$ for $\ell = 3/(1 - x^{(3)})$, but leave a_2 and a_3 to be specified. Take good partitions $\Gamma(4), \dots, \Gamma(\ell)$. Also recall from the general setting that $\lambda_3 - \lambda_2 = \gamma 3^{\gamma-1} > 0$. To proceed inductively, let $m = 1$, $E_1 = \{2, 3\}$, $j_1 = \ell + 1$, and proceed to step 2.
2. Given (m, E_m, j_m) , release the smallest evil partition $e_m = \min E_m$ by setting $a_{e_m} = j_m$. Start evil partitions $\Gamma(j_m), \dots, \Gamma(k_m)$ up to $k_m = \inf\{j \geq j_m : \lambda_j - \lambda_{j_m} \geq \lambda_{e_m} - \lambda_{e_m-1}\}$. Let
$$\ell_m = \inf \left\{ j \geq k_m : |\Gamma(e_m)_1^{[j]}| \approx x^{(e_m)} \text{ and } |\Gamma(i)_1^{[j]}| \approx 1 - x^{(i)}, j_m \leq i \leq k_m \right\},$$
and take good partitions $\Gamma(k_m + 1), \dots, \Gamma(\ell_m)$. Now set $E_{m+1} = (E_m \setminus \{e_m\}) \cup \{j_m, \dots, k_m\}$, $j_{m+1} = \ell_m + 1$ and repeat step 2. for $(m + 1, E_{m+1}, j_{m+1})$.

Now $|\Gamma(j)_1| = x^{(j)}$ for all $j \geq 2$ since $a_{e_m} < \infty$ for all evil partitions e_m . The criterion (3) of Theorem 4 for tree convergence holds, because the good partitions and the evil partitions that are either at frequency $x^{(j)}$ or $1 - x^{(j)}$ give convergence as in Example 18(a), while the evil partitions at intermediate frequencies have total weight $w_m = (\lambda_{e_m} - \lambda_{e_m-1}) + (\lambda_{k_m} - \lambda_{j_m-1}) \rightarrow 0$ as $m \rightarrow \infty$, so their contribution vanishes as $m \rightarrow \infty$.

The criterion (5) of Theorem 5 for residual mass process convergence is not satisfied, because for every $n \geq 3/(1 - x^{(3)})$, there are evil partitions of weight at least $\lambda_3 - \lambda_2$ which have a frequency $|\Gamma(j)_1^{[n]}| \approx 1 - x^{(j)}$ that is smaller by more than $1/4$ than their limit frequency $x^{(j)}$, since $x^{(j)} - (1 - x^{(j)}) > 1/4$ for all $j \geq 3$, and this cannot be offset by partitions that exceed their limit frequencies, by the argument in Example 18(a). \square

4 Background

In this section we present the background information needed to understand the statements of our results on scaling limits of random trees. Since the proofs of our results do not require any technical details about the constructions in this section we keep the discussion light and heuristic at times, referring to the existing literature for details.

4.1 Trees as metric measure spaces

The trees under discussion in this paper can naturally be considered as metric spaces with the graph metric. That is, the distance between two vertices is the number of edges on the path connecting them. Let (T, d, root) be a tree equipped with the graph metric. For $a > 0$, we define at to be the metric space (T, ad, root) , i.e. the metric is scaled by a . Moreover, the trees we are dealing with are rooted so we consider (T, d, root) as a pointed metric space with the root as the point. Additionally, we let μ_T be the uniform probability measure on the leaves of T . If we have a random tree \mathcal{T} , this gives rise to a random pointed metric measure space $(\mathcal{T}, d, \text{root}, \mu_{\mathcal{T}})$. For this last statement to be made rigorous, it is clear that we need to put a topology on pointed metric measure spaces. This is hard to do in general, but note that the pointed metric measure spaces that come from the trees we are discussing are compact and this simplifies matters.

Let \mathcal{M}_w be the set of equivalence classes of compact pointed metric measure spaces (equivalence here being up to point and measure preserving isometry). We endow \mathcal{M}_w with the pointed Gromov-Hausdorff-Prokhorov metric (see [16]). Fix $(X, d, \rho, \mu), (X', d', \rho, \mu') \in \mathcal{M}_w$ and define

$$d_{\text{GHP}}(X, X') = \inf_{(M, \delta)} \inf_{\substack{\phi: X \rightarrow M \\ \phi': X' \rightarrow M}} [\delta(\phi(\rho), \phi'(\rho')) \vee \delta_H(\phi(X), \phi'(X')) \vee \delta_P(\phi_*\mu, \phi'_*\mu')],$$

where the first infimum is over metric spaces (M, δ) , the second infimum if over isometric embeddings ϕ and ϕ' of X and X' into M , δ_H is the Hausdorff distance on compact subsets of M , and $\delta_P(\phi_*\mu, \phi'_*\mu')$ is the Prokhorov distance between the push-forward $\phi_*\mu$ of μ by ϕ and the push-forward $\phi'_*\mu'$ of μ' by ϕ' . It is worth noting briefly that the definitions of \mathcal{M}_w and d_{GHP} as just given do not make formal sense in Zermelo-Fraenkel set theory with the axiom of choice (ZFC); one might just as well try metrizing the set of all sets. Nonetheless, it is not hard to formalize the

heuristic definitions we have given. For example, one can use the fact that every separable metric space can be isometrically embedded in ℓ_∞ to find an honest set $\overline{\mathcal{M}}_w$ in ZFC such that every compact pointed metric measure space is isometric, by point and measure preserving isometry, to exactly one element of $\overline{\mathcal{M}}_w$ and then do everything internally in this set.

Proposition 20 (Proposition 1 in [16], see also [9, 12, 13, 24]) *The space (\mathcal{M}_w, d_{GHP}) is a complete separable metric space.*

Scaling limits of discrete trees are elements of \mathcal{M}_w that are tree-like metric spaces. An \mathbb{R} -tree is a complete metric space (T, d) with the following properties:

- For $v, w \in T$, there exists a unique isometry $\phi_{v,w}: [0, d(v, w)] \rightarrow T$ with $\phi_{v,w}(0) = v$ and $\phi_{v,w}(d(v, w)) = w$.
- For every continuous injective function $c: [0, 1] \rightarrow T$ such that $c(0) = v$ and $c(1) = w$, we have $c([0, 1]) = \phi_{v,w}([0, d(v, w)])$.

If (T, d) is a compact \mathbb{R} -tree, every choice of root $\rho \in T$ and probability measure μ on T yields an element (T, d, ρ, μ) of \mathcal{M}_w . With this choice of root also comes a height function $\text{ht}(v) = d(v, \rho)$. The leaves of T can then be defined as the points $v \in T$ such that v is not in $[[\rho, w[:= \phi_{\rho,w}([0, \text{ht}(w)))$ for any $w \in T$. The set of leaves is denoted $\mathcal{L}(T)$.

Definition 21 A *continuum tree* is an \mathbb{R} -tree (T, d, ρ, μ) with a choice of root and probability measure such that μ is non-atomic, $\mu(\mathcal{L}(T)) = 1$, and for every non-leaf vertex w , $\mu(\{v \in T: [[\rho, v] \cap [[\rho, w] = [[\rho, w]\}) > 0$.

A continuum random tree (CRT) is an (\mathcal{M}_w, d_{GHP}) -valued random variable that is almost surely a continuum tree. The continuum random trees we will be interested in are those associated with self-similar mass fragmentation processes.

4.2 Self-similar mass fragmentations

We are now prepared to introduce self-similar mass fragmentations and their genealogical trees. Suppose $\gamma > 0$ and let ν be a σ -finite measure on \mathcal{S}^\downarrow such that $\nu(\{(1, 0, 0, \dots)\}) = 0$ and $\int_{\mathcal{S}^\downarrow} (1 - s_1) \nu(ds) < \infty$ and $\nu(\sum_i s_i < 1) = 0$. Heuristically, a self-similar mass fragmentation with characteristics (γ, ν) is an \mathcal{S}^\downarrow -valued Markov process $(F(t), t \geq 0)$ such that $F(0) = (1, 0, 0, \dots)$ and such that a block of size x splits into blocks $xs = (xs_1, xs_2, \dots)$ at rate $x^{-\gamma} \nu(ds)$. A rigorous construction of such processes can be found in [4], though we remark that there is a slight difference in notation: our index γ of self-similarity corresponds to the index $-\gamma$ in [4]. The idea of the genealogical tree of a self-similar mass fragmentation is to construct an \mathbb{R} -tree that keeps track of the sizes of the blocks of the fragmentation process as time progresses.

For a continuum tree (T, μ) and $t \geq 0$, let $T_1(t), T_2(t), \dots$ be the tree components of $\{v \in T: \text{ht}(v) > t\}$, ranked in decreasing order of μ -mass (breaking ties uniformly). We call a continuum random tree (\mathcal{T}, μ) γ -self-similar if for every $t \geq 0$, conditionally on $(\mu(\mathcal{T}_i(t)), i \geq 1)$, $(\mathcal{T}_i(t), i \geq 1)$ has the same law as $(\mu(\mathcal{T}_i(t))^\gamma \mathcal{T}^{(i)}, i \geq 1)$ where the $\mathcal{T}^{(i)}$, $i \geq 1$, are independent copies of \mathcal{T} .

The following summarizes the parts of Theorem 1 and Lemma 5 in [14] that we will need.

Theorem 22 *Let F be a (γ, ν) -self-similar fragmentation with $\gamma > 0$ and ν as above. There exists a γ -self-similar CRT $(\mathcal{T}_{\gamma, \nu}, \mu_{\gamma, \nu})$ such that, writing $F'(t)$ for the decreasing sequence of masses of the connected components of $\{v \in \mathcal{T}_{\gamma, \nu}: \text{ht}(v) > t\}$, the process $(F'(t), t \geq 0)$ has the same law as F . Furthermore, $\mathcal{T}_{\gamma, \nu}$ is a.s. compact.*

The Brownian continuum random tree introduced by Aldous [2] as the scaling limit of conditioned Galton-Watson trees is an example of a self-similar fragmentation tree.

Definition 23 The Brownian CRT is the $1/2$ -self-similar random tree with dislocation measure ν given by

$$\int_{\mathcal{S}^\downarrow} f(s) \nu(ds) = \int_{1/2}^1 \sqrt{\frac{2}{\pi s_1^3 (1 - s_1)^3}} f(s_1, 1 - s_1, 0, 0, \dots) ds_1.$$

One of our main tools will be the general theory of scaling limits of unordered Markov branching trees. In particular, we make use of the following theorem.

Theorem 24 (Theorem 5 in [16]) *Let $(\mathcal{T}_n^\circ, n \geq 1)$ be a Markov branching model based on $(p_n^\circ, n \geq 2)$ as in Section 2.3. Suppose that there is a characteristic pair (γ, ν) with $\gamma > 0$, and ν satisfying the conditions at the start of Section 4.2 as well as a function $\ell: (0, \infty) \rightarrow (0, \infty)$, slowly varying at ∞ such that, in the sense of weak convergence of finite measures on \mathcal{S}^\downarrow , we have*

$$n^\gamma \ell(n) (1 - s_1) \bar{p}_n^\circ(ds) \rightarrow (1 - s_1) \nu(ds), \quad (13)$$

where \bar{p}_n° is the push-forward of the measure on \mathcal{P}_n° with probability function p_n° onto \mathcal{S}^\downarrow by the map

$$(n_1, \dots, n_p) \mapsto \left(\sum_{i=1}^p n_i \right)^{-1} (n_1, \dots, n_p, 0, 0, \dots).$$

If we view \mathcal{T}_n° as a random element of \mathcal{M}_w with the graph distance and the uniform probability measure its leaves, then we have the convergence in distribution

$$\frac{1}{n^\gamma \ell(n)} \mathcal{T}_n^\circ \rightarrow \mathcal{T}_{\gamma, \nu},$$

with respect to the rooted Gromov-Hausdorff-Prokhorov topology.

5 Scaling limits of regenerative tree growth processes

While every dislocation measure κ on \mathcal{P} gives rise to a regenerative tree growth process, not every such process has a scaling limit. Examples without scaling limit include the $(0, \theta)$ -tree growth process studied in [28, Proposition 13], where the growth is logarithmic and the branching structure degenerates under logarithmic scaling. In view of Proposition 11, it makes sense to try interpreting the hypotheses of Theorem 24 in terms of κ . In particular, let us examine the LHS of (13). From Proposition 11 we see that for a bounded continuous function f on \mathcal{S}^\downarrow ,

$$\begin{aligned} & \int_{\mathcal{S}^\downarrow} f(s) n^\gamma \ell(n) (1 - s_1) \bar{p}_n^\circ(ds) \\ &= \sum_{n_1 \geq \dots \geq n_k: n_1 + \dots + n_k = n} n^\gamma \ell(n) p_n^\circ(n_1, \dots, n_k) \left(1 - \frac{n_1}{n}\right) f\left(\frac{n_1}{n}, \dots, \frac{n_k}{n}, 0, \dots\right) \\ &= \frac{n^\gamma \ell(n)}{\lambda_n} \sum_{\pi \in \mathcal{P}_n \setminus \{1_{[n]}\}} \kappa(\mathcal{P}^\pi) \left(1 - \frac{(\#\pi)_1^\downarrow}{n}\right) f\left(\frac{(\#\pi)^\downarrow}{n}\right) \\ &= \frac{n^\gamma \ell(n)}{\lambda_n} \int_{\mathcal{P}} \left(1 - |\Gamma^{[n]}|_1^\downarrow\right) f\left(|\Gamma^{[n]}|^\downarrow\right) \kappa(d\Gamma), \end{aligned}$$

where now we write $(\#\pi)^\downarrow$ for the decreasing rearrangement of the block sizes of π , with an infinite string of zeros appended (whereas in our previous usage $(\#\pi)^\downarrow$ was a finite vector). Given this expression and the convergence (13) we need to establish, natural assumptions on κ become that $\lambda_n = n^\gamma \ell(n)$ for some $\gamma > 0$ and $\ell(n)$ slowly varying at ∞ and that

$$\lim_{n \rightarrow \infty} \int_{\mathcal{P}} \left(1 - |\Gamma^{[n]}|_1^\downarrow\right) f\left(|\Gamma^{[n]}|^\downarrow\right) \kappa(d\Gamma) = \int_{\mathcal{P}} \left(1 - |\Gamma|_1^\downarrow\right) f\left(|\Gamma|^\downarrow\right) \kappa(d\Gamma).$$

Of course, for this last equation to have hope of holding, we must assume that κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies. This holds for exchangeable and restricted exchangeable κ , and when κ is partially exchangeable in the sense of [26]. Let us, though, clarify the relationship between the existence of asymptotic frequencies and the existence of asymptotic ranked frequencies.

Lemma 25 *Existence of $|\Gamma|_i^\downarrow$ for all $i \geq 1$ holds if and only if $(|\Gamma_i|, i \geq 1)$ exists as a uniform limit. In this case, $(|\Gamma|_i^\downarrow, i \geq 1)$ is the decreasing rearrangement of $(|\Gamma_i|, i \geq 1)$, which we write as $|\Gamma|^\downarrow$.*

Lemma 25 is inessential to the remainder of our results, but for completeness we include a proof in Appendix B. Note that asymptotic (ranked) frequencies need not be in \mathcal{S}_1^\downarrow and that $|\Gamma_i|$ may vanish. The partition $\Gamma = (\{2^{i-1}, \dots, 2^i - 1\}, i \geq 1)$ is an example where $|\Gamma_i|$, $i \geq 1$, exists, but $|\Gamma|_1^\downarrow$ does not.

We can now give our main result on the existence of scaling limits of regenerative growth processes, which contains the statement of Theorem 4.

Theorem 26 *Let $(\mathcal{T}_n, n \geq 1)$ be a regenerative tree growth process associated with a dislocation measure κ . Assume that κ -a.e. $\Gamma \in \mathcal{P}$ has proper asymptotic ranked frequencies in $\mathcal{S}_1^\downarrow \setminus \{(1, 0, \dots)\}$, define ν to be the push-forward of κ under $\Gamma \mapsto |\Gamma|^\downarrow$ and suppose $\int_{\mathcal{S}^\downarrow} (1 - s_1) \nu(ds) < \infty$ and $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$. Then the following are equivalent:*

(i) *For all bounded continuous $f: \mathcal{S}^\downarrow \rightarrow [0, \infty)$,*

$$\int_{\mathcal{P}} \left(1 - |\Gamma^{[n]}|_1^\downarrow\right) f\left(|\Gamma^{[n]}|^\downarrow\right) \kappa(d\Gamma) \rightarrow \int_{\mathcal{S}^\downarrow} (1 - s_1) f(s) \nu(ds), \quad \text{as } n \rightarrow \infty;$$

(ii) *$\int_{\mathcal{P}} \left(|\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow\right) \kappa(d\Gamma) \rightarrow 0$, as $n \rightarrow \infty$;*

(iii) *$\int_{\mathcal{P}} \left||\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow\right| \kappa(d\Gamma) \rightarrow 0$, as $n \rightarrow \infty$, i.e. the convergence $|\Gamma^{[n]}|_1^\downarrow \rightarrow |\Gamma|_1^\downarrow$ holds in $L_1(\kappa)$.*

If condition (ii) holds, then $\frac{\mathcal{T}_n^\circ}{n^\gamma \ell(n)} \rightarrow \mathcal{T}_{\gamma, \nu}$ in distribution, as $n \rightarrow \infty$, in the rooted GHP sense.

Proof. If (i) holds, we obtain (ii) as a rearrangement of the special case $f = 1$. Now assume (ii). Let us prove (iii). The main difficulty arises from the possibility that $\kappa(\mathcal{P}^{[m]}) = \infty$. For all $m \geq 1$ we have

$$\int_{\mathcal{P}} \left||\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow\right| \kappa(d\Gamma) = \int_{\mathcal{P}^{[m]}} \left||\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow\right| \kappa(d\Gamma) + \int_{\mathcal{P} \setminus \mathcal{P}^{[m]}} \left||\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow\right| \kappa(d\Gamma).$$

Since $\kappa(\mathcal{P} \setminus \mathcal{P}^{[m]}) < \infty$ and κ -a.e. Γ has asymptotic ranked frequencies an application of dominated convergence shows that, for each fixed m , the second term vanishes as $n \rightarrow \infty$. From the triangle inequality, we see that

$$\begin{aligned} \int_{\mathcal{P}^{[m]}} \left||\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow\right| \kappa(d\Gamma) &= \int_{\mathcal{P}^{[m]}} \left|(1 - |\Gamma^{[n]}|_1^\downarrow) - (1 - |\Gamma|_1^\downarrow)\right| \kappa(d\Gamma) \\ &\leq \int_{\mathcal{P}^{[m]}} (1 - |\Gamma^{[n]}|_1^\downarrow) \kappa(d\Gamma) + \int_{\mathcal{P}^{[m]}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma) \\ &= \int_{\mathcal{P}} (1 - |\Gamma^{[n]}|_1^\downarrow) \kappa(d\Gamma) - \int_{\mathcal{P} \setminus \mathcal{P}^{[m]}} (1 - |\Gamma^{[n]}|_1^\downarrow) \kappa(d\Gamma) + \int_{\mathcal{P}^{[m]}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma). \end{aligned}$$

It follows from (ii) that

$$\lim_{n \rightarrow \infty} \int_{\mathcal{P}} (1 - |\Gamma^{[n]}|_1^\downarrow) \kappa(d\Gamma) = \int_{\mathcal{P}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma)$$

and, since $\kappa(\mathcal{P} \setminus \mathcal{P}^{[m]}) < \infty$, dominated convergence shows that

$$\lim_{n \rightarrow \infty} \int_{\mathcal{P} \setminus \mathcal{P}^{[m]}} (1 - |\Gamma^{[n]}|_1^\downarrow) \kappa(d\Gamma) = \int_{\mathcal{P} \setminus \mathcal{P}^{[m]}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma).$$

Consequently, for every $m \geq 1$ we have

$$\limsup_{n \rightarrow \infty} \int_{\mathcal{P}} \left| |\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow \right| \kappa(d\Gamma) \leq \limsup_{n \rightarrow \infty} \int_{\mathcal{P}^{[m]}} \left| |\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow \right| \kappa(d\Gamma) \leq 2 \int_{\mathcal{P}^{[m]}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma).$$

Since $\int_{\mathcal{P}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma) < \infty$, while $\bigcap_{m \geq 1} \mathcal{P}^{[m]} = \{1_{\mathbb{N}}\}$ and $\kappa(\{1_{\mathbb{N}}\}) = 0$, the infimum of these bounds over $m \geq 1$ vanishes and (iii) follows.

Now assume (iii). If $f: \mathcal{S}^\downarrow \rightarrow [0, \infty)$ is continuous and bounded, then

$$\begin{aligned} & \left| \int_{\mathcal{P}} (1 - |\Gamma^{[n]}|_1^\downarrow) f(|\Gamma^{[n]}|^\downarrow) \kappa(d\Gamma) - \int_{\mathcal{P}} (1 - |\Gamma|_1^\downarrow) f(|\Gamma|^\downarrow) \kappa(d\Gamma) \right| \\ & \leq \int_{\mathcal{P}} \left| |\Gamma^{[n]}|_1^\downarrow - |\Gamma|_1^\downarrow \right| \left| f(|\Gamma^{[n]}|^\downarrow) - f(|\Gamma|^\downarrow) \right| \kappa(d\Gamma) + \int_{\mathcal{P}} (1 - |\Gamma|_1^\downarrow) \left| f(|\Gamma^{[n]}|^\downarrow) - f(|\Gamma|^\downarrow) \right| \kappa(d\Gamma), \end{aligned}$$

and (i) follows from (iii) by dominated convergence, since κ -a.e. Γ has asymptotic ranked frequencies, since f is bounded and continuous, and since ν is the push-forward of κ under $\Gamma \mapsto |\Gamma|^\downarrow$.

The last part follows from Haas and Miermont [16, Theorem 5], which we formulated in Theorem 24 above. \square

6 Residual mass processes in regenerative tree growth processes

Let $(\mathcal{T}_n, n \geq 1)$ be a regenerative tree growth process and $(X_m^{(n)}, m \geq 0)$ the associated residual mass processes of label 1 in \mathcal{T}_n , $n \geq 1$, with transition probabilities

$$\mathbb{P}(X_1^{(n)} = k) = \sum_{\pi = (B_1, \dots, B_k) \in \mathcal{P}_n: \#B_1 = k} p_n(\pi) = \frac{1}{\lambda_n} \kappa \left(\left\{ \Gamma \in \mathcal{P}: \#\Gamma_1^{[n]} = k \right\} \right), \quad 1 \leq k < n,$$

as identified in Proposition 7, with $\lambda_n = \kappa(\{\Gamma \in \mathcal{P}: \Gamma_1^{[n]} \neq [n]\})$. The existence of a scaling limit for trees \mathcal{T}_n° as studied in Section 5 does not imply the existence of a scaling limit for associated residual mass processes $X^{(n)}$, in general (see Example 19). In this section, we study scaling limits $X_{\lfloor \lambda_n t \rfloor}^{(n)} / n \rightarrow X_t$, as $n \rightarrow \infty$. Since, for fixed $n \geq 1$, $(X_m^{(n)}, m \geq 0)$ is a non-increasing Markov chain with $X_0^{(n)} = n$, we can make use of the general theory of self-similar scaling limits for such chains that was recently developed in [15].

Theorem 27 (Theorems 1 and 2 in [15]) *Let $p = (p_{ij}, 0 \leq j \leq i)$ be a transition matrix, and for each $n \geq 1$ let $(Y_m^{(n)}, m \geq 0)$ be a Markov chain with transition matrix p such that $Y_0^{(n)} = n$. Define*

$$p_n^*(dx) = \sum_{k=0}^n p_{n,k} \delta_{k/n}(dx).$$

Suppose that there exists a sequence $(a_n, n \geq 0)$ of the form $a_n = n^\gamma \ell(n)$ for some $\gamma > 0$ and a slowly varying function ℓ as well as a non-zero finite measure μ on $[0, 1]$ such that

$$a_n(1 - x)p_n^*(dx) \rightarrow \mu(dx) \tag{14}$$

in the sense of weak convergence of finite measures on $[0, 1]$. Then we have the following convergence in distribution

$$\left(\frac{Y_{\lfloor a_n t \rfloor}^{(n)}}{n}, t \geq 0 \right) \rightarrow (X_t, t \geq 0)$$

in the Skorokhod sense, where X is a self-similar Markov process and in Lamperti's representation (4), we have

$$\mathbb{E}(e^{-s\xi_r}) = \exp(-r\psi(s)) \quad \text{with} \quad \psi(s) = \int_{[0,1]} \frac{1 - x^s}{1 - x} \mu(dx).$$

Moreover, letting A_n be the absorption time of $Y^{(n)}$ at 0, the above convergence in distribution happens jointly with the convergence of $a_n^{-1}A_n$ to the absorption time at 0 of the limiting process.

Proposition 7 shows that the residual mass process of the leaf $\{1\}$ falls into the scope of this theorem with $p_{ij} = \mathbb{P}(X_1^{(i)} = j)$. Given a dislocation measure κ with $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$ regularly varying, as $n \rightarrow \infty$, for some $\gamma > 0$, and taking $a_n = \lambda_n$, the LHS of condition (14) for the residual mass process of leaf $\{1\}$ becomes

$$\sum_{\pi=(B_1, \dots, B_k) \in \mathcal{P}_n} n^\gamma \ell(n) p_n(\pi) \left(1 - \frac{\#B_1}{n}\right) f\left(\frac{\#B_1}{n}\right) = \int_{\mathcal{P}} \left(1 - |\Gamma_1^{[n]}|\right) f\left(|\Gamma_1^{[n]}|\right) \kappa(d\Gamma).$$

Theorem 28 *Let $(\mathcal{T}_n, n \geq 1)$ be a regenerative tree growth process with dislocation measure κ and $X^{(n)}$ the residual mass process of $\{1\}$ in \mathcal{T}_n . Assume that the first block Γ_1 of κ -a.e. $\Gamma \in \mathcal{P}$ has an asymptotic frequency $|\Gamma_1| \in (0, 1)$, define Λ as push-forward of κ under $\Gamma \mapsto -\log(|\Gamma_1|)$. If $\int_{(0, \infty)} (1 - e^{-x}) \Lambda(dx) < \infty$ and $\lambda_n = \kappa(\mathcal{P} \setminus \mathcal{P}^{[n]}) = n^\gamma \ell(n)$, then the following are equivalent:*

(i) *For all bounded continuous $f: [0, 1] \rightarrow [0, \infty)$,*

$$\int_{\mathcal{P}} \left(1 - |\Gamma_1^{[n]}|\right) f\left(|\Gamma_1^{[n]}|\right) \kappa(d\Gamma) \rightarrow \int_{(0, \infty)} f(e^{-y})(1 - e^{-y}) \Lambda(dy), \text{ as } n \rightarrow \infty;$$

(ii) *$\int_{\mathcal{P}} \left(|\Gamma_1^{[n]}| - |\Gamma_1|\right) \kappa(d\Gamma) \rightarrow 0$, as $n \rightarrow \infty$;*

(iii) *$\int_{\mathcal{P}} \left||\Gamma_1^{[n]}| - |\Gamma_1|\right| \kappa(d\Gamma) \rightarrow 0$, as $n \rightarrow \infty$, i.e. the convergence $|\Gamma_1^{[n]}| \rightarrow |\Gamma_1|$ holds in $L_1(\kappa)$.*

If condition (ii) holds, $X_{[\lambda_n t]}^{(n)}/n \rightarrow X_t$ in distribution, as $n \rightarrow \infty$, in the Skorohod sense as functions of $t \geq 0$, where X is a self-similar Markov process and $\mathbb{E}(e^{-s\xi_r}) = \exp(-r \int_{(0, \infty)} (1 - e^{-sy}) \Lambda(dy))$ in Lamperti's representation (4). Moreover, letting A_n be the absorption time of $X^{(n)}$ at 0, the above convergence in distribution happens jointly with the convergence of $\lambda_n^{-1} A_n$ to the absorption time at 0 of the limiting process.

If, in addition, κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies, then (3) holds and we have $\frac{\mathcal{T}_n^o}{n^\gamma \ell(n)} \rightarrow \mathcal{T}_{\gamma, \nu}$ in distribution, as $n \rightarrow \infty$, in the rooted GHP sense.

We note that the statement of Theorem 5 is contained in the statement of this theorem.

Proof. The proof of the equivalences is the same as for Theorem 26, with $|\Gamma_1^{[n]}|_1^\downarrow$ and $|\Gamma_1|_1^\downarrow$ replaced by $|\Gamma_1^{[n]}|$ and $|\Gamma_1|$. Convergence of $X_{[\lambda_n t]}^{(n)}/n$ is now an application of [15, Theorem 1], which we formulated as Theorem 27 above. Finally, if κ -a.e. $\Gamma \in \mathcal{P}$ has asymptotic ranked frequencies, we first note that in the notation of Theorem 26

$$\int_{S^\downarrow} (1 - s_1) \nu(ds) = \int_{\mathcal{P}} (1 - |\Gamma_1|_1^\downarrow) \kappa(d\Gamma) \leq \int_{\mathcal{P}} (1 - |\Gamma_1|) \kappa(d\Gamma) = \int_{(0, \infty)} (1 - e^{-x}) \Lambda(dx) < \infty. \quad (15)$$

To apply Theorem 26, we verify condition (ii) of Theorem 26:

$$\begin{aligned} \int_{\mathcal{P}} \left(|\Gamma_1^{[n]}|_1^\downarrow - |\Gamma_1|_1^\downarrow\right) \kappa(d\Gamma) &= \int_{\{|\Gamma_1| = |\Gamma_1|_1^\downarrow\}} \left(|\Gamma_1^{[n]}| - |\Gamma_1|\right) \kappa(d\Gamma) + \int_{\{|\Gamma_1| \neq |\Gamma_1|_1^\downarrow\}} \left(|\Gamma_1^{[n]}|_1^\downarrow - |\Gamma_1|_1^\downarrow\right) \kappa(d\Gamma) \\ &\quad + \int_{\{|\Gamma_1| = |\Gamma_1|_1^\downarrow\}} \left(|\Gamma_1^{[n]}|_1^\downarrow - |\Gamma_1^{[n]}|\right) \kappa(d\Gamma) \end{aligned} \quad (16)$$

is a sum of three terms. The first term vanishes as $n \rightarrow \infty$ by (iii). The second term vanishes as $n \rightarrow \infty$ since $\kappa(|\Gamma_1| \neq |\Gamma_1|_1^\downarrow) < \infty$: if $|\Gamma_1| \neq |\Gamma_1|_1^\downarrow$, then one of them must be less than $1/2$, so $\kappa(|\Gamma_1| \neq |\Gamma_1|_1^\downarrow) \leq \kappa(|\Gamma_1| \leq 1/2) + \kappa(|\Gamma_1|_1^\downarrow \leq 1/2) < \infty$, by (15). The third term is non-negative, so that $\liminf_{n \rightarrow \infty} \text{LHS} \geq 0$ in (16). In

$$\int_{\mathcal{P}} \left(|\Gamma_1^{[n]}|_1^\downarrow - |\Gamma_1|_1^\downarrow\right) \kappa(d\Gamma) \leq \int_{\mathcal{P}^{[m]}} (1 - |\Gamma_1|) \kappa(d\Gamma) + \int_{\mathcal{P} \setminus \mathcal{P}^{[m]}} \left(|\Gamma_1^{[n]}|_1^\downarrow - |\Gamma_1|_1^\downarrow\right) \kappa(d\Gamma)$$

we can make the first term small by choosing m large and the second term vanishes as $n \rightarrow \infty$, for each fixed $m \geq 1$. Hence, $\limsup_{n \rightarrow \infty} \text{LHS} \leq 0$ and so $\lim_{n \rightarrow \infty} \text{LHS} = 0$. \square

Residual mass convergence and tree convergence are not equivalent. The last part of Theorem 28 finds that under the conditions for residual mass convergence in this theorem, we just need to assume the existence of asymptotic ranked frequencies to also obtain tree convergence. Example 19 demonstrates that residual mass convergence does not follow from tree convergence. In the following corollary we explore additional conditions in the tree convergence setting of Theorem 26, under which we also obtain residual mass convergence. Roughly speaking, condition (ii) below expresses the following intuition: we need label 1 in the asymptotically largest block most of the time, and on the corresponding set $\{|\Gamma|_1^\downarrow = |\Gamma_1|\}$ of infinite κ -measure, $|\Gamma_1^{[n]}|^\downarrow$ and $|\Gamma_1^{[n]}|$ approach their limit $|\Gamma|_1^\downarrow = |\Gamma_1|$ in a sufficiently regular way. The following statement includes Corollary 6.

Corollary 29 *In the setting of Theorem 26 (ii), the block Γ_1 containing 1 of κ -a.e. $\Gamma \in \mathcal{P}$ has an asymptotic frequency in $(0, 1)$. With Λ as in Theorem 28, the following are equivalent:*

- (i) $\int_{(0, \infty)} (1 - e^{-x}) \Lambda(dx) < \infty$ and $\int_{\mathcal{P}} (|\Gamma_1^{[n]}| - |\Gamma_1|) \kappa(d\Gamma) \rightarrow 0$;
- (ii) $\kappa(|\Gamma_1| \neq |\Gamma|_1^\downarrow) < \infty$ and $\int_{\{|\Gamma|_1^\downarrow = |\Gamma_1|\}} (|\Gamma_1^{[n]}|^\downarrow - |\Gamma_1^{[n]}|) \kappa(d\Gamma) \rightarrow 0$.

If condition (ii) holds, then $X_{[\lambda_n t]}^{(n)}/n \rightarrow X_t$ in distribution, as $n \rightarrow \infty$, in the Skorohod sense as functions of $t \geq 0$, and this convergence holds jointly with the convergence of $\lambda_n^{-1} A_n$ to the absorption time of X at 0, where our notation is as in Theorem 28.

Proof. Since Γ has asymptotic ranked frequencies, Γ_1 has an asymptotic frequency by Lemma 25, for κ -a.e. $\Gamma \in \mathcal{P}$. “(i) \Rightarrow (ii)” follows straight from the proof of Theorem 28, since (i) puts us into that setting; and also, the convergence of $X_{[\lambda_n t]}^{(n)}/n$ holds under (i). It remains to prove “(ii) \Rightarrow (i)”, so we note that under (ii),

$$\int_{(0, \infty)} (1 - e^{-x}) \Lambda(dx) = \int_{\mathcal{P}} (1 - |\Gamma_1|) \kappa(d\Gamma) \leq \kappa(|\Gamma_1| \neq |\Gamma|_1^\downarrow) + \int_{\mathcal{P}} (1 - |\Gamma|_1^\downarrow) \kappa(d\Gamma) < \infty \quad (17)$$

and $\int_{\mathcal{P}} (|\Gamma_1^{[n]}| - |\Gamma_1|) \kappa(d\Gamma) \rightarrow 0$ follows by the same argument as convergence in (16), with roles of $(|\Gamma_1|, |\Gamma_1^{[n]}|)$ and $(|\Gamma|_1^\downarrow, |\Gamma_1^{[n]}|^\downarrow)$ interchanged, using Theorem 26(iii) for the first term and using the second condition under (ii) here for the third term of the modification of (16). \square

7 Further problems and related work

Due to the coupling of $(\mathcal{T}_n, n \geq 1)$ in a regenerative tree growth process, the convergence in distribution in Theorems 26 and 28 should be strengthened to a convergence in probability or even to almost sure convergence in all cases discussed here. We have proved tree convergence in probability in the exchangeable case [17], and in the restricted exchangeable case [6] provided that $\nu_j = \nu_m$, $j \geq m$, but the general case including the alpha-theta model remains open.

In the alpha-theta model [28] and the (restricted) exchangeable [17, 6] cases, we have established a two-stage almost sure convergence to a self-similar tree \mathcal{T} by passing via reduced subtrees of \mathcal{T}_n and of \mathcal{T} spanned by the first k labelled leaves and letting first $n \rightarrow \infty$ and then $k \rightarrow \infty$. More specifically, we have embedded $(\mathcal{T}_n, n \geq 1)$ in \mathcal{T} as discrete trees with edge lengths.

The basic embedding problem is to find a random leaf in a self-similar tree (\mathcal{T}, μ) that induces a given decreasing self-similar Markov process as residual mass process, i.e. as the process that is parametrised by distance from the root on the path to the random leaf and that records for each point on the path the μ -mass in the subtree above the point. Another interesting structure is the joint distribution of two residual mass processes (see [28, 29]). When embedded in the same

tree, they coincide up to a branch point and then evolve independently. In [29], we use the terms *fragmenter* for exponential subordinators ($e^{-\xi_s}, s \geq 0$), which are time-changed in Lamperti's representation (4), and *bifurcator* for pairs of fragmenters that coincide up to an exponential time and then evolve independently. In [29] we investigate the fact that not all fragmenters appear as residual mass processes of typical (uniformly random) leaves. We introduce the notion of Markovian embedding in an exchangeable fragmentation process and show that for every (pure-jump) fragmenter X there is a unique exchangeable dislocation measure κ such that X has a Markovian embedding into an associated exchangeable fragmentation process.

In [28, 29], we study an autonomous description of the evolution of reduced subtrees, viewed as weighted trees equipped with an (atomic) measure on the branches. We refer to a single branch with an atomic measure as a *string of beads*, see also [25] for related structures. We refer to the evolution of reduced subtrees as *bead splitting*. In [29], we study certain binary bead splitting processes that evolve by size-biased branching, i.e. where an atom (a bead) is selected at random according to the measure on the branches and replaced by a (rescaled independent) copy of a given string of beads. We study the convergence of bead-splitting processes to self-similar CRTs.

A Proof of Proposition 2

First consider a regenerative tree growth rule, i.e. a sequence of transition probability matrices g_n from $\mathcal{P}_n \setminus \{1_{[n]}\}$ to $\{0, \dots, n+1\}$ with $g_n(\pi, 0)$ independent of π and $g_n(\pi, i) = 0$ if π has strictly fewer than $i-1$ blocks. For $n=1$ and $n=2$ the regenerative property is trivial. Consider the induction hypothesis that the growth rule gives rise to distributions Q_m on \mathbb{T}_m , $m \leq n$, and hence to Q_B on \mathbb{T}_B after relabelling via the increasing bijection $[m] \rightarrow B$, for all $B \subset \mathbb{N}$ with $\#B = m$, such that conditionally given a first split $\Pi_n = (B_1, \dots, B_k)$, the subtrees above the first split are independent, and the i th subtree \mathcal{T}_{n, B_i} has conditional distribution Q_{B_i} , $1 \leq i \leq k$. For the induction step, note that conditionally given $\Pi_n = (B_1, \dots, B_k)$, the tree growth step from n to $n+1$ specifies Q_{n+1} on each of the events $G_{n,i}$:

- $G_{n,0}$: here, $\Pi_{n+1} = ([n], \{n+1\})$ is not related to Π_n ; we will get back to this;
- $G_{n,i}$, $1 \leq i \leq k$: here, $\Pi_{n+1} = (B_1, \dots, B_{i-1}, B_i \cup \{n+1\}, B_{i+1}, \dots, B_k)$; regenerative growth in the i th subtree preserves the conditional independence of subtrees, and the induction hypothesis also yields that $\mathcal{T}_{n+1, B_j} = \mathcal{T}_{n, B_j}$ has conditional distribution Q_{B_j} for $j \neq i$, while $\mathcal{T}_{n+1, B_i \cup \{n+1\}}$ has conditional distribution $Q_{B_i \cup \{n+1\}}$ obtained from Q_{B_i} via the growth rule applied to B_i with $\#B_i \leq n-1$.
- $G_{n,k+1}$: here, $\Pi_{n+1} = (B_1, \dots, B_k, \{n+1\})$, and conditional independence of subtrees as well as conditional distributions follow from the induction hypothesis, with the addition of $\mathcal{T}_{n+1, \{n+1\}}$ with (degenerate) conditional distribution $Q_{\{n+1\}}$.

Conditionally given $\Pi_n = (B_1, \dots, B_k)$, the events $G_{n,i}$ for $i \in \{k+2, \dots, n+1\}$ have probability zero, since $g_n(\pi, i) = 0$ if π has strictly fewer than $i-1$ blocks. Hence, Q_{n+1} is fully specified and satisfies the regenerative property for each of the $i \geq 1$; for the remaining $i=0$ case, we cannot work conditionally given $\Pi_n = (B_1, \dots, B_k)$, because the regenerative property here is a statement conditionally given $\Pi_{n+1} = ([n], \{n+1\})$, and indeed, since $g_n(\pi, 0)$ does not depend on $\pi \in \mathcal{P}_n \setminus \{1_{[n]}\}$, the subtree $\mathcal{T}_{n+1, [n]} = \mathcal{T}_n$ has conditional distribution $Q_{[n]} = Q_n$, while $\mathcal{T}_{n+1, \{n+1\}}$ has conditional distribution $Q_{\{n+1\}}$. The induction proceeds.

To prove the other direction, let $(\mathcal{T}_n, n \geq 1)$ be a regenerative tree growth process. Consistency implies that the splitting rules $p_n(\pi) = \mathbb{P}(\Pi_n = \pi)$ satisfy, for $\pi = (B_1, \dots, B_k) \in \mathcal{P}_n \setminus \{1_{[n]}\}$,

$$p_n(B_1, \dots, B_k) = p_{n+1}([n], \{n+1\})p_n(B_1, \dots, B_k) + \sum_{i=1}^{k+1} p_{n+1}(B_1, \dots, B_i \cup \{n+1\}, \dots, B_k).$$

For $g_n(0)$ and $g_n(\pi, i)$ defined from $(p_n, n \geq 2)$ via (6), this implies that $g_n(\pi, i) \in [0, 1]$ and

$$\sum_{i=0}^{k+1} g_n(\pi, i) = p_{n+1}([n], \{n+1\}) + \sum_{i=1}^{k+1} \frac{p_{n+1}(B_1, \dots, B_i \cup \{n+1\}, \dots, B_k)}{p_n(B_1, \dots, B_k)} = 1.$$

Also, (8) holds and determines $\mathbb{P}(\mathcal{T}_n = \mathbf{t} \mid \mathcal{T}_{n-1})$ as required, since \mathcal{T}_n determines \mathcal{T}_{n-1} . \square

B Proof of Lemma 25

We consider the set

$$c_0 = \left\{ (s_1, s_2, \dots) \in [0, 1]^{\mathbb{N}} : \lim_{i \rightarrow \infty} s_i = 0 \right\},$$

which is equipped with the uniform norm $\|\cdot\|_{\infty}$. This set is clearly closed when considered as a subset of ℓ^{∞} and thus is a complete metric space. Let $F: c_0 \rightarrow c_0$ be the map defined by $F(\mathbf{s}) = \mathbf{s}^{\downarrow}$, that is, F is the map that takes a sequence to its non-increasing rearrangement. Our first step is to prove that F is continuous since this immediately implies that if $|\Gamma^{[n]}|$ converges uniformly, say to $(y_i)_{i \geq 1}$, then $|\Gamma^{[n]}|^{\downarrow}$ converges to the non-increasing rearrangement of $(y_i)_{i \geq 1}$.

Fix $\epsilon > 0$ and $\mathbf{s} \in c_0$. Without loss of generality, we may assume that $\epsilon < \sup_i s_i$ and $\epsilon \notin \{s_i, i \geq 1\}$. Let $B_{\mathbf{s}} = \{i \geq 1 : s_i > \epsilon\}$. The fact that $\mathbf{s} \in c_0$ implies that $\#B_{\mathbf{s}} < \infty$. Observe that $F(\mathbf{s})$ is equal to the sequence obtained by concatenating the non-increasing rearrangement of $(s_i : i \in B_{\mathbf{s}})$ with the non-increasing rearrangement of $(s_i : i \notin B_{\mathbf{s}})$.

Suppose that $\mathbf{s}_n \rightarrow \mathbf{s}$. For sufficiently large n we have $B_{\mathbf{s}_n} = B_{\mathbf{s}}$. Since ranking is continuous on $\mathbb{R}^{\#B_{\mathbf{s}}}$, it follows that

$$\lim_{n \rightarrow \infty} (s_{n,1}^{\downarrow}, \dots, s_{n,\#B_{\mathbf{s}}}^{\downarrow}) = (s_1^{\downarrow}, \dots, s_{\#B_{\mathbf{s}}}^{\downarrow})$$

and also that

$$\sup_{i > \#B_{\mathbf{s}}} s_i^{\downarrow} + \limsup_{n \rightarrow \infty} \sup_{i > \#B_{\mathbf{s}}} s_{n,i}^{\downarrow} \leq 2\epsilon.$$

As a result we have

$$\limsup_{n \rightarrow \infty} \|F(\mathbf{s}) - F(\mathbf{s}_n)\|_{\infty} \leq 2\epsilon,$$

and the continuity of F follows.

We now prove the opposite direction. To that end, assume that $|\Gamma^{[n]}|^{\downarrow} \rightarrow (x_i)_{i \geq 1}$ pointwise. We will prove that $|\Gamma^{[n]}|$ converges in c_0 and the proof of the previous part then identifies the limit. Since $|\Gamma^{[n]}|^{\downarrow}$ is non-increasing for each n with sums uniformly bounded by 1, this implies that $|\Gamma^{[n]}|^{\downarrow} \rightarrow (x_i)_{i \geq 1}$ uniformly. If $x_1 = 0$ we are done, so we assume that $x_1 > 0$. Let $\epsilon > 0$ be given, and without loss of generality suppose that $\epsilon < x_1$. By Fatou's lemma we have $\sum_{i \geq 1} x_i \leq 1 < \infty$ and, consequently, we can choose K so that $\sum_{i \geq K+1} x_i < \epsilon$. Let

$$\epsilon_1 = \epsilon \wedge \min \left\{ \frac{|x_i - x_j|}{3} : 1 \leq i, j \leq K+1 \text{ and } x_i \neq x_j \right\}.$$

Since $|\Gamma^{[n]}|^{\downarrow} \rightarrow (x_i)_{i \geq 1}$ uniformly, we can choose $N > 1/\epsilon_1$ such that $\sup_i \left| |\Gamma^{[n]}|^{\downarrow}_i - x_i \right| < \epsilon_1$ for all $n \geq N$. For each $n \geq N$ let $\sigma_n: \mathbb{N} \rightarrow \mathbb{N}$ be a bijection such that $(|\Gamma^{[n]}|_{\sigma_n(i)})_{i \geq 1} = |\Gamma^{[n]}|^{\downarrow}$. Note that we have used the fact that $|\Gamma^{[n]}|$ has only finitely many non-zero entries to obtain this bijection. Since $N > 1/\epsilon_1$, for all $i \geq 1$ and $n \geq N$ we have

$$\left| |\Gamma^{[n]}|_i - |\Gamma^{[n+1]}|_i \right| \leq 1/(n+1) < \epsilon_1.$$

It follows that for $n \geq N$

$$\sup_i \left| |\Gamma^{[n+1]}|_{\sigma_n(i)} - x_i \right| \leq 2\epsilon_1.$$

By our choice of ϵ_1 , for any $1 \leq j \leq K$ and any $i \geq 1$ such that $x_j \neq x_i$ we have

$$\left| |\Gamma^{[n+1]}|_{\sigma_n(i)} - x_j \right| \geq |x_i - x_j| - \left| |\Gamma^{[n+1]}|_{\sigma_n(i)} - x_i \right| \geq \epsilon_1.$$

However, since $\sup_i \left| |\Gamma^{[n]}|_i^\downarrow - x_i \right| < \epsilon_1$ for all $n \geq N$, this implies that

$$\sup_{1 \leq i \leq K} \left| |\Gamma^{[n+1]}|_{\sigma_n(i)} - x_i \right| < \epsilon_1 \quad \text{and} \quad \sup_{i \geq K+1} |\Gamma^{[n+1]}|_{\sigma_n(i)} < x_{K+1} + \epsilon_1.$$

Inductively, we conclude that for all $n \geq N$ and $k \geq 0$

$$\sup_{1 \leq i \leq K} \left| |\Gamma^{[n+k]}|_{\sigma_n(i)} - x_i \right| < \epsilon_1 \quad \text{and} \quad \sup_{i \geq K+1} |\Gamma^{[n+k]}|_{\sigma_n(i)} < x_{K+1} + \epsilon_1.$$

Combining these, we see that for $n \geq N$

$$\sup_{i \geq 1} \left| |\Gamma^{[n]}|_i - x_{\sigma_N^{-1}(i)} \right| = \sup_{i \geq 1} \left| |\Gamma^{[n]}|_{\sigma_N(i)} - x_i \right| < 2\epsilon.$$

We are not quite done since σ_N depends on ϵ_1 . Note, however, that the above inequality implies for $n, m \geq N$

$$\sup_{i \geq 1} \left| |\Gamma^{[n]}|_i - |\Gamma^{[m]}|_i \right| < 4\epsilon.$$

This shows that $(|\Gamma^{[n]}|)_{n \geq 1}$ is a Cauchy sequence in the complete metric space c_0 and, therefore, converges uniformly. \square

References

- [1] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.
- [2] D. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.
- [3] D. Aldous. Probability distributions on cladograms. In *Random discrete structures (Minneapolis, MN, 1993)*, volume 76 of *IMA Vol. Math. Appl.*, pages 1–18. Springer, New York, 1996.
- [4] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [5] B. Chen, D. Ford, and M. Winkel. A new family of Markov branching trees: the alpha-gamma model. *Electron. J. Probab.*, 14:no. 15, 400–430 (electronic), 2009.
- [6] B. Chen and M. Winkel. Restricted exchangeable partitions and embedding of associated hierarchies in continuum random trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):839–872, 2013.
- [7] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.
- [8] S. N. Evans, J. Pitman, and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.
- [9] S. N. Evans and A. Winter. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.*, 34(3):918–961, 2006.
- [10] D. J. Ford. Probabilities on cladograms: introduction to the alpha model. *arXiv:math.PR/0511246*, 2005.
- [11] A. Gnedin and J. Pitman. Regenerative composition structures. *Ann. Probab.*, 33(2):445–479, 2005.

- [12] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ -coalescent measure trees). *Probab. Theory Related Fields*, 145(1-2):285–322, 2009.
- [13] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999. Based on the 1981 French original, with appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.
- [14] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab.*, 9:no. 4, 57–97 (electronic), 2004.
- [15] B. Haas and G. Miermont. Self-similar scaling limits of non-increasing Markov chains. *Bernoulli*, 17(4):1217–1247, 2011.
- [16] B. Haas and G. Miermont. Scaling limits of Markov branching trees, with applications to Galton-Watson and random unordered trees. *Ann. Probab.*, 40(6):2589–2666, 2012.
- [17] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Probab.*, 36(5):1790–1837, 2008.
- [18] B. Haas, J. Pitman, and M. Winkel. Spinal partitions and invariance under re-rooting of continuum random trees. *Ann. Probab.*, 37(4):1381–1411, 2009.
- [19] C. Haulk and J. Pitman. A representation of exchangeable hierarchies by sampling from real trees. *arXiv:1101.5619 [math.PR]*, 2011.
- [20] J. Lamperti. Semi-stable Markov processes. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 22:205–225, 1972.
- [21] P. Marchal. A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*, volume AI, pages 489–500. Discrete Mathematics and Theoretical Computer Science, 2008.
- [22] P. McCullagh, J. Pitman, and M. Winkel. Gibbs fragmentation trees. *Bernoulli*, 14(4):988–1002, 2008.
- [23] G. Miermont. Self-similar fragmentations derived from the stable tree. I. Splitting at heights. *Probab. Theory Related Fields*, 127(3):423–454, 2003.
- [24] G. Miermont. Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)*, 42(5):725–781, 2009.
- [25] S. Pal. On the Aldous diffusion on Continuum Trees. I. *arXiv:1104.4186 [math.PR]*, 2011.
- [26] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields*, 102(2):145–158, 1995.
- [27] J. Pitman and D. Rizzolo. Schröder’s problems and scaling limits of random trees. *Trans. Amer. Math. Soc.*, to appear, preprint available at *arXiv:1107.1760 [math.PR]*, 2013.
- [28] J. Pitman and M. Winkel. Regenerative tree growth: binary self-similar continuum random trees and Poisson-Dirichlet compositions. *Ann. Probab.*, 37(5):1999–2042, 2009.
- [29] J. Pitman and M. Winkel. Regenerative tree growth: Markovian embedding of fragmenters, bifurcators and bead splitting processes. *arXiv:1304.0802 [math.PR]*, 2013.