The Annals of Probability
2015, Vol. 43, No. 5, 2611–2646
DOI: 10.1214/14-AOP945

REGENERATIVE TREE GROWTH: MARKOVIAN EMBEDDING OF FRAGMENTERS, BIFURCATORS, AND BEAD SPLITTING PROCESSES[1]

BY JIM PITMAN AND MATTHIAS WINKEL

University of California Berkeley and University of Oxford

Some, but not all processes of the form $M_t = \exp(-\xi_t)$ for a pure-jump subordinator ξ with Laplace exponent Φ arise as residual mass processes of particle 1 (tagged particle) in Bertoin's partition-valued exchangeable fragmentation processes. We introduce the notion of a *Markovian embedding* of $M = (M_t, t \geq 0)$ in a fragmentation process, and we show that for each Φ, there is a unique (in distribution) binary fragmentation process in which M has a Markovian embedding. The identification of the Laplace exponent Φ^* of its tagged particle process M^* gives rise to a symmetrisation operation $\Phi \mapsto \Phi^*$, which we investigate in a general study of pairs (M, M^*) that coincide up to a random time and then evolve independently. We call M a *fragmenter* and (M, M^*) a *bifurcator*.

For $\alpha > 0$, we equip the interval $R_1 = [0, \int_0^\infty M_t^\alpha \, dt]$ with a purely atomic probability measure μ_1, which captures the jump sizes of M suitably placed on R_1. We study binary tree growth processes that in the nth step sample an atom ("bead") from μ_n and build (R_{n+1}, μ_{n+1}) by replacing the atom by a rescaled independent copy of (R_1, μ_1) that we tie to the position of the atom. We show that any such *bead splitting process* $((R_n, \mu_n), n \geq 1)$ converges almost surely to an α-self-similar continuum random tree of Haas and Miermont, in the Gromov–Hausdorff–Prohorov sense. This generalises Aldous's line-breaking construction of the Brownian continuum random tree.

1. Introduction. We call a process $M := (M_t, t \geq 0)$ a *multiplicative subordinator*, or *fragmenter* for short, if

$$M_t = \exp(-\xi_t), \qquad t \geq 0,$$

for some subordinator $(\xi_t, t \geq 0)$. As shown by Pitman [27] and Bertoin [5], such processes arise naturally in the theory of continuous-time processes of coagulation and fragmentation. The process $(1 - M_t, t \geq 0)$ is the random cumulative distribution function of a random discrete probability measure on $(0, \infty)$. These random measures have been studied in the theory of Bayesian nonparametric statistics [12, 13, 24], not just for subordinators ξ, but more generally for increasing processes with independent increments which are not necessarily stationary.

Received April 2013; revised May 2014.

[1]Supported by the National Science Foundation Awards 0405779 and 0806118.

MSC2010 subject classifications. 60J80.

Key words and phrases. Fragmentation, self-similar tree, continuum random tree, $\mathbb{R}$-tree, weighted $\mathbb{R}$-tree.

We will use terminology based on the fragmentation interpretation of M_t as the residual mass of a block containing a particle at time t. Bertoin [5] showed that the mass containing particle 1 in an exchangeable homogeneous fragmentation process is a fragmenter. Let us recall the definition of a homogeneous fragmentation process (HFP). We denote by $\mathcal{P} = \mathcal{P}_{\mathbb{N}}$ the set of partitions of $\mathbb{N}$. An exchangeable HFP is a Markov process $\Pi = (\Pi(t), t \geq 0)$ in $\mathcal{P}$ such that:

- given $\Pi(t) = \{B_i, i \geq 1\} \in \mathcal{P}$, the partition $\Pi(t+s)$ is distributed as the collection of blocks of $B_i \cap \Pi^{(i)}(s)$, $i \geq 1$, for a family $\Pi^{(i)}$, $i \geq 1$, of independent copies of Π, and
- the distribution of Π is exchangeable, that is, invariant under all finite permutations of $\mathbb{N}$.

It is a well-known consequence of de Finetti's theorem that exchangeable partitions have asymptotic frequencies, so $|\Pi_i(t)| = \lim_{n\to\infty} n^{-1}\#(\Pi_i(t) \cap \{1, \ldots, n\})$ exists almost surely (which we abbreviate *a.s.*), in fact jointly for all $i \geq 1$ and $t \geq 0$; see [5]. Referring to an asymptotic frequency as *mass*, $M_t^* := |\Pi_1(t)|$ is the residual mass of the block $\Pi_1(t)$, which contains particle 1. Bertoin showed that for every exchangeable HFP Π, this process M^* is a fragmenter. We will call $M^* = |\Pi_1| := (|\Pi_1(t)|, t \geq 0)$ the *canonical fragmenter* of Π. As Haas [19] demonstrated, there are exchangeable HFPs with different distributions whose canonical fragmenters have the same distribution. On the other hand, in the subclass of binary models, where every infinitesimal split is into two parts (see Section 2.3), we show this cannot happen.

The starting point for this paper is the observation that not all fragmenters arise as canonical fragmenters in an exchangeable HFP. Furthermore, we have encountered a number of natural nonexchangeable models [10, 11, 16, 29, 30], in which masses of blocks can be defined as asymptotic frequencies, and the mass containing particle 1 is also a fragmenter. Via embedding of such residual mass processes into an exchangeable model or via limit considerations, we have found associated exchangeable models in all those examples. Our main result, Theorem 14, shows that for a suitable notion of "embedding," these examples generalise to a remarkably simple picture:

1. Every pure-jump fragmenter M can be embedded in an exchangeable binary HFP Π.
2. The distribution of the HFP Π is uniquely determined by that of the fragmenter M.
3. The canonical fragmenter M^* of Π is a symmetrised version of M defined in Section 2.2.

To prepare this result, Section 2.1 offers a systematic study of pairs of fragmenters $(M, \widehat{M})$ that coincide up to a random time, after which M and $\widehat{M}$ evolve independently. We call such pairs *bifurcators* and give several equivalent characterisations

(Propositions 4 and 5), which are of interest in their own right. Examples of bifurcators include pairs $((|A_t^1|, |A_t^2|), t \geq 0)$ of residual masses of the blocks A_t^1 and A_t^2 containing particles 1 and 2, respectively, in an exchangeable HFP; see also Proposition 13. Specifically, note that after the random time when 1 and 2 separate, the evolution of A_t^1 and A_t^2 is independent, since disjoint blocks evolve independently in a HFP. In Section 2.2 we focus more generally on bifurcators for which size-biased switching describes the separation time. This induces an idempotent transformation from distributions of M to M^*, which we call *symmetrisation*. This transformation is the key to finding the exchangeable binary HFP associated with M. In Section 2.3 we recall-known facts about exchangeable HFPs Π. We introduce the notion of a *Markovian embedding in* Π in Section 2.4 and show the existence of an embedding for M. We postpone the proof of uniqueness of Π to Section 3. In Section 2.5 we study the three-way mass split into the parts before and after the separation time, and in Section 2.6 associated lengths induced by a bifurcator.

To complete the proof of Theorem 14, we use Haas and Miermont's [20] α-self-similar continuum random trees (CRTs), which are certain random rooted compact metric space trees $(\mathcal{T}, d, 0)$ equipped with a probability measure μ. Specifically, a random element $\Sigma_1^* \in \mathcal{T}$ with distribution μ yields a path $\mathcal{R}_1^* = [\![0, \Sigma_1^*]\!]$ in $\mathcal{T}$. For $(\mathcal{T}, \mu)$ associated with Π (and some $\alpha > 0$), it is well known that the process of μ-masses in subtrees above points in $\mathcal{R}_1^*$ is related to a copy of M^* by a certain α-self-similar time change. Furthermore, $\mathcal{R}_n^* = \bigcup_{j=1}^n [\![0, \Sigma_j^*]\!]$ increases to $\mathcal{T}$ for a sample Σ_n^*, $n \geq 1$, from μ. See Section 2.3 for details.

We project μ onto $\mathcal{R}_n^*$ to equip $\mathcal{R}_n^*$ with a random discrete distribution μ_n^*. In particular, $(\mathcal{R}_1^*, \mu_1^*)$ is a *string of beads*, that is, an interval equipped with a purely atomic measure, and $(\mathcal{R}_1^*, \mu_1^*)$ can easily be completely expressed in terms of M^*; see Section 3.1. Note that $\mathcal{R}_{n+1}^*$ is a tree with one more branch than $\mathcal{R}_n^*$, and μ_n^* is the projection of μ_{n+1}^* onto $\mathcal{R}_n^*$. Since Σ_{n+1}^* is selected according to μ, we have an instance of the following general notion of a *bead splitting process*:

- Let (R_1, μ_1) be a string of beads.
- Given that R_n has been defined with a purely atomic probability measure μ_n, pick an atom ("bead") J_n according to μ_n. Given $\mu_n(\{J_n\}) = m$, remove the atom from μ_n, split it into smaller atoms, tie to J_n a string with these beads of total mass m to form (R_{n+1}, μ_{n+1}).

See Figure 1 for an illustration. A similar bead splitting process, but with different bead selection rules, was obtained for the alpha–theta model of [30] by exploiting properties of the Chinese restaurant process. The main developments in Section 3 culminate in Theorem 20:

4. We give an autonomous description meaningful outside a CRT for the evolution of the bead-splitting process $((\mathcal{R}_n^*, \mu_n^*), n \geq 1)$, which is associated with the canonical fragmenter.

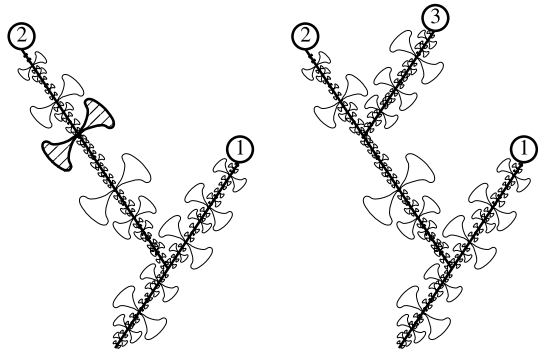

FIG. 1. *Binary trees equipped with strings of beads*; *the right tree was obtained from the left tree by performing the above operation on the shaded bead of the left tree*.

5. We generalise this description to start from a string of beads associated with an arbitrary strictly decreasing pure-jump fragmenter M rather than a canonical fragmenter M^*. At each growth step for (R_{n+1}, μ_{n+1}) we tie to J_n an independent rescaled copy of (R_1, μ_1).

6. We show that this bead splitting process $((R_n, \mu_n), n \geq 1)$ converges almost surely for the Gromov–Hausdorff–Prohorov metric to a CRT $(\mathcal{T}, \mu)$. This CRT is associated with a HFP whose canonical fragmenter is M^* as identified earlier.

As tools, we develop a general spinal decomposition of exchangeable HFPs along a Markovian path (Lemma 21) and show a CRT convergence result for bead splitting processes based on any Markovian path (Lemma 22), which we then also use to complete the proof of Theorem 14. The embedding for the existence part of the proof of Theorem 14 is not carried out in a CRT, but directly in an exchangeable HFP. A uniqueness proof entirely in the framework of HFP should be possible, but the construction is harder to formulate, and the compactness of CRTs would not be directly available. The transition kernel from (R_n, μ_n) to (R_{n+1}, μ_{n+1}) is simple for all fragmenters and gives an inductive description of the distribution of (R_n, μ_n) for every $n \geq 1$. Section 3.3 provides direct descriptions of the distribution of $(\mathcal{R}_n^*, \mu_n^*)$ in the special case of the Brownian CRT, exploiting relations to Aldous's line-breaking construction [2], Brownian path transformations [4] and Poisson–Dirichlet distributions [14, 28].

The CRT convergence result of bead splitting processes here can be complemented by scaling limit results of discrete tree shapes T_n of R_n and/or their reduced subtrees $T_{n,k}$ spanned by the first k leaves. Specifically, we applied methods of Haas and Miermont [21] in [29] to obtain convergence in distribution for trees like T_n, $n \geq 1$, suitably rescaled, to a limiting CRT.

We note in Proposition 19 that the numbers of leaves of $\mathcal{R}_{n+1}$ in subtrees of the spine from the root to leaf 1 form a strongly sampling consistent regenerative composition structure $\mathcal{C}_n$, $n \geq 1$, in the sense of Gnedin and Pitman [17],

also [30], Section 2.1. Gnedin, Pitman, and Yor [18] studied the number of blocks $\#\mathcal{C}_n$ and showed $\#\mathcal{C}_n/n^\alpha \to \int_0^\infty M_s^\alpha\, ds$ a.s., under a regular variation condition. We exploited this in [22], Proposition 7, for exchangeably labelled trees to see that reduced trees $T_{n,k}$ converge almost surely to R_k when rescaling all edge lengths by n^α. This result can be generalised to the present setup, which includes nonexchangeable cases. See also [30], Proposition 14, for the alpha–theta model, which adds projected uniform measures that converge to the limiting strings of beads. The bead splitting process we identified for the alpha–theta model develops by size-biased branching only for $\theta = \alpha$. For the other cases, we found different bead selection rules in connection with ordered Chinese restaurant processes.

2. Fragmenters and their embedding in fragmentation processes. This section studies a natural class of models for tracking two residual mass processes $(M_t, t \geq 0)$ and $(\widehat{M}_t, t \geq 0)$ that we can think of as parts in a fragmentation process. It is instructive and indeed a natural approach to fragmentation processes to first consider these models in their own right, not as parts in a fragmentation process. A systematic study does not appear to be available in the literature. This is provided in Section 2.1, before Section 2.2 focusses on the important special case of size-biased branching. In Section 2.3 we recall from the literature the concepts of a self-similar fragmentation process and associated CRTs. This enables us in Section 2.4 to apply results from Sections 2.1 and 2.2 to formalise and establish points 1 and 3 from the Introduction. Sections 2.5 and 2.6 study the three-way mass split and associated lengths before and after the separation time of a bifurcator under size-biased branching.

2.1. *Fragmenters, switching transformations, and bifurcators.* It will be assumed throughout this section that all fragmenters $M_t = \exp(-\xi_t)$, $t \geq 0$, are derived from subordinators ξ with zero drift and no killing. Furthermore, for most of our discussion, we will also assume an absolutely continuous Lévy density. This is just for convenience of presentation. We discuss general Lévy measures at the end of this section. The Lévy–Itô representation of ξ is then

$$\xi_t = \sum_{0<s\leq t} \Delta\xi_s, \qquad t \geq 0,$$

where $\{(s, \Delta\xi_s) : s > 0, \Delta\xi_s > 0\}$ is the set of points of a Poisson random measure on $(0,\infty)\times(0,\infty)$ with intensity measure $ds\,\lambda(x)\,dx$ where λ is the Lévy density of the subordinator, so

$$\mathbb{E}\big(e^{-\rho\xi_t}\big) = e^{-t\Phi(\rho)} \qquad \text{where } \Phi(\rho) = \int_0^\infty (1 - e^{-\rho x})\lambda(x)\,dx,\ \rho \geq 0, \tag{1}$$

is the Laplace exponent. Let $F_s := \exp(-\Delta\xi_s)$. Then the corresponding formulae for the fragmenter M are

$$M_t = \prod_{0<s\leq t} F_s, \qquad t \geq 0,$$

where $\{(s, F_s) : s > 0, F_s < 1\}$ is the set of points of a Poisson random measure on $(0, \infty) \times (0, 1)$ with intensity measure $ds\, uf(u)\, du$ on $(0, \infty) \times (0, 1)$, where $uf(u)\, du$ is the push-forward of $\lambda(x)\, dx$ via the transformation $u = e^{-x}$. So for all nonnegative Borel functions g,

$$\int_0^\infty g\big(e^{-x}\big)\lambda(x)\, dx = \int_0^1 g(u) uf(u)\, du. \tag{2}$$

We introduce the size-biasing factor u in the definition (2) of $f(u)$ to simplify applications to fragmenters associated with (binary) homogeneous fragmentations [7], which we define more formally in Section 2.3 and explain briefly in the next paragraph. We call f the *splitting density of the fragmenter*, which is related to λ by

$$f(u) = u^{-2}\lambda(-\log u), \qquad 0 < u < 1. \tag{3}$$

By (2) for $g(u) = 1 - u^\rho$, the Laplace exponent of ξ is then

$$\Phi(\rho) = \int_0^1 \big(1 - u^\rho\big) uf(u)\, du, \qquad \rho \geq 0. \tag{4}$$

Note that f is subject to the integrability condition that $\Phi(\rho) < \infty$ for some (hence all) $\rho > 0$, that is,

$$\Phi(1) = \int_0^1 u(1-u) f(u)\, du < \infty. \tag{5}$$

The Lévy–Khintchine formula (1) now provides a Mellin transform for the fragmenter,

$$\mathbb{E}\big(M_t^\rho\big) = e^{-t\Phi(\rho)}, \qquad t \geq 0.$$

If $(M_t, t \geq 0)$ is the mass of a randomly tagged fragment in a binary homogeneous fragmentation process with a dislocation measure ν_{ranked} concentrated on decreasing nonnegative sequences $\mathbf{s} = (s_1, s_2, \ldots)$ with $s_1 + s_2 = 1$, then M admits the above descriptions, assuming the existence of a density $\sum_{i=1}^2 \nu_{\text{ranked}}(s_i \in du) = f(u)\, du$. The size-biasing factor u then arises because f is necessarily *symmetric*, meaning $f(u) = f(1-u)$, and given a mass split $\mathbf{s} = (u, 1-u)$ with $u > 1 - u$ from $\nu_{\text{ranked}}(d\mathbf{s})$, the randomly tagged fragment will be found in component u with probability u and in component $1 - u$ with probability $1 - u$. This is mapping a Poisson point process of mass splits $\nu_{\text{ranked}}(d\mathbf{s})$ to a Poisson point process with intensity $uf(u)\, du$.

To further study fragmenters embedded in a homogeneous fragmentation process, we consider the following *switching transformation* of one fragmenter M into another fragmenter $\widehat{M}$. Let p be a nonnegative measurable function from $(0, 1)$ to $[0, 1]$. If M is a fragmenter with Lévy–Itô representation $M_t = \prod_{0<s\leq t} F_s$, consider the process

$$\widehat{M}_t := \prod_{0<s\leq t} \widehat{F}_s, \tag{6}$$

where conditionally given M, the factors $\widehat{F}_s$ are defined by $\widehat{F}_s = F_s$ with probability $1 - p(F_s)$, and $\widehat{F}_s = 1 - F_s$ with probability $p(F_s)$. Here the construction of the point process $((s, \widehat{F}_s), s > 0)$ from the point process $((s, F_s), s > 0)$ with intensity $ds\, u f(u)\, du$ is made rigorous in the usual way by some arbitrary indexing of these points by positive integers, and making independent choices for each of the countable number of F_s with $F_s < 1$. Here and below it is always assumed that all processes are defined on a rich enough probability space to admit all necessary auxiliary randomizations, as are involved in passing from $((s, F_s), s > 0)$ to $((s, \widehat{F}_s), s > 0)$. Standard transformation results for Poisson point processes imply that $((s, \widehat{F}_s), s > 0)$ is a Poisson point process with intensity $ds\, u \widehat{f}(u)\, du$ for $\widehat{f}$ determined by the formula

$$u\widehat{f}(u) = \big(1 - p(u)\big)uf(u) + p(1-u)(1-u)f(1-u), \qquad 0 < u < 1. \tag{7}$$

In particular, provided $\int_0^1 u(1-u)\widehat{f}(u)\, du < \infty$, that is,

$$\int_0^1 u^2 p(u) f(u)\, du < \infty, \tag{8}$$

the function $\widehat{f}$ serves as a splitting density, and (6) is the Lévy–Itô representation of a fragmenter $\widehat{M}$ with splitting density $\widehat{f}$. Call $\widehat{M}$ the *fragmenter derived from M by switching according to p*.

The following proposition provides a summary:

PROPOSITION 1. *If M is a fragmenter with splitting density f, and p is subject to (8), then $\widehat{M}$ derived from M by switching according to p is a fragmenter with splitting density $\widehat{f}$ as in (7). Moreover, M is then derived from $\widehat{M}$ by switching according to $\widehat{p}$, where*

$$\widehat{p}(u) = \frac{p(1-u)(1-u)f(1-u)}{u\widehat{f}(u)}, \qquad 0 < u < 1. \tag{9}$$

This generalises [30], Lemma 19(b), which treats the case of size-biased switching probabilities $p(u) = 1 - u$.

Observe that two fragmenters M and $\widehat{M}$ as above are by construction such that $M_t = \widehat{M}_t$ for $0 \le t < \tau$ where

$$\tau := \inf\{s : F_s \neq \widehat{F}_s\}$$

is the time of the first switch. It is clear from the Poisson construction that τ is exponential with rate

$$\phi := \int_0^1 p(u)uf(u)\, du = \int_0^1 \widehat{p}(u)u\widehat{f}(u)\, du =: \widehat{\phi} \in [0, \infty], \tag{10}$$

where we will usually exclude the trivial cases $\phi = 0$, that is, $\tau = \infty$, and $\phi = \infty$, that is, $\tau = 0$. The conditional distribution of τ given M is made explicit by the formula

$$\mathbb{P}(\tau > t|M) = \prod_{0<s\le t} \big(1 - p(F_s)\big), \tag{11}$$

where $F_s = M_s/M_{s-}$ and, by convention, $p(1) = 0$. Assuming further that $0 < \phi < \infty$, so $\mathbb{P}(0 < \tau < \infty) = 1$, it is clear by construction that M and $\widehat{M}$ satisfy the *identification rule*

$$M_t = \widehat{M}_t \qquad \text{for } 0 \le t < \tau, \tag{12}$$

hence $M_{\tau-} = \widehat{M}_{\tau-}$, and the *binary splitting condition* that the decrement of each fragmenter at time τ equals the value of the other fragmenter at time τ,

$$M_{\tau_-} - M_\tau = \widehat{M}_\tau \quad \text{and} \quad \widehat{M}_{\tau_-} - \widehat{M}_\tau = M_\tau. \tag{13}$$

Call (12) and (13) together the *binary junction conditions*. After time τ the random factors governing the evolution of M and $\widehat{M}$ are further coupled. We now modify this construction so that the two fragmenters continue independently after time τ:

DEFINITION 2 (Bifurcator). We call a pair of fragmenters $(M, \widetilde{M})$ a *bifurcator with switching time* τ if there are a splitting density f and a switching probability function p so that $\int_0^1 p(u)uf(u)\,du < \infty$ and $(M, \widetilde{M}, \tau)$ has the following joint distribution:

- M is a fragmenter with splitting density f,
- τ is the first switching time of an auxiliary fragmenter $\widehat{M}$ derived from M by switching according to p,
- $\widetilde{M}_t = M_t$, $t < \tau$, $\widetilde{M}_\tau = M_{\tau-} - M_\tau$, and $(\widetilde{M}_{\tau+t}/\widetilde{M}_\tau, t \ge 0)$ is a copy of $\widehat{M}$ independent of (M, τ).

See Propositions 4 and 5 for characterisations that may serve as alternative definitions.

Note that in our construction, $\widetilde{M} = \widehat{M}$ on $[0, \tau]$, so the binary junction conditions (12) and (13) hold just as well for $\widetilde{M}$ as for $\widehat{M}$. But after time τ the evolutions of M and $\widetilde{M}$ are decoupled. Dually, $(M_{\tau+t}/M_\tau, t \ge 0)$ is a copy of M which is independent of $(\widetilde{M}, \tau)$.

Henceforth we will no longer be concerned with any $\widehat{M}$ that is further coupled with M after τ, and we will instead use the generic notation $(M, \widehat{M})$ for a bifurcator. Then for some splitting time τ, whose joint law with M is determined by the switching probability function p,

$$\widehat{M}_t = M_t 1_{\{\tau > t\}} + (M_{\tau-} - M_\tau)\widehat{M}'_{t-\tau} 1_{\{\tau \le t\}}, \tag{14}$$

where $\widehat{M}' \stackrel{d}{=} \widehat{M}$ with $\widehat{M}'$ independent of (M, τ). Note the subtlety that $\widehat{M}$ is determined pathwise by M up to and including the splitting time τ, but thereafter the jumps of $\widehat{M}$ and M are decoupled: the distribution of how $\widehat{M}$ evolves after time τ is implicitly determined by M and p, but there is no pathwise coupling between M and $\widehat{M}$ after time τ. Rather, $(M, \widehat{M})$ satisfies

DEFINITION 3 (Asymmetric Markov branching property). We say that $(M, \widehat{M})$ has the *asymmetric Markov branching property* relative to the splitting time τ if:

- conditionally given $\tau > t$ the process $((M_{t+v}/M_t, \widehat{M}_{t+v}/\widehat{M}_t), v \geq 0)$ is a copy of $(M, \widehat{M})$, independent of $((M_s, \widehat{M}_s), 0 \leq s \leq t)$;
- conditionally given $\tau \leq t$ the two processes $(M_{t+v}/M_t, v \geq 0)$ and $((\widehat{M}_{t+v}/\widehat{M}_t), v \geq 0)$ are independent copies of M and $\widehat{M}$, respectively, independent of $((M_s, \widehat{M}_s), 0 \leq s \leq t)$.

The following variation of Proposition 1 follows easily from standard facts about Poisson point processes, and the above definitions:

PROPOSITION 4. (a) *The joint distribution of the bifurcator* $(M, \widehat{M})$ *is uniquely determined by the splitting density* f *of* M *and a switching probability function* p *subject to*

$$0 < \phi := \int_0^1 p(u) u f(u)\, du < \infty \tag{15}$$

or dually by the splitting density $\widehat{f}$ *of* $\widehat{M}$ *and the dual switching probability function* $\widehat{p}$, *subject to* $0 < \widehat{\phi} < \infty$, *as specified in* (7), (9) *and* (10). *Furthermore,* $\phi = \widehat{\phi}$ *is the rate of the exponentially distributed junction time* τ.

(b) *A bifurcator* $(M, \widehat{M})$ *as in* (a) *can also be constructed as follows from five independent ingredients*: *three fragmenters* M^0, M', *and* $\widehat{M}'$ *with splitting densities* $(1 - p(u)) f(u)$, $f(u)$, *and* $\widehat{f}(u)$, *respectively, an exponential time* τ *with rate* ϕ, *and a random variable* $U \in (0, 1)$ *with distribution*

$$\mathbb{P}(U \in du) = \phi^{-1} p(u) u f(u)\, du, \qquad 0 < u < 1. \tag{16}$$

Now define $(M, \widehat{M})$ *by* $M_t = \widehat{M}_t = M^0_t$ *for* $t < \tau$, *and let*

$$M_{\tau+v} = U M^0_\tau M'_v \quad \text{and} \quad \widehat{M}_{\tau+v} = (1 - U) M^0_\tau \widehat{M}'_v \qquad \text{for } v \geq 0. \tag{17}$$

We assumed for ease of exposition that M has a splitting density f. However, the operation of switching according to p and the notion of an associated bifurcator are meaningful when we replace $u f(u)\, du$ by a more general measure $\Lambda(du)$ satisfying $\int_{(0,1)} (1 - u) \Lambda(du) < \infty$. We generalise (7) and (9) to

$$\begin{aligned} \widehat{\Lambda}(du) &= \big(1 - p(u)\big)\Lambda(du) + p(1 - u)\overline{\Lambda}(du) \quad \text{and} \\ \widehat{p}(u)\widehat{\Lambda}(du) &= p(1 - u)\overline{\Lambda}(du), \end{aligned} \tag{18}$$

where $\overline{\Lambda}$ is the image measure of Λ under the switching operation $u \mapsto 1-u$. Then (18) defines $\widehat{\Lambda}(du)$ as a measure satisfying $\int_{(0,1)}(1-u)\widehat{\Lambda}(du) = 2\int_{(0,1)} p(u)\Lambda(du) < \infty$. Also, $p(1-u)\overline{\Lambda}(du)$ is, by definition of $\widehat{\Lambda}$, absolutely continuous with respect to $\widehat{\Lambda}$ with density taking values in $[0,1]$. This identifies $\widehat{p}(u)$ for $\widehat{\Lambda}$-a.e. $u \in (0,1)$. We define τ as the first switching time with distribution given in (11). If $p(1/2)\Lambda(\{1/2\}) > 0$, we can have $\tau \neq \inf\{s \geq 0 : F_s \neq \widehat{F}_s\}$, and if furthermore Λ is finite, also $\tau \neq \inf\{s \geq 0 : M_s \neq \widehat{M}_s\}$ for a bifurcator $(M, \widehat{M})$. A more satisfactory way to include those cases is to slightly re-model the switching transformation by marking $(F_s, s \geq 0)$ by a marking kernel K from $(0,1)$ to $\{0,1\}$, where $K(u,\{1\}) = p(u)$ and $K(u,\{0\}) = 1-p(u)$, with associated marked point process $((F_s, m_s), s \geq 0)$ mapping to $\widehat{F}_s = (1-m_s)F_s + m_s(1-F_s)$, and with $\tau = \inf\{s \geq 0 : m_s = 1\}$. In the following characterisation of bifurcators, it is more natural to exclude the cases when $\tau \neq \inf\{s \geq 0 : M_s \neq \widehat{M}_s\}$. Note that for those cases the analogue of (13) at $\inf\{s \geq 0 : M_s \neq \widehat{M}_s\}$ fails since this is the first of the jump times after τ, and the respective first jump times of M and $\widehat{M}$ after τ will be different a.s.

PROPOSITION 5. *Consider a pair of positive nonincreasing pure jump processes $(M, \widehat{M})$, and suppose that $\tau = \inf\{t \geq 0 : M_t \neq \widehat{M}_t\} \in (0,\infty)$ a.s. Then $(M, \widehat{M})$ is a bifurcator with splitting time τ if and only if the asymmetric Markov branching property of Definition 3 holds relative to τ, together with the binary junction conditions* (12) *and* (13).

PROOF. The *only if* part is straightforward. For the *if* part suppose that $(M, \widehat{M})$ satisfies the asymmetric Markov branching property. Then Definition 3 implies that $(M_t, \widehat{M}_t, 1_{\{\tau > t\}})$ is a Markov process in its natural filtration. Furthermore, each component is Markovian in its own right with respect to this filtration. More specifically, we see that M and $\widehat{M}$ are fragmenters with some Lévy measures Λ and $\widehat{\Lambda}$, and that τ is exponentially distributed with some rate ϕ. From (13), we have that τ is a common jump time of all three components, so we can consider τ as a marking time for each of the Poisson point processes $((s, F_s), s \geq 0)$ and $((s, \widehat{F}_s), s \geq 0)$, with mark 1 at τ, say. By Lemma 6 below, this yields marking kernels K and $\widehat{K}$ from $(0,1)$ to $\{0,1\}$, from which we define $p(u) = K(u,\{1\})$ and $\widehat{p}(u) = \widehat{K}(u,\{1\})$. By standard results for marking and thinning Poisson point processes, we find

$$\mathbb{P}(F_\tau \in du) = \phi^{-1} p(u)\Lambda(du) \quad \text{and} \quad \mathbb{P}(\widehat{F}_\tau \in du) = \phi^{-1}\widehat{p}(u)\widehat{\Lambda}(du),$$

and the points before τ, which by (12) are common to both processes, have equal thinned intensity measures

$$\big(1-p(u)\big)\Lambda(du) = \big(1-\widehat{p}(u)\big)\widehat{\Lambda}(du).$$

Together with (13), these equations are equivalent to (18), and we easily deduce that $(M, \widehat{M})$ is indeed a bifurcator with splitting time τ in the sense of Definition 2. □

LEMMA 6. *Consider a filtration $\mathcal{F}$, an $\mathcal{F}$-Poisson point process $(F_t, t \geq 0)$ with intensity measure Λ on $(0,1)$ and cemetery 1, an $\mathcal{F}$-stopping time τ such that $F_\tau \neq 1$ a.s. and such that conditionally given $\tau > t$, we have $(\tau - t, (F_{t+s}, s \geq 0)) \stackrel{d}{=} (\tau, (F_s, s \geq 0))$, for all $t \geq 0$. Then there exists a marking kernel K from $(0,1)$ to $\{0,1\}$ such that for a Poisson point process $((\widetilde{F}_t, \widetilde{m}_t), t \geq 0)$ with intensity measure $\Lambda^+(du, dm) = K(u, dm)\Lambda(du)$ and for $\widetilde{\tau} = \inf\{t \geq 0 : \widetilde{m}_t = 1\}$, we have $((F_t, t \geq 0), \tau) \stackrel{d}{=} ((\widetilde{F}_t, t \geq 0), \widetilde{\tau})$.*

We prove Lemma 6 in the Appendix.

REMARK 7. While a splitting density f and a switching probability function p together uniquely identify the distribution of a bifurcator, for two given splitting densities f and $\widehat{f}$, there may not be an associated bifurcator $(M, \widehat{M})$. Looking ahead at Theorem 14, this will, in fact, be the typical case. On the other hand, for fragmenters M and $\widehat{M}$ that can be coupled to form a bifurcator, there are typically many other couplings as different bifurcators. This can be seen from (7), which for each u leaves some choice of $p(u)$ and $p(1-u)$. We will see in Remark 15 that for any choice with both $p(u)$ and $p(1-u)$ in $[0,1]$, equation (7) for $1-u$ instead of u, which appears to give a second equation relating $p(u)$ and $p(1-u)$, will automatically hold if (and only if) the two fragmenters can be embedded in the same fragmentation process.

EXAMPLE 8. An extreme example of a switching probability function is $p(u) = 1$ for $u < 1/2$ and $p(u) = 0$ for $u > 1/2$. In words: switch if the other block is bigger. We then obtain from (7) that $u\widehat{f}(u) = uf(u) + (1-u)f(1-u)$ for $u > 1/2$ and $u\widehat{f}(u) = 0$ for $u < 1/2$. Note that in the context of Remark 7, there is only one bifurcator $(M, \widehat{M})$ which has a given f and this associated $\widehat{f}$ as splitting densities.

2.2. *Size-biased branching.* The instance of the bifurcator construction of the previous section with

$$p(u) = 1 - u, \qquad 0 < u < 1,$$

is of special interest. We then say that the bifurcator $(M, \widehat{M})$ is *derived from M by size-biased branching*, and use the notation (M, M^*) instead of $(M, \widehat{M})$ to indicate this special construction. Note that the "size" involved in the size-biasing is the size of the residual factor $1-u$ associated with decrements of M by a factor of u, that is, the size relative to the current value of M of the fragment that splits.

We note the following corollaries to the results obtained in the previous section.

COROLLARY 9. *If M is a fragmenter with splitting density f and Laplace exponent*

$$\Phi(\rho) = \int_0^1 (1 - u^\rho) u f(u)\, du,$$

then M^* derived from M by size-biased branching is a fragmenter with splitting density

$$f^*(u) = uf(u) + (1-u)f(1-u), \qquad 0 < u < 1, \tag{19}$$

and Laplace exponent

$$\begin{aligned}\Phi^*(\rho) &= \int_0^1 (1-u^\rho) u f^*(u)\, du \\ &= \Phi(\rho+1) - \Phi(\rho+1, \rho+1), \qquad \rho > 0,\end{aligned} \tag{20}$$

where $\Phi(\rho+1, \rho+1)$ is given by

$$\Phi(\rho+1, \rho+1) = \int_0^1 (1-u)^{\rho+1} u f(u)\, du.$$

Moreover, M is then derived from M^* by switching according to p^*, where

$$p^*(u) = \frac{(1-u)f(1-u)}{f^*(u)}, \qquad 0 < u < 1. \tag{21}$$

COROLLARY 10. *In the setting of the previous corollary, the following conditions are equivalent*:

(i) *f is symmetric*: $f(u) = f(1-u)$ *for all* $0 < u < 1$;
(ii) $f = f^*$;
(iii) $\Phi = \Phi^*$;
(iv) $p^*(u) = 1 - u$ *for all* $0 < u < 1$;
(v) $M \stackrel{d}{=} M^*$;
(vi) $(M, M^*) \stackrel{d}{=} (M^*, M)$.

Observe from (19) that whatever the splitting density f of M, the splitting density f^* of M^* is symmetric. Thus the operation of passing from the law of M to the law of M^* by size-biased branching is a kind of symmetrisation of laws of fragmenters corresponding to the elementary symmetrisation of density functions defined by formula (19). The operation is idempotent: $M^{**} \stackrel{d}{=} M^*$. So we make the following definition:

DEFINITION 11 (Symmetrised fragmenter). For a fragmenter M with splitting density f call the fragmenter M^* with splitting density f^* as in (19) the *symmetrisation* of M.

This notion of size-biased branching and symmetrisation can clearly be extended to fragmenters whose splitting measures do not have a density, as is achieved by formula (20) for Laplace exponents. Note, however, that the probabilistic meaning in terms of size-biased branching, and even the analytic fact that

$\Phi^{**} = \Phi^*$ is very much obscured by the Laplace exponents. Also, if $\Phi^*(\rho) = \int_{(0,1)}(1-u^\rho)\Lambda^*(du)$, we obtain from (18)

$$\Lambda^*(du) = u\Lambda(du) + u\overline{\Lambda}(du) \quad \text{and} \quad p^* = d\overline{\Lambda}/d(\Lambda + \overline{\Lambda}). \tag{22}$$

As can be seen in the symmetry discussion leading up to (23) below, this operation of symmetrisation of M projects the collection of laws of all fragmenters M onto the collection of laws of fragmenters M^* which are canonically associated with a binary homogeneous fragmentation process via the mass of a uniformly randomly tagged fragment. This raises the question: exactly how is a fragmenter M with splitting density f related to the binary homogeneous fragmentation process with splitting density f^*? We answer this question in Section 2.4, after development of the necessary framework in Section 2.3.

2.3. *Exchangeable fragmentation processes and self-similar CRTs.* In our context, we can express Bertoin's [5, 6] definitions of homogeneous and self-similar fragmentations as follows. For $\alpha \in \mathbb{R}$, we say that a family $\Pi_\alpha = (\Pi_\alpha(t), t \geq 0)$ of refining partitions in $\mathcal{P} = \mathcal{P}_\mathbb{N}$ is an *exchangeable α-self-similar fragmentation process* if both:

- Π_α is exchangeable in that the distribution of Π_α is invariant under permutations of $\mathbb{N}$;
- Π_α is a right-continuous strong Markov process whose transition kernel satisfies the branching property: for all $t \geq 0$, $s \geq 0$, conditionally given $\Pi_\alpha(t) = \{B_i, i \geq 1\}$, the partition $\Pi_\alpha(t+s)$ has the same distribution as the partition of $\mathbb{N}$ with blocks $B_i \cap \Pi_\alpha^{(i)}(|B_i|^{-\alpha}s)$, $i \geq 1$, where the $\Pi_\alpha^{(i)}$, $i \geq 1$, are independent copies of Π_α.

Usually we consider $\Pi_\alpha(0) = \{\mathbb{N}\}$, and we exclude the trivial case $\Pi_\alpha(t) = \{\mathbb{N}\}$ for all $t \geq 0$. Then $\Pi_\alpha(\infty) := \lim_{t\to\infty} \Pi_\alpha(t) = \{\{1\}, \{2\}, \ldots\}$. In the case $\alpha = 0$, the linear time-changes of $\Pi_\alpha^{(i)}$ by asymptotic frequencies $|B_i|^\alpha$ disappear; this case is called an *exchangeable homogeneous fragmentation process*. Bertoin [5] showed that the distribution of $\Pi = \Pi_0$ can be expressed in terms of an exchangeable σ-finite intensity measure $\kappa(d\Gamma)$ on $\mathcal{P} \setminus \{\{\mathbb{N}\}\}$ via a Lévy–Itô-type decomposition into elementary splits of blocks B by $\Gamma = \{\Gamma_i, i \geq 1\}$ into $B \cap \Gamma_i$, $i \geq 1$. The measure κ admits an integral representation

$$\kappa(d\Gamma) = c\sum_{n\geq 1} \delta_{\{\mathbb{N}\setminus\{i\},\{i\}\}}(d\Gamma) + \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(d\Gamma)\nu_{\text{ranked}}(d\mathbf{s}),$$

for an *erosion coefficient* $c \geq 0$ and a *ranked dislocation measure* ν_{ranked} on $\mathcal{S}^\downarrow \setminus \{(1,0,\ldots)\}$ satisfying $\int_{\mathcal{S}^\downarrow}(1-s_1)\nu_{\text{ranked}}(d\mathbf{s}) < \infty$, and where $\kappa_{\mathbf{s}}$ is Kingman's paintbox governing exchangeable partitions with asymptotic frequen-

cies $\mathbf{s} \in \mathcal{S}^{\downarrow} := \{(s_i)_{i\geq 1} : s_1 \geq s_2 \geq \cdots \geq 0, \sum_{i\geq 1} s_i \leq 1\}$. In the *binary* case, $\kappa(\Gamma \in \mathcal{P} \setminus \{\mathbb{N}\} : \Gamma_1 \cup \Gamma_2 \neq \mathbb{N}) = 0$, this representation can be written as

$$\kappa(d\Gamma) = c \sum_{n\geq 1} \delta_{\{\mathbb{N}\setminus\{i\},\{i\}\}}(d\Gamma) + \frac{1}{2}\int_{(0,1)} \kappa_{(s,1-s)}(d\Gamma)\nu(ds),$$

for a *symmetric dislocation measure* ν on $(0, 1)$ satisfying $\int_{(0,1)} s(1-s)\nu(ds) < \infty$ and $\nu = \overline{\nu}$, where $\overline{\nu}$ is the push-forward of ν under $u \mapsto 1-u$, so $\nu(du) = \nu_{\text{ranked}}(s_1 \in du) + \nu_{\text{ranked}}(s_2 \in du)$ and $\nu_{\text{ranked}}(s_1 \in \cdot) = \nu(\cdot \cap (\frac{1}{2}, 1)) + \frac{1}{2}\nu(\cdot \cap \{\frac{1}{2}\}) = \overline{\nu}(\cdot \cap (\frac{1}{2}, 1)) + \frac{1}{2}\overline{\nu}(\cdot \cap \{\frac{1}{2}\})$.

We denote by $|A^n_\alpha(t)|$ the asymptotic frequency of the block $A^n_\alpha(t)$ of $\Pi_\alpha(t)$ containing n. For $\alpha = 0$, the process $|A^n_0(t)|$ is a fragmenter, and $\xi_n(t) = -\log|A^n_0(t)|$ has Laplace exponent

$$\Phi^*(\rho) = c + c\rho + \int_{(0,1)} \bigl(1 - u^\rho\bigr)u\nu(du); \tag{23}$$

see [5]. Self-similar and homogeneous fragmentation processes are pathwise related by nonlinear time-change [6]. Specifically,

$$\begin{aligned} \bigl|A^n_\alpha(t)\bigr| &= \exp\bigl(-\xi_n\bigl(\eta_n(t)\bigr)\bigr) \\ &\text{where } \eta_n(t) = \inf\biggl\{u \geq 0 : \int_0^u e^{-\alpha\xi_n(w)}\,dw > t\biggr\}, \end{aligned} \tag{24}$$

is a self-similar Markov process, and for all $\alpha > 0$, we have $|A^n_\alpha(t)| = 0$ for $t \geq \int_0^\infty e^{-\alpha\xi_n(w)}\,dw$.

It was shown by Haas and Miermont [20] that for every exchangeable self-similar fragmentation process Π_α with index $\alpha > 0$, zero erosion $c = 0$ and infinite dislocation measure ν_{ranked} on $\mathcal{S}^{\downarrow}_1 = \{\mathbf{s} \in \mathcal{S}^{\downarrow} : \sum_{i\geq 1} s_i = 1\}$, there is an associated compact continuum random tree $(\mathcal{T}, d, 0, \mu)$. Vice versa, such a continuum random tree (CRT) allows an embedding of a self-similar fragmentation process. Specifically, a CRT is a random weighted and rooted $\mathbb{R}$-tree. A weighted and rooted $\mathbb{R}$-tree $(T, d, 0, \mu)$ is a complete, separable metric space (T, d) with a *root* $0 \in T$ and a probability measure μ on the Borel sets of (T, d), such that the following tree property holds:

- Any two points $\sigma, \sigma' \in T$ are connected by a unique injective path $[\![\sigma, \sigma']\!]$. Furthermore, this path can be uniquely parametrised $[\![\sigma, \sigma']\!] = \{g_{\sigma,\sigma'}(t), 0 \leq t \leq d(\sigma, \sigma')\}$ by an isometry $g_{\sigma,\sigma'} : [0, d(\sigma, \sigma)] \to T$ with $g_{\sigma,\sigma'}(0) = \sigma$ and $g_{\sigma,\sigma'}(d(\sigma, \sigma')) = \sigma'$.

We also write $]\!]\sigma, \sigma']\!] := [\![\sigma, \sigma']\!] \setminus \{\sigma\} = \{g_{\sigma,\sigma'}(t), 0 < t \leq d(\sigma, \sigma')\}$.

When there is no ambiguity about d, 0 or even μ, we simply write $(\mathcal{T}, \mu)$ or even $\mathcal{T}$ to refer to a CRT $(\mathcal{T}, d, 0, \mu)$. For the purpose of convergence of compact weighted and rooted $\mathbb{R}$-trees, we will identify $(T, d, 0, \mu)$ and $(T', d', 0', \mu')$ if there is an isometry between (T, d) and (T', d') that maps 0 to $0'$ and pushes μ

forward to μ'. The set $\mathbb{T}$ of such isometry classes can then be equipped with the so-called Gromov–Hausdorff–Prohorov distance d_{GHP}. Then $(\mathbb{T}, d_{\mathrm{GHP}})$ is a Polish metric space. See, for example, [15] for further background on the space $(\mathbb{T}, d_{\mathrm{GHP}})$.

For an $\mathbb{R}$-tree $(T, d, 0, \mu)$, let $T^t := \{\sigma \in T : d(0, \sigma) > t\}$, $t \geq 0$, and define fringe subtrees $T_\sigma := \{\sigma' \in T : \sigma \in [\![0, \sigma']\!]\}$, $\sigma \in T$. An α-self-similar CRT is a random weighted and rooted $\mathbb{R}$-tree $(\mathcal{T}, d, 0, \mu)$, or its isometry class with distribution on the Borel space of $(\mathbb{T}, d_{\mathrm{GHP}})$, such that:

- μ is nonatomic with dense support a.s., $\mu(\mathcal{T}_\sigma) > 0$ for all $\sigma \in \mathcal{T}$ with $\mathcal{T}_\sigma \neq \{\sigma\}$, while $\mu([\![0, \sigma]\!]) = 0$ for all $\sigma \in \mathcal{T}$, and
- for all $t \geq 0$, the connected components $(\mathcal{T}_i^t, i \geq 1)$ of $\mathcal{T}^t$, completed by a root vertex 0_i, are such that given $(\mu(\mathcal{T}_i^t), i \geq 1) = (m_i, i \geq 1)$, for some $m_1 \geq m_2 \geq \cdots \geq 0$, the trees

$$(\mathcal{T}_i^t, m_i^{-\alpha}\, d|_{\mathcal{T}_i^t}, 0_i, m_i^{-1}\mu|_{\mathcal{T}_i^t}), \qquad i \geq 1$$

  are independent and identically distributed isometric copies of $(\mathcal{T}, d, 0, \mu)$.

Recently, Stephenson [31] extended this class by relaxing the first bullet point to allow a support that is not dense, atoms of μ, and/or positive weights on branches, so as to include all dislocation measures ν_{ranked} and erosion $c > 0$. The CRT $(\mathcal{T}, \mu)$ constructed in [20, 31] is such that

$$\Pi_\alpha^*(t) = \big\{\{j \geq 1 : \Sigma_j^* \in \mathcal{T}_i^t\}, i \geq 1\big\} \cup \big\{\{j\}, j \geq 1 : \Sigma_j^* \notin \mathcal{T}^t\big\}, \qquad t \geq 0 \tag{25}$$

has the same distribution as $(\Pi_\alpha(t), t \geq 0)$, where conditionally given $(\mathcal{T}, \mu)$ the sequence Σ_n^*, $n \geq 1$, is independent and identically distributed according to μ. It was shown in [20, 31] that the subtrees $\mathcal{R}_k^* = \bigcup_{j=1}^k [\![0, \Sigma_j^*]\!] \subset \mathcal{T}$ converge a.s. in the Hausdorff sense for embeddings in $\ell_1(\mathbb{N})$, and this easily entails $d_{\mathrm{GHP}}((\mathcal{R}_k^*, \mu_k^*), (\mathcal{T}, \mu)) \to 0$ a.s., as $k \to \infty$, where μ_k^* is the push-forward of μ under the projection map $\pi^{\mathcal{R}_k^*} : \mathcal{T} \to \mathcal{R}_k^*$, $u \mapsto g_{0,\sigma}(\sup\{t \geq 0 : g_{0,\sigma}(t) \in \mathcal{R}_k^*\})$. Also, μ is then recovered in accordance with Aldous's theory of consistent leaf-exchangeable families of trees $(\mathcal{R}_k^*, k \geq 1)$ as the weak limit of the uniform distribution ν_k^* on $\Sigma_1^*, \ldots, \Sigma_k^*$, as $k \to \infty$.

2.4. *Embedding fragmenters in homogeneous binary fragmentation processes.* Let $\Pi = (\Pi(t), t \geq 0)$ be a binary homogeneous fragmentation process starting from $\Pi(0) = \{\mathbb{N}\}$, with absolutely continuous symmetric dislocation measure $\nu(du) = f^*(u)\, du$, for some *symmetric splitting density* f^* on $(0, 1)$, so Π takes values in the set of partitions of $\mathbb{N}$. Let $A = (A_t, t \geq 0)$ be a process with values in subsets of $\mathbb{N}$. Call A a *path in* Π if:

- $A_t \in \Pi(t)$ for all $t \geq 0$;
- A_t is decreasing in the inclusion partial order, as t increases.

DEFINITION 12 (Markovian path). We call a path A a *Markovian path in* Π if:

- $(A_t, t \geq 0)$ is adapted to some filtration $(\mathcal{F}_t, t \geq 0)$ with respect to which $(\Pi(t), t \geq 0)$ is Markovian, in such a way that for each $s \geq 0$ the process $((A_t, \Pi(t) \cap A_s), t \geq s)$ and the restriction of $(\Pi(t), t \geq s)$ to $\mathbb{N} - A_s$ are conditionally independent given $\mathcal{F}_s$.

To explain the terminology, think of Π as embedded by suitable time change in the α-self-similar continuum random tree (CRT) associated with Π for $\alpha > 0$ by Haas and Miermont [20], or by Stephenson [31] when ν is finite. Then A_t represents the set of leaf labels above some internal vertex v_t of the CRT in some path leading from the root to a leaf vertex v_∞ of the tree.

Let M be a fragmenter. Say that M admits a *Markovian embedding in* Π if it is possible to construct Π jointly with a Markovian path A such that

$$(M_t, t \geq 0) \stackrel{d}{=} \big(|A_t|, t \geq 0\big),$$

where $|A_t|$ is the asymptotic frequency of A_t, which is known to exist almost surely, simultaneously for all $t \geq 0$ and all sets $A_t \in \Pi(t)$. In terms of an associated CRT construction, the jumps of $|A| := (|A_t|, t \geq 0)$ would then describe the spinal partition of mass in the CRT along a spine leading from the root to some random leaf of the CRT. The most basic example is provided by the next proposition:

PROPOSITION 13. *For each positive integer n, let A_t^n be the block of $\Pi(t)$ containing n. Then*:

(i) (Bertoin [5]) $A^n = (A_t^n, t \geq 0)$ *is a Markovian path in* Π *such that* $|A^n|$ *is a fragmenter with splitting density* f^*.

(ii) *For each pair of positive integers m and n, the pair* $(|A^n|, |A^m|)$ *is a symmetric bifurcator, each derived by size-biased branching from the other.*

PROOF. Part (ii) follows easily from the (strong) homogeneous branching property of Π and Corollary 9. $\square$

Due to the natural embeddings provided by this proposition, we call the fragmenter M^* with splitting density f^* the *canonical fragmenter* associated with Π. In terms of the corresponding mass fragmentation $(|\Pi(t)|^\downarrow, t \geq 0)$ of asymptotic frequencies $|\Pi_i(t)|$, $i \geq 1$, ranked in decreasing order $|\Pi(t)|^\downarrow \in \mathcal{S}^\downarrow$, the process M^* describes the evolution of the mass of the fragment containing a randomly tagged particle. We state more generally without assuming the existence of the densities f and f^*:

THEOREM 14. *Every fragmenter M with Laplace exponent $\Phi(\rho) = \int_{(0,1)}(1 - u^\rho)\Lambda(du)$ admits a Markovian embedding in an exchangeable binary homogeneous fragmentation process* Π. *The distribution of* Π *is unique. Its symmetric*

dislocation measure is given by $\nu = \Lambda + \overline{\Lambda}$. *The canonical fragmenter of* Π *is the symmetrisation of* M.

This confirms points 1–3 of the Introduction. In the absolutely continuous case, we can rephrase as follows. M can be embedded in Π if and only if the splitting density of Π is f^*, the symmetrisation of the splitting density f of M.

REMARK 15. Consider any bifurcator $(M, \widehat{M})$. By adding equation (7) and the equation we obtain by substituting u by $1-u$ in (7), we see that f and $\widehat{f}$ have the same symmetrisation f^*. By Theorem 14, M and $\widehat{M}$ can each be embedded in the same binary homogeneous fragmentation process Π. In fact, the argument used to prove this theorem, can be adapted to prove that the bifurcator admits an embedding in Π. We leave the details of this to the reader.

To prepare for the proof of the theorem, we start with some remarks about paths A in Π. For $t \geq 0$, let $N_t := \min A_t$. Clearly, $N_0 = 1$. The fact that A_t decreases as t increases implies that $(N_t, t \geq 0)$ is some increasing process. Furthermore, $A_t \in \Pi(t)$ implies that

$$A_t = A_t^{N_t}, \qquad t \geq 0. \tag{26}$$

Assuming for simplicity that N_t tends to ∞ as $t \to \infty$, let $0 = \tau_0 < \tau_1 < \tau_2 < \cdots$ be the successive times of jumps of $(N_t, t \geq 0)$, and set $N(n) = N_{\tau_n}$, $n \geq 0$. Then $A_t = A_t^{N(n)}$ for $t \in [\tau_n, \tau_{n+1})$. Note that given the random sequence $1 = N(0) < N(1) < N(2) < \cdots$, the times τ_n can be recovered without further reference to A, from the family of paths A^n associated with Π, as $\tau_n = \inf\{t \geq 0 : A_t^{N(n-1)} \neq A_t^{N(n)}\}$ for each $n \geq 1$. Thus there is a natural correspondence between paths A in Π and increasing random sequences $(N(n), n \geq 1)$ subject to the constraint that $N(n) = \min A_{\tau_n}^{N(n)}$, where in general, the possibility of a finite increasing sequence of random length must also be allowed.

In connection with the α-self-similar CRT $\mathcal{T}$ derived from Π, notice that the random times τ_n are defined in a way which allows corresponding random times

$$\tau_{n,\alpha} := \int_0^{\tau_n} |A_t^{N(n-1)}|^\alpha \, dt$$

to be defined, and that in $\mathcal{T}$ there is a junction vertex V_n at height $\tau_{n,\alpha}$ at which the paths to leaves labelled $N(n-1)$ and $N(n)$ diverge. Here it is assumed that the CRT is equipped with a random sample $\Sigma_1^*, \Sigma_2^*, \ldots$ of its leaves according to its mass measure, and that the homogeneous fragmentation $\Pi = \Pi_0$ has been constructed by time changing the α-self-similar fragmentation associated with the tree and the leaf sample. The path from the root 0 to V_n in the CRT then passes through $V_1, \ldots, V_{n-1}$. Since the CRT is compact, we find a convergent subsequence of $(V_n, n \geq 1)$ with limit Σ, say. Because $(V_n, n \geq 1)$ is increasing for

the genealogical partial order $\prec$ that puts $\sigma \prec \sigma'$ if and only if $\sigma \in [\![0, \sigma'[\![$, the sequence converges to the same limit. Note that Σ must be a leaf almost surely because if Σ is not a leaf, then the fringe subtree $\mathcal{T}_\Sigma$ at Σ will have positive mass, but then $A_\infty = \bigcap_{t\geq 0} A_t$ will have positive limiting frequency, which contradicts $\Pi(\infty) = \{\{1\}, \{2\}, \ldots\}$. The path from 0 to Σ in $\mathcal{T}$ starts by following the path to 1, then branches off in the direction of $\Sigma_{N(1)}$, then branches again in the direction of $\Sigma_{N(2)}$, etc. This could be formalised to give a one-to-one correspondence between paths in Π and paths in $\mathcal{T}$.

PROOF OF THE EXISTENCE PART OF THEOREM 14. For a fragmenter with Laplace exponent $\Phi(\rho) = \int_{(0,1)}(1-u^\rho)\Lambda(du)$, consider its symmetrisation $\Phi^*(\rho) = \int_{(0,1)}(1-u^\rho)\Lambda^*(du)$ with Λ^* given in (22). By (23), this is the Laplace exponent of the canonical fragmenter of a fragmentation process Π with symmetric dislocation measure $\nu(du) = u^{-1}\Lambda^*(du) = \Lambda(du) + \overline{\Lambda}(du)$. Consider the canonical fragmenter $M^1 = |A^1|$ obtained from the blocks A^1_t, $t \geq 0$, of Π containing 1 as in Proposition 13. Construct a bifurcator (M^1, M) by switching from M^1 according to p^* in (22), as needed to create M up to some branching time τ_1. For $0 \leq t < \tau_1$ let $A_t = A^1_t$, and let $A_{\tau_1} = A^1_{\tau_1-} \setminus A^1_{\tau_1}$, that is, the block that splits off from the block containing 1 at time τ_1. To continue the construction of A after time τ_1, apply the strong Markov property of Π at τ_1, and let $N(1) = \min(A_{\tau_1})$. Set $A_t = A^{N(1)}_t$ for $\tau_1 \leq t < \tau_2$ where $\tau_2 - \tau_1$ is the branching time between $M^{(1)} := (|A^{N(1)}_{\tau_1+s}|/|A_{\tau_1}|, s \geq 0)$, which is another copy of M^1, and some further copy of M created conditionally given M^1 and $M^{(1)}$ by the same rule. It is clear that continuing like this creates time segments $\tau_i - \tau_{i-1}$ which are independent and identically distributed, and fresh copies of M^1 as needed. The process A with the desired feature that $|A| \stackrel{d}{=} M$ can be created for all times $t \geq 0$. Moreover, by construction this process A is a Markovian path in Π. $\square$

We postpone the uniqueness part of the proof. Specifically, points 1 and 3 of the Introduction have now been proved, while point 2 is postponed to Section 3, where we first establish points 4–6 in Theorem 20.

2.5. *Mass distributions.* The bifurcator (M, M^*) with M^* derived from M by size-biased branching plays a key role in following discussions. This section collects together some basic formulae for the joint distribution of the branching time $\tau := \inf\{t \geq 0 : M_t \neq M^*_t\}$ and the decrements

$$M_{\tau-} - M_\tau = M^*_\tau \quad \text{and} \quad M_\tau = M^*_{\tau-} - M^*_\tau,$$

where $M_{\tau-} = M^*_{\tau-}$. The triple of nonnegative variables $(1 - M_{\tau-}, M_{\tau^*}, M_\tau)$ with sum 1 is of special interest. In a suitably defined random $\mathbb{R}$-tree $\mathcal{R}_{\Sigma,\Sigma^*}$ spanned by a root 0 and two leaves Σ and Σ^*, this triple represents the masses of three

connected components of the tree formed by removal of a particular random junction vertex of the tree. As indicated in the previous section, this subtree $\mathcal{R}_{\Sigma,\Sigma^*}$ may be naturally embedded in a self-similar CRT $\mathcal{T}$ associated with a fragmentation process, whose canonical fragmenter is M^*. The joint distribution of this triple is determined by a formula for its joint moments provided by Gnedin and Pitman [17], page 477, where this triple is denoted (G, H, D), with the following more elementary interpretation: $(G, 1-D)$ is the interval component covering U in the complement of the range of $(M_t, t \geq 0)$, for U a uniform$(0,1)$ variable independent of M, and $H = 1 - D - G$ is the length of this interval component. The following formulae can be read either from the discussion of the previous sections, or from [17], page 477.

Recall first that for M with Lévy exponent Φ, Lévy measure $\lambda(x)\,dx$, and splitting density f, the branching time τ has exponential distribution with rate

$$\Phi(1) = \int_0^1 u(1-u)f(u)\,du = \int_0^\infty \bigl(1-e^{-x}\bigr)\lambda(x)\,dx.$$

The process $(M_t, 0 \leq t < \tau)$ is then the negative exponential of a killed subordinator with Lévy measure $e^{-x}\lambda(x)\,dx$ and killing at rate $\Phi(1)$.

Thus

$$\mathbb{E}\bigl[M_t^\rho \mid \tau > t\bigr] = \mathbb{E}\bigl[M_{\tau-}^\rho \mid \tau = t\bigr] = e^{-t\Phi_0(\rho)},$$

where

$$\Phi_0(\rho) = \int_0^1 \bigl(1-u^\rho\bigr)u^2 f(u)\,du = \Phi(\rho+1) - \Phi(1)$$

and hence by conditioning on τ

$$\mathbb{E}\bigl[M_{\tau-}^\rho\bigr] = \frac{\Phi(1)}{\Phi(1)+\Phi_0(\rho)} = \frac{\Phi(1)}{\Phi(\rho+1)}. \tag{27}$$

This is [17], formula (57) or (59). From [17], formula (28), or from (16),

$$\mathbb{P}\biggl(\frac{M_\tau}{M_{\tau-}} \in du\biggr) = \bigl(\Phi(1)\bigr)^{-1}(1-u)uf(u)\,du, \tag{28}$$

hence

$$\mathbb{E}\biggl[\biggl(\frac{M_\tau}{M_{\tau-}}\biggr)^\rho\biggr] = \frac{\Phi(\rho+1)-\Phi(\rho)}{\Phi(1)}. \tag{29}$$

Moreover, $M_{\tau-}$ and $M_\tau/M_{\tau-}$ are independent, so the last two formulae combine to give

$$\mathbb{E}\bigl(M_\tau^\rho\bigr) = \frac{\Phi(\rho+1)-\Phi(\rho)}{\Phi(\rho+1)}, \tag{30}$$

which is a simplification of [17], formula (60). Next, $M^*_\tau := M_{\tau-} - M_\tau$ is a size-biased pick from the decrements of M, whence

$$\mathbb{E}\big[(M^*_\tau)^\rho\big] = \frac{\Phi(\rho+1,\rho+1)}{\Phi(\rho+1)} \tag{31}$$
$$\text{where } \Phi(\rho+1,\rho+1) = \int_0^1 (1-u)^{\rho+1} u f(u)\,du.$$

Note that for positive integers $\rho = n$ say, this is a linear combination of evaluations of $\Phi(k)$ at integers $k \le n+1$, as indicated in [17], formula (25). In principle, these Mellin transforms determine the distributions of M_τ and M^*_τ, but there do not seem to be simple formulae for the densities of these variables except in special cases.

Observe that the expected masses of the three components in the junction split are

$$\mathbb{E}\big(M^*_\tau\big) = \big(2\Phi(1) - \Phi(2)\big)/\Phi(2), \tag{32}$$

$$\mathbb{E}(1 - M_{\tau-}) = \mathbb{E}(M_\tau) = \big(\Phi(2) - \Phi(1)\big)/\Phi(2). \tag{33}$$

It does not seem obvious intuitively why the expectations of $(1 - M_{\tau-})$ and M_τ are always equal.

Recall from (20) that $\Phi^*(\rho) := \Phi(\rho+1) - \Phi(\rho+1,\rho+1)$ is the Laplace exponent corresponding to M^* with splitting density $f^*(u) = uf(u) + (1-u)f(1-u)$. So we obtain the following extension of Corollary 10:

COROLLARY 16. *Each of the following two conditions is also equivalent to the symmetry of f, which we characterised in Corollary 10 as $f = f^*$, and as $\Phi = \Phi^*$:*

(vii) $M_\tau / M_{\tau-} \stackrel{d}{=} M^*_\tau / M_{\tau-}$;

(viii) $M_\tau \stackrel{d}{=} M^*_\tau$.

2.6. *Edge lengths and exponential functionals.* Continuing to suppose that (M, M^*) is the bifurcator derived from M by size-biased branching, as well as the basic triple of masses $(1 - M_{\tau-}, M^*_\tau, M_\tau)$ with sum 1, for each $\rho > 0$ we may consider the triple of *exponential functionals*

$$L_0 = \int_0^\tau M_t^\rho\,dt, \qquad L_\Sigma = \int_\tau^\infty M_t^\rho\,dt \quad \text{and} \quad L_* = \int_\tau^\infty (M^*_t)^\rho\,dt,$$

which can be interpreted as the lengths of branches in a suitably defined random ρ-self-similar $\mathbb{R}$-tree with three branches meeting at a junction point, these branches being labelled by 0 for the root, and Σ and $*$ for the two leaves associated with M and M^*, respectively. Note that the definition of the L_i depends on the parameter ρ, which is suppressed in the notation. In particular, suppose that the Lévy measure satisfies the regular variation condition $\int_x^\infty \lambda(x)\,dx = x^{-\alpha}\ell(1/x)$ as

$x \downarrow 0$, for some $\alpha \in (0,1)$ and some function $\ell : (0,\infty) \to (0,\infty)$ that is slowly varying at ∞. Then the above functionals for $\rho = \alpha$ are of special interest [18]. They govern the asymptotics of how numbers of new branch points grow along the three branches, as new branch points are selected by size-biased sampling from the mass distribution on the $\mathbb{R}$-tree $\mathcal{R}_{\Sigma,\Sigma^*}$, which assigns the decrements of M except $M_{\tau-} - M_\tau = M^*_\tau$ to the branch from the root 0 to leaf Σ^*, and the remaining decrements of M^* to the branch from the junction point to leaf Σ^*. Let

$$L_{0\Sigma} := L_0 + L_\Sigma = \int_0^\infty M_t^\rho \, dt = \int_0^\infty e^{-\rho\xi_t} \, dt,$$
$$L_{0*} := L_0 + L_* = \int_0^\infty (M_t^*)^\rho \, dt = \int_0^\infty e^{-\rho\xi^*_t} \, dt,$$

where ξ and ξ^* are the two subordinators associated with M and M^*. According to a known formula for subordinators [9],

$$\mathbb{E}(L_{0\Sigma}^n) = \frac{n!}{\Phi(\rho)\cdots\Phi(n\rho)} \qquad \text{for all } n \in \mathbb{N}, \tag{34}$$

where Φ is the Laplace exponent of ξ, and the same holds for L_{0*} instead of $L_{0\Sigma}$ with the Laplace exponent Φ^* of ξ^* instead of Φ. Now

$$L_0 = \int_0^\tau e^{-\rho\xi_t} \, dt = \int_0^\infty e^{-\rho\eta_t} \, dt,$$

where $\eta_t = \xi_t 1_{\{\tau > t\}} + \infty 1_{\{\tau \le t\}}$ is another subordinator, whose Lévy measure is $e^{-x}\lambda(x)\,dx + \Phi(1)\delta_\infty(dx)$ for $\lambda(x)\,dx$ the Lévy measure of ξ. It follows that the Laplace exponent of η at ρ is $\Phi_\eta(\rho) = \Phi(1) + \int_0^\infty (1 - e^{-\rho x}) e^{-x}\lambda(x)\,dx = \Phi(1) + \int_0^\infty (1 - e^{-(\rho+1)x})\lambda(x)\,dx - \int_0^\infty (1 - e^{-x})\lambda(x)\,dx = \Phi(\rho+1)$, and hence that

$$\mathbb{E}(L_0^n) = \frac{n!}{\Phi_\eta(\rho)\cdots\Phi_\eta(n\rho)} = \frac{n!}{\Phi(\rho+1)\cdots\Phi(n\rho+1)} \qquad \text{for all } n \in \mathbb{N}. \tag{35}$$

Moments of L_Σ and L_* can now be found using the distributional identities $L_\Sigma \stackrel{d}{=} M_\tau^\rho \widehat{L}_{0\Sigma}$ and $L_* \stackrel{d}{=} (M_\tau^*)^\rho \widehat{L}_{0*}$ where $\widehat{L}_{0\Sigma}$ is independent of M_τ with $\widehat{L}_{0\Sigma} \stackrel{d}{=} L_{0\Sigma}$, and $\widehat{L}_{0*}$ is independent of M^*_τ with $\widehat{L}_{0*} \stackrel{d}{=} L_{0*}$. Thus

$$\mathbb{E}(L_\Sigma^n) = \mathbb{E}(M_\tau^{\rho n})\mathbb{E}(L_{0\Sigma}^n) = \frac{\Phi(\rho n+1) - \Phi(\rho n)}{\Phi(\rho n+1)} \frac{n!}{\Phi(\rho)\cdots\Phi(n\rho)}, \tag{36}$$

$$\mathbb{E}(L_*^n) = \mathbb{E}\big((M_\tau^*)^{\rho n}\big)\mathbb{E}(L_{0*}^n) = \frac{\Phi(\rho n+1, \rho n+1)}{\Phi(\rho n+1)} \frac{n!}{\Phi^*(\rho)\cdots\Phi^*(n\rho)}. \tag{37}$$

Note the two identities in distribution

$$L_{0\Sigma} \stackrel{d}{=} L_0 + M_\tau^\rho \widehat{L}_{0\Sigma} \quad \text{and} \quad L_{0*} \stackrel{d}{=} L_0 + (M_\tau^*)^\rho \widehat{L}_{0*}, \tag{38}$$

where $L_{0\Sigma} \stackrel{d}{=} \widehat{L}_{0\Sigma}$ with $\widehat{L}_{0\Sigma}$ independent of (L_0, M_τ), and $L_{0*} \stackrel{d}{=} \widehat{L}_{0*}$ with $\widehat{L}_{0*}$ independent of (L_0, M^*_τ). As checks, the two equalities of means implied by (38) are easily seen to be consistent with previous formulae. The equalities of higher moments in (38) provide identities involving joint moments such as $\mathbb{E}(L_0^j M_\tau^{k\rho})$ for positive integers j and k. In particular, $\mathbb{E}(L_0 M_\tau^\rho)$ is determined by the second moment formula. But the third moment formula only gives access to a linear combination of $\mathbb{E}(L_0 M_\tau^{2\rho})$ and $\mathbb{E}(L_0^2 M_\tau^\rho)$, which is not so useful.

3. Bead splitting processes and continuum random trees. Recall from Section 2.3 that every self-similar CRT $(\mathcal{T}, \mu)$ gives rise to a growing family $(\mathcal{R}^*_k, \mu^*_k)$ of weighted $\mathbb{R}$-trees that converges to $(\mathcal{T}, \mu)$. As we will demonstrate more formally below, picking Σ^*_{k+1} from μ means that a junction point $J^*_k \in \mathcal{R}^*_k$ is picked from μ^*_k and that Σ^*_{k+1} is then picked within a subtree rooted at J^*_k, which is a rescaled copy of $(\mathcal{T}, \mu)$, by self-similarity. Then

$$\mathcal{R}^*_{k+1} = \mathcal{R}^*_k \cup [\![J^*_k, \Sigma^*_{k+1}]\!] \quad \text{and} \quad \mu^*_{k+1} = \mu^*_k - \mu^*_k(\{J^*_k\})\delta_{J^*_k} + \mu^*_{k,k+1},$$

where $\mu^*_{k,k+1}$ denotes the projection onto $[\![J^*_k, \Sigma^*_{k+1}]\!]$ of the restriction of μ to the subtree rooted at J^*_k, so that $([\![J^*_k, \Sigma^*_{k+1}]\!], \mu^*_{k,k+1})$ is a rescaled copy of $(\mathcal{R}^*_1, \mu^*_1)$. Since J^*_k is picked from μ^*_k, we say that $((\mathcal{R}^*_k, \mu^*_k), k \geq 1)$ develops by *size-biased branching*, generalising the case $k = 1$ that relates to Proposition 13 via the self-similar time change (24).

3.1. *Size-biased bead selection and strongly sampling consistent compositions.* The basic building block for the tree growth process $((\mathcal{R}^*_k, \mu^*_k), k \geq 1)$ is a family of independent copies of $(\mathcal{R}^*_1, \mu^*_1)$, or equivalently, a family of independent copies of a fragmenter M^*, related by the following general construction.

DEFINITION 17 (String of beads). Given a decreasing pure jump process M and two positive real parameters α and m, we construct a *string of beads of mass* m by placing a random discrete measure $\mu_{M,\alpha,m}$ on the interval $(0, Y_{M,\alpha,m}]$ of random length

$$Y_{M,\alpha,m} = m^\alpha \int_0^\infty M_s^\alpha \, ds$$

according to the formula

$$\mu_{M,\alpha,m}\left(m^\alpha \int_0^t M_s^\alpha \, ds, m^\alpha \int_0^\infty M_s^\alpha \, ds \right) = m M_t.$$

If M is a fragmenter with Laplace exponent Φ, we call $([0, Y_{M,\alpha,m}], \mu_{M,\alpha,m})$ an (α, Φ)-*string of beads of mass* m.

Note that for each t that is a jump time of M, the measure $\mu_{M,\alpha,m}$ puts mass $m(M_{t-} - M_t)$ at the location $m^\alpha \int_0^t M_s^\alpha \, ds$. Now, by repeated application of this scheme, we construct an increasing sequence of $\mathbb{R}$-trees $(R_n, n \geq 1)$, where each R_n is equipped with a random discrete distribution μ_n.

DEFINITION 18 (Bead splitting process). Let $\alpha > 0$ and M_n, $n \geq 1$, be a sequence of decreasing pure jump processes starting from 1:

- Let (R_1, μ_1) be the string of beads of mass 1 associated with M_1 and α. More specifically, let $R_1 = [0, \Sigma_1] := [0, Y_{M_1,\alpha,1}]$ be equipped with the usual distance, with root vertex 0 and with the random discrete distribution $\mu_1 = \mu_{M_1,\alpha,1}$.
- Given that R_n has been defined as an $\mathbb{R}$-tree with root vertex 0 and n leaves $\Sigma_1, \ldots, \Sigma_n$, and equipped with a mass measure μ_n with total mass 1, let R_{n+1} be defined as follows. Pick a junction point J_n from R_n according to μ_n. Given $\mu_n(\{J_n\}) = m$, distribute the mass m according to a copy $([\![J_n, \Sigma_{n+1}]\!], \mu_{n,n+1})$ of the string of beads $([0, Y_{M_{n+1},\alpha,m}], \mu_{M_{n+1},\alpha,m})$, and then attach this segment to $(R_n, \mu_n - m\delta_{J_n})$ at J_n to form (R_{n+1}, μ_{n+1}).

We refer to the projective sequence $((R_n, \mu_n), n \geq 1)$ of weighted $\mathbb{R}$-trees as a *bead splitting process* that develops by *size-biased branching*.

We think of (R_n, μ_n) as n pieces of string $[0, Y_{M_1,\alpha,1}]$, $[0, Y_{M_2,\alpha,\mu(\{J_1\})}]$, $\ldots$, $[0, Y_{M_n,\alpha,\mu(\{J_{n-1}\})}]$ tied at the junction points $J_1, \ldots, J_{n-1}$, with beads according to μ_n. The nth growth step selects bead J_n of size $\mu_n(\{J_n\})$ and splits it into smaller beads that are placed onto a new piece of string tied to J_n.

The growth process $((R_n, \mu_n), n \geq 1)$ gives rise to an ordered spinal partition of $\mathbb{N} \setminus \{1\}$ in the terminology of [23], which we can represent by a point process

$$\Pi_s^{\mathrm{ord}} = \{n \in \mathbb{N} \setminus \{1\} : J_{1,n} = g_{0,\Sigma_1}(s)\}, \qquad s \geq 0, \tag{39}$$

where $J_{1,n}$ is the branch point that has Σ_1 and Σ_n in two different subtrees, where $[\![0, J_{1,n}]\!] = [\![0, \Sigma_1]\!] \cap [\![0, \Sigma_n]\!]$, and $g_{0,\Sigma_1} : [0, d(0, \Sigma_1)] \to [\![0, \Sigma_1]\!]$ is the unique isometry with $g_{0,\Sigma_1}(0) = 0$.

PROPOSITION 19. *Given any bead splitting process* $((R_n, \mu_n), n \geq 1)$ *that develops by size-biased branching, the ordered spinal partition* Π^{ord} *defined in* (39) *is exchangeable. In particular, if we choose* M_1 *to be a fragmenter, then the spinal partition gives rise to a* (*strongly sampling consistent*) *regenerative composition structure* $(\mathcal{C}_n, n \geq 1)$, *which records for each* $n \geq 1$ *the vector* $\mathcal{C}_n$ *of nonzero block sizes* $\#(\Pi_s^{\mathrm{ord}} \cap [n+1])$, $s \geq 0$, *of* $\Pi^{\mathrm{ord}} \cap [n+1]$, *arranged in the spinal order of blocks given by the order of positions* $s \geq 0$ *on the spine.*

PROOF. The first statement holds since μ_1 is the projection of μ_n to R_1 for all $n \geq 1$, so the picks of J_n projected to R_1 are exchangeable picks from μ_1 by the use of size-biased branching. The second statement now follows directly from Gnedin and Pitman [17], Theorem 5.2; cf. also [30], Section 2.1, for the terminology of (weak and) strong sampling consistency. □

3.2. *Convergence of bead splitting processes to self-similar CRTs.* The next theorem establishes CRT convergence of bead splitting processes $((R_n, \mu_n), n \geq 1)$ in the sense of Definition 18, not just for the case of symmetric splitting rules f^* that relate directly to the growth process $((\mathcal{R}_n^*, \mu_n^*), n \geq 1)$ obtained by sampling from the measure μ of a CRT $(\mathcal{T}, \mu)$, but also for fragmenters $(M_n, n \geq 1)$ with nonsymmetric splitting rules f, with convergence to a CRT associated with the symmetrised splitting rule f^* associated with f. Again, the result holds without assuming the existence of densities f and f^*:

THEOREM 20. *For independent fragmenters M_n with Laplace exponent $\Phi(\rho) = \int_{(0,1)}(1-u^\rho)\Lambda(du)$ and $\alpha > 0$, the sequence of weighted random $\mathbb{R}$-trees (R_n, μ_n) converges almost surely in the Gromov–Hausdorff–Prohorov metric to a limit tree $(\mathcal{T}, \mu)$, which is a copy of the α-self-similar tree that is canonically associated with a binary fragmentation process with symmetric dislocation measure $\nu = \Lambda + \overline{\Lambda}$. In addition, we also have $(R_n, \nu_n) \to (T, \mu)$ almost surely in the Gromov–Hausdorff–Prohorov metric, where ν_n is the uniform measure on the n leaves of R_n.*

If $f(u) = \frac{1}{\sqrt{2\pi}} u^{-3/2}(1-u)^{-3/2}$, it follows easily from the work of Haas and Miermont [20] that the sequence $(R_n, n \geq 1)$ has the same distribution as the increasing sequence of trees provided by Aldous's [2] line-breaking construction of the Brownian CRT. Therefore, Theorem 20 can be seen as a generalisation of Aldous's line-breaking construction. We discuss this example of a bead splitting process $((R_n, \mu_n), n \geq 1)$ in Section 3.3.

To prove this theorem, we will embed (R_n, μ_n) in a CRT (T, μ), as has (essentially) been done for (R_1, μ_1) in Theorem 14. A key tool will be the following spinal decomposition result.

LEMMA 21 (Spinal decomposition). *Let $A = (A_t, t \geq 0)$ be a Markovian path in a homogeneous fragmentation process $\Pi = (\Pi(t), t \geq 0)$. For each $n \geq 1$, denote by $\Pi^{\{n\}}(t)$ the block of $\Pi(t)$ containing n, $t \geq 0$, consider $\sigma_n = \inf\{t \geq 0 : n \notin A_t\}$ and the associated spinal partition $\Pi^A(0) = \{\Pi^{\{n\}}(\sigma_n), n \geq 1\}$. Then conditionally given $\Pi^A(0)$ and $(\sigma_n, n \geq 1)$, the process*

$$\Pi^A(t) = \{\Pi^{\{n\}}(\sigma_n + t), n \geq 1\} = \bigcup_{i \geq 1} \Pi(\sigma_{\min \Pi_i^A(0)} + t) \cap \Pi_i^A(0), \qquad t \geq 0$$

is a fragmentation process starting from $\Pi^A(0)$, with the same transition kernel as Π.

This lemma says that the process Π can be decomposed along the path A into the partition $\Pi^A(0)$ of blocks that separate from A at any time $t \geq 0$. The blocks $\Pi_i^A(0)$, $i \geq 1$ then evolve independently and according the transition kernel of Π.

PROOF OF LEMMA 21. We extend the proof of [23], Proposition 4, to the higher generality here of decomposing along a Markovian path. The family of times $(\sigma_n, n \geq 1)$ is a stopping line for the filtration $\mathcal{F} = (\mathcal{F}_t, t \geq 0)$, with respect to which A is a Markovian path in Π. Here, we use the terminology of Bertoin [7], Definition 3.4, and seek to obtain from [7], Lemma 3.14, that the *extended branching property* holds, which yields precisely the result we need. Since Bertoin uses natural filtrations, and to demonstrate where the Markovian assumption on the path enters the argument, let us briefly retrace Bertoin's steps and sketch relevant parts of the proof of the extended branching property. Without loss of generality, $\mathcal{F}$ is the filtration generated by (Π, A). Also denote by $\mathcal{F}^{\{n\}}$ the filtrations generated by $(\Pi^{\{n\}}, A_{\cdot\wedge\sigma_n})$, for each $n \geq 1$. We consider approximations $\sigma_n^{(h)} = \inf\{t \in h\mathbb{N} : n \notin A_t\}$, $\sigma_n^{(h,k)} = \min(kh, \sigma_n^{(h)})$, and $\overline{\sigma}_n^{(h,k)} = \inf\{t \in \{h, 2h, \ldots, kh\} : n \notin A_t\}$ with $\inf \varnothing := \infty$, of σ_n. The branching property at the stopping line $(\sigma_n^{(h,1)}, n \geq 1)$ is just the branching property at $t = h$. At h, or by induction hypothesis at $(\sigma_n^{(h,k)}, n \geq 1)$, the assumption on the path to be Markovian ensures that $(\Pi(kh + t) \cap A_{kh}, A_{kh+t}, t \geq 0)$ is conditionally independent of $(\{\Pi^{\{n\}}(\sigma_n^{(h,k)} + t), n \notin A_{kh}\}, t \geq 0)$ given

$$\mathcal{F}_{(\sigma_n^{(k,h)}, n\geq 1)}, \text{ defined as the sigma-algebra generated by } \mathcal{F}^{\{n\}}_{\sigma_n^{(h,k)}}, n \geq 1.$$

To $(\Pi_{kh+t} \cap A_{kh}, A_{kh+t}, t \geq 0)$, we can apply the branching property at $t = h$ and trivially at $t = \infty$ to complete the induction step from k to $k + 1$. This establishes the extended branching property at $(\sigma_n^{(h,k)}, n \geq 1)$ and $(\overline{\sigma}_n^{(h,k)}, n \geq 1)$ for all $h > 0$ and $k \geq 1$. We omit the remainder of the proof, which uses the standard approximation $\overline{\sigma}_n^{2^{-k},2^{2k}} \downarrow \sigma_n$ as $k \to \infty$. $\square$

The next lemma and its proof demonstrate that we can iterate the embedding of a Markovian path in a homogeneous fragmentation process to embed a bead splitting process in an associated self-similar CRT to which the bead splitting process converges almost surely.

LEMMA 22. *Let $A = (A_t, t \geq 0)$ be a Markovian path in a binary fragmentation process $\Pi = (\Pi(t), t \geq 0)$ and $M_t = |A_t|$, $t \geq 0$ its residual mass process. If the M_n, $n \geq 1$ are independent copies of M, then for each $\alpha > 0$ the sequence of weighted random $\mathbb{R}$-trees (R_n, μ_n) converges almost surely in the Gromov–Hausdorff–Prohorov metric to a limit tree $(\mathcal{T}, \mu)$, which is a copy of the α-self-similar CRT that is canonically associated with Π. In addition, we also have $(R_n, \nu_n) \to (\mathcal{T}, \mu)$ almost surely in the Gromov–Hausdorff–Prohorov metric, where ν_n is the uniform measure on the n leaves of R_n.*

PROOF. We can consider the spinal partition $\Pi^A(0)$ of Π and use Lemma 21 to construct in a measurable way (see, e.g., [23], Corollary 3) a string of

beads $(R_1, \mu_1) = ([0, Y_{|A|,\alpha,1}], \mu_{|A|,\alpha,1})$ with a collection $(s_i, T^{(i)}, \mu^{(i)})$ of spinal subtrees constructed from $(\Pi(\sigma_{\min \Pi_i^A(0)} + t) \cap \Pi_i^A(0), t \geq 0)$, where $s_i = \int_0^{\sigma_{\min \Pi_i^A(0)}} |A_r|^\alpha \, dr$, $i \geq 1$, such that the tree $(\mathcal{T}, \mu)$ obtained by grafting $(T^{(i)}, \mu^{(i)})$ to $(R_1, 0)$ at s_i for all $i \geq 1$, is a self-similar CRT associated with Π. This construction gives rise to a family of regular conditional distributions of $(R_1, \mu_1; (s_i, T^{(i)}, \mu^{(i)}), i \geq 1)$ given $(\mathcal{T}, \mu)$, and we can use these via the Ionescu–Tulcea theorem to obtain a probability space that allows the following construction.

Suppose we have constructed $(R_n, \mu_n; (x_i, T^{(i)}, \mu^{(i)}), i \in I_n)$ with $R_n \subset \mathcal{T}$, μ_n the projection of μ onto R_n and, conditionally given (R_n, μ_n), a collection $((T^{(i)}, \mu^{(i)}), i \in I_n)$ of independent copies of $(\mathcal{T}, \mu)$ subjected to α-self-similar scaling by $\mu(\{x_i\})$, which when grafted at $x_i \in R_n$ for all $i \in I_n$ give $(\mathcal{T}, \mu)$. Now pick a junction point $J_n = x_{i_n}$ from R_n according to μ_n. Given that $\mu_n(\{J_n\}) = m$, remove J_n from μ_n and remove i_n from I_n. Use the regular conditional distribution given the rescaled chosen subtree $(T^{(i_n)}, \mu^{(i_n)})$ to obtain a string of beads with grafted spinal subtrees distributed as $(R_1, \mu_1; (s_i, T^{(i)}, \mu^{(i)}), i \geq 1)$, without modifying the chosen rescaled subtree. After α-self-similar scaling by m, graft the string of beads at J_n, add the new spinal subtrees to the collection to form $(R_{n+1}, \mu_{n+1}; (x_i, T^{(i)}, \mu^{(i)}), i \in I_{n+1})$. Then $R_{n+1} \subset \mathcal{T}$, μ_{n+1} is the projection of μ onto R_{n+1} and, conditionally given (R_{n+1}, μ_{n+1}), the collection $((T^{(i)}, \mu^{(i)}), i \in I_{n+1})$ consists of scaled independent copies of $(\mathcal{T}, \mu)$ that turn R_{n+1} into $(\mathcal{T}, \mu)$ when grafted at x_i, $i \in I_{n+1}$.

By induction, this gives a sequence $((R_n, \mu_n), n \geq 1)$ embedded in $(\mathcal{T}, \mu)$, which develops by size-biased branching and is based on independent copies of the string of beads associated with $M = |A|$. While constructed within a CRT, this sequence has the same (joint) distribution as the sequence described in the statement of the lemma. It therefore suffices to prove almost sure convergence for this embedded sequence.

First, consider the measures μ_n, $n \geq 1$. Denote by $|\mu_n|^\downarrow \in \mathcal{S}^\downarrow$ the decreasing rearrangement of bead sizes $\mu_n(\{x\})$, $x \in R_n$. Since A is embedded in $(\mathcal{T}, \mu)$, the measure μ_1 cannot have an atom of size 1; in particular there is $\lambda < 1$ such that $\mathbb{P}(|\mu_1|_1^\downarrow < \lambda) > 0$. Now let $\varepsilon = 1/K > 0$. By selecting the K largest beads in turn, we see that $\mathbb{P}(|\mu_{n+K}|_1^\downarrow < \lambda s_1 \vee \varepsilon || \mu_n|^\downarrow = \mathbf{s}) > 0$ for all $\mathbf{s} \in \mathcal{S}^\downarrow$. For m with $\lambda^m < \varepsilon$ this implies $p = \mathbb{P}(|\mu_m|_1^\downarrow < \varepsilon) > 0$, but then $|\mu_n|_1^\downarrow$ will be less than ε after a time that is bounded above by m times a geometric random variable with parameter p. In particular,

$$\mathbb{P}(\mu_n \text{ has an atom of size greater than } \varepsilon \text{ for all } n \geq 1) = 0. \tag{40}$$

Now denote by R_∞ the completion of the increasing union $\bigcup_{n \geq 1} R_n$ in $\mathcal{T}$, and assume that $\mathbb{P}(R_\infty \neq \mathcal{T}) > 0$. For $x \in \mathcal{T} \setminus R_\infty$, we find $y \in R_\infty$ such that $]\!]y, x]\!] \cap R_\infty = \varnothing$, but then $\mu(\mathcal{T}_y) > 0$, since μ assigns positive weight to all fringe subtrees. Since μ_n is the projection of μ onto $R_n \subset R_\infty$, this contradicts (40). Hence $\mathbb{P}(R_\infty = \mathcal{T}) = 1$. Similarly, assuming that R_n does not converge to $\mathcal{T}$ for the

Hausdorff distance on $\mathcal{T}$, we can use compactness to find $x \in \mathcal{T}$ with $d(x, R_n) > \varepsilon$ for all $n \geq 1$, so $x \notin R_\infty$ is a contradiction. Also, Hausdorff convergence of $R_n \subset \mathcal{T}$ to $\mathcal{T}$ with projected measures implies Gromov–Hausdorff–Prohorov convergence $d_{\mathrm{GHP}}(R_n, \mathcal{T}) \to 0$ almost surely as $n \to \infty$; see, for example, [30], Lemma 17.

Finally, the measure ν_n on the n leaves on R_n is more and more closely coupled with the measure ν_n^* on $\mathcal{T}$ associated with a sample $\Sigma_1^*, \ldots, \Sigma_n^*$ from μ, which is well known to converge weakly almost surely to μ. We can take as Σ_1^* an independent pick from μ and include Σ_{n+1}^* in the construction of (R_{n+1}, μ_{n+1}). Specifically, we can obtain the pick from μ_n for the junction point J_n as the junction point of the subtree containing Σ_{n+1}^*. Since $R_n \to \mathcal{T}$ almost surely in the Hausdorff sense, there is n_0 such that all subtrees of $\mathcal{T} \setminus R_{n_0}$ have height less than $\varepsilon/2$, but then the distance between Σ_{n+1}^* and the $(n+1)$st leaf of R_{n+1}, which are in the same subtree by construction, is at most ε for all $n \geq n_0$, which entails the result by standard arguments. $\square$

PROOF OF THEOREM 20. Let M be a fragmenter with Laplace exponent $\Phi(\rho) = \int_{(0,1)}(1 - u^\rho)\Lambda(du)$. By the existence part of Theorem 14, there is a Markovian embedding of M into Π for a binary homogeneous fragmentation process with symmetric dislocation measure $\nu = \Lambda + \overline{\Lambda}$. Hence, Lemma 22 applies and gives $(R_n, \mu_n) \to (\mathcal{T}, \mu)$ and $(R_n, \nu_n) \to (\mathcal{T}, \mu)$ almost surely in the Gromov–Hausdorff–Prohorov sense, for an α-self-similar CRT $(\mathcal{T}, \mu)$ with symmetric dislocation measure ν. $\square$

This establishes Theorem 20, and in particular shows that the bead-splitting processes of Definition 18 answer points 4–5 from the Introduction, and that point 6 then also holds. We can now use Theorem 20 and Lemma 22 to complete the proof of Theorem 14, and hence establish point 2, completing the programme of 6 points set out in the Introduction:

PROOF OF THE UNIQUENESS PART OF THEOREM 14. Let $\Phi(\rho) = \int_{(0,1)}(1 - u^\rho)\Lambda(du)$ be the Laplace exponent of a fragmenter. Consider the bead splitting process $((R_n, \mu_n), n \geq 1)$ based on a sequence M_n, $n \geq 1$, of such fragmenters. In Theorem 20 we showed that $(R_n, \mu_n) \to (\mathcal{T}, \mu)$ almost surely for a CRT $(\mathcal{T}, \mu)$ with symmetrised dislocation measure $\nu = \Lambda + \overline{\Lambda}$.

Now assume that a fragmenter with Laplace exponent Φ has a Markovian embedding A into an exchangeable binary homogeneous fragmentation process Π with any symmetric dislocation measure $\widetilde{\nu}$. By Lemma 22, $(R_n, \mu_n) \to (\widetilde{\mathcal{T}}, \widetilde{\mu})$ almost surely for a CRT with symmetrised dislocation measure $\widetilde{\nu}$. By uniqueness of limits, $(\widetilde{\mathcal{T}}, \widetilde{\mu}) = (\mathcal{T}, \mu)$. Since the distributions of CRTs for different dislocation measures are different, we find that $\nu = \widetilde{\nu}$. $\square$

In the remainder of the paper, we point out some further connections to related work.

REMARK 23. With the usual names $\Sigma_1, \dots, \Sigma_n$ of leaves of R_n, $n \geq 1$, any bead splitting process embedded in a CRT $(\mathcal{T}, \mu)$ gives rise to a, typically nonexchangeable, $\mathcal{P}$-valued process

$$\Pi_\alpha(t) = \{\{j \geq 1 : \Sigma_j \in \mathcal{T}_i^t\}, i \geq 1\} \cup \{\{j\}, j \geq 1 : \Sigma_j \notin \mathcal{T}^t\}, \qquad t \geq 0,$$

of the same form as the exchangeable special case Π_α^* in (25). Furthermore, if $(R_n, \mu_n) \to (\mathcal{T}, \mu)$ as in Lemma 22, we find equality of the decreasing rearrangements of asymptotic frequencies $|\Pi_\alpha(t)|^\downarrow = |(\mu(\mathcal{T}_i^t), i \geq 1)|^\downarrow = |\Pi_\alpha^*(t)|^\downarrow$ for all $t \geq 0$ a.s. So it is natural to perform the inverse of the self-similar time change (24) to construct a, typically nonexchangeable, homogeneous process $\Pi = \Pi_0$ from the consistently time-changed evolution of its blocks containing n, $n \geq 1$.

Our proof of the uniqueness part of Theorem 14 used the size-biased bead splitting process and the compactness of self-similar CRTs to show that the Markovian path A gives rise to an embedding that exhausts a CRT. The embedding for the existence part of Theorem 14 was not carried out in a CRT, but directly in an exchangeable homogeneous fragmentation process. Indeed, it should be possible to also prove the uniqueness in the framework of homogeneous fragmentation processes. We can rephrase our bead splitting argument for the uniqueness part here to directly construct a nonexchangeable process Π based on A by embedding into an exchangeable homogeneous fragmentation process Π^*, as indicated below. However, this is harder to formulate, and we lose natural compactness, so we do not attempt an alternative proof, but let us give the direct construction of Π.

Let A be a Markovian path in Π^*. Define branch times $J_{1,n} = \inf\{t \geq 0 : n \notin A_t\}$, $n \geq 2$, between 1 and n. Given $J_{i,n}$, $n \geq i+1$, have been constructed for all $i \in [k] := \{1, \dots, k\}$, consider the time $H_{k+1} = \max\{J_{i,k+1}, i \in [k]\}$ when $k+1$ separates from the last $i \in [k]$. Relabel the restriction of $(\Pi^*(H_{k+1} + t), t \geq 0)$ to the block B_{k+1} of $\Pi^*(H_{k+1})$ that contains $k+1$ by the increasing bijection $B_{k+1} \to \mathbb{N}$. Run a copy of A inside this process, relabel back $\mathbb{N} \to B_{k+1}$ to find a Markovian path $B^{(k+1)}$ that we specify to coincide with the canonical path A^{k+1} of Π^* up to H_{k+1} and to continue in $(\Pi^*(H_{k+1}+t) \cap B_{k+1}, t \geq 0)$, as constructed. Define $J_{k+1,n} = \inf\{t \geq 0 : n \notin B_t^{(k+1)}\}$, $n \geq k+2$. Finally set $J_{k,k} = \infty$, $J_{k,n} = J_{n,k}$ for $n < k$ and define the embedded $\mathcal{P}$-valued process

$$\Pi_t = \{\{n \geq 1 : J_{n,i} > t\}, i \geq 1\}, \qquad t \geq 0. \tag{41}$$

COROLLARY 24. *Let A be a Markovian path in an exchangeable homogeneous fragmentation process Π^*, and let Π be as in (41). Then $|\Pi|^\downarrow = |\Pi^*|^\downarrow$. Moreover, if $|A|$ is a fragmenter, then Π is a homogeneous fragmentation process with binary nonexchangeable κ-measure [29] of the form*

$$\kappa(\{\Gamma \in \mathcal{P} : \Gamma \cap [n] = (\pi_1, \pi_2)\}) = \int_{(0,1)} u^{\#\pi_1 - 1}(1-u)^{\#\pi_2} \Lambda(du),$$
$$\{\pi_1, \pi_2\} \in \mathcal{P}_n \setminus \{\{[n]\}\},$$

where $\mathcal{P}_n$ is the set of partitions of $[n] := \{1, \dots, n\}$, *with* $\kappa(\Gamma \in \mathcal{P} \setminus \{\{\mathbb{N}\}\} : \Gamma_1 \cup \Gamma_2 \neq \mathbb{N}) = 0$.

PROOF. We leave the equivalence of the two constructions of Π to the reader and just point out that the CRT construction of Remark 23 yields $|\Pi| = |\Pi^*|$. For the second claim, we note that the fragmenter has Laplace exponent $\Phi(\rho) = \int_{(0,1)}(1 - u^\rho)\Lambda(du)$, so, by standard thinning properties of the Poisson point process of jumps of the fragmenter M and size-biased branching, we identify the dislocation measure. □

REMARK 25. It may be observed from the form of the bead splitting process in the case of an independent and identically distributed sequence $(M_n, n \geq 1)$ that the size-biased bead selection rule is not crucial for convergence to a CRT since it mainly affects the (random) time n at which a particular bead is split. In the proof of Lemma 22, the main use of the size-biased selection rule was to establish (40). Indeed, as long as we split every bead eventually, we are quite free to choose the order in which we split the beads and may even contemplate rules like splitting all beads of μ_n at once at every stage of the bead splitting process.

The reader may also want to compare our bead splitting processes with Abraham's [1] construction of a version of the Brownian CRT. Let us rephrase Abraham's construction in our present framework. The construction is based on the distribution of the total height of the CRT and the decomposition of the CRT along the path from the root to the *highest* leaf. If we project the mass measure of the CRT onto the spine, we obtain a string of beads $(\overline{R}_1, \overline{\mu}_1)$. Abraham takes this string of beads and recursively replaces all beads of $\overline{\mu}_n$ by a rescaled copy of $(\overline{R}_1, \overline{\mu}_1)$ conditioned not to exceed the height of the branch of the bead. The path to the highest leaf does not correspond to a Markovian path, and the spinal subtrees are not rescaled copies of the CRT, but copies constrained in height, this falls outside the setting of Lemma 22.

Intuitively, the *Markovian* path A, whose leaf height $L_{0\Sigma} = \int_0^\infty |A_t|^\alpha \, dt$ in an associated α-self-similar CRT is most likely to be highest is the one based on the switching probabilities of Example 8, always choosing the bigger fragment. The homogeneous Poissonian structure for relative masses $(F_t, 1 - F_t)$ on the spine to the embedded leaf easily entails that this does not lead to the highest leaf a.s.

3.3. *The Brownian CRT.* Let $(\mathcal{T}, \mu)$ denote the Brownian Continuum Random Tree $\mathcal{T}$ equipped with its mass measure μ, which Aldous [3] constructed both as the tree embedded in (twice the standard) Brownian excursion, with μ corresponding to Lebesgue measure on $[0, 1]$, and as a limit as $n \to \infty$ of an increasing sequence of binary subtrees with edge-lengths R_n with n leaves labelled by $[n]$, in which case μ may be interpreted as the almost sure weak limit as $n \to \infty$ of the uniform probability distribution ν_n on the n leaves of R_n. The tree R_n may be constructed as the subtree of $\mathcal{T}$ spanned by n leaves of $\mathcal{T}$, which given $\mathcal{T}$ are picked

independently according to the mass measure μ. We recover this second construction in Theorem 20 for $f(u) = \frac{1}{\sqrt{2\pi}} u^{-3/2}(1-u)^{-3/2}$, enriched by the string of beads structure given to the branches of R_n by measures μ_n. According to a basic result of Aldous [3], the increasing sequence of lengths $(\lambda(R_1), \lambda(R_2), \ldots)$ of these subtrees can be constructed as $\lambda(R_n) = \sqrt{2\Gamma_n}$ where $\Gamma_n = \varepsilon_1 + \cdots + \varepsilon_n$ for a sequence of independent standard exponential variables ε_i. For $n = 1$ there is the identity in distribution

$$\lambda(R_1) \stackrel{d}{=} L_1^0(B^{\mathrm{br}}), \tag{42}$$

where B^{br} is a standard Brownian bridge, starting at 0 at time 0 and ending at 0 at time 1, and $(L_t^x(B^{\mathrm{br}}), 0 \le t \le 1, x \in \mathbb{R})$ is the jointly continuous process of local times of B^{br}, normalised so that $L_t^x(B^{\mathrm{br}})\,dx$ is the occupation measure of $(B^{\mathrm{br}}(s), 0 \le s \le t)$. The common distribution of both sides in (42) has the Rayleigh density $x\exp(-x^2/2)$ at $x \ge 0$. The work of Aldous, Miermont, and Pitman [4] yields a deeper connection between the Brownian CRT in a Brownian excursion on the one hand and Brownian bridge on the other. This establishes a spinal decomposition result for the Brownian CRT via path transformations rather than via Bertoin's extended branching property as in [23] or Lemma 21 here. See also Bertoin and Pitman [8], Theorem 3.2, for an expression in terms of paths rather than trees. To explore this spinal decomposition, let $(\tau_\ell, 0 \le \ell < L_1^0(B^{\mathrm{br}}))$ be the inverse local time process

$$\tau_\ell := \inf\bigl\{t \ge 0 : L_t^0(B^{\mathrm{br}}) > \ell\bigr\}$$

so that the collection of excursion intervals of B^{br} is exhausted by

$$\bigl\{(\tau_{\ell-}, \tau_\ell) : \ell > 0, \tau_{\ell-} < \tau_\ell\bigr\},$$

and let P be the random discrete distribution obtained by ranking these intervals by length. On the other hand, in the Brownian CRT $(\mathcal{T}, \mu)$, for $0 \le \ell < \lambda(R_1)$ let $\mu_1([0,\ell])$ denote the mass of all points $x \in \mathcal{T}$ such that the path from root to x in $\mathcal{T}$ branches from the path from root to leaf Σ_1 of $\mathcal{T}$ at a junction point on R_1 whose distance from the root of $\mathcal{T}$ is at most ℓ. Then according to the spinal decomposition of the Brownian CRT implied by [4], Lemma 9 and equation (12), the equality in distribution (42) extends to the equality in distribution of processes

$$\bigl(\mu_1([0,\ell]), 0 \le \ell < \lambda(R_1)\bigr) \stackrel{d}{=} \bigl(\tau_\ell, 0 \le \ell < L_1^0(B^{\mathrm{br}})\bigr). \tag{43}$$

Moreover, conditionally given the process on the left-hand side of (43), the Brownian CRT $(\mathcal{T}, \mu)$ decomposes into a countable collection of subtrees

$$\bigl\{\mathcal{T}_\ell, 0 < \ell < \lambda(R_1), \mu_1(\{\ell\}) > 0\bigr\},$$

where $\mathcal{T}_\ell$ is a Brownian CRT equipped with a mass measure μ_ℓ with total mass $\mu_\ell(\mathcal{T}_\ell) = \mu_1(\{\ell\}) > 0$. This decomposition corresponds on the right-hand side of (43) to (trees in) excursions of $|B^{\mathrm{br}}|$, the absolute value of B^{br}, excursions of

lengths $\{\tau_\ell - \tau_{\ell-}, 0 < \ell < L_0^1(B^{\mathrm{br}}), \tau_{\ell-} < \tau_\ell\}$. Indeed, the entire Brownian CRT $\mathcal{T}$ can be constructed from B^{br} so that the equality in distribution (43) holds almost surely, and for each ℓ with $\mu_1(\{\ell\}) = \tau_\ell - \tau_{\ell-} > 0$, the subtree $\mathcal{T}_\ell$ of $\mathcal{T}$ attached to the spine R_1 of $\mathcal{T}$ at distance ℓ from the root is constructed from the excursion of $|B^{\mathrm{br}}|$ on $(\tau_{\ell-}, \tau_\ell)$ in the same way that the entire tree $\mathcal{T}$ is constructed from a standard Brownian excursion. In particular, basic properties of Brownian excursions then imply the *spinal decomposition of* $\mathcal{T}$, that conditionally given the subtree masses $(\mu_1(\{\ell\}), 0 \le \ell < \lambda(R_1))$, the subtrees $(\mathcal{T}_\ell, \mu_\ell)$ associated with ℓ such that $\mu_1(\{\ell\}) > 0$ form a collection of independent random trees distributed like $(\sqrt{\mu_1(\{\ell\})}\mathcal{T}, \mu_1(\{\ell\})\mu)$, meaning that all edge-lengths in $\mathcal{T}$ are scaled by a factor of $\sqrt{\mu_1(\{\ell\})}$, while all masses are scaled by a factor of $\mu_1(\{\ell\})$.

The distribution of ranked masses of atoms of μ_1 is the Poisson–Dirichlet distribution $\mathrm{PD}(\frac{1}{2}, \frac{1}{2})$, which is the distribution of ranked lengths of excursions of Brownian bridge, and masses corresponding to these lengths are distributed along the spine of length $\lambda(R_1)$ in an exchangeable random order. Moreover, the length $\lambda(R_1)$ is itself a measurable functional of the $\mathrm{PD}(\frac{1}{2}, \frac{1}{2})$ random discrete distribution of masses along the spine, as discussed in [26, 28].

For $n \ge 1$, the bead splitting process $((R_n, \mu_n), n \ge 1)$ described in terms of decreasing mass processes in Definition 18 can be described in terms of Brownian bridges, as follows:

- Start from a segment

$$(R_1, \mu_1) = \big([0, L_1^0(B^{\mathrm{br}})], d\tau\big), \tag{44}$$

  where $d\tau$ denotes the Stieltjes measure with cumulative distribution function given in (43), associated with the inverse local time τ of a standard Brownian bridge B^{br} of length 1.
- Given (R_n, μ_n), pick a junction point J_n from R_n according to μ_n. Given that $\mu_n(\{J_n\}) = m$, remove the mass m from point J_n and attach as segment $(]\!]J_n, \Sigma_{n+1}]\!], \mu_{n,n+1})$ a copy of (44) derived from a Brownian bridge of length m.

Specifically, a Brownian bridge of length m may be constructed from the standard Brownian bridge B^{br} as $m^{1/2}B^{\mathrm{br}}(t/m)$, $0 \le t \le m$. That is to say, $R_{n+1} \setminus R_n =]\!]J_n, \Sigma_{n+1}]\!]$ is such that

$$\lambda(R_{n+1}) - \lambda(R_n) = \big(\mu_n(\{J_n\})\big)^{1/2} L_1^0\big(B_{(n)}^{\mathrm{br}}\big)$$

for some independent and identically distributed sequence of standard Brownian bridges

$$B_{(n)}^{\mathrm{br}} = \big(B_{(n)}^{\mathrm{br}}(t), 0 \le t \le 1\big),$$

and given $\mu_n(\{J_n\}) = m$, the mass m should be reallocated with a portion $(\tau_\ell - \tau_{\ell-})m$ placed at distance $\ell m^{1/2}$ from J_n along the branch of length $\lambda(R_{n+1}) - \lambda(R_n)$ from J_n to Σ_{n+1}, for each $\ell \in (0, L_1^0(B^{\mathrm{br}}_{(n)}))$ with $\tau_{\ell-} < \tau_\ell$.

The above prescription specifies the projective sequence of weighted $\mathbb{R}$-trees $((R_n, \mu_n), n \geq 1)$ as in Theorem 20, which is associated with the Brownian CRT as in [20] and as indicated at the beginning of Section 3. We wish to point out some special properties of this Brownian tree growth sequence:

PROPOSITION 26. *Let $((R_n, \mu_n), n \geq 1)$ be the bead splitting process derived from a sequence of Brownian bridges, as above. Then we have the following description of the law of (R_n, μ_n):*

(i) *The sequence $P_n = (P_{n,1}, P_{n,2}, \ldots)$ of sizes of ranked atoms of μ_n has* $\mathrm{PD}(\frac{1}{2}, (2n-1)/2)$ *distribution.*

(ii) *The total length $\lambda(R_n)$ can be represented as both $\lambda(R_n) = \sqrt{2\Gamma_n}$ where $\Gamma_n = \varepsilon_1 + \cdots + \varepsilon_n$ for a sequence of independent standard exponential variables ε_i, and as $\lambda(R_n) = S_{1/2}(P_n)/\sqrt{2}$, where $S_{1/2}(P_n)$ is the $\frac{1}{2}$-diversity of P_n, which may be recovered from P_n as*

$$S_{1/2}(P_n) = \sqrt{\pi} \lim_{k\to\infty} k P_{n,k}^{1/2} \ a.s.,$$

where $P_{n,k}$ is the kth largest μ-measure of the collection of all fringe subtrees of $\mathcal{T}$ attached to R_n, or again as

$$S_{1/2}(P_n) = \lim_{m\to\infty} m^{-1/2} K_{n,m},$$

where $K_{n,m}$ is the number of junction vertices J_i with $i \leq m$ such that $J_i \in R_n$.

(iii) *Conditionally given $\lambda(R_n)$, the tree R_n consists of $2n-1$ segments, whose relative lengths, when listed in order of depth-first search of R_n, passing first along $[\![0, \Sigma_1]\!] = R_1$, then along $]\!]J_1, \Sigma_2]\!] = R_2 \setminus R_1$, and so on, is independent of R_n, with the same Dirichlet distribution with $2n-1$ parameters equal to 1 as the sequence of $2n-1$ spacings generated by a sequence of $2n-2$ independent uniform variables on $[0,1]$.*

(iv) *For $n \geq 2$ the combinatorial shape of R_n is equally likely to be any of the $1 \times 3 \times \cdots \times (2n-3)$ possible shapes of binary trees with root 0 and leaves labeled by $[n]$, independently of $\lambda(R_n)$ and of the sequence of relative lengths of segments.*

(v) *Conditionally given $\lambda(R_n) = \ell_1$, the combinatorial shape of R_n and the sequence of relative lengths of segments, let $(\sigma_v, 0 \leq v \leq \ell_1)$ be a path which traverses R_n at unit speed, passing first along $[\![0, L_1]\!] = R_1$, then along $]\!]J_1, L_2]\!] = R_2 \setminus R_1$, and so on, and let $R_{n,\ell}$ be the range of $(\sigma_v, 0 \leq v \leq \ell)$, so that by construction $\lambda(R_{n,\ell}) = \ell$ for all $0 \leq \ell \leq \lambda(R_n) = \ell_1$. Then the cumulative mass process $(\mu_n(R_{n,\ell}), 0 \leq \ell \leq \lambda(R_n) | \lambda(R_n) = \ell_1)$ has the same distribution with exchangeable increments as the inverse of the local time process at 0 of a Brownian bridge B^{br} conditionally on $L_1^0(B^{\mathrm{br}}) = \ell_1$.*

PROOF. This can largely be read from the preceding discussion and known descriptions of $(R_n, n \geq 1)$ and properties of Poisson–Dirichlet distributions. Partial results appear in many places, including [3], Section 4.3, [22], Proposition 18, and [30], Proposition 14.

In the terminology of [14], part (i) is a particular case of their result that crushing a size-biased pick from a ranked list with distribution $\mathrm{PD}(\alpha, \theta)$ into $\mathrm{PD}(\alpha, 1-\alpha)$-distributed proportions yields a $\mathrm{PD}(\alpha, \theta+1)$ ranked vector. This can also be read from Aldous's sequential description of the growth of $(R_n, n \geq 1)$. That description implies part (v) quite easily. To deduce (ii) from (v), observe that P_n is the sequence of ranked jumps of the cumulative mass process $(\mu_n(R_{n,\ell}), 0 \leq \ell \leq \lambda(R_n))$ which given $\lambda(R_n) = \ell$ is distributed like the ranked lengths of excursion intervals of a standard Brownian bridge. As n changes, the distribution of P_n is therefore obtained from that of P_1 by tilting the distribution by the density factor between the exponential distribution of $\lambda(R_1)^2/2$ and the $\mathrm{Gamma}(n, 1)$ distribution of $\lambda(R_n)^2/2$. But this factor is just $(\lambda(R_1)^2/2)^{n-1}/\Gamma(n)$ where $\lambda(R_1)^2/2 = (S_{1/2}(P_1))^2/4$, which is precisely the density factor between $\mathrm{PD}(\frac{1}{2}, (2n-1)/2)$ and $\mathrm{PD}(\frac{1}{2}, \frac{1}{2})$; see [28], Theorem 3.13.

As another check of the consistency of the two different descriptions of $\lambda(R_n)$, observe that the description in terms of independent exponential variables gives

$$\mathbb{E}\big(\lambda(R_n)^\rho\big) = 2^{\rho/2}\mathbb{E}\big(\Gamma_n^{\rho/2}\big) = 2^{p/2}\Gamma(n+\rho/2)/\Gamma(n), \tag{45}$$

whereas for $S_{\alpha,\theta}$ the α-diversity of a $\mathrm{PD}(\alpha, \theta)$ random discrete distribution it is known [28] that

$$\mathbb{E}\big(S_{\alpha,\theta}^\rho\big) = \frac{\Gamma(\theta/\alpha+\rho+1)\Gamma(\theta+1)}{\Gamma(\theta+p\alpha+1)\Gamma(\theta/\alpha+1)} \tag{46}$$

so in particular

$$\mathbb{E}\big(S_{1/2,(2n-1)/2}^\rho\big) = \frac{\Gamma(2n+\rho)\Gamma(n+(1/2))}{\Gamma(n+\rho/2+(1/2))\Gamma(2n)} = \frac{2^\rho\Gamma(n+\rho/2)}{\Gamma(n)}, \tag{47}$$

where the second equality uses the Gamma duplication formula $\Gamma(2z) = 2^{2z-1} \times \Gamma(z)\Gamma(z+\frac{1}{2})/\Gamma(\frac{1}{2})$. $\square$

APPENDIX: PROOF OF LEMMA 6

First note that the assumptions imply that τ is exponentially distributed and, furthermore, that for all nonnegative Borel functions g and all $t \geq 0$

$$\mathbb{E}\big(g(F_\tau)1_{\{\tau>t\}}\big) = \mathbb{P}(\tau > t)\mathbb{E}\big(g(F_{t+\tau-t})|\tau > t\big) = \mathbb{P}(\tau > t)\mathbb{E}\big(g(F_\tau)\big),$$

so that τ is independent of F_τ. Since τ is a stopping time, $(F_{\tau+s}, s \geq 0)$ is independent of $\mathcal{F}_\tau$ and has the same distribution as $(F_s, s \geq 0)$. Consider a sequence of

independent copies $(F_s^{(n)}, 0 \le s \le \tau_n)$, $n \ge 1$, of $(F_s, 0 \le s \le \tau)$ and splice them together as

$$(\widetilde{F}_{\tau_1+\cdots+\tau_{n-1}+s}, \widetilde{m}_{\tau_1+\cdots+\tau_{n-1}+s}) = \big(F_s^{(i)}, 1_{\{s=\tau_n\}}\big), \qquad 0 < s \le \tau_n, n \ge 1.$$

Then $((F_t, t \ge 0), \tau) \stackrel{d}{=} ((\widetilde{F}_t, t \ge 0), \widetilde{\tau})$, by construction, where $\widetilde{\tau} = \tau_1$. Also by construction, $\widetilde{F} = (\widetilde{F}_t, t \ge 0)$ and $\widetilde{F}^\bullet = ((\widetilde{F}_t \widetilde{m}_t + (1 - \widetilde{m}_t)), t \ge 0)$ are $\widetilde{\mathcal{F}}$-Poisson point processes in their joint natural filtration $\widetilde{\mathcal{F}}$. The intensity measure of $\widetilde{F}^\bullet$ is the distribution of F_τ times the rate of τ. Since $F_\tau \ne 1$ a.s., this intensity measure is absolutely continuous with respect to Λ as otherwise, we could find a Borel set A with $\mathbb{P}(F_\tau \in A) > 0$ and $\Lambda(A) = 0$, so the restrictions of $\widetilde{F}$ and $\widetilde{F}^\bullet$ would reveal points of $\widetilde{F}^\bullet$ that are not points of $\widetilde{F}$, a contradiction. We denote the Radon–Nikodym derivative by $K(u, dk)$ and set $p(u) = K(u, \{1\})$, so that $\widetilde{F}^\bullet$ is a Poisson point process with intensity measure $p(u)\Lambda(du)$. We note that $p(u) \le 1$ for Λ-a.e. $u \in (0, 1)$ as otherwise restrictions to the Borel set $A = \{u \in (0, 1) : p(u) > 1\}$ would reveal points of $\widetilde{F}^\bullet$ that are not points of $\widetilde{F}$, a contradiction.

The proof is not complete yet because we have not yet shown that $((\widetilde{F}_t, \widetilde{m}_t), t \ge 0)$ is an $\widetilde{\mathcal{F}}$-Poisson point process, or equivalently that the unmarked points $\widetilde{F}^\circ = ((\widetilde{F}_t(1 - \widetilde{m}_t) + \widetilde{m}_t), t \ge 0)$ form an $\widetilde{\mathcal{F}}$-Poisson point process. We can represent the point process $\widetilde{F}^\circ$ as a random measure

$$N_t^\circ(A) = \#\big\{s \le t : \widetilde{F}_s^\circ \in A\big\}, \qquad t \ge 0 \text{ and } A \text{ Borel subset of } (0, 1).$$

For A with $\Lambda(A) < \infty$ and $\int_A p(u)\Lambda(du) > 0$, we now show that $(N_t^\circ(A), t \ge 0)$ is an $\widetilde{\mathcal{F}}$-Poisson process. Consider $\overline{N}_t(A) = \#\{s \le t : \widetilde{F}_s \in A\}$, which we know is an $\widetilde{\mathcal{F}}$-Poisson process with rate $\Lambda(A)$. The time $\tau^A = \inf\{t \ge 0 : \widetilde{F}_t^\bullet \in A\}$ is an $\widetilde{\mathcal{F}}$-stopping time, since $\widetilde{F}^\bullet$ is an $\widetilde{\mathcal{F}}$-Poisson point process. Denote by $(T_n, n \ge 1)$ the times of the points of $\overline{N}(A)$, also $\widetilde{\mathcal{F}}$-stopping times. Let $A_n = \{\tau^A = T_n\}$ and set $q = \mathbb{P}(\tau^A = T_1)$. Then, by the strong Markov property of $\overline{N}(A)$ at T_n, we find that also for $n \ge 2$

$$\mathbb{P}(A_n) = \mathbb{P}\big(\tau^A \ne T_1\big)\mathbb{P}\big(A_n | \tau^A > T_1\big) = (1 - q)\mathbb{P}(A_{n-1}) = (1 - q)^{n-1} q. \tag{48}$$

Hence, we find that $\tau^A = T_G$ for a G geometric with parameter q, which is a stopping time in the discrete filtration $(\widetilde{\mathcal{F}}_{T_n}, n \ge 1)$. By Wald's equation,

$$\mathbb{E}\big(\tau^A\big) = \mathbb{E}(G)\mathbb{E}(T_1) \qquad \text{hence } \int_A p(u)\Lambda(du) = q\Lambda(A).$$

From (48) and the strong Markov property at τ^A, we deduce that $(N^\bullet(A), \overline{N}(A))$ are such that $N^\bullet(A)$ is a q-thinning of $\overline{N}(A)$, so $N^\circ(A) = \overline{N}(A) - N^\bullet(A)$ is also a Poisson process with rate

$$(1 - q)\Lambda(du) = \int_A \big(1 - p(u)\big)\Lambda(du).$$

By [25], Theorem 12.8, this suffices to identify N° as a Poisson random measure with intensity measure $(1 - p(u))\Lambda(du)$.

REFERENCES

[1] ABRAHAM, R. (1992). Un arbre aléatoire infini associé à l'excursion brownienne. In *Séminaire de Probabilités*, *XXVI*. *Lecture Notes in Math*. **1526** 374–397. Springer, Berlin. MR1232004

[2] ALDOUS, D. (1991). The continuum random tree. I. *Ann. Probab.* **19** 1–28. MR1085326

[3] ALDOUS, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[4] ALDOUS, D., MIERMONT, G. and PITMAN, J. (2004). Brownian bridge asymptotics for random p-mappings. *Electron. J. Probab.* **9** 37–56 (electronic). MR2041828

[5] BERTOIN, J. (2001). Homogeneous fragmentation processes. *Probab. Theory Related Fields* **121** 301–318. MR1867425

[6] BERTOIN, J. (2002). Self-similar fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 319–340. MR1899456

[7] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics* **102**. Cambridge Univ. Press, Cambridge. MR2253162

[8] BERTOIN, J. and PITMAN, J. (1994). Path transformations connecting Brownian bridge, excursion and meander. *Bull. Sci. Math.* **118** 147–166. MR1268525

[9] BERTOIN, J. and YOR, M. (2005). Exponential functionals of Lévy processes. *Probab. Surv.* **2** 191–212. MR2178044

[10] CHEN, B., FORD, D. and WINKEL, M. (2009). A new family of Markov branching trees: The alpha-gamma model. *Electron. J. Probab.* **14** 400–430. MR2480547

[11] CHEN, B. and WINKEL, M. (2013). Restricted exchangeable partitions and embedding of associated hierarchies in continuum random trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 839–872. MR3112436

[12] DOKSUM, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2** 183–201. MR0373081

[13] DOKSUM, K. A. and JAMES, L. F. (2004). On spatial neutral to the right processes and their posterior distributions. In *Mathematical Reliability*: *An Expository Perspective*. *Internat. Ser. Oper. Res. Management Sci.* **67** 87–103. Kluwer Academic, Boston, MA. MR2065000

[14] DONG, R., GOLDSCHMIDT, C. and MARTIN, J. B. (2006). Coagulation-fragmentation duality, Poisson–Dirichlet distributions and random recursive trees. *Ann. Appl. Probab.* **16** 1733–1750. MR2288702

[15] EVANS, S. N. (2008). *Probability and Real Trees*. *Lecture Notes in Math*. **1920**. Springer, Berlin. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587

[16] FORD, D. J. (2005). Probabilities on cladograms: Introduction to the alpha model. Ph.D. thesis, Stanford Univ. Available at arXiv:math.PR/0511246. MR2708802

[17] GNEDIN, A. and PITMAN, J. (2005). Regenerative composition structures. *Ann. Probab.* **33** 445–479. MR2122798

[18] GNEDIN, A., PITMAN, J. and YOR, M. (2006). Asymptotic laws for compositions derived from transformed subordinators. *Ann. Probab.* **34** 468–492. MR2223948

[19] HAAS, B. (2003). Loss of mass in deterministic and random fragmentations. *Stochastic Process. Appl.* **106** 245–277. MR1989629

[20] HAAS, B. and MIERMONT, G. (2004). The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab.* **9** 57–97 (electronic). MR2041829

[21] HAAS, B. and MIERMONT, G. (2012). Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *Ann. Probab.* **40** 2589–2666. MR3050512

[22] HAAS, B., MIERMONT, G., PITMAN, J. and WINKEL, M. (2008). Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Probab.* **36** 1790–1837. MR2440924

[23] HAAS, B., PITMAN, J. and WINKEL, M. (2009). Spinal partitions and invariance under rerooting of continuum random trees. *Ann. Probab.* **37** 1381–1411. MR2546748

[24] JAMES, L. F. (2006). Poisson calculus for spatial neutral to the right processes. *Ann. Statist.* **34** 416–440. MR2275248

[25] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York. MR1876169

[26] PERMAN, M., PITMAN, J. and YOR, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields* **92** 21–39. MR1156448

[27] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892

[28] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Springer, Berlin. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR2245368

[29] PITMAN, J., RIZZOLO, D. and WINKEL, M. (2014). Regenerative tree growth: Structural results and convergence. *Electron. J. Probab.* **19** no. 70, 1–29. Also available at arXiv:1207.3551. MR3256870

[30] PITMAN, J. and WINKEL, M. (2009). Regenerative tree growth: Binary self-similar continuum random trees and Poisson–Dirichlet compositions. *Ann. Probab.* **37** 1999–2041. MR2561439

[31] STEPHENSON, R. (2013). General fragmentation trees. *Electron. J. Probab.* **18** 1–45. MR3141802

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELEY, CALIFORNIA 94720
USA
E-MAIL: pitman@stat.berkeley.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF OXFORD
1 SOUTH PARKS ROAD
OXFORD OX1 3TG
UNITED KINGDOM
E-MAIL: winkel@stats.ox.ac.uk