CONSISTENT MARKOV BRANCHING TREES WITH DISCRETE EDGE LENGTHS

HARRY CRANE

Abstract. We study consistent collections of random fragmentation trees with random integer-valued edge lengths. We prove several equivalent necessary and sufficient conditions under which geometrically distributed edge lengths can be consistently assigned to a Markov branching tree. Among these conditions is a characterization by a unique probability measure, which plays a role similar to the dislocation measure for homogeneous fragmentation processes. We discuss this and other connections to previous work on Markov branching trees and homogeneous fragmentation processes.

1. Introduction

Random tree models arise in population genetics when inferring unknown phylogenetic relationships among extant species. Phylogenetic trees are often used to represent these relationships, with leaves labeled by extant species and branch points corresponding to speciation events. The root of the tree corresponds to the most recent common ancestor of the species under consideration. In [1], Aldous provides some modeling axioms for phylogenetic trees; among these axioms are exchangeability and consistency (under subsampling). Typically, the species labeling the leaves of a tree are represented by distinct elements of $[n] := \{1, \ldots, n\}$, and the exchangeability axiom reflects the assumption that our model should be invariant under arbitrary relabeling of the species. In a statistical setting, consistency reflects the assumption that the finite phylogenetic tree we observe is sampled from the (possibly infinite) phylogenetic tree for all species. An admissible statistical model, therefore, corresponds to a family of probability measures on the space of infinite phylogenetic trees, that is, trees with leaves labeled in the natural numbers $\mathbb{N}$.

Along with these axioms, Aldous introduced the beta-splitting model for random trees, which is a specific family of Markov branching model for phylogenetic trees. In general, a Markov branching tree is a random tree with the property that non-overlapping subtrees are conditionally independent. Within the phylogenetic framework, it is natural to consider random trees with edge lengths or weights (weighted Markov branching trees), where edge lengths are interpretted as time between speciation events. Previous authors [4, 6] have considered the task of assigning continuous (exponentially distributed) edge lengths to Markov branching trees in a consistent way as the size of the initial mass varies. In this paper, we undertake the related question of assigning discrete (geometrically distributed) edge lengths to Markov branching trees. In a phylogenetic context, trees with discrete edge lengths correspond to evolution occurring in discrete-time and, therefore, reflects the assumption that generations are nonoverlapping, an assumption shared by some classical

---

Date: July 26, 2013.

1991 *Mathematics Subject Classification.* Primary 60J80, secondary 60G09, 60C05.

Key words and phrases. Markov branching model, homogeneous fragmentation process, splitting rule, dislocation measure, sampling consistency, exchangeable random partition, weighted tree, random tree.

population genetics models; see [7] for an extensive treatment of probability models in population genetics.

Aside from applications to phylogenetics, random tree models are of their own mathematical interest. Particularly, part of the treatment in [4] relates weighted Markov branching trees to homogeneous fragmentation processes [2], a type of continuous-time Feller process on partitions of $\mathbb{N}$. In our main theorem, we give precise conditions under which discrete edges can be consistently attached to a Markov branching tree; and we characterize these trees by a unique probability measure on the space of ranked-mass partitions.

We point out at least one novelty that distinguishes this paper from previous work on weighted Markov branching trees. In contrast to the main theorems in [4], our theory does not appeal to Bertoin's theory of homogeneous fragmentations; rather, our proofs rely on a construction of discrete-weighted Markov branching trees as the projective limit of a sequence of finite weighted trees. At least some of the conclusions in [4] could be derived using our methods; however, as we explicitly consider trees with *integer-valued* edge lengths, we cannot appeal to the theory of homogeneous fragmentations, which evolve in continuous-time. Nevertheless, our characterization of discrete-weighted Markov branching models also ties into previous work on homogeneous fragmentations, which we discuss in Sections 3.1 and 3.4.

Probabilistically, the discrete-weighted Markov branching models are complementary to previous work on continuous-weighted Markov branching models. Taken together, these weighted tree models illustrate a fundamental aspect of the memoryless property: the exponential and geometric distributions are, respectively, the unique memoryless distributions on the positive real numbers and positive integers. An interesting twist, however, is that, unlike the continuous weight case, it is not always possible to attach geometric random edge weights consistently for all $n \in \mathbb{N}$. Our main theorem states precisely when this embedding is possible.

An overview of the paper is as follows: in Section 2, we state our main theorem as well as give some preliminary definitions and notation; in Section 3, we discuss the components of the main theorem in detail, putting our observations in the context of previous literature on the topic; in Section 4, we formally define some concepts introduced in previous sections, which we need to prove the main theorem; in Section 5, we prove the main theorem.

2. Preliminaries and statement of main theorem

Definition 2.1. *A* fragmentation *of a finite set $A \subset \mathbb{N}$ is a collection $\mathbf{t}_A$ of subsets of A such that*

(i) *$A \in \mathbf{t}_A$ and*
(ii) *if $\#A \geq 2$, then there exists a (root) partition $\pi_A := \{A_1, \ldots, A_k\}$ of A such that*

$$\mathbf{t}_A := \{A\} \cup \mathbf{t}_1 \cup \cdots \cup \mathbf{t}_k,$$

where $\mathbf{t}_i$ is a fragmentation of A_i for each $i = 1, \ldots, k$.

We call the elements of π_b, for $b \in \mathbf{t}_A$, the children *of b and write $\Pi_{\mathbf{t}_A} = \pi_A$ to denote the* root partition *of $\mathbf{t}_A$. We identify the set $A \in \mathbf{t}_A$ as the* root *of A and we write $\mathcal{T}_A$ to denote the collection of all fragmentations with root A. Alternatively, we may refer to a fragmentation as a* fragmentation tree *or, simply, a* tree.

Remark 2.2. *Definition 2.1 is initialized by taking $\mathbf{t}_{\{i\}} := \{\{i\}, \emptyset\}$ for each singleton $\{i\} \subset \mathbb{N}$. Inclusion of the empty set in the definition of $\mathbf{t}_A$ is done for notational convenience, which arises when taking restrictions of weighted trees in the sequel.*

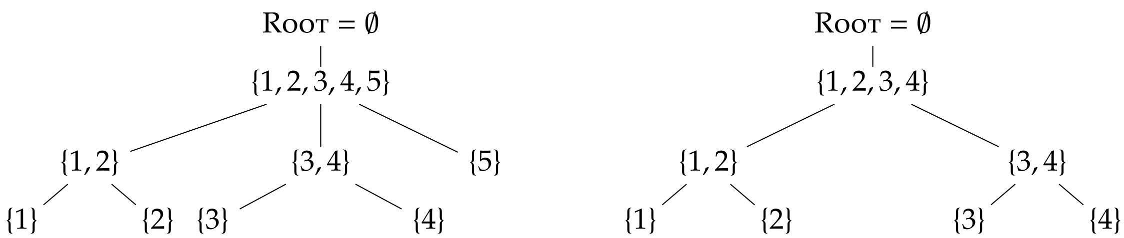
To any subset $A' \subset A$, there is a natural restriction of any $\mathbf{t} \in \mathcal{T}_A$ to $\mathcal{T}_{A'}$ by

$$\mathbf{R}_{A',A}\mathbf{t} = \mathbf{t}_{|A'} := \{b \cap A' : b \in \mathbf{t}\}, \quad \mathbf{t} \in \mathcal{T}_A, \tag{1}$$

called the *reduced subtree*. For $m \leq n$, we write $\mathbf{R}_{m,n} := \mathbf{R}_{[m],[n]}$. The projective limit of $\{\mathcal{T}_{[n]}\}_{n\in\mathbb{N}}$ under the restriction maps $\{\mathbf{R}_{m,n}\}_{m\leq n}$ is denoted $\mathcal{T}_{\mathbb{N}}$ and corresponds to the space of fragmentation trees with root $\mathbb{N}$. For $n \in \mathbb{N}$, we write $\mathbf{R}_n : \mathcal{T}_{\mathbb{N}} \to \mathcal{T}_{[n]}$ to denote the restriction to $\mathcal{T}_{[n]}$ of an infinite tree, as defined in (1) with $A' = [n]$ and $A = \mathbb{N}$. We equip $\mathcal{T}_{\mathbb{N}}$ with the σ-field $\sigma\langle \mathbf{R}_n \rangle_{n\in\mathbb{N}}$ so that these maps are measurable.

We illustrate the action of the restriction map $\mathbf{R}_{4,5}$ in (2) below. Note that, in the left panel, $\mathbf{t}$ is a tree with root $\{1,2,3,4,5\}$ and root partition $\{\{1,2\},\{3,4\},\{5\}\}$. Also, relating to Definition 2.1, $\mathbf{t}$ corresponds to the collection of subsets $\{\emptyset, \{1,2,3,4,5\}, \{1,2\}, \{3,4\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$ that label its vertices.

(2)



$\mathbf{t} \quad \mapsto \quad \mathbf{R}_{4,5}(\mathbf{t}) = \mathbf{t}_{|[4]}$

We are specifically interested in probability models for fragmentation trees with integer-valued edge lengths. From any $\mathbf{t} \in \mathcal{T}_A$, we obtain a *discrete-weighted tree* $\mathbf{t}^{\bullet}$ by assigning a positive integer weight $w_b > 0$ to every $b \in \mathbf{t}$. The pair $\mathbf{t}^{\bullet} := (\mathbf{t}, w)$ then determines a tree with edge lengths. We write $\mathcal{T}_A^{\bullet}$ to denote the space of discrete-weighted trees with root A, for which there is also a natural restriction map $\mathbf{R}^{\bullet}_{A',A}$, for every $A' \subseteq A$, defined by removing elements and elongating edges as needed. These restrictions make the collection $\{\mathcal{T}^{\bullet}_{[n]}\}_{n\in\mathbb{N}}$ of finite discrete-weighted trees projective with limit $\mathcal{T}^{\bullet}_{\mathbb{N}}$. Weighted fragmentations are formally introduced in Section 4.2; a pictorial representation of a discrete-weighted tree is given in (15).

The probability models we consider are extensions of the Markov branching models on $\mathcal{T}_{\mathbb{N}}$. By the projective structure of $\mathcal{T}_{\mathbb{N}}$, any probability measure Q on $\mathcal{T}_{\mathbb{N}}$ is determined by its finite-dimensional restrictions $Q^{[n]} := Q\mathbf{R}_n^{-1}$ to $\mathcal{T}_{[n]}$, for every $n \in \mathbb{N}$. Specifically, we consider the task of assigning random geometrically distributed edge lengths to exchangeable Markov branching trees.

In general, the collection $Q := (Q^{[n]})_{n\in\mathbb{N}}$ determines an *exchangeable Markov branching model* if, for every $n \in \mathbb{N}$, $\mathbf{T} \sim Q^{[n]}$ is

- *exchangeable*: the law of $\mathbf{T}$ is invariant under the obvious action of relabeling its leaves by an arbitrary permutation $\sigma : [n] \to [n]$;
- *consistent*: $\mathbf{R}_{m,n}\,\mathbf{T} \sim Q^{[m]}$ for every $m \leq n$; and,
- *Markovian*: given any collection $\{A_1, \ldots, A_k\}$ of non-overlapping subsets in $\mathbf{T}$, the collection $\{\mathbf{T}_{|A_1}, \ldots, \mathbf{T}_{|A_k}\}$ of reduced subtrees is conditionally independent and distributed according to $Q^{[n_1]}, \ldots, Q^{[n_k]}$, respectively, where $n_j := \#A_j$, $j = 1, \ldots, k$.

Any exchangeable Markov branching model Q is determined by a family of exchangeable *splitting rules* $p := (p_n)_{n\geq 2}$, where each p_n is a probability measure on the space $\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}$ of partitions of the set $[n]$ with the trivial partition $\mathbf{1}_{[n]} := \{[n]\}$ removed. For $m \leq n$, there is an obvious deletion operation $\mathbf{D}_{m,n} : \mathcal{P}_{[n]} \to \mathcal{P}_{[m]}$ defined by removing elements in $[n] \setminus [m]$,

$$\mathbf{D}_{m,n}(\pi) := \{b \cap [m] : b \in \pi\} \setminus \{\emptyset\}, \quad \pi \in \mathcal{P}_{[n]}. \tag{3}$$

It has been shown, e.g. in [1, 4, 6], that $p := (p_n)_{n\geq 2}$ determines an exchangeable Markov branching model if and only if p_n is exchangeable and

$$p_n(\pi) = p_{n+1}(\mathbf{D}_{n,n+1}^{-1}(\pi)) + p_{n+1}(\mathbf{e}_{n+1}^{(n+1)})p_n(\pi), \quad \pi \in \mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}, \quad \text{for every } n \geq 2, \tag{4}$$

where $\mathbf{e}_{n+1}^{(n+1)} := \{[n], \{n+1\}\}$. We write $Q_p := (Q_p^{[n]})_{n\in\mathbb{N}}$ to denote the Markov branching model determined by the consistent splitting rule p. Note that (4) is merely the requirement that the marginal distribution of the root partition of $\mathbf{T} \sim Q_p^{[n+1]}$, after removal of element $n+1$, is the same as the distribution of the root partition under $Q_p^{[n]}$, for every $n \geq 2$.

Given a Markov branching tree $\mathbf{T}_n \sim Q_p^{[n]}$, we randomly assign edge lengths to $\mathbf{T}_n$ as follows. First, we specify $\tau := (\tau_n)_{n\geq 0}$, with $\tau_0 = \tau_1 = 0$ and $\tau_n \in (0, 1]$ for all $n \geq 2$. Given $\mathbf{T}_n = \mathbf{t}$, we take $W_n := \{W_n(b)\}_{b\in\mathbf{t}}$ independent random variables, where $W_n(b) \sim \text{Geo}(\tau_{\#b})$ has the geometric distribution with parameter $\tau_{\#b}$. (We define $\text{Geo}(0)$ to be the point mass at ∞.) We write $Q_{p,\tau}^{[n]}$ to denote the distribution of $\mathbf{T}_n^\bullet := (\mathbf{T}_n, W_n)$ obtained in this way. Our main theorem considers the question of when the collection $(Q_{p,\tau}^{[n]})_{n\in\mathbb{N}}$ of finite-dimensional distributions determines a unique probability measure $Q_{p,\tau}^\bullet$ on the limit space $\mathcal{T}_{\mathbb{N}}^\bullet$.

We now state our main theorem.

Theorem 2.3. *Let $p := (p_n)_{n\geq 2}$ be a family of exchangeable splitting rules satisfying* (4). *The following are equivalent.*

(i) *There exists a collection $\tau := (\tau_n)_{n\geq 0}$ of geometric success probabilities such that $(Q_{p,\tau}^{[n]})_{n\in\mathbb{N}}$ are the finite-dimensional restrictions of a unique probability measure $Q_{p,\tau}^\bullet$ on $\mathcal{T}_{\mathbb{N}}^\bullet$.*

(ii) *The family $\tau := (\tau_n)_{n\geq 0}$ satisfies $\tau_0 = \tau_1 = 0$ and*

$$\tau_n = \tau_{n+1}(1 - p_{n+1}(\mathbf{e}_{n+1}^{(n+1)})) \quad \text{for all } n \geq 2. \tag{5}$$

(iii) *There is a unique probability measure ν^* on $\Delta^\downarrow := \{(s_1, s_2, \ldots) : s_1 \geq s_2 \geq \cdots \geq 0, \sum_i s_i \leq 1\}$ satisfying*

$$\nu^*(\{(1, 0, \ldots)\}) < 1 \tag{6}$$

so that (p, τ) is given by $p := (p_n^{\nu^*})_{n\geq 2}$ and $\tau := (\tau_n^{\nu^*})_{n\geq 0}$ in (8) *and* (9)*, respectively.*

(iv) *The pair (p, τ) is determined by (ν, τ_∞), where $\tau_\infty := \lim_{n\to\infty} \tau_n > 0$ and ν is a finite measure on $\Delta^\downarrow$ that satisfies $\nu(\{(1, 0, \ldots)\}) = 0$.*

(v) *Q_p-almost every $\mathbf{t} \in \mathcal{T}_{\mathbb{N}}$ possesses a root partition.*

(vi) *$\lambda_\infty := \lim_{n\to\infty} \lambda_n < \infty$, where $\lambda := (\lambda_n)_{n\geq 2}$ is defined recursively by $\lambda_2 = 1$ and*

$$\lambda_{n+1} = \lambda_n / (1 - p_{n+1}(\mathbf{e}_{n+1}^{(n+1)})), \quad n \geq 2. \tag{7}$$

2.1. **The paintbox measure.** The paintbox measure plays a key role in our discussion in the next section as well as in our proof of uniqueness of ν^* in Theorem 2.3(iii). For $s \in \Delta^\downarrow$, we write $s_0 := 1 - \sum_{i=1}^\infty s_i$ to denote the amount of *dust* in s and we define the *paintbox measure* ϱ_s directed by s as the distribution of a random partition Π generated as follows. First, we take $X_1, X_2, \ldots$ independent random variables with distribution

$$\mathbb{P}_s(X_i = j) := \begin{cases} s_j, & j \geq 1 \\ s_0, & j = -i. \end{cases}$$

Given $(X_1, X_2, \ldots)$, we define Π by

$$i \text{ and } j \text{ are in the same block of } \Pi \quad \Longleftrightarrow \quad X_i = X_j.$$

We write $\Pi \sim \varrho_s$ to denote that Π is distributed as a paintbox directed by s. Given a measure ν on $\Delta^{\downarrow}$, the paintbox measure directed by ν is the mixture of paintboxes:

$$\varrho_\nu(d\pi) := \int_{\Delta^{\downarrow}} \varrho_s(d\pi)\nu(ds), \quad \pi \in \mathcal{P}_{\mathbb{N}}.$$

According to Kingman's correspondence [5], to any exchangeable random partition Π of $\mathbb{N}$ there corresponds a unique probability measure ν^* on $\Delta^{\downarrow}$ such that $\Pi \sim \varrho_{\nu^*}$.

3. Discussion of Theorem 2.3

We now discuss the components of Theorem 2.3 in some detail, paying attention to the interplay among (i)-(vi) as well as connections to previous literature. Roughly speaking, the six parts of the theorem can be broken up into three motifs: (i)-(ii) is a condition in the vein of Markov branching trees with exponential edge lengths; (iii)-(iv) gives a structure result that is reminiscent of the characterization of homogeneous fragmentations; (v)-(vi) describes the existence of $Q^{\bullet}_{p,\tau}$ without explicit reference to τ; in particular, both (v) and (vi) depend only on p.

3.1. The characteristic measure ν^*.

Theorem 2.3(iii) establishes a bijection between probability laws $Q^{\bullet}_{p,\tau}$ of infinite Markov branching trees with integer edge lengths and probability measures ν^* satisfying (6). Given such a ν^*, we define (p, τ) by $p := (p_n^{\nu^*})_{n\geq 2}$ and $\tau := (\tau_n^{\nu^*})_{n\geq 0}$, where

$$p_n^{\nu^*}(\pi) := \frac{\varrho_{\nu^*}^{(n)}(\pi)}{1 - \varrho_{\nu^*}^{(n)}(\mathbf{1}_{[n]})}, \quad \pi \in \mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}, \quad n \geq 2, \tag{8}$$

$\tau_0^{\nu^*} = \tau_1^{\nu^*} = 0$ and

$$\tau_n^{\nu^*} := 1 - \varrho_{\nu^*}^{(n)}(\mathbf{1}_{[n]}), \quad n \geq 2. \tag{9}$$

Note that we have written $\varrho_{\nu^*}^{(n)}$ to denote the image of ϱ_{ν^*} by the obvious restriction map $\mathcal{P}_{\mathbb{N}} \to \mathcal{P}_{[n]}$. Condition (6) ensures that (8) is a well-defined probability distribution on $\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}$ and the success probabilities τ_n are strictly positive for every $n \geq 2$.

A further consequence of the characterization by ν^* ties into part (iv) of the theorem. In particular, from (9), the sequence τ is monotonically nondecreasing and bounded above by 1; hence, the limit $\tau_\infty := \lim_{n\to\infty} \tau_n^{\nu^*}$ exists and equals

$$\tau_\infty = 1 - \varrho_{\nu^*}(\mathbf{1}_{[\infty]}) = 1 - \nu^*(\{(1, 0, \ldots)\}) > 0.$$

From ν^*, we can define a finite measure ν_K, for any $K \in (0, \infty)$, by

$$\nu_K(ds) := K\nu^*(ds)(1 - \delta_{(1,0,\ldots)}(s)), \quad s \in \Delta^{\downarrow}, \tag{10}$$

where $\delta_\bullet(\cdot)$ is the point mass at $\bullet$. Note that ν_K is finite and satisfies $\nu_K(\{(1, 0, \ldots)\}) = 0$. Since trivial partitions are assigned zero probability by any splitting rule, the measures ν^* and ν_K determine the same splitting rule through the generalization to (8):

$$p_n^{\nu_K}(\pi) := \frac{\varrho_{\nu_K}^{(n)}(\pi)}{\nu_K(\Delta^{\downarrow}) - \varrho_{\nu_K}^{(n)}(\mathbf{1}_{[n]})}, \quad \pi \in \mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}, \quad n \geq 2.$$

Indeed, from (10), we have, for $\pi \in \mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}$,

$$\begin{aligned} p_n^{\nu_K}(\pi) &= \frac{\varrho_{\nu_K}^{(n)}(\pi)}{\nu_K(\Delta^{\downarrow}) - \varrho_{\nu_K}^{(n)}(\mathbf{1}_{[n]})} \\ &= \frac{K\varrho_{\nu^*}^{(n)}(\pi)}{K(1-\nu^*(\{(1,0,\ldots)\})) - \varrho_{\nu_K}^{(n)}(\mathbf{1}_{[n]})} \\ &= \frac{\varrho_{\nu^*}^{(n)}(\pi)}{1-\varrho_{\nu^*}^{(n)}(\mathbf{1}_{[n]})}, \end{aligned}$$

which coincides with (8).

Conversely, given $\tau_\infty \in (0,1]$ and a finite measure ν satisfying $\nu(\{(1,0,\ldots)\}) = 0$, we obtain a measure μ^* satisfying (6) by

$$\mu^*(ds) := \frac{\nu(ds)}{\nu(\Delta^{\downarrow})}\tau_\infty + (1-\tau_\infty)\delta_{(1,0,\ldots)}(s), \quad s \in \Delta^{\downarrow}. \tag{11}$$

We see that for ν_K in (10) and $\tau_\infty := 1 - \nu^*(\{(1,0,\ldots)\})$, μ^* in (11) coincides with

$$\frac{\nu_K(ds)}{\nu_K(\Delta^{\downarrow})}(1-\nu^*(\{(1,0,\ldots)\})) + \nu^*(\{(1,0,\ldots)\})\delta_{(1,0,\ldots)}(s) = \nu^*(ds).$$

3.2. **The role of τ_∞.** The quantity $\tau_\infty := \lim_{n\to\infty} \tau_n$ plays an important role in the description of the limiting tree $\mathbf{T}^\bullet \sim Q_{p,\tau}^\bullet$ in that it parameterizes its edge lengths. That is, the limiting object $\mathbf{T}^\bullet$ is an infinite Markov branching tree with independent geometrically distributed edge lengths, all with success probability τ_∞. Moreover, the special case $\tau_\infty = 1$ corresponds to geometric edge lengths all with success probability 1. Hence, almost surely, the edge lengths of the limiting tree $\mathbf{T}^\bullet$ are all identically 1. In this case, the randomness of the edge lengths disappears in the limiting object. Viewed another way, from (11), we notice that $1 - \tau_\infty = \nu^*(\{(1,0,\ldots)\})$ corresponds to the probability that a random partition of $\mathbb{N}$ is trivial. Since only non-trivial partitions correspond to dislocations in a fragmentation tree, $\tau_\infty = 1 - \nu^*(\{(1,0,\ldots)\})$ naturally corresponds to a success probability in our geometric weighting scheme.

3.3. **The success probabilities τ.** Previous authors, e.g. [4, 6], have considered the task of consistently appending random exponentially distributed edges lengths to Markov branching trees. Given a splitting rule $p = (p_n)_{n\geq 2}$ and a collection $\lambda := (\lambda_n)_{n\geq 0}$ with $\lambda_0 = \lambda_1 = 0$ and $\lambda_n > 0$ for all $n \geq 2$, we assign independent random lengths $W_n(b) \sim \textsc{Exp}(\lambda_{\#b})$ to each $b \in \mathbf{T}_n$, where $\textsc{Exp}(\lambda)$ denotes the exponential distribution with rate parameter λ. (The $\textsc{Exp}(0)$ distribution corresponds to the point mass at ∞.) We write $Q_{p,\lambda}^{[n]}$ to denote the law of a $Q_p^{[n]}$-distributed Markov branching tree with exponential edge lengths parameterized by λ. By Proposition 3 of [4] the collection $(Q_{p,\lambda}^{[n]})_{n\in\mathbb{N}}$ is consistent if and only if p satisfies (4) and λ satisfies

$$\lambda_n = \lambda_{n+1}(1 - p_{n+1}(\mathbf{e}_{n+1}^{(n+1)})) \quad \text{for every } n \geq 2. \tag{12}$$

We write $Q_{p,\lambda}^\circ$ to denote the distribution $(Q_{p,\lambda}^{[n]})_{n\in\mathbb{N}}$ induces on the space of infinite trees with continuous edge lengths.

Note that (12) is identical to condition (5) of Theorem 2.3(ii); however, in the discrete case we encounter the additional constraint $0 \leq \tau_n \leq 1$ for all $n \geq 0$. Moreover, while continuous embedding is always possible for an infinitely exchangeable family of splitting rules, discrete embedding is not. Conditions (5) and (12) seem intimately tied to the memoryless property of the exponential and geometric distributions. Both (5) and (12) can be proven using the same strategy as in Theorem 5.1, with the modification that to prove (12) we use characteristic functions rather than probability generating functions.

3.4. **Relation to homogeneous fragmentations.** The definition of ν in (10) connects the characteristic measure ν^* to a collection of dislocation measures ν of homogeneous fragmentation processes. From Theorem 1 of [4], any exchangeable splitting rule $p = (p_n)_{n\geq 2}$ satisfying (4) is associated to a pair (c, ν) (see equations (2) and (3) of [4]), where $c \geq 0$ is the *erosion coefficient* and ν is the *dislocation measure* of a homogeneous fragmentation process. To ensure that each finite restriction of $\mathbf{T}^\circ \sim Q^\circ_{p,\lambda}$ determines a fragmentation of a finite set with strictly positive edge lengths, the dislocation measure ν is subject to the constraint

$$\nu(\{(1,0,\ldots)\}) = 0 \quad \text{and} \quad \int_{\Delta^\downarrow}(1-s_1)\nu(ds) < \infty; \tag{13}$$

see also, Bertoin [3] (Theorem 3.1). The measure ν constructed in (10) trivially satisfies (13) and, therefore, is the dislocation measure of some homogeneous fragmentation. Moreover, by Theorem 1 of [4], the pair (c, ν) is unique only up to constant multiples. So (c, ν) and $(Kc, K\nu)$, $K > 0$, determine the same splitting rule p. Similarly, the measure ν in Theorem 2.3(iv) is only unique up to constant multiples, as discussed in Section 3.1.

3.5. **Root partitions.** The erosion coefficient $c \geq 0$ also relates to (v) and (vi) of our theorem. In particular, the erosion coefficient is the rate at which "erosion" of a single element occurs, that is, the event that the initial split of the mass $[n]$ is into $\mathbf{e}^{(n)}_n := \{[n-1]\}, \{n\}\}$. Each split into $\mathbf{e}^{(n)}_n$ occurs at rate $c \geq 0$. Hence, assuming the dislocation measure ν is finite, the total rate at which a (c, ν)-fragmentation process experiences dislocation is $\lambda_n = cn + \varrho^{(n)}_\nu\left(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}\right)$. As a result, we see that $\lambda_n \to \infty$ whenever $c > 0$ and $\lambda_n \to \nu(\Delta^\downarrow) < \infty$ when $c = 0$. Therefore, (iv) and (vi) together imply that discrete-weighted fragmentations correspond to homogeneous fragmentations with zero erosion coefficient and finite dislocation measure.

Furthermore, Theorem 2.3(v) asserts that the existence of a collection τ for which $Q^\bullet_{p,\tau}$ exists depends on whether $\mathbf{T} \sim Q_p$ possesses a well-defined root partition. Intuitively, there will be such a root partition only if λ_∞ is finite because if $\lambda_\infty = \infty$ then the root edges of the finite trees $(\mathbf{T}^\circ_n)_{n\in\mathbb{N}}$ must be getting shorter as n grows. Thus, Theorem 2.3(v) divides Markov branching trees into two classes, those with root partition and those without. By (v), Markov branching trees with a root partition can be assigned geometrically distributed edge lengths, while those without a root partition cannot. To be explicit, given $\lambda_\infty < \infty$, we can choose any $\lambda^* \in [\lambda_\infty, \infty)$ and put $\tau_n = \lambda_n/\lambda^*$ for each $n \geq 2$. By (7), $(\tau_n)_{n\geq 2}$ chosen this way satisfies (5). Moreover, relating to Section 3.2, we have $\tau_\infty = \lambda_\infty/\lambda^* \in (0, 1]$.

3.6. **Beta-splitting model.** We conclude this section with an illustration of Theorem 2.3 in the special case of the beta-splitting model. For $-2 < \beta < \infty$, we define the splitting rule

$$p^\beta_n(\pi) := 2\kappa_n^{-1}\frac{\beta^{\uparrow\#\pi_1}\beta^{\uparrow\#\pi_2}}{(2\beta)^{\uparrow n}}, \tag{14}$$

where $\pi = \{\pi_1, \pi_2\}$ is a partition of $[n]$ with exactly two blocks, $\kappa_n := 1 - 2\beta^{\uparrow n}/(2\beta)^{\uparrow n}$ and $\beta^{\uparrow n} := \beta(\beta+1)\cdots(\beta+n-1)$. (The limiting cases $\beta \to -2$ and $\beta \to \infty$ are also defined: $\beta = -2$

corresponds to the exchangeable distribution on "combs" and $\beta = \infty$ corresponds to the "symmetric binary trie." For simplicity, we ignore these cases.)

These splitting rules are based on the dislocation measure $\nu_\beta(dx) := 2x^\beta(1-x)^\beta \mathbf{1}_{[1/2,1]}(x)dx$ for $-2 < \beta < \infty$, which is a measure on the subspace of binary mass partitions. Note that ν satisfies (13) and is, therefore, a dislocation measure for a sub-family homogeneous fragmentation processes. In particular, for $\beta > -1$, ν is a finite measure and, for $-2 < \beta \leq -1$, ν is infinite. Therefore, even when $c = 0$, $\lambda_\infty \to \nu_\beta(\Delta^\downarrow) < \infty$ only for $\beta > -1$, and so these are the only β for which $(p_n^\beta)_{n\geq 2}$ in (14) determines a distribution $Q^\bullet_{p,\tau}$ on discrete-weighted trees. In fact, in the case $\beta > -1$, the splitting rule $(p_n^\beta)_{n\geq 2}$ is determined by the Beta distribution with parameter (β, β). In particular, $\nu_\beta(dx) := 2x^\beta(1-x)^\beta \mathbf{1}_{[1/2,1]}(x)dx$ is the kernel of the probability measure ν^*_β governing $\max(X, 1-X)$ for $X \sim \text{Beta}(\beta, \beta)$. Note that $\nu^*_\beta(\{(1, 0, \ldots)\}) = 0$ in this case and so we are in the situation $\tau_\infty = 1$. Alternatively, given ν^*_β, $\beta > -1$, we can define ν^* with arbitrary $\tau_\infty \in (0, 1]$ through (11). Through (8) and (9), the resulting probability measure ν^* determines a unique pair (p, τ) that parameterizes $Q^\bullet_{p,\tau}$.

4. Some formalities

We now introduce some concepts from previous sections more formally in preparation for the proof of Theorem 2.3 in the next section.

4.1. **Root partitions.** With $A \subset_f \mathbb{N}$ denoting that $A \subset \mathbb{N}$ is finite, a *partition* of A is a collection $\{A_1, \ldots, A_k\}$ of non-empty, disjoint subsets for which $\bigcup_{i=1}^k A_i = A$. We write $\mathcal{P}_A$ to denote the collection of all partitions of A. The collection $\{\mathcal{P}_{[n]}\}_{n\in\mathbb{N}}$ of spaces of finite set partitions is projective under the deletion maps (3). We write $\mathcal{P}_\mathbb{N}$ to denote the projective limit of partitions of $\mathbb{N}$, which we furnish with the discrete σ-algebra $\sigma\left\langle \bigcup_{n\in\mathbb{N}} \mathcal{P}_{[n]} \right\rangle$. For each $n \in \mathbb{N}$, we write $\mathbf{D}_n := \mathbf{D}_{n,\infty}$ to denote the deletion operation $\mathcal{P}_\mathbb{N} \to \mathcal{P}_{[n]}$, where $[\infty] := \mathbb{N}$ in (3). Partitions factor into the study of Markov branching trees through their splitting rules, which are distributions on $\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}$ and determine the law of the branching below a child of size n in a random fragmentation.

Also, in Theorem 2.3(v), partitions of $\mathbb{N}$ arise in the notion of a limiting root partition. For any $A \subset_f \mathbb{N}$, $\#A \geq 2$, every $\mathbf{t} \in \mathcal{T}_A$ has a well-defined *root partition* denoted by $\Pi_\mathbf{t}$ and defined by the partition π_A in Definition 2.1(ii). In general, for any $A \subseteq \mathbb{N}$, we say that $\mathbf{t} \in \mathcal{T}_A$ possesses a root partition if there exists $N \in \mathbb{N}$ such that the sequence $(\Pi_{\mathbf{t}_{|[m]}})_{m\geq N}$ has a projective limit in $\mathcal{P}_\mathbb{N}$, that is, if for all $n \geq m \geq N$, $\Pi_{\mathbf{t}_{|[m]}} = \mathbf{D}_{m,n}\,\Pi_{\mathbf{t}_{|[n]}}$. We denote this root partition by $\Pi_\mathbf{t} := \lim_{n\to\infty} \Pi_{\mathbf{t}_{|[n]}}$.

Example 4.1. *An infinite tree need not possess a well-defined root partition. For example, the infinite* comb $\mathbf{c}$ *is defined by the collection* $\mathbf{c} := (\mathbf{c}_n)_{n\geq 2}$ *where* $\Pi_{\mathbf{c}_n} = \mathbf{e}_n^{(n)}$ *for every* $n \geq 2$. *In this case, the sequence of finite root partitions is* $(\mathbf{e}_n^{(n)})_{n\geq 2}$, *for which* $\mathbf{D}_{m,n}\,\mathbf{e}_n^{(n)} = \mathbf{1}_{[m]} \neq \mathbf{e}_m^{(m)}$ *for any* $m < n$; *hence,* $\lim_{n\to\infty} \Pi_{\mathbf{c}_n}$ *does not exist.*

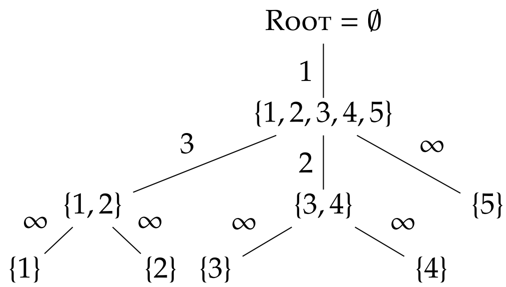
4.2. **Weighted fragmentation trees.** We define a *weighted fragmentation* of $A \subset_f \mathbb{N}$ as a pair $\mathbf{t}^\circ := (\mathbf{t}, w)$ such that $\mathbf{t} \in \mathcal{T}_A$ and $w = \{w_b\}_{b\subseteq A}$, with $w_b \in [0, \infty]$ for all $b \subseteq A$ and

(i)$^\mathbf{w}$ $w_b = \infty$ if and only if b is a singleton or the empty set;

(ii)$^\mathbf{w}$ $w_b = 0$ if and only if $b \notin \mathbf{t}$.

Remark 4.2. *Item (i)*$^\mathbf{w}$ *is not necessary for the above definition to make sense; however, we are interested in constructing consistent collections of weighted fragmentations of* $\mathbb{N}$, *and (i)*$^\mathbf{w}$ *is the convention that works best in this context.*

Pictorially, we interpret w_b as the length of the edge *above* $b \in \mathbf{t}$, although we suppress the edge of infinite length associated to $\emptyset$. For example, for the tree $\mathbf{t}$ in (2), if we specify $w_{\{1,2,3,4,5\}} = 1$, $w_{\{1,2\}} = 3$ and $w_{\{3,4\}} = 2$, then we obtain

(15)


,

where edge lengths are not drawn to scale. We write $\mathcal{T}_A^\circ$ to denote the collection of weighted fragmentations of $A \subset_f \mathbb{N}$.

For non-empty subsets $A' \subseteq A \subset_f \mathbb{N}$, we define $\mathbf{R}^\circ_{A',A} : \mathcal{T}^\circ_A \to \mathcal{T}^\circ_{A'}$ by $\mathbf{t}^\circ \mapsto \mathbf{t}^\circ_{|A'} := (\mathbf{R}_{A',A}(\mathbf{t}), w')$, with $\mathbf{R}_{A',A}$ defined in (1) and $w' := \{w'_b\}_{b \subseteq A'}$, where

$$w'_b := \sum_{\{b' \subseteq A : b' \cap A' = b\}} w_{b'}, \quad b \subseteq A', \tag{16}$$

the sum of all weights associated to b by restriction of A to A'. In particular, for $m \leq n < \infty$, we write $\mathbf{R}^\circ_{m,n} := \mathbf{R}^\circ_{[m],[n]}$ and denote the projective limit of $\{\mathcal{T}^\circ_{[n]}\}_{n \in \mathbb{N}}$ under these restrictions by $\mathcal{T}^\circ_{\mathbb{N}}$, the space of weighted fragmentations of $\mathbb{N}$. Any $\mathbf{t}^\circ \in \mathcal{T}^\circ_{\mathbb{N}}$ is determined by a sequence $(\mathbf{t}^\circ_n)_{n \in \mathbb{N}}$ satisfying $\mathbf{R}^\circ_{m,n}\, \mathbf{t}^\circ_n = \mathbf{t}^\circ_m$ for all $m \leq n$, for every $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, we define $\mathbf{R}^\circ_n : \mathcal{T}^\circ_{\mathbb{N}} \to \mathcal{T}^\circ_{[n]}$ by the projection of $\mathcal{T}^\circ_{\mathbb{N}}$ into $\mathcal{T}_{[n]}$, $\mathbf{t}^\circ \mapsto \mathbf{t}^\circ_n$.

4.2.1. *Integer-valued edge weights.* For each $n \in \mathbb{N}$, we write $\mathcal{T}^\bullet_{[n]} \subset \mathcal{T}^\circ_{[n]}$ to denote the subspace of all $\mathbf{t}^\circ \in \mathcal{T}^\circ_{[n]}$ such that $w_n(b) \in \{0, 1, \ldots, \infty\}$ for all $b \subseteq [n]$. For $m \leq n$, we let $\mathbf{R}^\bullet_{m,n}$ be the restriction of $\mathbf{R}^\circ_{m,n}$ to $\mathcal{T}^\bullet_{[n]}$ and we define $\mathcal{T}^\bullet_{\mathbb{N}}$ as the projective limit of $\{\mathcal{T}^\bullet_{[n]}\}_{n \in \mathbb{N}}$ under these restriction maps. The space $\mathcal{T}^\bullet_{\mathbb{N}}$ comes equipped with $\mathbf{R}^\bullet_n : \mathcal{T}^\bullet_{\mathbb{N}} \to \mathcal{T}^\bullet_{[n]}$, the restriction of $\mathbf{R}^\circ_n$ to $\mathcal{T}^\bullet_{\mathbb{N}}$ for each $n \in \mathbb{N}$. Writing $\mathcal{D}_n := \bigotimes_{b \subseteq [n]} 2^{\{0,1,\ldots,\infty\}}$ to denote the product of discrete σ-fields on subsets of $\{0, 1, \ldots, \infty\}$, we equip $\mathcal{T}^\bullet_{[n]}$ with the σ-field $\mathcal{T}_{[n]} \otimes \mathcal{D}_n$ and $\mathcal{T}^\bullet_{\mathbb{N}}$ with the σ-field $\sigma\langle \mathbf{R}^\bullet_n \rangle_{n \in \mathbb{N}}$ so that the restriction maps are measurable.

4.3. **Random weighted fragmentations of $\mathbb{N}$.** Let $p := (p_n)_{n \geq 2}$ be a collection of splitting rules satisfying (4) and let $\tau := (\tau_n)_{n \geq 0}$ satisfy $\tau_0 = \tau_1 = 0$ and $\tau_n \in (0, 1]$ for all $n \geq 2$. Formally, we define $Q^{[n]}_{p,\tau}$ as the law of $\mathbf{T}^\bullet_n := (\mathbf{T}_n, W_n)$ where $\mathbf{T}_n \sim Q^{[n]}_p$ is a Markov branching tree with splitting rules p and W_n are discrete edge weights defined as follows. We generate $\Upsilon_n := \{\Upsilon_n(b)\}_{b \subseteq [n]}$, independent geometric random variables with $\Upsilon_n(b) \sim \textsc{Geo}(\tau_{\#b})$ and, given $\mathbf{T}_n = \mathbf{t}$ and Υ_n, we define a discrete weighted tree $\mathbf{T}^\bullet_n := (\mathbf{T}_n, W_n)$ in $\mathcal{T}^\bullet_{[n]}$, where $W_n := \{W_n(b)\}_{b \subseteq [n]}$ is defined from Υ_n by

$$W_n(b) := \begin{cases} \Upsilon_n(b), & b \in \mathbf{T}_n \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

We can express $Q^{[n]}_{p,\tau}$ explicitly by

$$Q^{[n]}_{p,\tau}(\mathbf{t}^\bullet) = \prod_{b\in\mathbf{t}:\#b\geq 2} p_b(\Pi_{\mathbf{t}_{|b}})\tau_{\#b}(1-\tau_{\#b})^{w(b)-1}, \quad \mathbf{t}^\bullet := (\mathbf{t}, w) \in \mathcal{T}^\bullet_{[n]}, \tag{18}$$

where $p_b(\cdot)$ denotes the splitting measure induced on $\mathcal{P}_b\setminus\{\mathbf{1}_b\}$ by $p_{\#b}$ through exchangeability.

5. Proof of Theorem 2.3

Theorem 2.3 summarizes the conclusions of a series of theorems and propositions that we prove in this section. Throughout this section, assume $p := (p_n)_{n\geq 2}$ is a collection of splitting rules satisfying (4) and $\tau := (\tau_n)_{n\geq 0}$ is a collection of geometric success probabilities. The pair (p,τ) determines a family $(Q^{[n]}_{p,\tau})_{n\in\mathbb{N}}$ of finite-dimensional probability distributions through (18). By Kolmogorov's extension theorem, $(Q^{[n]}_{p,\tau})_{n\in\mathbb{N}}$ determines a unique probability measure $Q^\bullet_{p,\tau}$ on $\mathcal{T}^\bullet_{\mathbb{N}}$ if and only if

$$Q^{[m]}_{p,\tau} = Q^{[n]}_{p,\tau}\mathbf{R}^{-1}_{m,n} \quad \text{for all } m\leq n, \quad \text{for every } n\in\mathbb{N}. \tag{19}$$

Theorem 5.1. *The family $(Q^{[n]}_{p,\tau})_{n\in\mathbb{N}}$ satisfies* (19) *if and only if $\tau_0 = \tau_1 = 0$ and*

$$\tau_n = \tau_{n+1}(1 - p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})) \quad \textit{for every } n\geq 2. \tag{20}$$

Proof. Clearly, $\tau_0=\tau_1=0$ is both necessary and sufficient for $Q^{[n]}_{p,\tau}$-almost every $\mathbf{t}\in\mathcal{T}^\bullet_{[n]}$ to satisfy (i)$^{\mathbf{w}}$ in the definition of a weighted fragmentation tree, for every $n\in\mathbb{N}$. Henceforth, we fix $n\geq 2$ and examine condition (20) for τ_n.

For $Q^{[n]}_{p,\tau}$ defined as in (18), let $\mathbf{T}^\bullet_{n+1} = (\mathbf{T}_{n+1}, W_{n+1}) \sim Q^{[n+1]}_{p,\tau}$ and define $\mathbf{T}^\bullet_n = (\mathbf{T}_n, W_n) := \mathbf{R}^\bullet_{n,n+1}\mathbf{T}^\bullet_{n+1}$. By (19), we must show that $\mathbf{T}^\bullet_n \sim Q^{[n]}_{p,\tau}$.

In general, for any pair $(\mathbf{t},\mathbf{t}')$, with $\mathbf{t}'\in\mathcal{T}_{[n+1]}$ and $\mathbf{t} := \mathbf{R}_{n,n+1}\mathbf{t}'$, there is a unique element $b\in\mathbf{t}$ such that $b\cup\{n+1\}$, b and $\{n+1\}$ are all elements of $\mathbf{t}'$. We denote this unique element by $b^*\in\mathbf{t}$ and we say that $\{n+1\}$ is *attached below* b^* in $\mathbf{t}'$. Now, by construction, $\mathbf{R}^\bullet_{n,n+1}\mathbf{T}^\bullet_{n+1} = \mathbf{T}^\bullet_n$ and, therefore, $\mathbf{R}_{n,n+1}\mathbf{T}_{n+1} = \mathbf{T}_n$. Hence, we can define $b^*\in\mathbf{T}_n$ for the unique b^* below which $n+1$ is attached in $\mathbf{T}_{n+1}$. By definition of $\mathbf{R}^\bullet_{n,n+1}$ in (16),

$$W_n(b) = \max(W_{n+1}(b), W_{n+1}(b\cup\{n+1\})) \quad \text{for all } b\in\mathbf{t}\setminus\{b^*\}$$

and

$$W_n(b^*) = W_{n+1}(b^*) + W_{n+1}(b^*\cup\{n+1\}) > \max(W_{n+1}(b^*), W_{n+1}(b^*\cup\{n+1\})) \quad \text{a.s.}$$

By assumption (4) on p, the finite-dimensional distributions $(Q^{[n]}_p)_{n\in\mathbb{N}}$ on $\{\mathcal{T}_{[n]}\}_{n\in\mathbb{N}}$ are consistent and, therefore, $\mathbf{T}_n$ is distributed according to $Q^{[n]}_p$. The Markov property of the $\mathbf{T}_n$, together with conditional independence of the edge lengths, implies that $\mathbf{T}^\bullet_n \sim Q^{[n]}_{p,\tau}$ if and only if, for every $n\geq 0$,

$$X + X'I_E =_{\mathcal{L}} X',$$

where $X\sim \textsc{Geo}(\tau_{n+1})$, $X'\sim\textsc{Geo}(\tau_n)$, E is an event with probability $p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})$ and X, X', E are mutually independent. (This plainly holds for $n\in\{0,1\}$ by assumption $\tau_0=\tau_1=0$, and so we consider the case $n\geq 2$.) The probability generating function $G_Y(s) := \mathbb{E}s^Y$ of a geometric variable Y with success probability $p\in(0,1)$ is

$$G_Y(s) := \frac{sp}{1-s(1-p)};$$

and so, condition $X + X'I_E =_{\mathcal{L}} X'$ implies that

$$\mathbb{E}s^{X+X'I_E} = \frac{s\tau_n}{1 - s(1-\tau_n)}, \quad \text{for all } n \in \mathbb{N}.$$

Fixing $s > 0$ and writing $\sigma_n := 1 - \tau_n$, we have

$$\begin{aligned}
\mathbb{E}s^{X+X'I_E} &= \frac{s\tau_{n+1}}{1-s\sigma_{n+1}}\left[p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})\frac{s\tau_n}{1-s\sigma_n} + 1 - p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})\right] \\
&= \frac{s\tau_n}{1-s\sigma_n}\left\{\frac{s\tau_{n+1}}{1-s\sigma_{n+1}}\left[\frac{p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})s\tau_n + (1-p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})) - s\sigma_n(1-p_{n+1}(\mathbf{e}^{(n+1)}_{n+1}))}{s\tau_n}\right]\right\} \\
&= \frac{s\tau_n}{1-s\sigma_n}\left\{\frac{s\tau_{n+1}}{1-s\sigma_{n+1}}\left[\frac{(1-s)(1-p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})) + s\tau_n}{s\tau_n}\right]\right\}.
\end{aligned}$$

It follows that $X + X'I_E =_{\mathcal{L}} X'$ if and only if

$$\frac{\tau_{n+1}}{1-s\sigma_{n+1}} = \frac{\tau_n}{(1-s)(1-p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})) + s\tau_n}.$$

By assumption, both τ_n and τ_{n+1} are strictly positive. Hence, there exists a unique $\alpha > 0$ such that $\alpha\tau_n = \tau_{n+1}$. We must have

$$\frac{\tau_{n+1}}{1-s\sigma_{n+1}} = \frac{\alpha}{\alpha}\frac{\tau_n}{(1-s)(1-p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})) + s\tau_n} = \frac{\tau_{n+1}}{(1-s)(1-p_{n+1}(\mathbf{e}^{(n+1)}_{n+1}))\alpha + s\tau_{n+1}}.$$

Because $s > 0$, it follows that $\alpha(1 - p_{n+1}(\mathbf{e}^{(n+1)}_{n+1})) = 1$. This completes the proof. □

Our next step is to show the correspondence between probability measures ν^* satisfying (6) and pairs (p, τ) satisfying (4) and (5). In this direction, let ν^* be a probability measure on $\Delta^{\downarrow}$ satisfying (6). By Kingman's correspondence, ν^* determines a unique exchangeable paintbox measure ϱ_{ν^*} on $\mathcal{P}_{\mathbb{N}}$. As before, we write $\varrho^{(n)}_{\nu^*} := \varrho_{\nu^*}\mathbf{D}_n^{-1}$ to denote the distribution ϱ_{ν^*} induces on $\mathcal{P}_{[n]}$ through deletion. Furthermore, for any $b \subset_f \mathbb{N}$, we write $\varrho^b_{\nu^*}$ to denote the measure ϱ_{ν^*} induces on $\mathcal{P}_b$. By construction $(\varrho^{(n)}_{\nu^*})_{n\in\mathbb{N}}$ is exchangeable and satisfies the consistency condition

$$\varrho^{(m)}_{\nu^*}(\pi) = \varrho^{(n)}_{\nu^*}(\mathbf{D}^{-1}_{m,n}(\pi)), \quad \pi \in \mathcal{P}_{[m]}, \quad \text{for every } m \le n < \infty. \tag{21}$$

Given ν^*, define $p := (p^{\nu^*}_n)_{n\ge 2}$ as in (8). It is clear, by assumption (6), that $p^{\nu^*}_n$ is a probability distribution on $\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}$ for every $n \ge 2$. Exchangeability and consistency (4) of p follows easily from properties of ϱ_{ν^*}.

Theorem 5.2. *The identities* (8) *and* (9) *establish a bijection between pairs* (p, τ) *satisfying* (4) *and* (5) *and probability measures* ν^* *on* $\Delta^{\downarrow}$ *satisfying* (6)*. Therefore, to any such* (p, τ)*, there is a unique measure* ν^* *such that* $Q^{\bullet}_{p,\tau}$ *has finite-dimensional marginal distributions*

$$Q^{[n]}_{\nu^*}(\mathbf{t}^{\bullet}) = \prod_{b\in\mathbf{t}:\#b\ge 2} \varrho^b_{\nu^*}(\mathbf{1}_b)^{w(b)-1}\varrho^b_{\nu^*}(\Pi_{\mathbf{t}_{|b}}), \quad \mathbf{t}^{\bullet} := (\mathbf{t}, w) \in \mathcal{T}^{\bullet}_{[n]}, \quad n \in \mathbb{N}. \tag{22}$$

Proof. First, suppose (p, τ) satisfies (4) and (5). For each $n \in \mathbb{N}$, we define a probability measure $P_n(\cdot)$ on $\mathcal{P}_{[n]}$ by

$$P_n(\pi) := \begin{cases} \tau_n p_n(\pi), & \pi \ne \mathbf{1}_{[n]} \\ 1 - \tau_n, & \pi = \mathbf{1}_{[n]}. \end{cases}$$

Putting $P_1(\mathbf{1}_{[1]}) = 1$, we have a collection $(P_n)_{n\in\mathbb{N}}$ of exchangeable marginal distributions on $\{\mathcal{P}_{[n]}\}_{n\in\mathbb{N}}$ that corresponds to p through (8). It is easy to check that $(P_n)_{n\in\mathbb{N}}$ is consistent from the assumptions (4) and (5). Therefore, by Kolmogorov's extension theorem, $(P_n)_{n\in\mathbb{N}}$ determines a unique exchangeable probability measure on $\mathcal{P}_{\mathbb{N}}$ which, by Kingman's correspondence, is a paintbox measure ϱ_{ν^*} for some unique probability measure ν^* on $\Delta^{\downarrow}$. Moreover, by assumption, $\tau_{n+1} \geq \tau_n > 0$ for all $n \geq 2$ and so $\tau_n \to \tau_\infty > 0$. By monotone convergence, we have

$$\varrho_{\nu^*}(\mathbf{1}_{[\infty]}) = \lim_{n\to\infty} \downarrow \varrho_{\nu^*}\mathbf{D}_n^{-1}(\mathbf{1}_{[n]}) = \lim_{n\to\infty} \downarrow \varrho_{\nu^*}^{(n)}(\mathbf{1}_{[n]}) = 1 - \lim_{n\to} \tau_n = 1 - \tau_\infty < 1.$$

Hence, ν^* must satisfy (6).

Conversely, let ν^* be a probability measure on $\Delta^{\downarrow}$ satisfying (6) and define $p^* := (p_n^{\nu^*})_{n\in\mathbb{N}}$ by (8) and $\tau^* := (\tau_n^{\nu^*})_{n\geq 0}$ by (9). Plainly, p^* satisfies (4). We also see that, for every $n \geq 2$,

$$\begin{aligned}
\tau_{n+1}^{\nu^*}(1 - p_{n+1}^{\nu^*}(\mathbf{e}_{n+1}^{(n+1)})) &= (1 - \varrho_{\nu^*}^{(n+1)}(\mathbf{1}_{[n+1]}))\left(1 - \frac{\varrho_{\nu^*}^{(n+1)}(\mathbf{e}_{n+1}^{(n+1)})}{1 - \varrho_{\nu^*}^{(n+1)}(\mathbf{1}_{[n+1]})}\right) \\
&= 1 - \varrho_{\nu^*}^{(n+1)}(\mathbf{1}_{[n+1]}) - \varrho_{\nu^*}^{(n+1)}(\mathbf{e}_{n+1}^{(n+1)}) \\
&= 1 - \varrho_{\nu^*}^{(n)}(\mathbf{1}_{[n]}) \\
&= \tau_n^{\nu^*},
\end{aligned}$$

where the above expression simplifies because $\varrho_{\nu^*}^{(n+1)}(\mathbf{1}_{[n+1]}) + \varrho_{\nu^*}^{(n+1)}(\mathbf{e}_{n+1}^{(n+1)}) = \varrho_{\nu^*}^{(n)}(\mathbf{1}_{[n]})$ by consistency (21) of $(\varrho_{\nu^*}^{(n)})_{n\in\mathbb{N}}$. Hence, (5) is satisfied.

Equation (22) follows immediately from (18). This completes the proof. □

Theorem 5.3. *Let $p := (p_n)_{n\geq 2}$ be a family of splitting rules satisfying* (4) *and let $\lambda := (\lambda_n)_{n\geq 2}$ be defined in* (7) *with respect to p. Then Q_p-almost every $\mathbf{t} \in \mathcal{T}_{\mathbb{N}}$ possesses a root partition if and only $\lambda_\infty := \lim_{n\to\infty} \lambda_n < \infty$.*

Proof. First, suppose that Q_p-almost every $\mathbf{t} \in \mathcal{T}_{\mathbb{N}}$ possesses a root partition. Then, by our definition of root partition in Section 4.1,

$$\mathbb{P}(\{\Pi_{\mathbf{T}} \text{ exists}\}) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{\Pi_{\mathbf{T}} \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}\right) = 1.$$

On the other hand, by (7), we have

$$\lambda_n/\lambda_{n+1} = 1 - p_{n+1}(\mathbf{e}_{n+1}^{(n+1)}) \quad \text{for all } n \geq 2.$$

Now, $p_n(\mathbf{e}_n^{(n)}) \in [0,1]$ for every $n \in \mathbb{N}$, and so the sequence $\lambda := (\lambda_n)_{n\geq 2}$ is monotonically non-decreasing and $\lambda_\infty := \lim_{n\to\infty} \lambda_n$ exists. For fixed $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\}$,

$$\mathbb{P}(\{\Pi_{\mathbf{T}} \in \mathbf{D}_n^{-1}(\pi)\}) = p_n(\pi) \prod_{j=1}^{\infty} (1 - p_{n+j}(\mathbf{e}_{n+j}^{(n+j)})) = p_n(\pi)\lambda_n \lim_{j\to\infty} \lambda_{n+j}^{-1};$$

hence,

$$\mathbb{P}(\{\Pi_{\mathbf{T}} \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}) = \lambda_n \lim_{j\to\infty} \lambda_{n+j}^{-1} = \lambda_n/\lambda_\infty. \tag{23}$$

Now, either $\lambda_\infty = \infty$ or $0 < \lambda_\infty < \infty$. On the one hand, if $\lambda_\infty = \infty$, then $\mathbb{P}(\{\Pi_{\mathbf{T}} \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}) = \lambda_n/\lambda_\infty = 0$ for all $n \in \mathbb{N}$; whence,

$$1 = \mathbb{P}(\{\Pi_{\mathbf{T}} \text{ exists}\}) = \mathbb{P}\left(\bigcup_{n=1}^{\infty}\{\Pi_T \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(\{\Pi_T \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}\right) = 0,$$

a contradiction. On the other hand, if $\lambda_\infty < \infty$, then $\lambda_n/\lambda_\infty \to 1$ as $n \to \infty$ and, therefore, $\mathbb{P}\left(\{\Pi_{\mathbf{T}} \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}\right) \to 1$ as $n \to \infty$. Consequently,

$$1 = \mathbb{P}(\{\Pi_{\mathbf{T}} \text{ exists}\}) = \mathbb{P}\left(\bigcup_{n=1}^{\infty}\{\Pi_T \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(\{\Pi_T \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\}\right) = \infty,$$

establishing the first claim.

Conversely, suppose $\lambda_\infty := \lim_{n\to\infty} \lambda_n < \infty$. For each $n \geq 2$, define the event $A_n := \{\Pi_{\mathbf{T}_{|[n]}} = \mathbf{e}_n^{(n)}\}$. By the Markov branching property and consistency (4), the events $\{A_n\}_{n\geq 2}$ are independent; hence, the random variables $\{\mathbf{1}_{A_n}\}_{n\geq 2}$ are independent Bernoulli random variables with parameter $p_n(\mathbf{e}_n^{(n)})$ for each $n \geq 2$. Moreover, $\{\Pi_{\mathbf{T}} \text{ exists}\} = \{\sum \mathbf{1}_{A_n} < \infty\}$. Clearly, the event $\{\sum \mathbf{1}_{A_n} < \infty\}$ is in the tail σ-field generated by $\{A_n\}_{n\geq 2}$. Hence, the event $\{\Pi_{\mathbf{T}} \text{ exists}\}$ has probability 0 or 1 by Kolmogorov's 0-1 law. However, by (23),

$$\mathbb{P}\{\Pi_{\mathbf{T}} \in \mathbf{D}_n^{-1}(\mathcal{P}_{[n]} \setminus \{\mathbf{1}_{[n]}\})\} = \lambda_n \lim_{j\to\infty} \lambda_{n+j}^{-1} = \lambda_n/\lambda_\infty > 0 \quad \text{for every } n \geq 2.$$

Therefore, $\mathbb{P}(\{\Pi_{\mathbf{T}} \text{ exists}\}) \geq \lambda_n/\lambda_\infty > 0$ and we conclude $\{\Pi_{\mathbf{T}} \text{ exists}\}$ has probability one. □

Proposition 5.4. *Let $p := (p_n)_{n\geq 2}$ be a family of splitting rules satisfying* (4) *and define $\lambda := (\lambda_n)_{n\geq 2}$ as in (vi) of Theorem 2.3. Then there exists a collection $\tau := (\tau_n)_{n\geq 0}$ of success probabilities satisfying* (5) *with respect to p if and only if* $\lim_{n\to\infty} \lambda_n < \infty$.

Proof. We have already noted that $(\lambda_n)_{n\geq 2}$ defined in (7) is monotonically non-decreasing, and so $\lim_{n\to\infty} \lambda_n$ exists. Suppose there exists τ satisfying (5) with respect to p. Then $\lambda := (\lambda_n)_{n\geq 2}$, as defined in (7), satisfies (12), which is the same condition as (5); hence, there exists $\alpha \in (0, \infty)$ such that $\lambda_n = \alpha\tau_n$ for every $n \in \mathbb{N}$. Since $\tau_n \leq 1$ for all $n \in \mathbb{N}$, we conclude $\lim_{n\to\infty} \lambda_n = \alpha \lim_{n\to\infty} \tau_n \leq \alpha < \infty$. Conversely, if $\lambda_n \to \lambda_\infty < \infty$, we can define $\tau_n := \lambda_n/\lambda_\infty$ for $n \geq 2$, which satisfies (5).

In fact, we could take any $\lambda_\infty \leq \lambda^* < \infty$ and put $\tau_n := \lambda_n/\lambda^*$. The choice $\lambda^* = \lambda_\infty$ coincides with the case $\tau_\infty = 1$; in general, to specify $\tau_\infty \in (0, 1]$, we choose $\lambda^* = \lambda_\infty/\tau_\infty \geq \lambda_\infty$ and we have

$$\lim_{n\to\infty} \tau_n = \lim_{n\to\infty} \lambda_n/\lambda^* = \frac{\tau_\infty}{\lambda_\infty} \lim_{n\to\infty} \lambda_n = \tau_\infty.$$

□

Proposition 5.5. *There is a bijection between probability measures ν^* satisfying* (6) *and pairs (ν, τ_∞) with $\nu(\{(1, 0, \ldots)\}) = 0$, $\nu(\Delta^{\downarrow}) < \infty$ and $0 < \tau_\infty \leq 1$.*

Proof. This follows by the discussion in Section 3.1: Given ν^* satisfying (6), we define ν as in (10) and put $\tau_\infty := 1 - \nu^*(\{(1, 0, \ldots)\})$. Conversely, given (ν, τ_∞), we define ν^* by (11). This completes the proof.

□

The equivalence of Parts (i)-(vi) of Theorem 2.3 have now been proven according to the following scheme.

(i)⇔(ii): Theorem 5.1

(ii)⇔(iii): Theorem 5.2
(v)⇔(vi): Theorem 5.3
(ii)⇔(vi): Proposition 5.4
(iii)⇔(iv): Proposition 5.5

This completes the proof.

References

[1] D. Aldous. Probability distributions on cladograms. In *Random Discrete Structures*, pages 1–18. Springer, 1996.

[2] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Related Fields*, 121(3):301–318, 2001.

[3] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[4] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Probab.*, 36(5):1790–1837, 2008.

[5] J. F. C. Kingman. The representation of partition structures. *J. London Math. Soc. (2)*, 18(2):374–380, 1978.

[6] P. McCullagh, J. Pitman, and M. Winkel. Gibbs fragmentation trees. *Bernoulli*, 14(4):988–1002, 2008.

[7] S. Tavaré. *Ancestral Inference in Population Genetics*, volume 1837 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2004. Lectures from the 31st Summer School on Probability Theory held in Saint-Flour, 2001.

Rutgers University, Department of Statistics, 110 Frelinghuysen Road, Piscataway, NJ 08854.