Partly exchangeable fragmentations

Bo Chen

Jesus College

University of Oxford

A dissertation submitted for the degree

Doctor of Philosophy

Trinity Term 2009

I certify that this thesis is my own work (except where otherwise indicated). It is being submitted to the University of Oxford for the degree Doctor of Philosophy. It has not been submitted before to any other University for any degree or examination .

Candidate: Bo Chen

Signed ....

Dated ....

Acknowledgements

First and most importantly, I would like to thank my supervisor Matthias Winkel. All of my research in University of Oxford was under his supervision. His excellent guidance, invaluable support and creative ideas for the subject were absolutely essential for me. I am grateful to my examiners Bob Griffiths and Andreas E. Kyprianou for their careful reading and constructive comments. Furthermore, I acknowledge the contribution of K C Wong Scholarship granted by K C Wong Education foundation and University of Oxford.

Finally, I would like to thank my parents and wife for their generosity in all respects.

Partly exchangeable fragmentations

Bo Chen

Jesus College

A dissertation submitted for the degree

Doctor of Philosophy

in Trinity Term 2008

Abstract

We introduce a simple tree growth process that gives rise to a new two-parameter family of discrete fragmentation trees that extends Ford's alpha model to multifurcating trees and includes the trees obtained by uniform sampling from Duquesne and Le Gall's stable continuum random tree. We call these new trees the alpha-gamma trees. In this thesis, we obtain their splitting rules, dislocation measures both in ranked order and in sized-biased order, and we study their limiting behaviour. We further extend the underlying exchangeable fragmentation processes of such trees into partly exchangeable fragmentation processes by weakening the exchangeability. We obtain the integral representations for the measures associated with partly exchangeable fragmentation processes and subordinator of the tagged fragments. We also embed the trees associated with such processes into continuum random trees and study their limiting behaviour. In the end, we generate a three-parameter family of partly exchangeable trees which contains the family of the alpha-gamma trees and another important two-parameter family based on Poisson-Dirichlet distributions.

Key words: Alpha-gamma tree, splitting rule, sampling consistency, self-similar fragmentation, dislocation measure, continuum random tree, $\mathbb{R}$-tree, Markov branching model, exchangeability, part exchangeability, Poisson-Dirichlet distribution

Contents

1 Introduction **1**

1.1 Background 1

1.2 Ford's alpha model 2

1.2.1 Sequential growth rule 2

1.2.2 Markov branching property 2

1.2.3 Fragmentation processes 3

1.2.4 Dislocation measure and scaling limits for the alpha model 7

1.3 Outline of the Thesis 7

1.3.1 Chapter 2: A new family of Markov branching trees: the alpha-gamma model 8

1.3.2 Chapter 3: Continuum tree asymptotics of partly exchangeable fragmentations 10

2 A new family of Markov branching trees: the alpha-gamma model **15**

2.1 Introduction 16

2.2 Sampling consistent splitting rules for the alpha-gamma trees 23

2.2.1 Notation and terminology of partitions and discrete fragmentation trees 23

2.2.2 Colour-marked trees and the proof of Proposition 2.6 25

2.2.3 The Chinese Restaurant Process 26

2.2.4 The splitting rule of alpha-gamma trees and the proof of Proposition 2.1 28

2.2.5 Sampling consistency and strong sampling consistency 33

2.3 Dislocation measures and asymptotics of alpha-gamma trees 35

2.3.1 Dislocation measures associated with the alpha-gamma-splitting rules 35
2.3.2 The alpha-gamma model when $\alpha = 1$, spine with bushes of singleton-trees . . . 37
2.3.3 Continuum random trees and self-similar trees . . . 40
2.3.4 The alpha-gamma model when $\gamma = 1 - \alpha$, sampling from the stable CRT . . . 42
2.3.5 Dislocation measures in size-biased order . . . 43
2.3.6 Convergence of alpha-gamma trees to self-similar CRTs . . . 46
2.4 Limiting results for labelled alpha-gamma trees . . . 48
2.4.1 The scaling limits of reduced alpha-gamma trees . . . 48
2.4.2 Further limiting results . . . 50

3 Continuum tree asymptotics of partly exchangeable fragmentation trees 53
3.1 Introduction . . . 54
3.2 Homogeneous partly exchangeable fragmentation processes . . . 59
3.2.1 Exchangeable partitions and fragmentation processes . . . 59
3.2.2 Partially exchangeable random partitions and constrained exchangeable partitions . . . 61
3.2.3 Splitting rates and Poisson construction for partly exchangeable fragmentation processes . . . 62
3.2.4 Auxiliary σ-finite measures . . . 66
3.2.5 Decomposition of partly exchangeable measures . . . 71
3.2.6 Subordinator representation of the tagged partly exchangeable fragment . . . 74
3.2.7 Self-similar partly exchangeable fragmentations . . . 79
3.2.8 Discrete partly exchangeable fragmentation trees . . . 81
3.2.9 Consistency . . . 85
3.2.10 Building discrete trees with edge lengths . . . 87
3.2.11 $\mathbb{R}$-trees and reduced trees . . . 88
3.2.12 Continuum random trees . . . 90

3.3 Embedding partly exchangeable fragmentation trees into CRTs 92
3.3.1 Mass-fragmentation and CRT 92
3.3.2 A generic procedure to sample a leaf from a CRT 92
3.3.3 Sample $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_k)$ from $(\mathcal{T}, \mu)$. 96
3.4 Some applications of embedding procedure 99
3.4.1 Special branch points estimation 99
3.4.2 The convergence of reduced discrete partly exchangeable fragmentation tree 106
3.4.3 Height estimation for discrete fragmentation trees 112
3.4.4 Tightness estimation and Proof of Theorem 3.2 119
3.5 Three examples of partly exchangeable fragmentation trees 121
3.5.1 Ford's alpha models 121
3.5.2 Alpha-gamma models 122
3.5.3 The three-factor model 124

Chapter 1

Introduction

1.1 Background

For a number of years there has been an increased interest in phylogenetic trees in mathematical literatures. They are used in biological systems to represent the evolutionary relationship between species and called cladograms in some literature on the binary trees. The fundamental classes of phylogenetic trees are connected graphs with no degree-2 vertices and $n+1$ degree-1 vertices, of which one degree-1 vertex is distinguished as the ROOT and the rest are considered as leaves and labelled 1 up to n.The number of leaves is called size of the tree. We denote by $\mathbb{T}_n$ the space of such trees with n leaves.

Three popular and fundamental probability models on binary phylogenetic trees are the Yule, uniform and comb models [41]. The Yule model is also referred to as the neutral evolution model. The uniform model assigns the uniform probability measure to phylogenetic trees of each size. The comb model assigns probability 1 to the most asymmetric tree of each size. Aldous [4, 5, 6] introduced a one dimensional continuous family of models, called the beta-splitting model, which interpolates between the Yule, uniform and comb models. McCullagh et al. [35] extend the beta-splitting model to the multifurcating case as Poisson-Dirichlet model. Ford [21] gave another family of models on binary phylogenetic trees, called the alpha model which also interpolates between the Yule, uniform and comb models. The most important model of multifurcating phylogenetic trees is that related to stable trees. Stable trees have been introduced by Duquesne and Le Gall [14] and implicitly considered in the work of Kersting [30].

1.2 Ford's alpha model

1.2.1 Sequential growth rule

Ford's alpha model is a binary model, which is built by simple sequential growth rules starting from the unique elements $T_1 \in \mathbb{T}_1$ and $T_2 \in \mathbb{T}_2$ as follows:

(i)$^{\mathrm{F}}$ given T_n for $n \geq 2$, assign a weight $1-\alpha$ to each of the n edges adjacent to a leaf, and a weight α to each of the $n-1$ other edges;

(ii)$^{\mathrm{F}}$ select at random with probabilities proportional to the weights assigned by step (i), an edge of T_n, say $a_n \to c_n$ directed away from the ROOT;

(iii)$^{\mathrm{F}}$ to create T_{n+1} from T_n, replace $a_n \to c_n$ by three edges $a_n \to b_n$, $b_n \to c_n$ and $b_n \to n+1$ so that two new edges connect the two vertices a_n and c_n to a new branch point b_n and a further edge connects b_n to a new leaf labelled $n+1$.

The trees generated by the growth rule of the alpha model are labelled, where labels come from the natural insertion order. We can obtain unlabelled alpha-trees by removing their labels. Mathematically, the unlabelled trees are equivalent classes of labelled trees.

1.2.2 Markov branching property

Informally, the Markov branching property means that given the split at the branch point adjacent to the root in a random tree, the subtrees above it are distributed independently from each other and they are distributed according to the same model. Mathematically, a sequence $(T_n^\circ, n \geq 1)$ of unlabelled trees has Markov branching property if for all $n \geq 2$ conditionally given that at the branch point adjacent to the ROOT, the tree is split into a number of tree components with $n_1, \ldots, n_k$ leaves, these tree components are the independent copies of $T_{n_i}^\circ$. The probability mass functions of the split at the branch point adjacent to the ROOT of $T_n^\circ, n \geq 2$ are denoted by $q_n(n_1, \ldots, n_k)$, $n_1 \geq \ldots \geq n_k \geq 1$, $k \geq 2$, $n_1 + \ldots + n_k = n$ and referred to as the *splitting rules* of $(T_n^\circ, n \geq 1)$.

Ford showed that unlabelled alpha trees are Markov branching and splitting rules

$$q_n(m, n-m) = \left(\frac{\alpha}{2} \binom{n}{m} + (1-2\alpha) \binom{n-2}{m-1} \right) \frac{\Gamma(m-\alpha)\Gamma(n-m-\alpha)}{\Gamma(1-\alpha)\Gamma(n-\alpha)}.$$

1.2.3 Fragmentation processes

As each leaf is uniquely labelled in the alpha model, each subtree can be uniquely distinguished by a set of labels. Consider the first branch point that an n-leaf alpha tree splits into several subtrees characterized by their sets of leaf labels, which are disjoint subsets of $\{1, \ldots, n\}$ and their union is exactly $\{1, \ldots, n\}$. This gives us a natural relationship between the split at the branch point adjacent to the root in alpha trees and partitions. Further the alpha trees are connected with fragmentation process. In this section, we review some basic definitions and results on partition-valued fragmentation processes.

Partitions

Denote by $[n]$ the set $\{1, \ldots, n\}$ for all $n \in \mathbb{N}$. Let $B \subseteq \mathbb{N}$, a *partition* of B is a countable collection $\pi = \{\pi_i, i \in \mathbb{N}\}$ of pairwise disjoint subsets of B such that $\cup_{i\in\mathbb{N}}\pi_i = B$. These disjoint subsets $\pi_i, i \in \mathbb{N}$ are called the *blocks* of π. We write $\mathcal{P}_B$ for the set of partitions of B. In the special case when $B = [n]$, we simply write $\mathcal{P}_n := \mathcal{P}_{[n]}$; in particular $\mathcal{P} := \mathcal{P}_{\mathbb{N}}$. If $\pi \in \mathcal{P}_B$, and $B' \subseteq \mathbb{N}$, we let $\pi|_{B'} = B' \cap \pi$ be the partition of $B' \cap B$ obtained by restricting π to the elements of $B' \cap B$. Let $\pi|_n := \pi|_{[n]}$ for every $n \geq 1$. By convention, we let $\mathbf{1}_B$ be the trivial partition $(B, \emptyset, \ldots)$ of B, and $\mathbf{0}_B = (\{i_1\}, \{i_2\}, \ldots)$ the partition of B into singletons, where $i_1 < i_2 < \ldots$ is the ranked list of elements of B.

We say that a block B of some partition $\pi \in \mathcal{P}$ possesses an *asymptotic frequency* if the limit

$$|B| := \lim_{n\to\infty} \frac{\#(B \cap [n])}{n}$$

exists. If each block of π has an asymptotic frequency, then we say that π possesses asymptotic frequencies and write $|\pi|^{\downarrow} = (|\pi_1|, \ldots)^{\downarrow}$ as the ranked order of asymptotic frequencies of blocks. Fatou's lemma implies that $\sum_{i=1}^{\infty} |\pi_i| \leq 1$. We say that a partition π has *proper* asymptotic frequencies if π possesses asymptotic frequencies with $\sum_{i=1}^{\infty} |\pi_i| = 1$.

The asymptotic frequencies of a partition actually can be viewed as a decreasing sequence of masses with a sum no larger than 1, which was referred to as mass-partition by Ferguson [20] and Kingman [32]. A *mass-partition* is an infinite numerical sequence

$\mathbf{s} = (s_1, \ldots)$ with

$$s_1 \geq s_2 \geq \ldots \geq 0 \text{ and } \sum_{i=1}^{\infty} s_i \leq 1.$$

We denote the space of mass-partitions by $\mathcal{S}^{\downarrow}$.

We endow the space $\mathcal{S}^{\downarrow}$ with the uniform distance $d(\mathbf{s}, \mathbf{s}') = \max\{|s_i - s'_i|, i \in \mathbb{N}\}$, for any $\mathbf{s}, \mathbf{s}' \in \mathcal{S}^{\downarrow}$. Then the space is compact and the induced topology coincides with that of pointwise convergence.

Homogeneous fragmentations

Bertoin [8] introduced right-continuous partition-valued Markov processes called fragmentation processes. The random partitions considered in Bertoin's fragmentation processes are *exchangeable*, which means that their laws are invariant under the natural action of the permutations of $\mathbb{N}$.

Let $B \subset \mathbb{N}$, and consider a $\mathbb{P}_B$-valued Markov process $(\Pi(t), t \geq 0)$. Π is called a *homogeneous fragmentation process* if its semigroup can be described as follows. For every $t, t' \geq 0$, the conditional distribution of $\Pi(t + t')$ given $\Pi(t) = \pi$ is the same as that of the random partition whose blocks are given by $\Pi_i(t) \cap \pi_j^{(i)}$, where $\pi^{(\cdot)} = (\pi^{(i)} : i = 1, \ldots)$ is an i.i.d sequence of exchangeable random partitions whose law only depend on t' and $\Pi_i(t)$ and $\pi_j^{(i)}$ are the ith and jth block of $\Pi(t)$ and $\pi^{(i)}$, respectively. If a homogeneous fragmentation of B starts from the trivial partition $\mathbf{1}_B$ of B, we say that the process is *standard*.

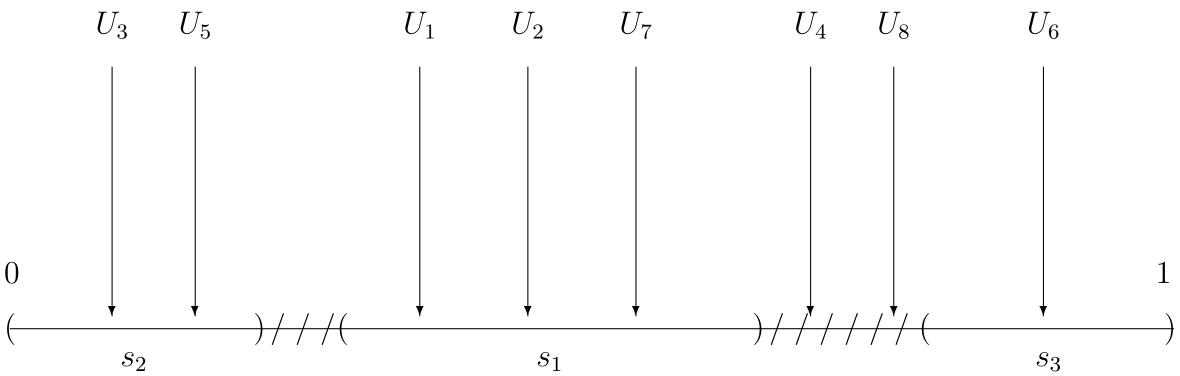
As shown by Bertoin [8], the laws of standard homogeneous fragmentations of $\mathbb{N}$ are in one-to-one correspondence with σ-finite measures κ on $\mathcal{P} \setminus \{\mathbf{1}_{\mathbb{N}}\}$, that satisfy

$$\kappa(\{\pi \in \mathcal{P} : \pi|_n \neq \mathbf{1}_{[n]}\}) < \infty. \tag{1.1}$$

We call such measures *dislocation measures* on $\mathcal{P}$. Such a measure admits a simple representation called paintbox construction, which was originally shown by Kingman [31] for probability measures and was extended by Bertoin [8] to σ-finite measures.

Kingman's paintbox construction

Let $\mathcal{S}^{\downarrow} = \{\mathbf{s} = (s_1, \ldots), s_1 \geq s_2 \geq \ldots, \sum_{i=1}^{\infty} s_i \leq 1\}$. Fix $\mathbf{s} \in \mathcal{S}^{\downarrow}$ and consider an interval representation ϑ of $\mathbf{s}$ as follows: ϑ is an open subset of $(0, 1)$ such that the ranked sequence



An example for $\Pi = (\{1,2,7\},\{3,5\},\{4\},\{6\},\{8\})$ *based on* $\mathbf{s} = (s_1, s_2, s_3)$

of the lengths of its interval components is given by $\mathbf{s}$. Let $U_1, \ldots$ be an i.i.d sequence of uniform variables on $(0,1)$. Consider the random partition Π induced by the following equivalence relation:

$$i \overset{\Pi}{\sim} j \Leftrightarrow i = j \text{ or } U_i, U_j \text{ belong to the same component of } \vartheta.$$

If U_i does not belong to ϑ, $\{i\}$ is a singleton of Π. We refer to the random partition Π defined above as the paintbox based on $\mathbf{s}$.

As shown by Kingman [31], the paintbox based on $\mathbf{s} \in \mathcal{S}^{\downarrow}$ is a random exchangeable partition. Its law does not depend on the choice of the interval representation ϑ of $\mathbf{s}$. We denote the law by $\kappa_{\mathbf{s}}$. A key result obtained by Kingman [31] through martingale argument and also by Aldous [1] with a simpler approach through de Finetti's theorem is as follows. Let Π be an exchangeable random partition of $\mathbb{N}$. Then Π possesses asymptotic frequencies a.s. and the law of Π can be expressed as a mixture of paint-boxes:

$$\mathbb{P}(\Pi \in \cdot) = \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot)\mathbb{P}(|\Pi|^{\downarrow} \in d\mathbf{s}).$$

The idea of using Kingman's theory to investigate homogeneous fragmentation processes was suggested by Pitman [38] and developed by Bertoin [8, 10]. They represent the dislocation measure κ of a homogeneous fragmentation process by an integral representation of some σ-finite measure ν on $\mathcal{S}^{\downarrow}$ such that

$$\nu(\{(1,0,\ldots)\}) = 0 \text{ and } \int_{\mathcal{S}^{\downarrow}} (1 - s_1)\nu(d\mathbf{s}) < \infty \tag{1.2}$$

The result is as follows. Let κ be an exchangeable measure on $\mathcal{P}$ which fulfills (1.1). Then there exists a unique $c \geq 0$ and a unique measure ν on $\mathcal{S}^{\downarrow}$ that fulfills (1.2) such that

$$\kappa(\cdot) = c\epsilon(\cdot) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot)\nu(d\mathbf{s}), \tag{1.3}$$

where $\epsilon := \sum_{n=1}^{\infty} \delta_{\{\mathbb{N}\setminus\{n\},\{n\}\}}$ and δ_π stands for the Dirac point mass at $\pi \in \mathcal{P}$. We refer to c as the *erosion coefficient.*

Trees and fragmentation processes

Fragmentation processes are naturally related to trees. Suppose B is a finite subset of $\mathbb{N}$ with n elements and $\mathbf{t}$ is a collection of subsets of B with an additional member called "ROOT" such that,

- $B \in \mathbf{t}$; B is called common ancestor of $\mathbf{t}$;
- $\{i\} \in \mathbf{t}$ for all $i \in B$; $\{i\}$ is called a leaf of $\mathbf{t}$;
- for all A, $C \in \mathbf{t}$, either $A \cap C = \emptyset$, or $A \subseteq C$ or $C \subseteq A$.

If $A \subset C$, A is called a descendant of C and C is then called an ancestor of A. If for all $D \in \mathbf{t}$ with $A \subseteq D \subseteq C$ either $A = D$ or $D = C$, we call A a child of C and C the parent of A. If we equip $\mathbf{t}$ with the parent-child relation and also relate ROOT with B, then $\mathbf{t}$ is a rooted connected acyclic graph i.e. a tree. We denote the space of such trees $\mathbf{t}$ by $\mathbb{T}_B$ and simply denote $T_n = T_{[n]}$.

Let $(\pi(t), t \geq 0)$ be a fragmentation process. Assume further for some finite $B \subset \mathbb{N}$, $\pi(0)|_B = \mathbf{1}_B$ and $\pi(t)|_B = \mathbf{0}_B$ for some finite $t > 0$, where $\mathbf{1}_B$ is the trivial partition into a single block B and $\mathbf{0}_B$ is the partition of B into singletons. We define $\mathbf{t}_{\pi,B} = \{\text{ROOT}\} \cup \{A \subset B : A \in \pi(t)|_B \text{ for some } t \geq 0\}$ as the associated *fragmentation tree.*

As argued by Bertoin, a $\mathcal{P}$-valued process Π is a homogeneous fragmentation if and only if its restrictions to $[n]$ are homogeneous fragmentations of $[n]$, $n \geq 1$. In other words, homogeneous fragmentations of $\mathbb{N}$ are the same as consistent families of homogeneous fragmentations of $[n]$. Obviously, this amounts to a consistency property for the associated fragmentation trees $(T_n, n \geq 1)$. As the leaf labels are exchangeable, if we delabel the sequence $(T_{[n]}, n \geq 1)$ to be $(T_n^\circ, n \geq 1)$, the consistency property will be pushed forward as follows, which is called *sampling consistency*: T_n° can be obtained by removing a leaf uniformly chosen from T_{n+1}°.

As shown in [28], sampling consistent splitting rules are in one-to-one correspondence with dislocation measures κ on $\mathcal{P}$ of homogeneous fragmentation processes and eventually can be expressed as integral representations for some pair (c, ν) of the erosion coefficient c and σ-finite measure ν satisfying (1.2).

1.2.4 Dislocation measure and scaling limits for the alpha model

Ford showed that the unlabelled alpha trees are sampling consistent. Hence, there is a family of fragmentation processes whose associated unlabelled trees are alpha trees. For such processes, there erosion coefficient is 0 and the dislocation measure is binary (i.e. only assign mass on set $\{\mathbf{s} : s_2 = 1 - s_1\}$) with,

$$\nu_{\text{Ford-}\alpha}(s_1 \in dx) = \frac{1}{\Gamma(1-\alpha)} \left(\alpha (x(1-x))^{-\alpha-1} + (2-4\alpha)(x(1-x))^{-\alpha} \right) \mathbf{1}_{\{1/2 \le x \le 1\}} dx.$$

We can assign unit edge lengths to alpha trees and hence turn them into random metric spaces. Then as an application of the work of Haas et al. [28], the properly scaled alpha tree will converge to a continuum object, i.e

$$\frac{T_n^\circ}{n^\alpha} \xrightarrow{(d)} \mathcal{T}_{\alpha, \nu_{\text{Ford-}\alpha}} \tag{1.4}$$

in the so-called Gromov-Hausdorff topology as $n \to \infty$. $\mathcal{T}_{\alpha,\nu}$ is a self-similar continuum random tree with parameter (α, ν) constructed by Haas and Miermont [27].

1.3 Outline of the Thesis

In this section, we give an outline of the thesis. It consists of two independent and self-contained research papers. Chapter 2 is joint work with Matthias Winkel and Daniel J. Ford. I made major contributions of the theoretical work except for Proposition 12. This chapter focuses on developing a multifurcating extension of Ford's alpha model and its splitting rules, dislocation measures and asymptotics. Chapter 3 is joint work with Matthias Winkel who has been supervising my work on it. This chapter is devoted to weakening the exchangeability of fragmentations and to building a large family of new models which include the alpha model and the model developed in Chapter 2. The key properties and convergence results of such models are developed there.

1.3.1 Chapter 2: A new family of Markov branching trees: the alpha-gamma model

This paper has appeared in the Electronic Journal of Probability, see [12]. As we discussed in the preceding section, the alpha model is a binary model, whose law is determined by its sequential growth rules. In this section, we introduce a new model by extending the simple sequential growth rules to allow *multifurcation*. Specifically, we also assign weights to vertices as follows, cf. Figure 2.1:

(i) given T_n for $n \geq 2$, assign a weight $1-\alpha$ to each of the n edges adjacent to a leaf, a weight γ to each of the $n-1$ other edges, and a weight $(k-1)\alpha-\gamma$ to each vertex of degree $k+1 \geq 3$;

(ii) select at random with probabilities proportional to the weights assigned by step (i),

- an edge of T_n, say $a_n \to c_n$ directed away from the ROOT,
- or, as the case may be, a vertex of T_n, say v_n;

(iii) to create T_{n+1} from T_n, do the following:

- if an edge $a_n \to c_n$ was selected, replace it by three edges $a_n \to b_n$, $b_n \to c_n$ and $b_n \to n+1$ so that two new edges connect the two vertices a_n and c_n to a new branch point b_n and a further edge connects b_n to a new leaf labeled $n+1$;
- if a vertex v_n was selected, add an edge $v_n \to n+1$ to a new leaf labeled $n+1$.

The resulting model is called the *alpha-gamma model*. These growth rules satisfy the rules of probability for all $0 \leq \alpha \leq 1$ and $0 \leq \gamma \leq \alpha$. When $\alpha = \gamma$, the alpha-gamma model degenerates to the alpha model.

We obtain the splitting rules and their integral representation.

Theorem 1.1. *Let* $(T_n, n \geq 1)$ *be alpha-gamma trees for some* $0 \leq \alpha \leq 1$ *and* $0 \leq \gamma \leq \alpha$. *Then*

(i) *the delabelled trees* $T_n^\circ, n \geq 1$ *have the Markov branching property;*

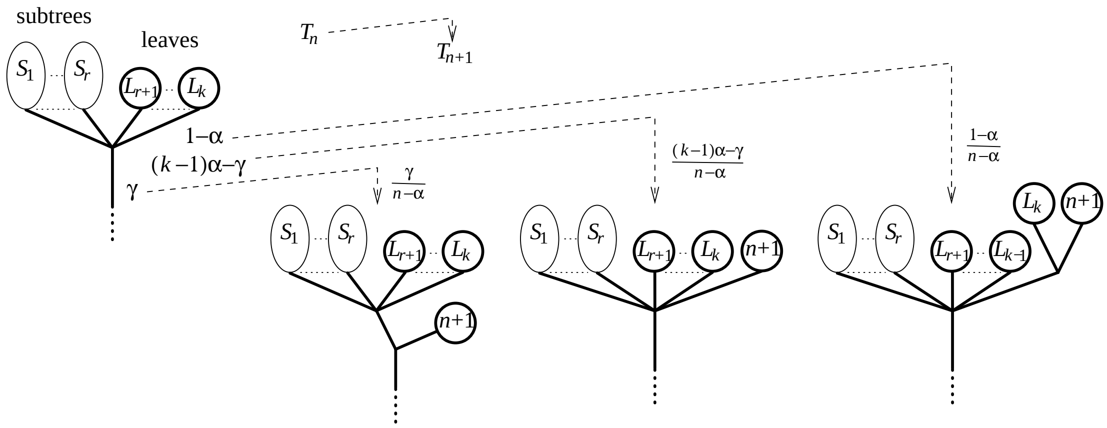


Figure 1.1: Sequential growth rule: displayed is one branch point of T_n with degree $k+1$, hence vertex weight $(k-1)\alpha-\gamma$, with $k-r$ leaves $L_{r+1},\ldots,L_k\in[n]$ and r bigger subtrees $S_1,\ldots,S_r$ attached to it; all edges also carry weights, weight $1-\alpha$ and γ are displayed here for one leaf edge and one inner edge only; the three associated possibilities for T_{n+1} are displayed.

(ii) *the splitting rules are sampling consistent and of following form: for* $n_1\ge\ldots\ge n_k$ *and* $n_1+\ldots+n_k=n$,

$$q^{\rm seq}_{\alpha,\gamma}(n_1,\ldots,n_k)\propto C_{n_1,\ldots,n_k}\left(\gamma+(1-\alpha-\gamma)\frac{\sum_{i\neq j}n_in_j}{n(n-1)}\right)p^{\rm PD^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k)$$

in the case $0\le\alpha<1$, *where* $p^{\rm PD^*}_{\alpha,-\alpha-\gamma}$ *is the Ewens-Pitman-type exchangeable partition probability function (EPPF) given by*

$$p^{\rm PD^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k)=\frac{\alpha^{k-2}\Gamma(k-1-\gamma/\alpha)\prod_{i=1}^k\Gamma(n_i-\alpha)}{\Gamma(1-\gamma/\alpha)\Gamma(n-\alpha)\Gamma^{k-1}(1-\alpha)},$$

$C_{n_1,\ldots,n_k}=\binom{n}{n_1,\ldots,n_k}/(m_1!\ldots m_n!)$ *and* m_r *is the number of* r*s of the sequence* $(n_1,\ldots,n_k)$*;*

(iii) *the measure* ν *in the integral representation* (1.3) *can be chosen as*

$$\nu_{\alpha,\gamma}(ds)=\left(\gamma+(1-\alpha-\gamma)\sum_{i\neq j}s_is_j\right){\rm PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}),\qquad(1.5)$$

where ${\rm PD}^*_{\alpha,-\alpha-\gamma}$ *is the Poisson-Dirichlet measure on* $\mathcal{S}^\downarrow$.

The case $\alpha=1$ is degenerate, we will discuss it in Section 2.3.2.

We observe that every dislocation measure ν on $\mathcal{S}^\downarrow$ gives rise to a measure $\nu^{\rm sb}$ on the space of summable sequences under which fragment sizes are in size-biased random order. One of the advantages of size-biased versions is that we can calculate marginal distributions explicitly.

Proposition 1.2. *For $0 < \alpha < 1$, distributions ν_k^{sb} of the first $k \geq 1$ marginals of the size-biased form of $\nu_{\alpha,\gamma}$ are given, for $\mathbf{x} = (x_1, \ldots, x_k)$, by*

$$\nu_k^{\mathrm{sb}}(d\mathbf{x}) = \left(\gamma + (1-\alpha-\gamma)\left(1 - \sum_{i=1}^{k} x_i^2 - \frac{(1-\alpha)(1-\sum_{i=1}^{k} x_i)^2}{1+(k-1)\alpha-\gamma}\right)\right) \frac{(1-\sum_{i=1}^{k} x_i)^{(k-1)\alpha-\gamma} \prod_{i=1}^{k} x_i^{-\alpha}}{\prod_{i=1}^{k}(1-\sum_{j=1}^{i} x_j) \prod_{i=1}^{k-1} B(1-\alpha, i\alpha-\gamma)} d\mathbf{x},$$

where $B(1-\alpha, i\alpha-\gamma) = \int_0^1 x^{-\alpha}(1-x)^{i\alpha-\gamma-1}dx$ is the beta function.

Similar to the alpha model, the unlabelled alpha-gamma trees have the same distribution as the discrete fragmentation tree associated with a fragmentation process with no erosion and dislocation measure (1.5). As in (1.4), we obtain the scaling limits for alpha-gamma trees from [28].

Corollary 1.3. *For some $0 < \alpha < 1$ and $0 < \gamma \leq \alpha$, let $(T_n^\circ, n \geq 1)$ be delabelled alpha-gamma trees, represented as discrete $\mathbb{R}$-trees with unit edge lengths. Then*

$$\frac{T_n^\circ}{n^\gamma} \xrightarrow{(d)} \mathcal{T}^{\alpha,\gamma} \qquad \text{for the Gromov-Hausdorff topology,}$$

as $n \to \infty$, where the scaling n^γ is applied to all edge lengths, and $\mathcal{T}^{\alpha,\gamma}$ is a γ-self-similar CRT whose dislocation measure is a multiple of $\nu_{\alpha,\gamma}$.

1.3.2 Chapter 3: Continuum tree asymptotics of partly exchangeable fragmentations

A version of this work will be submitted probably to the Annales de l'Institut Henri Poincaré, shortly. As we showed in the last section, both alpha trees and alpha-gamma trees are discrete fragmentation trees except that they do not possess exchangeable leaf labels. However, as the unlabelled trees have the same distribution as some discrete fragmentation trees, properly scaled alpha trees and alpha-gamma trees will converge to CRTs in distribution. In Chapter 3, one of our goals is to show that their convergence in distribution can be strengthened to convergence in probability. In fact, we develop this for a much larger class of trees.

Consider partitions induced by the labelled alpha-gamma trees. We will show in Chapter 3 that if 1 and 2 are in the same block, the probability mass of the partition will not change under any permutation of $\mathbb{N}$ that keeps 1 and 2 in the same block; similarly, if 1 and 2 are in different blocks, the probability mass of the partition will not change under any permutation of $\mathbb{N}$ that keep 1 and 2 in the different blocks. To describe such property of trees, we have to modify the fragmentation process by alternating its exchangeability.

Given two partitions $\pi^{[m]}, \pi^{[n]}$ of $[m], [n]$, where $m < n$, $m \in \mathbb{N}$, and $n \in \mathbb{N}$, we say that $\pi^{[m]}$ and $\pi^{[n]}$ are *compatible* if $\pi^{[m]}$ coincides with the restriction of $\pi^{[n]}$ to $[m]$. In this terminology, we can cut $\mathcal{P}_n, n \in \mathbb{N}$ into two subsets as follows:

$$\begin{aligned}\mathcal{P}_n^1 &:= \{\pi \in \mathcal{P}_n : \pi \text{ is compatible with } \mathbf{1}_{[2]}\};\\ \mathcal{P}_n^2 &:= \{\pi \in \mathcal{P}_n : \pi \text{ is compatible with } \mathbf{0}_{[2]}\}.\end{aligned}$$

A *partly exchangeable* measure μ on $\mathcal{P}_n$ is one for which $\mu(\pi_1) = \mu(\pi_2)$ if the block sizes of π_1 and π_2 are the same and both of $\pi_1, \pi_2 \in \mathcal{P}_n^1$ or both of $\pi_1, \pi_2 \in \mathcal{P}_n^2$; a measure μ on $\mathcal{P}$ is partly exchangeable if its restriction μ_n on $\mathcal{P}_n$ for each n is partly exchangeable. Similar to the Kingman paint-box construction for exchangeable partitions, we also have a paintbox construction for partly exchangeable partitions in Section 3.2.2.

Now we use the partly exchangeable partitions to build partly exchangeable fragmentation processes.

Definition 1.1. Let $B \subset \mathbb{N}$, and consider a $\mathcal{P}_B$-valued Markov process $(\Pi(t), t \geq 0)$. We assume that for every t, $t' \geq 0$, the distribution of $\Pi(t + t')$ given $\Pi(t) = \pi$ is the same as that of the random partition whose blocks are given by

$$\sigma_{\Pi_i(t)}(\pi_j^{(i)})\ i, j \geq 1,$$

where $\pi^{(i)}, i = 1, \ldots$ is an i.i.d. sequence of partly exchangeable partitions of $\mathbb{N}$ and $\sigma_{\Pi_i(t)} : \mathbb{N} \to \Pi_i(t)$ is a map so that $\sigma_{\Pi_i(t)}(m)$ is the mth smallest element of $\Pi_i(t)$ for $m \leq \#\Pi_i(t)$ and the largest element for $m > \#\Pi_i(t)$. Then the process Π is called a homogeneous partly exchangeable fragmentation of B.

We show that such processes are in one-to-one correspondence with σ-finite partly exchangeable measures κ on $\mathcal{P}$ called *splitting rates*, which fulfil

$$\kappa(\{1_{\mathbb{N}}\}) = 0 \text{ and } \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_{[n]} \neq 1_{[n]}\}) < \infty \text{ for every } n \geq 2. \tag{1.6}$$

Such a measure has a unique decomposition as follows.

Theorem 1.4. *Let κ be a partly exchangeable measure on $\mathcal{P}$ which fulfils* (3.6). *Then there are unique constants $c_1, c_2, c_3 \geq 0$ and unique measures ν_1, ν_2 on $\mathcal{S}^{\downarrow}$, where ν_1 fulfils*

$$\nu_1(\{(0,0,\ldots)\}) = 0 \text{ and } \nu_1(\{(1,0,\ldots)\}) = 0 \text{ and } \int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{\infty} s_i^2(1-s_i)\nu_1(d\mathbf{s}) < \infty,$$

and ν_2 fulfils

$$\nu_2(\{(1,0,\ldots)\}) = 0 \text{ and } \int_{\mathcal{S}^{\downarrow}} (1-s_1)\nu_2(d\mathbf{s}) < \infty,$$

such that for every $n \in \mathbb{N}$,

$$\begin{aligned}\kappa(\cdot) &= c_1\delta_{\omega^{(1,2)}}(\cdot) + c_2(\delta_{\epsilon^{(1)}}(\cdot) + \delta_{\epsilon^{(2)}}(\cdot)) + c_3 \sum_{i=3}^{\infty} \delta_{\epsilon^{(i)}}(\cdot) \\ &\quad + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^1)\nu_1(d\mathbf{s}) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^2)\nu_2(d\mathbf{s})\end{aligned}$$

where $\omega^{(1,2)} = (\{1,2\},\{3\},\{4\},\ldots)$, $\epsilon^{(i)} = (\{i\}, \mathbb{N} \setminus \{i\})$, δ_π stands for the Dirac point mass on π and $\kappa_{\mathbf{s}}$ is the law of Kingman's paintbox construction [31]. *In particular, κ is exchangeable if and only if $\nu_1 = \nu_2$.*

As refining partition-valued processes, partly exchangeable fragmentation processes are in natural correspondence with random labelled trees. We call such trees *discrete partly exchangeable fragmentation trees.*

Generally speaking, discrete partly exchangeable fragmentation trees can be embedded into a CRT with dislocation measure

$$\nu(d\mathbf{s}) = (\sum_{i\geq 1} s_i^2)\nu_1(d\mathbf{s}) + (s_0 + \sum_{i\geq 1} s_i(1-s_i))\nu_2(d\mathbf{s}). \tag{1.7}$$

Furthermore, the scaling limits of delabelled discrete partly exchangeable fragmentation trees will be CRTs as well.

Theorem 1.5. *Let Π be a partly exchangeable fragmentation process with $c_1 = c_2 = c_3 = 0$ and two dislocation measures ν_1 and ν_2 that fulfill the condition specified in Theorem 1.4 and further*

$$\nu_1(\{\mathbf{s} \in \mathcal{S}^{\downarrow} : \sum_{j=1}^{\infty} s_j < 1\}) = \nu_2(\{\mathbf{s} \in \mathcal{S}^{\downarrow} : \sum_{j=1}^{\infty} s_j < 1\}) = 0.$$

ν_1 *fulfills the following two further conditions: for some* $\alpha \in (0,1)$, $\rho > 0$ *and slowly varying function* $\ell(x)$ *as* $x \to \infty$,

$$\nu_1(s_1 \leq 1-\varepsilon) = \varepsilon^{-\alpha}\ell(\frac{1}{\varepsilon});$$
$$\int_{\mathcal{S}^{\downarrow}} \sum_{j\geq 2} s_j |\ln(s_j)|^{\varrho} \nu_1(d\mathbf{s}) < \infty.$$

Let $(T_n^\circ, n \geq 1)$ *be the associated sequence of unlabelled discrete partly exchangeable fragmentation trees. Then*

$$\frac{T_n^\circ}{n^\alpha \ell(n)\Gamma(1-\alpha)} \xrightarrow[n\to\infty]{(p)} \mathcal{T}_{(\alpha,\nu)},$$

for the Gromov-Hausdorff metric, where ν *is specified in* (1.7).

As an application of the above theorem, we show that alpha trees and alpha-gamma trees are partly exchangeable trees and converge to CRTs in probability.

Corollary 1.6. *Let* $(T_n, n \geq 1)$ *be a family of labelled alpha-gamma trees. Then*

(i) T_n *is a partly exchangeable discrete tree with no erosion and two dislocation measures*

$$\begin{aligned} \nu_1(d\mathbf{s}) &= \gamma \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}), \\ \nu_2(d\mathbf{s}) &= (1-\alpha)\mathrm{PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}). \end{aligned}$$

(ii)

$$\frac{T_n^\circ}{n^\gamma} \xrightarrow{(p)} \mathcal{T}_{\alpha,\nu_{\alpha,\gamma}}$$

in the Gromov-Hausdorff sense, in probability as $n \to \infty$, *where*

$$\nu_{\alpha,\gamma}(d\mathbf{s}) = \left(\gamma + (1-\alpha-\gamma)\sum_{i\neq j} s_i s_j\right) \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}).$$

When $\gamma = \alpha$, $(T_n, n \geq 1)$ *degenerate to a family of alpha trees, the two dislocation measures are binary with:*

$$\begin{aligned} \nu_1(s_1 \in dx) &= \alpha \frac{(x(1-x))^{-\alpha-1}}{\Gamma(1-\alpha)} \mathbf{1}_{\{1/2\leq x\leq 1\}} dx, \\ \nu_2(s_1 \in dx) &= (1-\alpha)\frac{(x(1-x))^{-\alpha-1}}{\Gamma(1-\alpha)} \mathbf{1}_{\{1/2\leq x\leq 1\}} dx. \end{aligned}$$

At the end of Chapter 3, we introduce another new family of Markov branching trees that we call *three-factor model* which is a subfamily of partly exchangeable fragmentation trees by setting the two dislocation measure as follows:

$$\begin{aligned}\nu_1(d\mathbf{s}) &= \lambda \mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s}),\\ \nu_2(d\mathbf{s}) &= (1-\lambda)\mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s}),\end{aligned}$$

for $\alpha \in (0,1), \theta \in [-2\alpha, \alpha]$ and $\lambda \in [0,1]$. When $\lambda = \frac{\alpha+\theta}{2\alpha+\theta-1}$, the three-factor model is the alpha-gamma model up to a linear scaling coefficient; when $\lambda = 1/2$, it is the Poisson-Dirichlet model. The three-factor model is sampling consistent if and only if it is an alpha-gamma model or Poisson-Dirichlet model. As a family of partly exchangeable fragmentation trees, the properly scaled three-factor model will nevertheless converge to a CRT in probability.

Chapter 2

A new family of Markov branching trees: the alpha-gamma model

Abstract

We introduce a simple tree growth process that gives rise to a new two-parameter family of discrete fragmentation trees that extends Ford's alpha model to multifurcating trees and includes the trees obtained by uniform sampling from Duquesne and Le Gall's stable continuum random tree. We call these new trees the alpha-gamma trees. In this chapter, we obtain their splitting rules, dislocation measures both in ranked order and in sized-biased order, and we study their limiting behaviour.

Key words: Alpha-gamma tree, splitting rule, sampling consistency, self-similar fragmentation, dislocation measure, continuum random tree, $\mathbb{R}$-tree, Markov branching model

AMS 2000 Subject Classifications: 60J80.

Submitted to EJP on July 3, 2008, final version accepted January 20, 2009.

2.1 Introduction

Markov branching trees were introduced by Aldous [4] as a class of random binary phylogenetic models and extended to the multifurcating case in [28]. Consider the space $\mathbb{T}_n$ of combinatorial trees without degree-2 vertices, one degree-1 vertex called the ROOT and exactly n further degree-1 vertices labelled by $[n] = \{1, \ldots, n\}$ and called the *leaves*; we call the other vertices *branch points*. Distributions on $\mathbb{T}_n$ of random trees T_n^* are determined by distributions of the delabelled tree T_n° on the space $\mathbb{T}_n^\circ$ of *unlabelled trees* and conditional label distributions, e.g. *exchangeable* labels. A sequence $(T_n^\circ, n \geq 1)$ of unlabelled trees has the *Markov branching property* if for all $n \geq 2$ conditionally given that the branching adjacent to the ROOT is into tree components whose numbers of leaves are $n_1, \ldots, n_k$, these tree components are independent copies of $T_{n_i}^\circ$, $1 \leq i \leq k$. The distributions of the sizes in the first branching of T_n°, $n \geq 2$, are denoted by

$$q(n_1, \ldots, n_k), \qquad n_1 \geq \ldots \geq n_k \geq 1, \quad k \geq 2: \quad n_1 + \ldots + n_k = n,$$

and referred to as the *splitting rule* of $(T_n^\circ, n \geq 1)$.

Aldous [4] studied in particular a one-parameter family ($\beta \geq -2$) that interpolates between several models known in various biology and computer science contexts (e.g. $\beta = -2$ comb, $\beta = -3/2$ uniform, $\beta = 0$ Yule) and that he called the *beta-splitting model*, he sets for $\beta > -2$:

$$q_\beta^{\text{Aldous}}(n-m, m) = \frac{1}{Z_n}\binom{n}{m} B(m+1+\beta, n-m+1+\beta), \qquad \text{for } 1 \leq m < n/2,$$
$$q_\beta^{\text{Aldous}}(n/2, n/2) = \frac{1}{2Z_n}\binom{n}{n/2} B(n/2+1+\beta, n/2+1+\beta), \qquad \text{if } n \text{ even},$$

where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the Beta function and Z_n, $n \geq 2$, are normalisation constants; this extends to $\beta = -2$ by continuity, i.e. $q_{-2}^{\text{Aldous}}(n-1, 1) = 1$, $n \geq 2$.

For exchangeably labelled Markov branching models $(T_n, n \geq 1)$ it is convenient to set

$$p(n_1, \ldots, n_k) := \frac{m_1! \ldots m_n!}{\binom{n}{n_1, \ldots, n_k}} q((n_1, \ldots, n_k)^\downarrow), \quad n_j \geq 1, j \in [k]; k \geq 2: \; n = n_1 + \ldots + n_k, \tag{2.1}$$

where $(n_1, \ldots, n_k)^\downarrow$ is the decreasing rearrangement and m_r the number of rs of the sequence $(n_1, \ldots, n_k)$. The function p is called *exchangeable partition probability function*

(EPPF) and gives the probability that the branching adjacent to the ROOT splits into tree components with label sets $\{A_1, \ldots, A_k\}$ partitioning $[n]$, with *block sizes* $n_j = \#A_j$. Note that p is invariant under permutations of its arguments. It was shown in [35] that Aldous's beta-splitting models for $\beta > -2$ are the only *binary* Markov branching models for which the EPPF is of Gibbs type

$$p^{\text{Aldous}}_{-1-\alpha}(n_1, n_2) = \frac{w_{n_1} w_{n_2}}{Z_{n_1+n_2}}, \quad n_1 \geq 1, n_2 \geq 1, \qquad \text{in particular } w_n = \frac{\Gamma(n-\alpha)}{\Gamma(1-\alpha)},$$

and that the *multifurcating* Gibbs models are an *extended* Ewens-Pitman two-parameter family of random partitions, $0 \leq \alpha \leq 1$, $\theta \geq -2\alpha$, or $-\infty \leq \alpha < 0$, $\theta = -m\alpha$ for some integer $m \geq 2$,

$$p^{\text{PD}^*}_{\alpha,\theta}(n_1, \ldots, n_k) = \frac{a_k}{Z_n} \prod_{j=1}^{k} w_{n_j}, \quad \text{where } w_n = \frac{\Gamma(n-\alpha)}{\Gamma(1-\alpha)} \text{ and } a_k = \alpha^{k-2} \frac{\Gamma(k+\theta/\alpha)}{\Gamma(2+\theta/\alpha)}, \tag{2.2}$$

boundary cases by continuity (cf. p. 20), including Aldous's binary models for $\theta = -2\alpha$. Ford [21] introduced a different one-parameter *binary* model, the *alpha model* for $0 \leq \alpha \leq 1$, using simple sequential growth rules starting from the unique elements $T_1 \in \mathbb{T}_1$ and $T_2 \in \mathbb{T}_2$:

(i)$^{\text{F}}$ given T_n for $n \geq 2$, assign a weight $1 - \alpha$ to each of the n edges adjacent to a leaf, and a weight α to each of the $n - 1$ other edges;

(ii)$^{\text{F}}$ select at random with probabilities proportional to the weights assigned by step (i)$^{\text{F}}$, an edge of T_n, say $a_n \to c_n$ directed away from the ROOT;

(iii)$^{\text{F}}$ to create T_{n+1} from T_n, replace $a_n \to c_n$ by three edges $a_n \to b_n$, $b_n \to c_n$ and $b_n \to n+1$ so that two new edges connect the two vertices a_n and c_n to a new branch point b_n and a further edge connects b_n to a new leaf labelled $n + 1$.

It was shown in [21] that these trees are Markov branching trees but that the labelling is not exchangeable. The splitting rule was calculated and shown to coincide with Aldous's beta-splitting rules if and only if $\alpha = 0$, $\alpha = 1/2$ or $\alpha = 1$, interpolating differently between Aldous's corresponding models for $\beta = 0$, $\beta = -3/2$ and $\beta = -2$. This study was taken further in [28, 40].

In this chapter, we introduce a new model by extending the simple sequential growth rules to allow *multifurcation*. Specifically, we also assign weights to *vertices* depending on two parameters $0 \leq \alpha \leq 1$ and $0 \leq \gamma \leq \alpha$ as follows, cf. Figure 2.1:

(i) given T_n for $n \geq 2$, assign a weight $1-\alpha$ to each of the n edges adjacent to a leaf, a weight γ to each of the other edges, and a weight $(k-1)\alpha - \gamma$ to each vertex of degree $k+1 \geq 3$; this distributes a total weight of $n-\alpha$;

(ii) select at random with probabilities proportional to the weights assigned by step (i),

- an edge of T_n, say $a_n \to c_n$ directed away from the ROOT,
- or, as the case may be, a vertex of T_n, say v_n;

(iii) to create T_{n+1} from T_n, do the following:

- if an edge $a_n \to c_n$ was selected, replace it by three edges $a_n \to b_n$, $b_n \to c_n$ and $b_n \to n+1$ so that two new edges connect the two vertices a_n and c_n to a new branch point b_n and a further edge connects b_n to a new leaf labelled $n+1$;
- if a vertex v_n was selected, add an edge $v_n \to n+1$ to a new leaf labelled $n+1$.

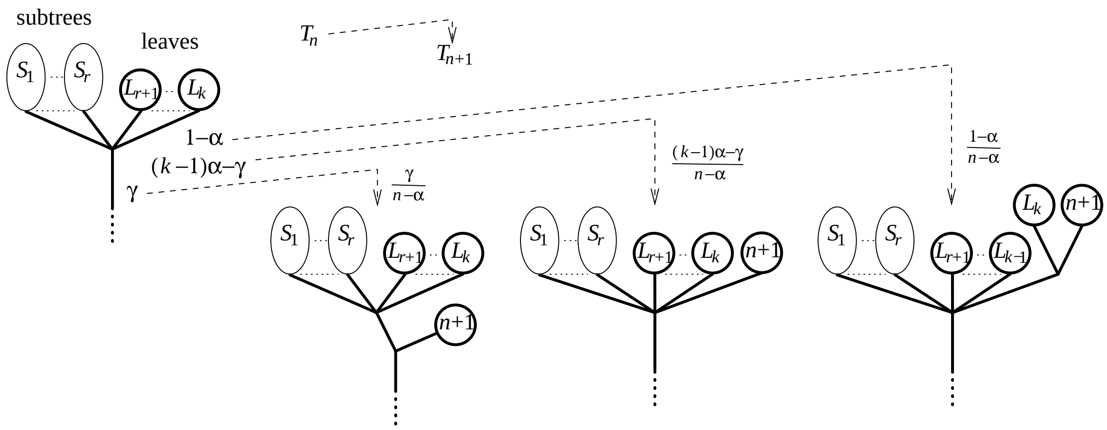


Figure 2.1: Sequential growth rule: displayed is one branch point of T_n with degree $k+1$, hence vertex weight $(k-1)\alpha-\gamma$, with $k-r$ leaves $L_{r+1}, \ldots, L_k \in [n]$ and r bigger subtrees $S_1, \ldots, S_r$ attached to it; all edges also carry weights, weight $1-\alpha$ and γ are displayed here for one leaf edge and one inner edge only; the three associated possibilities for T_{n+1} are displayed.

We call this model the *alpha-gamma model*. It contains the binary alpha model for $\gamma = \alpha$. We show here that the cases $\gamma = 1-\alpha$, $1/2 \leq \alpha \leq 1$, and $\alpha = \gamma = 0$ form the intersection with the extended Ewens-Pitman-type two-parameter family of models (2.2). The growth rules for $\gamma = 1-\alpha$, when all edges have the same weight, was studied recently by Marchal [34]. It is related to the stable tree of Duquesne and Le Gall [14], see also [36] and Section 2.3.4 here.

Proposition 2.1. *Let* $(T_n, n \geq 1)$ *be alpha-gamma trees with distributions as implied by the sequential growth rules* (i)-(iii) *for some* $0 \leq \alpha \leq 1$ *and* $0 \leq \gamma \leq \alpha$. *Then*

(a) *the delabelled trees* T_n°, $n \geq 1$, *have the Markov branching property. The splitting rules are*

$$q^{\mathrm{seq}}_{\alpha,\gamma}(n_1, \ldots, n_k) \quad \propto \quad \left(\gamma + (1-\alpha-\gamma)\frac{1}{n(n-1)}\sum_{i \neq j} n_i n_j\right) q^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1, \ldots, n_k), \quad (2.3)$$

in the case $0 \leq \alpha < 1$, *where* $q^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}$ *is the splitting rule associated via (2.1) with* $p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}$, *the Ewens-Pitman-type EPPF given in (2.2), and LHS* $\propto$ *RHS means equality up to a multiplicative constant depending on* n *and* (α, γ) *that makes the LHS a probability function;*

(b) *the labelling of* T_n *is exchangeable for all* $n \geq 1$ *if and only if* $\gamma = 1-\alpha$, $1/2 \leq \alpha \leq 1$.

The normalisation constants in (2.2) and (2.3) can be expressed in terms of Gamma functions, see Section 2.2.4. The case $\alpha = 1$ is discussed in Section 2.3.2.

For any function $(n_1, \ldots, n_k) \mapsto q(n_1, \ldots, n_k)$ that is a probability function for all fixed $n = n_1 + \ldots + n_k$, $n \geq 2$, we can construct a Markov branching model $(T_n^\circ, n \geq 1)$. A condition called *sampling consistency* [4] is to require that the tree $T_{n,-1}^\circ$ constructed from T_n° by removal of a uniformly chosen leaf (and the adjacent branch point if its degree is reduced to 2) has the same distribution as T_{n-1}°, for all $n \geq 2$. This is appealing for applications with incomplete observations. It was shown in [28] that all sampling consistent splitting rules admit an integral representation (c, ν) for an erosion coefficient $c \geq 0$ and a dislocation measure ν on $\mathcal{S}^\downarrow = \{s = (s_i)_{i\geq 1} : s_1 \geq s_2 \geq \ldots \geq 0, s_1 + s_2 + \ldots \leq 1\}$ with $\nu(\{(1, 0, 0, \ldots)\}) = 0$ and $\int_{\mathcal{S}^\downarrow}(1-s_1)\nu(ds) < \infty$ as in Bertoin's continuous-time fragmentation theory [8, 9, 10]. In the most relevant case for us when $c = 0$ and

$\nu(\{s \in \mathcal{S}^{\downarrow} : s_1 + s_2 + \ldots < 1\}) = 0$, this representation is

$$p(n_1, \ldots, n_k) = \frac{1}{\widetilde{Z}_n} \int_{\mathcal{S}^{\downarrow}} \sum_{\substack{i_1,\ldots,i_k \geq 1 \\ \text{distinct}}} \prod_{j=1}^{k} s_{i_j}^{n_j} \nu(ds), \quad n_j \geq 1, j \in [k]; k \geq 2: \ n = n_1 + \ldots + n_k, \tag{2.4}$$

where $\widetilde{Z}_n = \int_{\mathcal{S}^{\downarrow}} (1 - \sum_{i \geq 1} s_i^n) \nu(ds)$, $n \geq 2$, are the normalisation constants. The measure ν is unique up to a multiplicative constant. In particular, it can be shown [36, 29] that for the Ewens-Pitman EPPFs $p_{\alpha,\theta}^{\mathrm{PD}^*}$ we obtain $\nu = \mathrm{PD}^*_{\alpha,\theta}(ds)$ of Poisson-Dirichlet type (hence our superscript PD^* for the Ewens-Pitman type EPPF), where for $0 < \alpha < 1$ and $\theta > -2\alpha$ we can express

$$\int_{\mathcal{S}^{\downarrow}} f(s) \mathrm{PD}^*_{\alpha,\theta}(ds) = \mathbb{E}\left(\sigma_1^{-\theta} f\left(\Delta\sigma_{[0,1]}/\sigma_1\right)\right),$$

for an α-stable subordinator σ with Laplace exponent $-\log(\mathbb{E}(e^{-\lambda\sigma_1})) = \lambda^{\alpha}$ and with ranked sequence of jumps $\Delta\sigma_{[0,1]} = (\Delta\sigma_t, t \in [0,1])^{\downarrow}$. For $\alpha < 1$ and $\theta = -2\alpha$, we have

$$\int_{\mathcal{S}^{\downarrow}} f(s) \mathrm{PD}^*_{\alpha,-2\alpha}(ds) = \int_{1/2}^{1} f(x, 1-x, 0, 0, \ldots) x^{-\alpha-1}(1-x)^{-\alpha-1} dx.$$

Note that $\nu = \mathrm{PD}^*_{\alpha,\theta}$ is infinite but σ-finite with $\int_{\mathcal{S}^{\downarrow}} (1-s_1)\nu(ds) < \infty$ for $-2\alpha \leq \theta \leq -\alpha$. This is the relevant range for this chapter. For $\theta > -\alpha$, the measure $\mathrm{PD}^*_{\alpha,\theta}$ just defined is a multiple of the usual Poisson-Dirichlet probability measure $\mathrm{PD}_{\alpha,\theta}$ on $\mathcal{S}^{\downarrow}$, so for the integral representation of $p_{\alpha,\theta}^{\mathrm{PD}^*}$ we could also take $\nu = \mathrm{PD}_{\alpha,\theta}$ in this case, and this is also an appropriate choice for the two cases $\alpha = 0$ and $m \geq 3$; the case $\alpha = 1$ is degenerate $q_{\alpha,\theta}^{\mathrm{PD}^*}(1, 1, \ldots, 1) = 1$ (for all θ) and can be associated with $\nu = \mathrm{PD}^*_{1,\theta} = \delta_{(0,0,\ldots)}$, see [35].

Theorem 2.2. *The alpha-gamma-splitting rules $q_{\alpha,\gamma}^{\mathrm{seq}}$ are sampling consistent. For $0 \leq \alpha < 1$ and $0 \leq \gamma \leq \alpha$ we have no erosion ($c = 0$) and the measure in the integral representation (2.4) can be chosen as*

$$\nu_{\alpha,\gamma}(ds) = \left(\gamma + (1 - \alpha - \gamma) \sum_{i \neq j} s_i s_j\right) \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds). \tag{2.5}$$

The case $\alpha = 1$ is discussed in Section 2.3.2. We refer to Griffiths [25] who used discounting of Poisson-Dirichlet measures by quantities involving $\sum_{i \neq j} s_i s_j$ to model genic selection.

In [28], Haas and Miermont's self-similar continuum random trees (CRTs) [27] are shown to be scaling limits for a wide class of Markov branching models. See Sections 2.3.3 and 2.3.6 for details. This theory applies here to yield:

Corollary 2.3. *Let $(T_n^\circ, n \geq 1)$ be delabelled alpha-gamma trees, represented as discrete $\mathbb{R}$-trees with unit edge lengths, for some $0 < \alpha < 1$ and $0 < \gamma \leq \alpha$. Then*

$$\frac{T_n^\circ}{n^\gamma} \to \mathcal{T}^{\alpha,\gamma} \qquad \text{in distribution for the Gromov-Hausdorff topology,}$$

where the scaling n^γ is applied to all edge lengths, and $\mathcal{T}^{\alpha,\gamma}$ is a γ-self-similar CRT whose dislocation measure is a multiple of $\nu_{\alpha,\gamma}$.

We observe that every dislocation measure ν on $\mathcal{S}^\downarrow$ gives rise to a measure ν^{sb} on the space of summable sequences under which fragment sizes are in a size-biased random order, just as the $\mathrm{GEM}_{\alpha,\theta}$ distribution can be defined as the distribution of a $\mathrm{PD}_{\alpha,\theta}$ sequence re-arranged in size-biased random order [39]. We similarly define $\mathrm{GEM}^*_{\alpha,\theta}$ from $\mathrm{PD}^*_{\alpha,\theta}$. One of the advantages of size-biased versions is that, as for $\mathrm{GEM}_{\alpha,\theta}$, we can calculate marginal distributions explicitly.

Proposition 2.4. *For $0 < \alpha < 1$ and $0 \leq \gamma < \alpha$, distributions ν_k^{sb} of the first $k \geq 1$ marginals of the size-biased form $\nu_{\alpha,\gamma}^{\mathrm{sb}}$ of $\nu_{\alpha,\gamma}$ are given, for $x = (x_1, \ldots, x_k)$, by*

$$\nu_k^{\mathrm{sb}}(dx) = \left(\gamma + (1-\alpha-\gamma)\left(1 - \sum_{i=1}^k x_i^2 - \frac{1-\alpha}{1+(k-1)\alpha-\gamma}\left(1 - \sum_{i=1}^k x_i\right)^2\right)\right)\mathrm{GEM}^*_{\alpha,-\alpha-\gamma}(dx).$$

The other boundary values of parameters are trivial here – there are at most two non-zero parts.

We can investigate the convergence of Corollary 2.3 when labels are retained. Since labels are non-exchangeable, in general, it is not clear how to nicely represent a continuum tree with infinitely many labels other than by a consistent sequence $\mathcal{R}_k$ of trees with k leaves labelled $[k]$, $k \geq 1$. See however [40] for developments in the binary case $\gamma = \alpha$ on how to embed $\mathcal{R}_k$, $k \geq 1$, in a CRT $\mathcal{T}^{\alpha,\alpha}$. The following theorem extends Proposition 18 of [28] to the multifurcating case.

Theorem 2.5. *Let $(T_n, n \geq 1)$ be a sequence of trees resulting from the alpha-gamma-tree growth rules for some $0 < \alpha < 1$ and $0 < \gamma \leq \alpha$. Denote by $R(T_n, [k])$ the subtree of T_n*

spanned by the ROOT *and leaves* $[k]$, *reduced by removing degree-2 vertices, represented as discrete* $\mathbb{R}$*-tree with graph distances in* T_n *as edge lengths. Then*

$$\frac{R(T_n,[k])}{n^\gamma} \to \mathcal{R}_k \qquad \text{a.s. in the sense that all edge lengths converge,}$$

for some discrete tree $\mathcal{R}_k$ *with shape* T_k *and edge lengths specified in terms of three random variables, conditionally independent given that* T_k *has* $k+\ell$ *edges, as* $L_k W_k^\gamma D_k$ *with*

- $W_k \sim \text{beta}(k(1-\alpha)+\ell\gamma,(k-1)\alpha-\ell\gamma)$, *where* $\text{beta}(a,b)$ *is the beta distribution with density* $B(a,b)^{-1}x^{a-1}(1-x)^{b-1}1_{(0,1)}(x)$;
- L_k *with density* $\dfrac{\Gamma(1+k(1-\alpha)+\ell\gamma)}{\Gamma(1+\ell+k(1-\alpha)/\gamma)}s^{\ell+k(1-\alpha)/\gamma}g_\gamma(s)$, *where* g_γ *is the Mittag-Leffler density, the density of* $\sigma_1^{-\gamma}$ *for a subordinator* σ *with Laplace exponent* λ^γ;
- $D_k \sim \text{Dirichlet}((1-\alpha)/\gamma,\ldots,(1-\alpha)/\gamma,1,\ldots,1)$, *where* $\text{Dirichlet}(a_1,\ldots,a_m)$ *is the Dirichlet distribution on* $\Delta_m = \{(x_1,\ldots,x_m)\in[0,1]^m : x_1+\ldots+x_m = 1\}$ *with density of the first* $m-1$ *marginals proportional to* $x_1^{a_1-1}\ldots x_{m-1}^{a_{m-1}-1}(1-x_1-\ldots-x_{m-1})^{a_m-1}$; *here* D_k *contains edge length proportions, first with parameter* $(1-\alpha)/\gamma$ *for edges adjacent to leaves and then with parameter* 1 *for the other edges, each enumerated e.g. by depth first search [33] (see Section 2.4.2).*

In fact, $1-W_k$ captures the total limiting leaf proportions of subtrees that are attached on the vertices of T_k, and we can study further how this is distributed between the branch points, see Section 2.4.2.

We conclude this introduction by giving an alternative description of the alpha-gamma model obtained by adding colouring rules to the alpha model growth rules (i)$^\text{F}$-(iii)$^\text{F}$, so that in T_n^col each edge except those adjacent to leaves has either a blue or a red colour mark.

(iv)$^\text{col}$ To turn T_{n+1} into a colour-marked tree T_{n+1}^col, keep the colours of T_n^col and do the following:

- if an edge $a_n \to c_n$ adjacent to a leaf was selected, mark $a_n \to b_n$ blue;
- if a red edge $a_n \to c_n$ was selected, mark both $a_n \to b_n$ and $b_n \to c_n$ red;

- if a blue edge $a_n \to c_n$ was selected, mark $a_n \to b_n$ blue; mark $b_n \to c_n$ red with probability c and blue with probability $1-c$;

When $(T_n^{\mathrm{col}}, n \geq 1)$ has been grown according to (i)$^{\mathrm{F}}$-(iii)$^{\mathrm{F}}$ and (iv)$^{\mathrm{col}}$, crush all red edges, i.e.

(cr) identify all vertices connected via red edges, remove all red edges and remove the remaining colour marks; denote the resulting sequence of trees by $(\widetilde{T}_n, n \geq 1)$;

Proposition 2.6. *Let* $(\widetilde{T}_n, n \geq 1)$ *be a sequence of trees according to growth rules* (i)$^{\mathrm{F}}$-(iii)$^{\mathrm{F}}$,(iv)$^{\mathrm{col}}$ *and crushing rule* (cr)*. Then* $(\widetilde{T}_n, n \geq 1)$ *is a sequence of alpha-gamma trees with* $\gamma = \alpha(1-c)$*.*

The structure of this chapter is as follows. In Section 2.2 we study the discrete trees grown according to the growth rules (i)-(iii) and establish Proposition 2.6 and Proposition 2.1 as well as the sampling consistency claimed in Theorem 2.2. Section 2.3 is devoted to the limiting CRTs, we obtain the dislocation measure stated in Theorem 2.2 and deduce Corollary 2.3 and Proposition 2.4. In Section 2.4 we study the convergence of labelled trees and prove Theorem 2.5.

2.2 Sampling consistent splitting rules for the alpha-gamma trees

2.2.1 Notation and terminology of partitions and discrete fragmentation trees

For $B \subseteq \mathbb{N}$, let $\mathcal{P}_B$ be the *set of partitions of* B into disjoint non-empty subsets called *blocks*. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which supports a $\mathcal{P}_B$-valued random partition Π_B for some *finite* $B \subset \mathbb{N}$. If the probability function of Π_B only depends on its block sizes, we call it *exchangeable*. Then

$$\mathbb{P}(\Pi_B = \{A_1, \ldots, A_k\}) = p(\#A_1, \ldots, \#A_k) \qquad \text{for each partition } \pi = \{A_1, \ldots, A_k\} \in \mathcal{P}_B,$$

where $\#A_j$ denotes the block size, i.e. the number of elements of A_j. This function p is called the *exchangeable partition probability function* (EPPF) of Π_B. Alternatively, a random partition Π_B is exchangeable if its distribution is invariant under the natural action on partitions of B by the symmetric group of permutations of B.

Let $B \subseteq \mathbb{N}$, we say that a partition $\pi \in \mathcal{P}_B$ is *finer than* $\pi' \in \mathcal{P}_B$, and write $\pi \preceq \pi'$, if any block of π is included in some block of π'. This defines a partial order $\preceq$ on $\mathcal{P}_B$. A process or a sequence with values in $\mathcal{P}_B$ is called refining if it is decreasing for this partial order. Refining partition-valued processes are naturally related to trees. Suppose that B is a finite subset of $\mathbb{N}$ and $\mathbf{t}$ is a collection of subsets of B with an additional member called the ROOT such that

- we have $B \in \mathbf{t}$; we call B the *common ancestor* of $\mathbf{t}$;
- we have $\{i\} \in \mathbf{t}$ for all $i \in B$; we call $\{i\}$ a *leaf* of $\mathbf{t}$;
- for all $A \in \mathbf{t}$ and $C \in \mathbf{t}$, we have either $A \cap C = \varnothing$, or $A \subseteq C$ or $C \subseteq A$.

If $A \subset C$, then A is called a *descendant* of C, or C an *ancestor* of A. If for all $D \in \mathbf{t}$ with $A \subseteq D \subseteq C$ either $A = D$ or $D = C$, we call A a *child* of C, or C the *parent* of A and denote $C \to A$. If we equip $\mathbf{t}$ with the parent-child relation and also ROOT $\to B$, then $\mathbf{t}$ is a rooted connected acyclic graph, i.e. a combinatorial tree. We denote the space of such trees $\mathbf{t}$ by $\mathbb{T}_B$ and also $\mathbb{T}_n = \mathbb{T}_{[n]}$. For $\mathbf{t} \in \mathbb{T}_B$ and $A \in \mathbf{t}$, the rooted subtree $\mathbf{s}_A$ of $\mathbf{t}$ with common ancestor A is given by $\mathbf{s}_A = \{\text{ROOT}\} \cup \{C \in \mathbf{t} : C \subseteq A\} \in \mathbb{T}_A$. In particular, we consider the *subtrees* $\mathbf{s}_j = \mathbf{s}_{A_j}$ *of the common ancestor* B *of* $\mathbf{t}$, i.e. the subtrees whose common ancestors A_j, $j \in [k]$, are the children of B. In other words, $\mathbf{s}_1, \ldots, \mathbf{s}_k$ are the rooted connected components of $\mathbf{t} \setminus \{B\}$.

Let $(\pi(t), t \geq 0)$ be a $\mathcal{P}_B$-valued refining process for some finite $B \subset \mathbb{N}$ with $\pi(0) = \mathbf{1}_B$ and $\pi(t) = \mathbf{0}_B$ for some $t > 0$, where $\mathbf{1}_B$ is the trivial partition into a single block B and $\mathbf{0}_B$ is the partition of B into singletons. We define $\mathbf{t}_\pi = \{\text{ROOT}\} \cup \{A \subset B : A \in \pi(t) \text{ for some } t \geq 0\}$ as the associated *labelled fragmentation tree.*

Definition 2.1. Let $B \subset \mathbb{N}$ with $\#B = n$ and $\mathbf{t} \in \mathbb{T}_B$. We associate the relabelled tree

$$\mathbf{t}^\sigma = \{\text{ROOT}\} \cup \{\sigma(A) : A \in \mathbf{t}\} \in \mathbb{T}_n,$$

for any bijection $\sigma : B \to [n]$, and the combinatorial tree shape of $\mathbf{t}$ as the equivalence class

$$\mathbf{t}^\circ = \{\mathbf{t}^\sigma | \sigma : B \to [n] \text{ bijection}\} \subset \mathbb{T}_n.$$

We denote by $\mathbb{T}_n^\circ = \{\mathbf{t}^\circ : \mathbf{t} \in \mathbb{T}_n\} = \{\mathbf{t}^\circ : \mathbf{t} \in \mathbb{T}_B\}$ the collection of all tree shapes with n leaves, which we will also refer to in their own right as *unlabelled fragmentation trees.*

Note that the number of subtrees of the common ancestor of $\mathbf{t} \in \mathbb{T}_n$ and the numbers of leaves in these subtrees are invariants of the equivalence class $\mathbf{t}^\circ \subset \mathbb{T}_n$. If $\mathbf{t}^\circ \in \mathbb{T}_n^\circ$ has subtrees $\mathbf{s}_1^\circ, \ldots, \mathbf{s}_k^\circ$ with $n_1 \geq \ldots \geq n_k \geq 1$ leaves, we say that $\mathbf{t}^\circ$ is formed by *joining together* $\mathbf{s}_1^\circ, \ldots, \mathbf{s}_k^\circ$, denoted by $\mathbf{t}^\circ = \mathbf{s}_1^\circ * \ldots * \mathbf{s}_k^\circ$. We call the *composition* $(n_1, \ldots, n_k)$ of n the *first split* of $\mathbf{t}_n^\circ$.

With this notation and terminology, a sequence of random trees $T_n^\circ \in \mathbb{T}_n^\circ$, $n \geq 1$, has the *Markov branching property* if, for all $n \geq 2$, the tree T_n° has the same distribution as $S_1^\circ * \ldots * S_{K_n}^\circ$, where $N_1 \geq \ldots \geq N_{K_n} \geq 1$ form a random composition of n with $K_n \geq 2$ parts, and conditionally given $K_n = k$ and $N_j = n_j$, the trees S_j°, $j \in [k]$, are independent and distributed as $T_{n_j}^\circ$, $j \in [k]$.

2.2.2 Colour-marked trees and the proof of Proposition 2.6

The growth rules (i)$^{\mathrm{F}}$-(iii)$^{\mathrm{F}}$ construct binary combinatorial trees T_n^{bin} with vertex set

$$V = \{\textsc{root}\} \cup [n] \cup \{b_1, \ldots, b_{n-1}\}$$

and an edge set $E \subset V \times V$. We write $v \to w$ if $(v, w) \in E$. In Section 2.2.1, we identify leaf i with the set $\{i\}$ and vertex b_i with $\{j \in [n] : b_i \to \ldots \to j\}$, the edge set E then being identified by the parent-child relation. In this framework, a *colour mark* for an edge $v \to b_i$ can be assigned to the vertex b_i, so that a *coloured binary tree* as constructed in (iv)$^{\mathrm{col}}$ can be represented by

$$V^{\mathrm{col}} = \{\textsc{root}\} \cup [n] \cup \{(b_1, \chi_n(b_1)), \ldots, (b_{n-1}, \chi_n(b_{n-1}))\}$$

for some $\chi_n(b_i) \in \{0, 1\}$, $i \in [n-1]$, where 0 represents red and 1 represents blue.

Proof of Proposition 2.6. We only need to check that the growth rules (i)$^{\mathrm{F}}$-(iii)$^{\mathrm{F}}$ and (iv)$^{\mathrm{col}}$ for $(T_n^{\mathrm{col}}, n \geq 1)$ imply that the uncoloured multifurcating trees $(\widetilde{T}_n, n \geq 1)$ obtained from

$(T_n^{\rm col}, n \geq 1)$ via crushing (cr) satisfy the growth rules (i)-(iii). Let therefore $\mathbf{t}_{n+1}^{\rm col}$ be a tree with $\mathbb{P}(T_{n+1}^{\rm col} = \mathbf{t}_{n+1}^{\rm col}) > 0$. It is easily seen that there is a unique tree $\mathbf{t}_n^{\rm col}$, a unique insertion edge $a_n^{\rm col} \to c_n^{\rm col}$ in $\mathbf{t}_n^{\rm col}$ and, if any, a unique colour $\chi_{n+1}(c_n^{\rm col})$ to create $\mathbf{t}_{n+1}^{\rm col}$ from $\mathbf{t}_n^{\rm col}$. Denote the trees obtained from $\mathbf{t}_n^{\rm col}$ and $\mathbf{t}_{n+1}^{\rm col}$ via crushing (cr) by $\mathbf{t}_n$ and $\mathbf{t}_{n+1}$. If $\chi_{n+1}(c_n^{\rm col}) = 0$, denote by $k+1 \geq 3$ the degree of the branch point of $\mathbf{t}_n$ with which $c_n^{\rm col}$ is identified in the first step of the crushing (cr).

- If the insertion edge is a leaf edge ($c_n^{\rm col} = i$ for some $i \in [n]$), we obtain

$$\mathbb{P}(\widetilde{T}_{n+1} = \mathbf{t}_{n+1} | \widetilde{T}_n = \mathbf{t}_n, T_n^{\rm col} = \mathbf{t}_n^{\rm col}) = (1-\alpha)/(n-\alpha).$$

- If the insertion edge has colour blue ($\chi_n(c_n^{\rm col}) = 1$) and also $\chi_{n+1}(c_n^{\rm col}) = 1$, we obtain

$$\mathbb{P}(\widetilde{T}_{n+1} = \mathbf{t}_{n+1} | \widetilde{T}_n = \mathbf{t}_n, T_n^{\rm col} = \mathbf{t}_n^{\rm col}) = \alpha(1-c)/(n-\alpha).$$

- If the insertion edge has colour blue ($\chi_n(c_n^{\rm col}) = 1$), but $\chi_{n+1}(c_n^{\rm col}) = 0$, or if the insertion edge has colour red ($\chi_n(c_n^{\rm col}) = 0$, and then necessarily $\chi_{n+1}(c_n^{\rm col}) = 0$ also), we obtain

$$\mathbb{P}(\widetilde{T}_{n+1} = \mathbf{t}_{n+1} | \widetilde{T}_n = \mathbf{t}_n, T_n^{\rm col} = \mathbf{t}_n^{\rm col}) = (c\alpha + (k-2)\alpha)/(n-\alpha),$$

  because in addition to $a_n^{\rm col} \to c_n^{\rm col}$, there are $k-2$ other edges in $\mathbf{t}_n^{\rm col}$, where insertion and crushing also create $\mathbf{t}_{n+1}$.

Because these conditional probabilities do not depend on $\mathbf{t}_n^{\rm col}$ and have the form required, we conclude that $(\widetilde{T}_n, n \geq 1)$ obeys the growth rules (i)-(iii) with $\gamma = \alpha(1-c)$. □

2.2.3 The Chinese Restaurant Process

An important tool in this paper is the Chinese Restaurant Process (CRP), a partition-valued process $(\Pi_n, n \geq 1)$ due to Dubins and Pitman, see [39], which generates the Ewens-Pitman two-parameter family of exchangeable random partitions Π_∞ of $\mathbb{N}$. In the restaurant framework, each block of a partition is represented by a *table* and each element of a block by a *customer* at a table. The construction rules are the following. The first customer sits at the first table and the following customers will be seated at an occupied table or a new one. Given n customers at k tables with $n_j \geq 1$ customers at the jth table, customer $n+1$ will be placed at the jth table with probability $(n_j - \alpha)/(n+\theta)$, and at a new table with probability $(\theta + k\alpha)/(n+\theta)$. The parameters α and θ can be chosen as either $\alpha < 0$ and $\theta = -m\alpha$ for some $m \in \mathbb{N}$ or $0 \leq \alpha \leq 1$ and $\theta > -\alpha$. We refer to this process as the CRP with (α, θ)-*seating plan*.

In the CRP $(\Pi_n, n \geq 1)$ with $\Pi_n \in \mathcal{P}_{[n]}$, we can study the block sizes, which leads us to consider the proportion of each table relative to the total number of customers. These proportions converge to *limiting frequencies* as follows.

Lemma 2.7 (Theorem 3.2 in [39]). *For each pair of parameters (α, θ) subject to the constraints above, the Chinese restaurant with the (α, θ)-seating plan generates an exchangeable random partition Π_∞ of $\mathbb{N}$. The corresponding EPPF is*

$$p^{\mathrm{PD}}_{\alpha,\theta}(n_1, \ldots, n_k) = \frac{\alpha^{k-1}\Gamma(k+\theta/\alpha)\Gamma(1+\theta)}{\Gamma(1+\theta/\alpha)\Gamma(n+\theta)} \prod_{i=1}^{k} \frac{\Gamma(n_i - \alpha)}{\Gamma(1-\alpha)}, \quad n_i \geq 1, i \in [k]; k \geq 1 : \sum n_i = n,$$

boundary cases by continuity. The corresponding limiting frequencies of block sizes, in size-biased order of least elements, are $\mathrm{GEM}_{\alpha,\theta}$ *and can be represented as*

$$(\tilde{P}_1, \tilde{P}_2, \ldots) = (W_1, \overline{W}_1 W_2, \overline{W}_1 \overline{W}_2 W_3, \ldots)$$

where the W_i are independent, W_i has $\mathrm{beta}(1-\alpha, \theta + i\alpha)$ *distribution, and $\overline{W}_i := 1 - W_i$. The distribution of the associated ranked sequence of limiting frequencies is Poisson-Dirichlet* $\mathrm{PD}_{\alpha,\theta}$.

We also associate with the EPPF $p^{\mathrm{PD}}_{\alpha,\theta}$ the distribution $q^{\mathrm{PD}}_{\alpha,\theta}$ of block sizes in decreasing order via (2.1) and, because the Chinese restaurant EPPF is *not* the EPPF of a splitting rule leading to $k \geq 2$ block (we use notation $q^{\mathrm{PD}^*}_{\alpha,\theta}$ for the splitting rules induced by conditioning on $k \geq 2$ blocks), but can lead to a single block, we also set $q^{\mathrm{PD}}_{\alpha,\theta}(n) = p^{\mathrm{PD}}_{\alpha,\theta}(n)$.

The asymptotic properties of the number K_n of blocks of Π_n under the (α, θ)-seating plan depend on α: if $\alpha < 0$ and $\theta = -m\alpha$ for some $m \in \mathbb{N}$, then $K_n = m$ for all sufficiently large n a.s.; if $\alpha = 0$ and $\theta > 0$, then $\lim_{n\to\infty} K_n / \log n = \theta$ a.s. The most relevant case for us is $\alpha > 0$.

Lemma 2.8 (Theorem 3.8 in [39]). *For $0 < \alpha < 1$, $\theta > -\alpha$, ,*

$$\frac{K_n}{n^\alpha} \to S \qquad a.s. \text{ as } n \to \infty,$$

where S has a continuous density on $(0, \infty)$ given by

$$\frac{d}{ds}\mathbb{P}(S \in ds) = \frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)} s^{-\theta/\alpha} g_\alpha(s),$$

and g_α is the density of the Mittag-Leffler distribution with pth moment $\Gamma(p+1)/\Gamma(p\alpha+1)$.

As an extension of the CRP, Pitman and Winkel in [40] introduced the *ordered* CRP. Its seating plan is as follows. The tables are ordered from left to right. Put the second table to the right of the first with probability $\theta/(\alpha+\theta)$ and to the left with probability $\alpha/(\alpha+\theta)$. Given k tables, put the $(k+1)$st table to the right of the right-most table with probability $\theta/(k\alpha+\theta)$ and to the left of the left-most or between two adjacent tables with probability $\alpha/(k\alpha+\theta)$ each.

A composition of n is a sequence $(n_1, \ldots, n_k)$ of positive numbers with sum n. A sequence of random compositions $\mathcal{C}_n$ of n is called *regenerative* if conditionally given that the first part of $\mathcal{C}_n$ is n_1, the remaining parts of $\mathcal{C}_n$ form a composition of $n - n_1$ with the same distribution as $\mathcal{C}_{n-n_1}$. Given any decrement matrix $(q^{\mathrm{dec}}(n,m), 1 \leq m \leq n)$, there is an associated sequence $\mathcal{C}_n$ of regenerative random compositions of n defined by specifying that $q^{\mathrm{dec}}(n,\cdot)$ is the distribution of the first part of $\mathcal{C}_n$. Thus for each composition $(n_1, \ldots, n_k)$ of n,

$$\mathbb{P}(\mathcal{C}_n = (n_1, \ldots, n_k)) = q^{\mathrm{dec}}(n, n_1) q^{\mathrm{dec}}(n-n_1, n_2) \ldots q^{\mathrm{dec}}(n_{k-1}+n_k, n_{k-1}) q^{\mathrm{dec}}(n_k, n_k).$$

Lemma 2.9 (Proposition 6 (i) in [40]). *For each (α, θ) with $0 < \alpha < 1$ and $\theta \geq 0$, denote by $\mathcal{C}_n$ the composition of block sizes in the ordered Chinese restaurant partition with parameters (α, θ). Then $(\mathcal{C}_n, n \geq 1)$ is regenerative, with decrement matrix*

$$q^{\mathrm{dec}}_{\alpha,\theta}(n,m) = \binom{n}{m} \frac{(n-m)\alpha + m\theta}{n} \frac{\Gamma(m-\alpha)\Gamma(n-m+\theta)}{\Gamma(1-\alpha)\Gamma(n+\theta)} \quad (1 \leq m \leq n). \tag{2.6}$$

2.2.4 The splitting rule of alpha-gamma trees and the proof of Proposition 2.1

Proposition 2.1 claims that the unlabelled alpha-gamma trees $(T_n^\circ, n \geq 1)$ have the Markov branching property, identifies the splitting rule and studies the exchangeability of labels.

In preparation of the proof of the Markov branching property, we use CRPs to compute the probability function of the first split of T_n° in Proposition 2.10. We will then establish the Markov branching property from a spinal decomposition result (Lemma 2.11) for T_n°.

Proposition 2.10. *Let T_n° be an unlabelled alpha-gamma tree for some $0 \le \alpha < 1$ and $0 \le \gamma \le \alpha$, then the probability function of the first split of T_n° is*

$$q_{\alpha,\gamma}^{\mathrm{seq}}(n_1, \ldots, n_k) = \frac{Z_n \Gamma(1-\alpha)}{\Gamma(n-\alpha)} \left(\gamma + (1-\alpha-\gamma) \frac{1}{n(n-1)} \sum_{i \neq j} n_i n_j \right) q_{\alpha,-\alpha-\gamma}^{\mathrm{PD}^*}(n_1, \ldots, n_k),$$

$n_1 \ge \ldots \ge n_k \ge 1$, $k \ge 2$: $n_1 + \ldots + n_k = n$, where Z_n is the normalisation constant in (2.2).

In fact, we can express explicitly Z_n in (2.2) as follows (see formula (22) in [29])

$$Z_n = \frac{\Gamma(1+\theta/\alpha)}{\Gamma(1+\theta)} \left(1 - \frac{\Gamma(n-\alpha)\Gamma(1+\theta)}{\Gamma(1-\alpha)\Gamma(n+\theta)} \right)$$

in the first instance for $0 < \alpha < 1$ and $\theta > -\alpha$, and then by analytic continuation and by continuity to the full parameter range.

Proof. In the binary case $\gamma = \alpha$, the expression simplifies and the result follows from Ford [21], see also [28, Section 5.2]. For the remainder of the proof, let us consider the multifurcating case $\gamma < \alpha$. We start from the growth rules of the labelled alpha-gamma trees T_n. Consider the *spine* ROOT $\to v_1 \to \ldots \to v_{L_{n-1}} \to 1$ of T_n, and the *spinal subtrees* S_{ij}^{sp}, $1 \le i \le L_{n-1}$, $1 \le j \le K_{n,i}$, not containing 1 of the spinal vertices v_i, $i \in [L_{n-1}]$. By joining together the subtrees of the spinal vertex v_i we form the ith *spinal bush* $S_i^{\mathrm{sp}} = S_{i1}^{\mathrm{sp}} * \ldots * S_{iK_{n,i}}^{\mathrm{sp}}$. Suppose a bush S_i^{sp} consists of k subtrees with m leaves in total, then its weight will be $m - k\alpha - \gamma + k\alpha = m - \gamma$ according to growth rule (i) – recall that the total weight of the tree T_n is $n - \alpha$.

Now we consider each bush as a table, each leaf $n = 2, 3, \ldots$ as a customer, 2 being the first customer. Adding a new leaf to a bush or to an edge on the spine corresponds to adding a new customer to an existing or to a new table. The weights are such that we construct an ordered Chinese restaurant partition of $\mathbb{N} \setminus \{1\}$ with parameters $(\gamma, 1-\alpha)$.

Suppose that the first split of T_n is into tree components with numbers of leaves $n_1 \ge \ldots \ge n_k \ge 1$. Now suppose further that leaf 1 is in a subtree with n_i leaves in the first split, then the first spinal bush S_1^{sp} will have $n-n_i$ leaves. Notice that this event is equivalent to that of $n-n_i$ customers sitting at the first table with a total of $n-1$ customers present, in the terminology of the ordered CRP. According to Lemma 2.9, the probability of this is

$$\begin{aligned} q^{\mathrm{dec}}_{\gamma,1-\alpha}(n-1,n-n_i) &= \binom{n-1}{n-n_i}\frac{(n_i-1)\gamma+(n-n_i)(1-\alpha)}{n-1}\frac{\Gamma(n_i-\alpha)\Gamma(n-n_i-\gamma)}{\Gamma(n-\alpha)\Gamma(1-\gamma)} \\ &= \binom{n}{n-n_i}\left(\frac{n_i}{n}\gamma+\frac{n_i(n-n_i)}{n(n-1)}(1-\alpha-\gamma)\right)\frac{\Gamma(n_i-\alpha)\Gamma(n-n_i-\gamma)}{\Gamma(n-\alpha)\Gamma(1-\gamma)}. \end{aligned}$$

Next consider the probability that the first bush S_1^{sp} joins together subtrees with $n_1 \geq \ldots \geq n_{i-1} \geq n_{i+1} \geq \ldots n_k \geq 1$ leaves conditional on the event that leaf 1 is in a subtree with n_i leaves. The first bush has a weight of $n-n_i-\gamma$ and each subtree in it has a weight of $n_j-\alpha, j\neq i$. Consider these $k-1$ subtrees as tables and the leaves in the first bush as customers. According to the growth procedure, they form a second (unordered, this time) Chinese restaurant partition with parameters $(\alpha,-\gamma)$, whose EPPF is

$$p^{\mathrm{PD}}_{\alpha,-\gamma}(n_1,\ldots,n_{i-1},n_{i+1},\ldots,n_k)=\frac{\alpha^{k-2}\Gamma(k-1-\gamma/\alpha)\Gamma(1-\gamma)}{\Gamma(1-\gamma/\alpha)\Gamma(n-n_i-\gamma)}\prod_{j\in[k]\setminus\{i\}}\frac{\Gamma(n_j-\alpha)}{\Gamma(1-\alpha)}.$$

Let m_j be the number of js in the sequence of $(n_1,\ldots,n_k)$. Based on the exchangeability of the second Chinese restaurant partition, the probability that the first bush consists of subtrees with $n_1\geq\ldots\geq n_{i-1}\geq n_{i+1}\geq\ldots\geq n_k\geq 1$ leaves conditional on the event that leaf 1 is in one of the m_{n_i} subtrees with n_i leaves will be

$$\frac{m_{n_i}}{m_1!\ldots m_n!}\binom{n-n_i}{n_1,\ldots,n_{i-1},n_{i+1},\ldots,n_k}p^{\mathrm{PD}}_{\alpha,-\gamma}(n_1,\ldots,n_{i-1},n_{i+1},\ldots,n_k).$$

Thus the joint probability that the first split is $(n_1,\ldots,n_k)$ and that leaf 1 is in a subtree with n_i leaves is,

$$\begin{aligned} &\frac{m_{n_i}}{m_1!\ldots m_n!}\binom{n-n_i}{n_1,\ldots,n_{i-1},n_{i+1},\ldots,n_k}q^{\mathrm{dec}}_{\gamma,1-\alpha}(n-1,n-n_i)p^{\mathrm{PD}}_{\alpha,-\gamma}(n_1,\ldots,n_{i-1},n_{i+1},\ldots,n_k) \\ &= m_{n_i}\left(\frac{n_i}{n}\gamma+\frac{n_i(n-n_i)}{n(n-1)}(1-\alpha-\gamma)\right)\frac{Z_n\Gamma(1-\alpha)}{\Gamma(n-\alpha)}q^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k). \qquad (2.7) \end{aligned}$$

Hence the splitting rule will be the sum of (2.7) for all *different* n_i (not i) in $(n_1,\ldots,n_k)$, but they contain factors m_{n_i}, so we can write it as sum over $i\in[k]$:

$$\begin{aligned} q^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k) &= \left(\sum_{i=1}^k\left(\frac{n_i}{n}\gamma+\frac{n_i(n-n_i)}{n(n-1)}(1-\alpha-\gamma)\right)\right)\frac{Z_n\Gamma(1-\alpha)}{\Gamma(n-\alpha)}q^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k) \\ &= \left(\gamma+(1-\alpha-\gamma)\frac{1}{n(n-1)}\sum_{i\neq j}n_in_j\right)\frac{Z_n\Gamma(1-\alpha)}{\Gamma(n-\alpha)}q^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k). \end{aligned}$$

□

We can use the nested Chinese restaurants described in the proof to study the subtrees of the spine of T_n. We have decomposed T_n into the subtrees S_{ij}^{sp} of the spine from the ROOT to 1 and can, conversely, build T_n from S_{ij}^{sp}, for which we now introduce notation

$$T_n = \coprod_{i,j} S_{ij}^{\mathrm{sp}}.$$

We will also write $\coprod_{i,j} S_{ij}^{\circ}$ when we join together unlabelled trees S_{ij}° along a spine. The following unlabelled version of a spinal decomposition theorem will entail the Markov branching property.

Lemma 2.11 (Spinal decomposition)**.** *Let $(T_n^{\circ 1}, n \geq 1)$ be alpha-gamma trees, delabelled apart from label 1. For all $n \geq 2$, the tree $T_n^{\circ 1}$ has the same distribution as $\coprod_{i,j} S_{ij}^{\circ}$, where*

- *$\mathcal{C}_{n-1} = (N_1, \ldots, N_{L_{n-1}})$ is a regenerative composition with decrement matrix $q_{\gamma,1-\alpha}^{\mathrm{dec}}$,*
- *conditionally given $L_{n-1} = \ell$ and $N_i = n_i$, $i \in [\ell]$, the sizes $N_{i1} \geq \ldots \geq N_{iK_{n,i}} \geq 1$ form random compositions of n_i with distribution $q_{\alpha,-\gamma}^{PD}$, independently for $i \in [\ell]$,*
- *conditionally given also $K_{n,i} = k_i$ and $N_{ij} = n_{ij}$, the trees S_{ij}°, $j \in [k_i]$, $i \in [\ell]$, are independent and distributed as $T_{n_{ij}}^{\circ}$.*

Proof. For an induction on n, note that the claim is true for $n = 2$, since $T_n^{\circ 1}$ and $\coprod_{i,j} S_{ij}^{\circ}$ are deterministic for $n = 2$. Suppose then that the claim is true for some $n \geq 2$ and consider T_{n+1}°.

The growth rules (i)-(iii) of the labelled alpha-gamma tree T_n are such that, for $0 \leq \gamma < \alpha \leq 1$

- leaf $n+1$ is inserted into a new bush or any of the bushes S_i^{sp} selected according to the rules of the ordered CRP with $(\gamma, 1-\alpha)$-seating plan,
- further into a new subtree or any of the subtrees S_{ij}^{sp} of the selected bush S_i^{sp} according to the rules of a CRP with $(\alpha, -\gamma)$-seating plan,
- and further within the subtree S_{ij}^{sp} according to the weights assigned by (i) and growth rules (ii)-(iii).

These selections do not depend on T_n except via $T_n^{\circ 1}$. In fact, since labels do not feature in the growth rules (i)-(iii), they are easily seen to induce growth rules for partially labelled alpha-gamma trees $T_n^{\circ 1}$, and also for unlabelled alpha-gamma trees such as S_{ij}°.

From these observations and the induction hypothesis, we deduce the claim for T_{n+1}°. In the multifurcating case $\gamma < \alpha$, the conditional independence of compositions $(N_{i1}, \ldots, N_{iK_{n+1,i}})$, $i \in [\ell]$, given $L_{n-1} = \ell$ and $N_i = n_i$ can be checked by explicit calculation of the conditional probability function. Similarly, the conditional independence of the trees S_{ij}° follows, because conditional probabilities such as the following factorise and do not depend on (i_0, j_0):

$$\mathbb{P}\left(S_{\bullet\bullet}^{(n+1)\circ} = \mathbf{t}_{\bullet\bullet}^{(n+1)\circ} \,\middle|\, L_n = L_{n-1} = \ell, N_{\bullet}^{(n)} = n_{\bullet}, N_{\bullet\bullet}^{(n)} = n_{\bullet\bullet}, N_{i_0 j_0}^{(n+1)} = n_{i_0 j_0} + 1\right),$$

where $n_{\bullet} = (n_i, 1 \leq i \leq \ell)$ and $n_{\bullet\bullet} = (n_{ij}, 1 \leq i \leq \ell, 1 \leq j \leq k_i)$ etc.; superscripts ${}^{(n)}$ and ${}^{(n+1)}$ refer to the respective stage of the growth process. In the binary case $\gamma = \alpha$, the argument is simpler, because each spinal bush consists of a single tree. □

Proof of Proposition 2.1. (a) Firstly, the distributions of the first splits of the unlabelled alpha-gamma trees T_n° were calculated in Proposition 2.10, for $0 \leq \alpha < 1$ and $0 \leq \gamma \leq \alpha$.

Secondly, let $0 \leq \alpha \leq 1$ and $0 \leq \gamma \leq \alpha$. By the regenerative property of the spinal composition $\mathcal{C}_{n-1}$ and the conditional distribution of $T_n^{\circ 1}$ given $\mathcal{C}_{n-1}$ identified in Lemma 2.11, we obtain that given $N_1 = m$, $K_{n,1} = k_1$ and $N_{1j} = n_{1j}$, $j \in [k_1]$, the subtrees S_{1j}°, $j \in [k_1]$, are independent alpha-gamma trees distributed as $T_{n_{1j}}^{\circ}$, also independent of the remaining tree $S_{1,0} := \coprod_{i \geq 2, j} S_{ij}^{\circ}$, which, by Lemma 2.11, has the same distribution as T_{n-m}°.

This is equivalent to saying that conditionally given that the first split is into subtrees with $n_1 \geq \ldots \geq n_i \geq \ldots \geq n_k \geq 1$ leaves and that leaf 1 is in a subtree with n_i leaves, the delabelled subtrees $S_1^{\circ}, \ldots, S_k^{\circ}$ of the common ancestor are independent and distributed as $T_{n_j}^{\circ}$ respectively, $j \in [k]$. Since this conditional distribution does not depend on i, we have established the Markov branching property of T_n°.

(b) Notice that if $\gamma = 1 - \alpha$, the alpha-gamma model is the model related to stable trees, the labelling of which is known to be exchangeable, see Section 2.3.4.

On the other hand, if $\gamma \neq 1 - \alpha$, let us turn to look at the distribution of T_3.

Probability: $\frac{\gamma}{2-\alpha}$ Probability: $\frac{1-\alpha}{2-\alpha}$

We can see the probabilities of the two labelled trees in the above picture are different although they have the same unlabelled tree. So if $\gamma \neq 1-\alpha$, T_n is not exchangeable. □

2.2.5 Sampling consistency and strong sampling consistency

Recall that an unlabelled Markov branching tree T_n°, $n \geq 2$ has the property of *sampling consistency*, if when we select a leaf uniformly and delete it (together with the adjacent branch point if its degree is reduced to 2), then the new tree, denoted by $T_{n,-1}^\circ$, is distributed as T_{n-1}°. Denote by $d : \mathbb{D}_n \to \mathbb{D}_{n-1}$ the induced deletion operator on the space $\mathbb{D}_n$ of probability measures on $\mathbb{T}_n^\circ$, so that for the distribution P_n of T_n°, we define $d(P_n)$ as the distribution of $T_{n,-1}^\circ$. Sampling consistency is equivalent to $d(P_n) = P_{n-1}$. This property is also called *deletion stability* in [21].

Proposition 2.12. *The unlabelled alpha-gamma trees for $0 \leq \alpha \leq 1$ and $0 \leq \gamma \leq \alpha$ are sampling consistent.*

Proof. The sampling consistency formula (14) in [28] states that $d(P_n) = P_{n-1}$ is equivalent to

$$\begin{aligned} q(n_1, \ldots, n_k) \;\; = \;\; & \sum_{i=1}^{k} \frac{(n_i+1)(m_{n_i+1}+1)}{(n+1)m_{n_i}} q((n_1, \ldots, n_i+1, \ldots, n_k)^{\downarrow}) \\ & + \frac{m_1+1}{n+1} q(n_1, \ldots, n_k, 1) + \frac{1}{n+1} q(n,1) q(n_1, \ldots, n_k) \end{aligned} \tag{2.8}$$

for all $n_1 \geq \ldots \geq n_k \geq 1$ with $n_1 + \ldots + n_k = n \geq 2$, where m_j is the number of n_i, $i \in [k]$, that equal j, and where q is the splitting rule of $T_n^\circ \sim P_n$. In terms of EPPFs (2.1), formula (2.8) is equivalent to

$$(1 - p(n,1))\, p(n_1, \ldots, n_k) = \sum_{i=1}^{k} p(n_1, \ldots, n_i+1, \ldots, n_k) + p(n_1, \ldots, n_k, 1). \tag{2.9}$$

Now according to Proposition 2.10, the EPPF of the alpha-gamma model with $\alpha < 1$

is

$$p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k) = \frac{Z_n}{\Gamma_\alpha(n)}\left(\gamma + (1-\alpha-\gamma)\frac{1}{n(n-1)}\sum_{u\neq v} n_u n_v\right) p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k), \tag{2.10}$$

where $\Gamma_\alpha(n) = \Gamma(n-\alpha)/\Gamma(1-\alpha)$. Therefore, we can write $p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_i+1,\ldots,n_k)$ using (2.2)

$$\begin{aligned}
&\frac{Z_{n+1}}{\Gamma_\alpha(n+1)}\left(\gamma + (1-\alpha-\gamma)\frac{1}{(n+1)n}\left(\sum_{u\neq v} n_u n_v + 2(n-n_i)\right)\right)\frac{a_k}{Z_{n+1}}\left(\prod_{j:j\neq i} w_{n_j}\right) w_{n_i+1}\\
&= \left(p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k) + 2(1-\alpha-\gamma)\frac{(n-1)(n-n_i) - \sum_{u\neq v} n_u n_v}{(n+1)n(n-1)}\frac{Z_n}{\Gamma_\alpha(n)} p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k)\right)\\
&\quad\times \frac{n_i-\alpha}{n-\alpha}
\end{aligned}$$

and $p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k,1)$ as

$$\begin{aligned}
&\frac{Z_{n+1}}{\Gamma_\alpha(n+1)}\left(\gamma + (1-\alpha-\gamma)\frac{1}{(n+1)n}\left(\sum_{u\neq v} n_u n_v + 2n\right)\right)\frac{a_{k+1}}{Z_{n+1}}\left(\prod_{j=1}^{k} w_{n_j}\right) w_1\\
&= \left(p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k) + 2(1-\alpha-\gamma)\frac{(n-1)n - \sum_{u\neq v} n_u n_v}{(n+1)n(n-1)}\frac{Z_n}{\Gamma_\alpha(n)} p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k)\right)\\
&\quad\times \frac{(k-1)\alpha-\gamma}{n-\alpha}.
\end{aligned}$$

Sum over the above formulas, then the right-hand side of (2.9) is

$$\left(1 - \frac{1}{n-\alpha}\left(\gamma + \frac{2}{n+1}(1-\alpha-\gamma)\right)\right) p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k).$$

Notice that the factor is indeed $p^{\mathrm{seq}}_{\alpha,\gamma}(n,1)$. Hence, the splitting rules of the alpha-gamma model satisfy (2.9), which implies sampling consistency for $\alpha < 1$. The case $\alpha = 1$ is postponed to Section 2.3.2. □

Moreover, sampling consistency can be enhanced to *strong sampling consistency* [28] by requiring that $(T^\circ_{n-1}, T^\circ_n)$ has the same distribution as $(T^\circ_{n,-1}, T^\circ_n)$.

Proposition 2.13. *The alpha-gamma model is strongly sampling consistent if and only if* $\gamma = 1-\alpha$.

Proof. For $\gamma = 1 - \alpha$, the model is known to be strongly sampling consistent, cf. Section

2.3.4. $\mathbf{t}_3^\circ$ $\mathbf{t}_4^\circ$

If $\gamma \neq 1 - \alpha$, consider the above two deterministic unlabelled trees.

$$\mathbb{P}(T_4^\circ = \mathbf{t}_4^\circ) = q_{\alpha,\gamma}^{\mathrm{seq}}(2,1,1)q_{\alpha,\gamma}^{\mathrm{seq}}(1,1) = (\alpha-\gamma)(5-5\alpha+\gamma)/((2-\alpha)(3-\alpha)).$$

Then we delete one of the two leaves at the first branch point of $\mathbf{t}_4^\circ$ to get $\mathbf{t}_3^\circ$. Therefore

$$\mathbb{P}((T_{4,-1}^\circ, T_4^\circ) = (\mathbf{t}_3^\circ, \mathbf{t}_4^\circ)) = \frac{1}{2}\mathbb{P}(T_4^\circ = \mathbf{t}_4^\circ) = \frac{(\alpha-\gamma)(5-5\alpha+\gamma)}{2(2-\alpha)(3-\alpha)}.$$

On the other hand, if $T_3^\circ = \mathbf{t}_3^\circ$, we have to add the new leaf to the first branch point to get $\mathbf{t}_4^\circ$. Thus

$$\mathbb{P}((T_3^\circ, T_4^\circ) = (\mathbf{t}_3^\circ, \mathbf{t}_4^\circ)) = \frac{\alpha-\gamma}{3-\alpha}\mathbb{P}(T_3^\circ = \mathbf{t}_3^\circ) = \frac{(\alpha-\gamma)(2-2\alpha+\gamma)}{(2-\alpha)(3-\alpha)}.$$

It is easy to check that $\mathbb{P}((T_{4,-1}^\circ, T_4^\circ) = (\mathbf{t}_3^\circ, \mathbf{t}_4^\circ)) \neq \mathbb{P}((T_3^\circ, T_4^\circ) = (\mathbf{t}_3^\circ, \mathbf{t}_4^\circ))$ if $\gamma \neq 1 - \alpha$, which means that the alpha-gamma model is then not strongly sampling consistent. □

2.3 Dislocation measures and asymptotics of alpha-gamma trees

2.3.1 Dislocation measures associated with the alpha-gamma-splitting rules

Theorem 2.2 claims that the alpha-gamma trees are sampling consistent, which we proved in Section 2.2.5, and identifies the integral representation of the splitting rule in terms of a dislocation measure, which we will now establish.

Proof of Theorem 2.2. In the binary case $\gamma = \alpha$, the expression simplifies and the result follows from Ford [21], see also [28, Section 5.2].

In the multifurcating case $\gamma < \alpha$, we first make some rearrangement for the coefficient of the sampling consistent splitting rules of alpha-gamma trees identified in Proposition

2.10:

$$\gamma+(1-\alpha-\gamma)\frac{1}{n(n-1)}\sum_{i\neq j}n_i n_j$$
$$=\frac{(n+1-\alpha-\gamma)(n-\alpha-\gamma)}{n(n-1)}\left(\gamma+(1-\alpha-\gamma)\left(\sum_{i\neq j}A_{ij}+2\sum_{i=1}^{k}B_i+C\right)\right),$$

where

$$\begin{aligned}A_{ij} &= \frac{(n_i-\alpha)(n_j-\alpha)}{(n+1-\alpha-\gamma)(n-\alpha-\gamma)},\\ B_i &= \frac{(n_i-\alpha)((k-1)\alpha-\gamma)}{(n+1-\alpha-\gamma)(n-\alpha-\gamma)},\\ C &= \frac{((k-1)\alpha-\gamma)(k\alpha-\gamma)}{(n+1-\alpha-\gamma)(n-\alpha-\gamma)}.\end{aligned}$$

Notice that $B_i p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\dots,n_k)$ simplifies to

$$\frac{(n_i-\alpha)((k-1)\alpha-\gamma)}{(n+1-\alpha-\gamma)(n-\alpha-\gamma)}\frac{\alpha^{k-2}\Gamma(k-1-\gamma/\alpha)}{Z_n\Gamma(1-\gamma/\alpha)}\Gamma_\alpha(n_1)\dots\Gamma_\alpha(n_k)$$
$$=\frac{Z_{n+2}}{Z_n(n+1-\alpha-\gamma)(n-\alpha-\gamma)}\frac{\alpha^{k-1}\Gamma(k-\gamma/\alpha)}{Z_{n+2}\Gamma(1-\gamma/\alpha)}\Gamma_\alpha(n_1)\dots\Gamma_\alpha(n_i+1)\dots\Gamma_\alpha(n_k)$$
$$=\frac{\widetilde{Z}_{n+2}}{\widetilde{Z}_n}p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\dots,n_i+1,\dots,n_k,1),$$

where $\Gamma_\alpha(n)=\Gamma(n-\alpha)/\Gamma(1-\alpha)$ and $\widetilde{Z}_n=Z_n\alpha\Gamma(1-\gamma/\alpha)/\Gamma(n-\alpha-\gamma)$ is the normalisation constant in (2.4) for $\nu=\mathrm{PD}^*_{\alpha,-\gamma-\alpha}$. The latter can be seen from [29, Formula (17)], which yields

$$\widetilde{Z}_n=\sum_{\{A_1,\dots,A_k\}\in\mathcal{P}_{[n]}\setminus\{[n]\}}\frac{\alpha^{k-1}\Gamma(k-1-\gamma/\alpha)}{\Gamma(n-\alpha-\gamma)}\prod_{i=1}^{k}\frac{\Gamma(\#A_i-\alpha)}{\Gamma(1-\alpha)},$$

whereas Z_n is the normalisation constant in (2.2) and hence satisfies

$$Z_n=\sum_{\{A_1,\dots,A_k\}\in\mathcal{P}_{[n]}\setminus\{[n]\}}\frac{\alpha^{k-2}\Gamma(k-1-\gamma/\alpha)}{\Gamma(1-\gamma/\alpha)}\prod_{i=1}^{k}\frac{\Gamma(\#A_i-\alpha)}{\Gamma(1-\alpha)}.$$

According to (2.4),

$$p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\dots,n_k)=\frac{1}{\widetilde{Z}_n}\int_{\mathcal{S}^\downarrow}\sum_{\substack{i_1,\dots,i_k\geq 1\\ \text{distinct}}}\prod_{l=1}^{k}s_{i_l}^{n_l}\mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds).$$

Thus,

$$\sum_{i=1}^{k}B_i p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\dots,n_k) = \frac{1}{\widetilde{Z}_n}\int_{\mathcal{S}^\downarrow}\sum_{\substack{i_1,\dots,i_k\geq 1\\ \text{distinct}}}\prod_{l=1}^{k}s_{i_l}^{n_l}\left(\sum_{u\in\{i_1,\dots,i_k\},v\notin\{i_1,\dots,i_k\}}s_u s_v\right)\mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds)$$

Similarly,

$$\sum_{i\neq j} A_{ij} p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k) = \frac{1}{\widetilde{Z}_n}\int_{\mathcal{S}^\downarrow} \sum_{\substack{i_1,\ldots,i_k\geq 1\\ \text{distinct}}} \prod_{l=1}^k s_{i_l}^{n_l} \left(\sum_{u,v\in\{i_1,\ldots,i_k\}:u\neq v} s_u s_v\right) \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds)$$

$$C p^{\mathrm{PD}^*}_{\alpha,-\alpha-\gamma}(n_1,\ldots,n_k) = \frac{1}{\widetilde{Z}_n}\int_{\mathcal{S}^\downarrow} \sum_{\substack{i_1,\ldots,i_k\geq 1\\ \text{distinct}}} \prod_{l=1}^k s_{i_l}^{n_l} \left(\sum_{u,v\notin\{i_1,\ldots,i_k\}:u\neq v} s_u s_v\right) \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds),$$

Hence, the EPPF $p^{\mathrm{seq}}_{\alpha,\gamma}(n_1,\ldots,n_k)$ of the sampling consistent splitting rule takes the following form:

$$\frac{(n+1-\alpha-\gamma)(n-\alpha-\gamma)Z_n}{n(n-1)\Gamma_\alpha(n)}\left(\gamma+(1-\alpha-\gamma)\left(\sum_{i\neq j}A_{ij}+2\sum_{i=1}^k B_i + C\right)\right)$$
$$\times p^{\mathrm{PD}^*}_{\alpha,\gamma}(n_1,\ldots,n_k)$$
$$= \frac{1}{Y_n}\int_{\mathcal{S}^\downarrow} \sum_{\substack{i_1,\ldots,i_k\geq 1\\ \text{distinct}}} \prod_{l=1}^k s_{i_l}^{n_l}\left(\gamma+(1-\alpha-\gamma)\sum_{i\neq j} s_i s_j\right)\mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds), \tag{2.11}$$

where $Y_n = n(n-1)\Gamma_\alpha(n)\alpha\Gamma(1-\gamma/\alpha)/\Gamma(n+2-\alpha-\gamma)$ is the normalisation constant. Hence, we have $\nu_{\alpha,\gamma}(ds) = \left(\gamma + (1-\alpha-\gamma)\sum_{i\neq j} s_i s_j\right)\mathrm{PD^*}_{\alpha,-\alpha-\gamma}(ds)$. □

2.3.2 The alpha-gamma model when $\alpha = 1$, spine with bushes of singleton-trees

Within the discussion of the alpha-gamma model so far, we restricted to $0 \leq \alpha < 1$. In fact, we can still get some interesting results when $\alpha = 1$. The weight of each leaf edge is $1-\alpha$ in the growth procedure of the alpha-gamma model. If $\alpha = 1$, the weight of each leaf edge becomes zero, which means that the new leaf can only be inserted to internal edges or branch points. Starting from the two leaf tree, leaf 3 must be inserted into the root edge or the branch point. Similarly, any new leaf must be inserted into the spine leading from the root to the common ancestor of leaf 1 and leaf 2. Hence, the shape of the tree is just a spine with some bushes of one-leaf subtrees rooted on it. Moreover, the first split of an n-leaf tree will be into k parts $(n-k+1, 1, \ldots, 1)$ for some $2 \leq k \leq n$. The cases $\gamma = 0$ and $\gamma = 1$ lead to degenerate trees with, respectively, all leaves connected to a single branch point and all leaves connected to a spine of binary branch points (comb).

Proposition 2.14. *Consider the alpha-gamma model with $\alpha = 1$ and $0 < \gamma < 1$.*

(a) *The model is sampling consistent with splitting rules*

$$q^{\mathrm{seq}}_{1,\gamma}(n_1,\ldots,n_k) = \begin{cases} \gamma\Gamma_\gamma(k-1)/(k-1)!, & \text{if } 2\le k\le n-1 \text{ and } (n_1,\ldots,n_k)=(n-k+1,1,\ldots,1);\\ \Gamma_\gamma(n-1)/(n-2)!, & \text{if } k=n \text{ and } (n_1,\ldots,n_k)=(1,\ldots,1);\\ 0, & \text{otherwise,}\end{cases} \tag{2.12}$$

where $n_1\ge\ldots\ge n_k\ge 1$ *and* $n_1+\ldots+n_k=n$.

(b) *The dislocation measure associated with the splitting rules can be expressed as follows*

$$\int_{\mathcal{S}^\downarrow} f(s_1,s_2,\ldots)\nu_{1,\gamma}(ds) = \int_0^1 f(s_1,0,\ldots)\left(\gamma(1-s_1)^{-1-\gamma}ds_1+\delta_0(ds_1)\right). \tag{2.13}$$

In particular, it does not *satisfy* $\nu(\{s\in\mathcal{S}^\downarrow : s_1+s_2+\ldots<1\})=0$. *The erosion coefficient* c *vanishes.*

The presence of the Dirac measure δ_0 in the dislocation measure means that the associated fragmentation process exhibits dislocation events that split a fragment of positive mass into infinitesimal fragments of zero mass, often referred to as *dust* in the fragmentation literature. Dust is also produced by the other part of $\nu_{1,\gamma}$, where also a fraction s_1 of the fragment of positive mass is retained.

Proof. (a) We start from the growth procedure of the alpha-gamma model when $\alpha=1$. Consider a first split into k parts $(n-k+1,1,\ldots,1)$ for some labelled n-leaf tree for some $2\le k\le n$. Suppose $k\le n-1$ and that the branch point adjacent to the root is created when leaf l is inserted to the root edge, where $l\ge 3$. This insertion happens with probability $\gamma/(l-2)$, as $\alpha=1$. At stage l, the first split is $(l-1,1)$. In the following insertions, leaves $l+1,\ldots,n$ have to be added either to the first branch point or to the subtree with $l-1$ leaves at stage l. Hence the probability that the first split of this tree is $(n-k+1,1,\ldots,1)$ is

$$\frac{(n-k-1)!}{(n-2)!}\gamma\Gamma_\gamma(k-1),$$

which does not depend on l. Notice that the growth rules imply that if the first split of $[n]$ is $(n-k+1,1,\ldots,1)$ with $k\le n-1$, then leaves 1 and 2 will be located in the subtree with $n-k+1$ leaves. There are $\binom{n-2}{n-k-1}$ labelled trees with the above first split.

Therefore,

$$q_{1,\gamma}^{\mathrm{seq}}(n-k+1,1,\ldots,1)=\binom{n-2}{n-k-1}\frac{(n-k-1)!}{(n-2)!}\gamma\Gamma_\gamma(k-1)=\gamma\Gamma_\gamma(k-1)/(k-1)!.$$

On the other hand, for $k=n$, there is only one n-leaf labelled tree with the corresponding first split $(1,\ldots,1)$ and in this case, all leaves have to be added to the only branch point. Hence

$$q_{1,\gamma}^{\mathrm{seq}}(1,\ldots,1)=\Gamma_\gamma(n-1)/(n-2)!.$$

For sampling consistency, we check criterion (2.8), which reduces to the two formulas for $2\le k\le n-1$ and $k=n$, respectively,

$$\begin{aligned}\left(1-\frac{1}{n+1}q_{1,\gamma}^{\mathrm{seq}}(n,1)\right)q_{1,\gamma}^{\mathrm{seq}}(n-k+1,1,\ldots,1) &= \frac{n-k+2}{n+1}q_{1,\gamma}^{\mathrm{seq}}(n-k+2,1,\ldots,1)\\ &\quad+\frac{k}{n+1}q_{1,\gamma}^{\mathrm{seq}}(n-k+1,1,\ldots,1)\\ \left(1-\frac{1}{n+1}q_{1,\gamma}^{\mathrm{seq}}(n,1)\right)q_{1,\gamma}^{\mathrm{seq}}(1,\ldots,1) &= \frac{2}{n+1}q_{1,\gamma}^{\mathrm{seq}}(2,1,\ldots,1)+q_{1,\gamma}^{\mathrm{seq}}(1,\ldots,1),\end{aligned}$$

where the right-hand term on the left-hand side is a split of n (into k parts), all others are splits of $n+1$.

(b) According to (2.12), we have for $2\le k\le n-1$

$$\begin{aligned}&q_{1,\gamma}^{\mathrm{seq}}(n-k+1,1,\ldots,1)\\ &=\binom{n}{n-k+1}\frac{\Gamma_\gamma(n+1)}{n!}\gamma B(n-k+2,k-1-\gamma)\\ &=\frac{1}{Y_n}\binom{n}{n-k+1}\int_0^1 s_1^{n-k+1}(1-s_1)^{k-1}\left(\gamma(1-s_1)^{-1-\gamma}ds_1\right)\\ &=\frac{1}{Y_n}\binom{n}{n-k+1}\int_0^1 s_1^{n-k+1}(1-s_1)^{k-1}\left(\gamma(1-s_1)^{-1-\gamma}ds_1+\delta_0(ds_1)\right),\end{aligned}\tag{2.14}$$

where $Y_n=n!/\Gamma_\gamma(n+1)$. Similarly, for $k=n$,

$$\begin{aligned}q_{1,\gamma}^{\mathrm{seq}}(1,\ldots,1) &= \frac{1}{Y_n}\int_0^1\left(n(1-s_1)^{n-1}s_1+(1-s_1)^n\right)\\ &\quad\times\left((\gamma(1-s_1)^{-1-\gamma}ds_1+\delta_0(ds_1)\right).\end{aligned}\tag{2.15}$$

Formulas (2.14) and (2.15) are of the form of [28, Formula (2)], which generalises (2.4) to the case where ν does not necessarily satisfy $\nu(\{s\in\mathcal{S}^\downarrow : s_1+s_2+\ldots<1\})=0$, hence $\nu_{1,\gamma}$ is identified. □

2.3.3 Continuum random trees and self-similar trees

Let $B \subset \mathbb{N}$ finite. A *labelled tree with edge lengths* is a pair $\vartheta = (\mathbf{t}, \eta)$, where $\mathbf{t} \in \mathbb{T}_B$ is a labelled tree, $\eta = (\eta_A, A \in \mathbf{t} \setminus \{\text{ROOT}\})$ is a collection of marks, and every edge $C \to A$ of $\mathbf{t}$ is associated with mark $\eta_A \in (0, \infty)$ at vertex A, which we interpret as the *edge length* of $C \to A$. Let Θ_B be the set of such trees $(\mathbf{t}, \eta)$ with $\mathbf{t} \in \mathbb{T}_B$.

We now introduce continuum trees, following the construction by Evans et al. in [17]. A complete separable metric space (τ, d) is called an $\mathbb{R}$-tree, if it satisfies the following two conditions:

1. for all $x, y \in \tau$, there is an isometry $\varphi_{x,y} : [0, d(x, y)] \to \tau$ such that $\varphi_{x,y}(0) = x$ and $\varphi_{x,y}(d(x, y)) = y$,
2. for every injective path $c : [0, 1] \to \tau$ with $c(0) = x$ and $c(1) = y$, one has $c([0, 1]) = \varphi_{x,y}([0, d(x, y)])$.

We will consider rooted $\mathbb{R}$-trees (τ, d, ρ), where $\rho \in \tau$ is a distinguished element, the *root*. We think of the root as the lowest element of the tree.

We denote the range of $\varphi_{x,y}$ by $[[x, y]]$ and call the quantity $d(\rho, x)$ the *height* of x. We say that x is an ancestor of y whenever $x \in [[\rho, y]]$. We let $x \wedge y$ be the unique element in τ such that $[[\rho, x]] \cap [[\rho, y]] = [[\rho, x \wedge y]]$, and call it the *highest common ancestor* of x and y in τ. Denoted by $(\tau_x, d|_{\tau_x}, x)$ the set of $y \in \tau$ such that x is an ancestor of y, which is an $\mathbb{R}$-tree rooted at x that we call the *fringe subtree* of τ above x.

Two rooted $\mathbb{R}$-trees $(\tau, d, \rho), (\tau', d', \rho')$ are called *equivalent* if there is a bijective isometry between the two metric spaces that maps the root of one to the root of the other. We also denote by Θ the set of equivalence classes of compact rooted $\mathbb{R}$-trees. We define the *Gromov-Hausdorff distance* between two rooted $\mathbb{R}$-trees (or their equivalence classes) as

$$d_{\mathrm{GH}}(\tau, \tau') = \inf\{d_{\mathrm{H}}(\widetilde{\tau}, \widetilde{\tau}')\}$$

where the infimum is over all metric spaces E and isometric embeddings $\widetilde{\tau} \subset E$ of τ and $\widetilde{\tau}' \subset E$ of τ' with common root $\widetilde{\rho} \in E$; the Hausdorff distance on compact subsets of E

is denoted by d_{H}. Evans et al. [17] showed that $(\Theta, d_{\mathrm{GH}})$ is a complete separable metric space.

We call an element $x \in \tau$, $x \neq \rho$, in a rooted $\mathbb{R}$-tree τ, a *leaf* if its removal does not disconnect τ, and let $\mathcal{L}(\tau)$ be the set of leaves of τ. On the other hand, we call an element of τ a *branch point*, if it has the form $x \wedge y$ where x is neither an ancestor of y nor vice-visa. Equivalently, we can define branch points as points disconnecting τ into three or more connected components when removed. We let $\mathcal{B}(\tau)$ be the set of branch points of τ.

A *weighted* $\mathbb{R}$*-tree* (τ, μ) is called a *continuum tree* [2], if μ is a probability measure on τ and

1. μ is supported by the set $\mathcal{L}(\tau)$,
2. μ has no atom,
3. for every $x \in \tau \backslash \mathcal{L}(\tau)$, $\mu(\tau_x) > 0$.

A *continuum random tree (CRT)* is a random variable whose values are continuum trees, defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Several methods to formalize this have been developed [3, 18, 24]. For technical simplicity, we use the method of Aldous [3]. Let the space $\ell_1 = \ell_1(\mathbb{N})$ be the base space for defining CRTs. We endow the set of compact subsets of ℓ_1 with the Hausdorff metric, and the set of probability measures on ℓ_1 with any metric inducing the topology of weak convergence, so that the set of pairs (T, μ) where T is a rooted $\mathbb{R}$-tree embedded as a subset of ℓ_1 and μ is a measure on T, is endowed with the product σ-algebra.

An exchangeable $\mathcal{P}_{\mathbb{N}}$*-valued fragmentation process* $(\Pi(t), t \geq 0)$ is called *self-similar* with index $a \in \mathbb{R}$ if given $\Pi(t) = \pi = \{\pi_i, i \geq 1\}$ with asymptotic frequencies $|\pi_i| = \lim_{n\to\infty} n^{-1}\#[n] \cap \pi_j$, the random variable $\Pi(t+s)$ has the same law as the random partition whose blocks are those of $\pi_i \cap \Pi^{(i)}(|\pi_i|^a s), i \geq 1$, where $(\Pi^{(i)}, i \geq 1)$ is a sequence of i.i.d. copies of $(\Pi(t), t \geq 0)$. The process $(|\Pi(t)|^{\downarrow}, t \geq 0)$ is an $S^{\downarrow}$*-valued self-similar fragmentation process*. Bertoin [9] proved that the distribution of a $\mathcal{P}_{\mathbb{N}}$-valued self-similar fragmentation process is determined by a triple (a, c, ν), where $a \in \mathbb{R}$, $c \geq 0$ and ν is a dislocation measure on $S^{\downarrow}$. For this article, we are only interested in the case $c = 0$ and when $\nu(s_1 + s_2 + \ldots < 1) = 0$. We call (a, ν) the characteristic pair. When $a = 0$, the process $(\Pi(t), t \geq 0)$ is also called *homogeneous fragmentation process*.

A CRT $(\mathcal{T}, \mu)$ is a *self-similar CRT* with index $a = -\gamma < 0$ if for every $t \geq 0$, given $(\mu(\mathcal{T}_t^i), i \geq 1))$ where $\mathcal{T}_t^i, i \geq 1$ is the ranked order of connected components of the open set $\{x \in \tau : d(x, \rho(\tau)) > t\}$, the continuum random trees

$$\left(\mu(\mathcal{T}_t^1)^{-\gamma}\mathcal{T}_t^1, \frac{\mu(\cdot \cap \mathcal{T}_t^1)}{\mu(\mathcal{T}_t^1)}\right), \left(\mu(\mathcal{T}_t^2)^{-\gamma}\mathcal{T}_t^2, \frac{\mu(\cdot \cap \mathcal{T}_t^2)}{\mu(\mathcal{T}_t^2)}\right), \ldots$$

are i.i.d copies of $(\mathcal{T}, \mu)$, where $\mu(\mathcal{T}_t^i)^{-\gamma}\mathcal{T}_t^i$ is the tree that has the same set of points as $\mathcal{T}_t^i$, but whose distance function is divided by $\mu(\mathcal{T}_t^i)^{\gamma}$. Haas and Miermont in [27] have shown that there exists a self-similar continuum random tree $\mathcal{T}_{(\gamma,\nu)}$ characterized by such a pair (γ, ν), which can be constructed from a self-similar fragmentation process with characteristic pair (γ, ν).

2.3.4 The alpha-gamma model when $\gamma = 1 - \alpha$, sampling from the stable CRT

Let $(\mathcal{T}, \rho, \mu)$ be the stable tree of Duquesne and Le Gall [14]. The distribution on Θ of any CRT is determined by its so-called finite-dimensional marginals: the distributions of $\mathcal{R}_k$, $k \geq 1$, the subtrees $\mathcal{R}_k \subset \mathcal{T}$ defined as the discrete trees with edge lengths spanned by $\rho, U_1, \ldots, U_k$, where given $(\mathcal{T}, \mu)$, the sequence $U_i \in \mathcal{T}$, $i \geq 1$, of leaves is sampled independently from μ. See also [37, 15, 28, 29, 34] for various approaches to stable trees. Let us denote the discrete tree without edge lengths associated with $\mathcal{R}_k$ by T_k and note the Markov branching structure.

Lemma 2.15 (Corollary 22 in [28]). *Let $1/\alpha \in (1, 2]$. The trees T_n, $n \geq 1$, sampled from the $(1/\alpha)$-stable CRT are Markov branching trees, whose splitting rule has EPPF*

$$p_{1/\alpha}^{\text{stable}}(n_1, \ldots, n_k) = \frac{\alpha^{k-2}\Gamma(k - 1/\alpha)\Gamma(2 - \alpha)}{\Gamma(2 - 1/\alpha)\Gamma(n - \alpha)} \prod_{j=1}^{k} \frac{\Gamma(n_j - \alpha)}{\Gamma(1 - \alpha)}$$

for any $k \geq 2$, $n_1 \geq 1, \ldots, n_k \geq 1$, $n = n_1 + \ldots + n_k$.

We recognise $p_{1/\alpha}^{\text{stable}} = p_{\alpha,-1}^{\text{PD}^*}$ in (2.2), and by Proposition 2.1, we have $p_{\alpha,-1}^{\text{PD}^*} = p_{\alpha,1-\alpha}^{\text{seq}}$. The full distribution of $\mathcal{R}_n$, $n \geq 1$, is displayed in Theorem 2.5, which in the stable case was first obtained by [14, Theorem 3.3.3]. Furthermore, it can be shown that the trees

$(T_k, k \geq 1)$ obtained by sampling from the stable CRT follow the alpha-gamma growth rules for $\gamma = 1 - \alpha$, see e.g. Marchal [34]. This observation yields the following corollary:

Corollary 2.16. *The alpha-gamma trees with $\gamma = 1 - \alpha$ are strongly sampling consistent and exchangeable.*

Proof. These properties follow from the representation by sampling from the stable CRT, particularly the exchangeability of the sequence U_i, $i \geq 1$. Specifically, since U_i, $i \geq 1$, are conditionally independent and identically distributed given $(\mathcal{T}, \mu)$, they are exchangeable. If we denote by $\mathcal{L}_{n,-1}$ the random set of leaves $\mathcal{L}_n = \{U_1, \ldots, U_n\}$ with a uniformly chosen member removed, then $(\mathcal{L}_{n,-1}, \mathcal{L}_n)$ has the same conditional distribution as $(\mathcal{L}_{n-1}, \mathcal{L}_n)$. Hence the pairs of (unlabelled) tree shapes spanned by ρ and these sets of leaves have the same distribution – this is strong sampling consistency as defined before Proposition 2.13. □

2.3.5 Dislocation measures in size-biased order

In actual calculations, we may find that the splitting rules in Proposition 2.1 are quite difficult and the corresponding dislocation measure ν is always inexplicit, which leads us to transform ν to a more explicit form. For simplicity, let us assume that ν satisfies $\nu(\{s \in \mathcal{S}^{\downarrow} : s_1 + s_2 + \ldots < 1\}) = 0$. The method proposed here is to change the space $\mathcal{S}^{\downarrow}$ into the space $[0,1]^{\mathbb{N}}$ and to rearrange the elements $s \in \mathcal{S}^{\downarrow}$ under ν into the *size-biased random order* that places s_{i_1} first with probability s_{i_1} (its *size*) and, successively, the remaining ones with probabilities $s_{i_j}/(1 - s_{i_1} - \ldots - s_{i_{j-1}})$ proportional to their sizes s_{i_j} into the following positions, $j \geq 2$.

Definition 2.2. We call a measure ν^{sb} on the space $[0,1]^{\mathbb{N}}$ the size-biased dislocation measure associated with dislocation measure ν, if for any subset $A_1 \times A_2 \times \ldots \times A_k \times [0,1]^{\mathbb{N}}$ of $[0,1]^{\mathbb{N}}$,

$$\begin{aligned} &\nu^{\mathrm{sb}}(A_1 \times A_2 \times \ldots \times A_k \, [0,1]^{\mathbb{N}}) \\ = &\sum_{\substack{i_1,\ldots,i_k \geq 1 \\ \text{distinct}}} \int_{\{s \in \mathcal{S}^{\downarrow} : s_{i_1} \in A_1, \ldots, s_{i_k} \in A_k\}} \frac{s_{i_1} \ldots s_{i_k}}{\prod_{j=1}^{k-1}(1 - \sum_{l=1}^{j} s_{i_l})} \nu(ds) \end{aligned} \tag{2.16}$$

for any $k \in \mathbb{N}$, where ν is a dislocation measure on $\mathcal{S}^{\downarrow}$ satisfying $\nu(s \in \mathcal{S}^{\downarrow} : s_1 + s_2 + \ldots < 1) = 0$. We also denote by $\nu_k^{\mathrm{sb}}(A_1 \times A_2 \times \ldots \times A_k) = \nu^{\mathrm{sb}}(A_1 \times A_2 \times \ldots \times A_k \times [0,1]^{\mathbb{N}})$ the distribution of the first k marginals.

The sum in (2.16) is over all possible rank sequences $(i_1, \ldots, i_k)$ to determine the first k entries of the size-biased vector. The integral in (2.16) is over the decreasing sequences that have the jth entry of the re-ordered vector fall into A_j, $j \in [k]$. Notice that the support of such a size-biased dislocation measure ν^{sb} is a subset of $\mathcal{S}^{\mathrm{sb}} := \{s \in [0,1]^{\mathbb{N}} : \sum_{i=1}^{\infty} s_i = 1\}$. If we denote by $s^{\downarrow}$ the sequence $s \in \mathcal{S}^{\mathrm{sb}}$ rearranged into ranked order, taking (2.16) into formula (2.4), we obtain

Proposition 2.17. *The EPPF associated with a dislocation measure ν can be represented as:*

$$p(n_1, \ldots, n_k) = \frac{1}{\widetilde{Z}_n} \int_{[0,1]^k} x_1^{n_1-1} \ldots x_k^{n_k-1} \prod_{j=1}^{k-1} (1 - \sum_{l=1}^{j} x_l) \nu_k^{\mathrm{sb}}(dx),$$

where ν^{sb} is the size-biased dislocation measure associated with ν, where $n_1 \geq \ldots \geq n_k \geq 1, k \geq 2, n = n_1 + \ldots + n_k$ and $x = (x_1, \ldots, x_k)$.

Now turn to see the case of Poisson-Dirichlet measures $\mathrm{PD}^*_{\alpha,\theta}$ to then study $\nu^{\mathrm{sb}}_{\alpha,\gamma}$.

Lemma 2.18. *If we define $\mathrm{GEM}^*_{\alpha,\theta}$ as the size-biased dislocation measure associated with $\mathrm{PD}^*_{\alpha,\theta}$ for $0 < \alpha < 1$ and $\theta > -2\alpha$, then the first k marginals have joint density*

$$\begin{aligned} \mathrm{gem}^*_{\alpha,\theta}(x_1, \ldots, x_k) &= \frac{\alpha \Gamma(2 + \theta/\alpha)}{\Gamma(1-\alpha)\Gamma(\theta + \alpha + 1) \prod_{j=2}^{k} B(1-\alpha, \theta + j\alpha)} \\ &\quad \times \frac{(1 - \sum_{i=1}^{k} x_i)^{\theta + k\alpha} \prod_{j=1}^{k} x_j^{-\alpha}}{\prod_{j=1}^{k} (1 - \sum_{i=1}^{j} x_i)}, \end{aligned} \tag{2.17}$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$ is the beta function.

This is a simple σ-finite extension of the GEM distribution and (2.17) can be derived analogously to Lemma 2.7. Applying Proposition 2.17, we can get an explicit form of the size-biased dislocation measure associated with the alpha-gamma model.

Proof of Proposition 2.4. We start our proof from the dislocation measure associated with the alpha-gamma model. According to (2.5) and (2.16), the first k marginals of $\nu^{\mathrm{sb}}_{\alpha,\gamma}$ are given by

$$
\begin{aligned}
\nu_k^{\mathrm{sb}}(A_1 \times \ldots \times A_k) &= \sum_{\substack{i_1,\ldots,i_k \geq 1 \\ \text{distinct}}} \int_{\{s \in \mathcal{S}^{\downarrow}: s_{i_j} \in A_j, j \in [k]\}} \frac{s_{i_1} \ldots s_{i_k}}{\prod_{j=1}^{k-1}(1 - \sum_{l=1}^{j} s_{i_l})} \\
&\quad \times \left(\gamma + (1-\alpha-\gamma)\sum_{i \neq j} s_i s_j\right) \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds) \\
&= \gamma D + (1-\alpha-\gamma)(E - F),
\end{aligned}
$$

where

$$
\begin{aligned}
D &= \sum_{\substack{i_1,\ldots,i_k \geq 1 \\ \text{distinct}}} \int_{\{s \in \mathcal{S}^{\downarrow}:\ s_{i_1} \in A_1,\ldots,s_{i_k} \in A_k\}} \frac{s_{i_1} \ldots s_{i_k}}{\prod_{j=1}^{k-1}(1 - \sum_{l=1}^{j} s_{i_l})} \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds) \\
&= \mathrm{GEM}^*_{\alpha,-\alpha-\gamma}(A_1 \times \ldots \times A_k), \\
E &= \sum_{\substack{i_1,\ldots,i_k \geq 1 \\ \text{distinct}}} \int_{\{s \in \mathcal{S}^{\downarrow}:\ s_{i_1} \in A_1,\ldots,s_{i_k} \in A_k\}} \left(1 - \sum_{u=1}^{k} s_{i_u}^2\right) \frac{s_{i_1} \ldots s_{i_k}}{\prod_{j=1}^{k-1}(1 - \sum_{l=1}^{j} s_{i_l})} \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds) \\
&= \int_{A_1 \times \ldots \times A_k} \left(1 - \sum_{i=1}^{k} x_i^2\right) \mathrm{GEM}^*_{\alpha,-\alpha-\gamma}(dx) \\
F &= \sum_{\substack{i_1,\ldots,i_k \geq 1 \\ \text{distinct}}} \int_{\{s \in \mathcal{S}^{\downarrow}:\ s_{i_1} \in A_1,\ldots,s_{i_k} \in A_k\}} \left(\sum_{v \notin \{i_1,\ldots,i_k\}} s_v^2\right) \frac{s_{i_1} \ldots s_{i_k}}{\prod_{j=1}^{k-1}(1 - \sum_{l=1}^{j} s_{i_l})} \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds) \\
&= \sum_{\substack{i_1,\ldots,i_{k+1} \geq 1 \\ \text{distinct}}} \int_{\{s \in \mathcal{S}^{\downarrow}:\ s_{i_1} \in A_1,\ldots,s_{i_k} \in A_k\}} \frac{s_{i_{k+1}}^2}{1 - \sum_{l=1}^{k} s_{i_l}} \frac{s_{i_1} \ldots s_{i_{k+1}}}{\prod_{j=1}^{k}(1 - \sum_{l=1}^{j} s_{i_l})} \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(ds) \\
&= \int_{A_1 \times \ldots \times A_k \times [0,1]} \frac{x_{k+1}}{1 - \sum_{i=1}^{k} x_i} \mathrm{GEM}^*_{\alpha,-\alpha-\gamma}(d(x_1,\ldots,x_{k+1})).
\end{aligned}
$$

Applying (2.17) to F (and setting $\theta = -\alpha - \gamma$), then integrating out x_{k+1}, we get:

$$
F = \int_{A_1 \times \ldots \times A_k} \frac{1-\alpha}{1 + (k-1)\alpha - \gamma} \left(1 - \sum_{i=1}^{k} x_i\right)^2 \mathrm{GEM}^*_{\alpha,-\alpha-\gamma}(dx).
$$

Summing over D, E, F, we obtain the formula stated in Proposition 2.4. □

As the model related to stable trees is a special case of the alpha-gamma model when $\gamma = 1 - \alpha$, the sized-biased dislocation measure for it is

$$
\nu^{\mathrm{sb}}_{\alpha,1-\alpha}(ds) = \gamma \mathrm{GEM}^*_{\alpha,-1}(ds).
$$

For general (α, γ), the explicit form of the dislocation measure in size-biased order, specifically the density $g_{\alpha,\gamma}$ of the first marginal of $\nu^{\mathrm{sb}}_{\alpha,\gamma}$, yields immediately the tagged particle [8] Lévy measure associated with a fragmentation process with alpha-gamma dislocation measure.

Corollary 2.19. *Let* $(\Pi^{\alpha,\gamma}(t), t \geq 0)$ *be an exchangeable homogeneous* $\mathcal{P}_{\mathbb{N}}$*-valued fragmentation process with dislocation measure* $\nu_{\alpha,\gamma}$ *for some* $0 < \alpha < 1$ *and* $0 \leq \gamma < \alpha$*. Then, for the size* $|\Pi^{\alpha,\gamma}_{(i)}(t)|$ *of the block containing* $i \geq 1$*, the process* $\xi_{(i)}(t) = -\log|\Pi^{\alpha,\gamma}_{(i)}(t)|$*,* $t \geq 0$*, is a pure-jump subordinator with Lévy measure*

$$\begin{aligned}\Lambda_{\alpha,\gamma}(dx) &= \mathrm{e}^{-x} g_{\alpha,\gamma}(\mathrm{e}^{-x})dx \\ &= \frac{\alpha\Gamma(1-\gamma/\alpha)}{\Gamma(1-\alpha)\Gamma(1-\gamma)}\left(1-\mathrm{e}^{-x}\right)^{-1-\gamma}\left(\mathrm{e}^{-x}\right)^{1-\alpha} \\ &\quad \times \left(\gamma + (1-\alpha-\gamma)\left(2\mathrm{e}^{-x}(1-\mathrm{e}^{-x}) + \frac{\alpha-\gamma}{1-\gamma}(1-\mathrm{e}^{-x})^2\right)\right) dx.\end{aligned}$$

A similar result holds for the binary case $\gamma = \alpha$, see [40, Equation (10), also Section 4.2].

2.3.6 Convergence of alpha-gamma trees to self-similar CRTs

In this subsection, we will prove that the delabelled alpha-gamma trees T_n°, represented as $\mathbb{R}$-trees with unit edge lengths and suitably rescaled converge to fragmentation CRTs $\mathcal{T}^{\alpha,\gamma}$ as n tends to infinity, where $\mathcal{T}^{\alpha,\gamma}$ is a γ-selfsimilar fragmentation CRT whose dislocation measure is a multiple of $\nu_{\alpha,\gamma}$, as in Corollary 2.3, cf. Section 2.3.3.

Lemma 2.20. *If* $(\widetilde{T}_n^\circ)_{n\geq 1}$ *are strongly sampling consistent discrete fragmentation trees in the sense that* $(T_{n-1}^\circ, T_n^\circ)$ *has the same distribution as* $(T_{n,-1}^\circ, T_n^\circ)$ *for all* $n \geq 2$*, cf. Section 2.2.5, associated with dislocation measure* $\nu_{\alpha,\gamma}$ *for some* $0 < \alpha < 1$ *and* $0 < \gamma \leq \alpha$*, then*

$$\frac{\widetilde{T}_n^\circ}{n^\gamma} \to \mathcal{T}^{\alpha,\gamma}$$

in the Gromov-Hausdorff sense, in probability as $n \to \infty$*.*

Proof. For $\gamma = \alpha$ this is [28, Corollary 17]. For $\gamma < \alpha$, we apply Theorem 2 in [28], which says that a strongly sampling consistent family of discrete fragmentation trees $(\widetilde{T}_n^\circ)_{n\geq 1}$ converges in probability to a CRT

$$\frac{\widetilde{T}_n^\circ}{n^{\gamma_\nu}\ell(n)\Gamma(1-\gamma_\nu)} \to \mathcal{T}_{(\gamma_\nu,\nu)}$$

for the Gromov-Hausdorff metric if the dislocation measure ν satisfies following two conditions:

$$\nu(s_1 \leq 1-\varepsilon) = \varepsilon^{-\gamma_\nu}\ell(1/\varepsilon); \tag{2.18}$$

$$\int_{\mathcal{S}^\downarrow}\sum_{i\geq 2} s_i|\ln s_i|^\rho \nu(ds) < \infty, \tag{2.19}$$

where ρ is some positive real number, $\gamma_\nu \in (0,1)$, and $x \mapsto \ell(x)$ is slowly varying as $x \to \infty$.

By virtue of (19) in [28], we know that (2.18) is equivalent to

$$\Lambda([x,\infty)) = x^{-\gamma_\nu}\ell^*(1/x), \qquad \text{as } x \downarrow 0,$$

where Λ is the Lévy measure of the tagged particle subordinator as in Corollary 2.19. Specifically, the slowly varying functions ℓ and ℓ^* are asymptotically equivalent since

$$\begin{aligned}\Lambda([x,\infty)) &= \int_{\mathcal{S}^\downarrow}(1-s_1)\nu(ds) + \nu(s_1 \leq e^{-x}) \\ &= \int_{\mathcal{S}^\downarrow}(1-s_1)\nu(ds) + (1-e^{-x})^{-\gamma_\nu}\ell\left(\frac{1}{1-e^{-x}}\right)\end{aligned}$$

implies that

$$\frac{\ell^*(1/x)}{\ell(1/x)} = \frac{\Lambda([x,\infty)}{x^{-\gamma_\nu}\ell(1/x)} \sim \left(\frac{1-e^{-x}}{x}\right)^{-\gamma_\nu}\frac{\ell(x+\frac{1-x+xe^{-x}}{1-e^{-x}})}{\ell(x)} \to 1.$$

So, the dislocation measure $\nu_{\alpha,\gamma}$ satisfies (2.18) with $\ell(x) \to \alpha\Gamma(1-\gamma/\alpha)/\Gamma(1-\alpha)\Gamma(1-\gamma)$ and $\gamma_{\nu_{\alpha,\gamma}} = \gamma$. Notice that

$$\int_{\mathcal{S}^\downarrow}\sum_{i\geq 2} s_i|\ln s_i|^\rho \nu_{\alpha,\gamma}(ds) \leq \int_0^\infty x^\rho \Lambda_{\alpha,\gamma}(dx).$$

As $x \to \infty$, $\Lambda_{\alpha,\gamma}$ decays exponentially, so $\nu_{\alpha,\gamma}$ satisfies condition (2.19). This completes the proof. □

Proof of Corollary 2.3. The splitting rules of T_n° are the same as those of $\widetilde{T}_n^\circ$, which leads to the identity in distribution for the whole trees. The preceding lemma yields convergence in distribution for T_n°. □

2.4 Limiting results for labelled alpha-gamma trees

In this section we suppose $0 < \alpha < 1$ and $0 < \gamma \leq \alpha$. In the boundary case $\gamma = 0$ trees grow logarithmically and do not possess non-degenerate scaling limits; for $\alpha = 1$ the study in Section 2.3.2 can be refined to give results analogous to the ones below, but with degenerate tree shapes.

2.4.1 The scaling limits of reduced alpha-gamma trees

For τ a rooted $\mathbb{R}$-tree and $x_1, \ldots, x_n \in \tau$, we call $R(\tau, x_1, \ldots, x_n) = \bigcup_{i=1}^{n}[[\rho, x_i]]$ the reduced subtree associated with $\tau, x_1, \ldots, x_n$, where ρ is the root of τ.

As a fragmentation CRT, the limiting CRT $(\mathcal{T}^{\alpha,\gamma}, \mu)$ is naturally equipped with a mass measure μ and contains subtrees $\widetilde{\mathcal{R}}_k, k \geq 1$ spanned by k leaves chosen independently according to μ. Denote the discrete tree without edge lengths by $\tilde{T}_n$ – it has *exchangeable* leaf labels. Then $\widetilde{\mathcal{R}}_n$ is the almost sure scaling limit of the reduced trees $R(\widetilde{T}_n, [k])$, by Proposition 7 in [28].

On the other hand, if we denote by T_n the (non-exchangeably) labelled trees obtained via the alpha-gamma growth rules, the above result will not apply, but, similarly to the result for the alpha model shown in Proposition 18 in [28], we can still establish a.s. convergence of the reduced subtrees in the alpha-gamma model as stated in Theorem 2.5, and the convergence result can be strengthened as follows.

Proposition 2.21. *In the setting of Theorem 2.5*

$$(n^{-\gamma} R(T_n, [k]), n^{-1} W_{n,k}) \to (\mathcal{R}_k, W_k) \qquad \text{a.s. as } n \to \infty,$$

in the sense of Gromov-Hausdorff convergence, where $W_{n,k}$ *is the total number of leaves in subtrees of* $T_n \backslash R(T_n, [k])$ *that are linked to the present branch points of* $R(T_n, [k])$.

Proof of Theorem 2.5 and Proposition 2.21. Actually, the labelled discrete tree $R(T_n, [k])$ with edge lengths removed is T_k for all n. Thus, it suffices to prove the convergence of its total length and of its edge length proportions.

Let us consider a first urn model, cf. [19], where at level n the urn contains a black ball for each leaf in a subtree that is directly connected to a branch point of $R(T_n, [k])$,

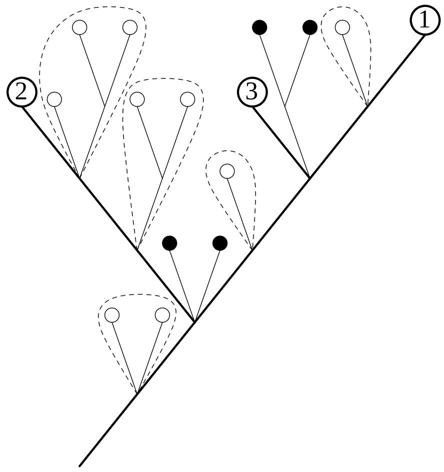


Figure 2.2: We display an example of $S(T_{16},[3])$, seven skeletal subtrees in five skeletal bushes (within the dashed lines, white leaves) and further subtrees in the branch points of $S(T_{16},[3])$ (with black leaves).

and a white ball for each leaf in one of the remaining subtrees connected to the edges of $R(T_n,[k])$. Suppose that the balls are labelled like the leaves they represent. If the urn then contains $W_{n,k}=m$ black balls and $n-k-m$ white balls, the induced partition of $\{k+1,\dots,n\}$ has probability function

$$p(m,n-k-m)=\frac{\Gamma(n-m-\alpha-w)\Gamma(w+m)\Gamma(k-\alpha)}{\Gamma(k-\alpha-w)\Gamma(w)\Gamma(n-\alpha)}=\frac{B(n-m-\alpha-w,w+m)}{B(k-\alpha-w,w)}$$

where $w=k(1-\alpha)+\ell\gamma$ is the total weight on the k leaf edges and ℓ other edges of T_k. As $n\to\infty$, the urn is such that $W_{n,k}/n\to W_k$ a.s., where $W_k\sim\text{beta}((k-1)\alpha-l\gamma,k(1-\alpha)+l\gamma)$.

We will partition the white balls further. Extending the notions of spine, spinal subtrees and spinal bushes from Proposition 2.10 ($k=1$), we call, for $k\geq 2$, *skeleton* the tree $S(T_n,[k])$ of T_n spanned by the ROOT and leaves $[k]$ including the degree-2 vertices, for each such degree-2 vertex $v\in S(T_n,[k])$, we consider the skeletal subtrees S_{vj}^{sk} that we join together into a *skeletal bush* S_v^{sk}, cf. Figure 2.2. Note that the total length $L_k^{(n)}$ of the skeleton $S(T_n,[k])$ will increase by 1 if leaf $n+1$ in T_{n+1} is added to any of the edges of $S(T_n,[k])$; also, $L_k^{(n)}$ is equal to the number of skeletal bushes (denoted by $\overline{K}_n$) plus the original total length $k+\ell$ of T_k. Hence, as $n\to\infty$

$$\frac{L_k^{(n)}}{n^\gamma}\sim\frac{\overline{K}_n}{W_{n,k}^\gamma}\left(\frac{W_{n,k}}{n}\right)^\gamma\sim\frac{\overline{K}_n}{W_{n,k}^\gamma}W_k^\gamma. \tag{2.20}$$

The partition of leaves (associated with white balls), where each skeletal bush gives rise to a block, follows the dynamics of a Chinese Restaurant Process with (γ, w)-seating plan: given that the number of white balls in the first urn is m and that there are $K_m := \overline{K}_n$ skeletal bushes on the edges of $S(T_n, [k])$ with n_i leaves on the ith bush, the next leaf associated with a white ball will be inserted into any particular bush with n_i leaves with probability proportional to $n_i - \gamma$ and will create a new bush with probability proportional to $w + K_m\gamma$. Hence, the EPPF of this partition of the white balls is

$$p_{\gamma,w}(n_1, \ldots, n_{K_m}) = \frac{\gamma^{K_m-1}\Gamma(K_m + w/\gamma)\Gamma(1+w)}{\Gamma(1+w/\gamma)\Gamma(m+w)} \prod_{i=1}^{K_m} \Gamma_\gamma(n_i).$$

Applying Lemma 2.8 in connection with (2.20), we get the probability density of L_k/W_k^γ as specified.

Finally, we set up another urn model that is updated whenever a new skeletal bush is created. This model records the edge lengths of $R(T_n, [k])$. The alpha-gamma growth rules assign weights $1 - \alpha + (n_i - 1)\gamma$ to leaf edges of $R(T_n, [k])$ and weights $n_i\gamma$ to other edges of length n_i, and each new skeletal bush makes one of the weights increase by γ. Hence, the conditional probability that the length of each edge is $(n_1, \ldots, n_{k+l})$ at stage n is that

$$\frac{\prod_{i=1}^{k} \Gamma_{1-\alpha}(n_i) \prod_{i=k+1}^{k+\ell} \Gamma_\gamma(n_i)}{\Gamma_{k\alpha+\ell\gamma}(n-k)}.$$

Then $D_k^{(n)}$ converge a.s. to the Dirichlet limit as specified. Moreover, $L_k^{(n)}D_k^{(n)} \to L_kD_k$ a.s., and it is easily seen that this implies convergence in the Gromov-Hausdorff sense.

The above argument actually gives us the conditional distribution of L_k/W_k^γ given T_k and W_k, which does not depend on W_k. Similarly, the conditional distribution of D_k given given T_k, W_k and L_k does not depend on W_k and L_k. Hence, the conditional independence of W_k, L_k/W_k^γ and D_k given T_k follows. □

2.4.2 Further limiting results

Alpha-gamma trees not only have edge weights but also vertex weights, and the latter are in correspondence with the vertex degrees. We can get a result on the limiting ratio between the degree of each vertex and the total number of leaves. To be specific, it is useful to enumerate all vertices in a unique way, e.g. in the order they are visited by depth first search [33], where beginning from the root each subtree is visited recursively, in the order of least labels.

Proposition 2.22. *Let $(c_1+1, \ldots, c_\ell+1)$ be the degree of each vertex in T_k, listed by depth first search. The ratio between the degrees in T_n of these vertices and n^α will converge to*

$C_k = (C_{k,1}, \ldots, C_{k,\ell}) = \overline{W}_k^\alpha M_k D_k'$, *where* $D_k' \sim \mathrm{Dirichlet}(c_1 - 1 - \gamma/\alpha, \ldots, c_\ell - 1 - \gamma/\alpha)$,

M_k and W_k are conditionally independent given T_k, where $\overline{W}_k = 1 - W_k$, and M_k has density

$$\frac{\Gamma(\overline{w}+1)}{\Gamma(\overline{w}/\alpha+1)} s^{\overline{w}/\alpha} g_\alpha(s), \qquad s \in (0, \infty),$$

$\overline{w} = (k-1)\alpha - \ell\gamma$ is total branch point weight in T_k and $g_\alpha(s)$ is the Mittag-Leffler density.

Proof. Recall the first urn model in the preceding proof which assigns colour black to leaves attached in subtrees of branch points of T_k. We will partition the black balls further. The partition of leaves (associated with black balls), where each *subtree* S_{vj}^{sk} of a branch point $v \in R(T_n, [k])$ gives rise to a block, follows the dynamics of a Chinese Restaurant Process with $(\alpha, \overline{w})$-seating plan. Hence, the total degree $C_k^{\mathrm{tot}}(n)/\overline{W}_{n,k}^\alpha \to M_k$ a.s., where $C_k^{\mathrm{tot}}(n)$ is the sum of degrees in T_n of the branch points of T_k, and $\overline{W}_{n,k} = n - k - W_{n,k}$ is the total number of leaves of T_n that are in subtrees directly connected to the branch points of T_k.

Similarly to the discussion of edge length proportions, we now see that the sequence of degree proportions will converge a.s. to the Dirichlet limit as specified. Since $1 - W_k$ is the a.s. limiting proportion of leaves in subtrees connected to the vertices of T_k. □

Given an alpha-gamma tree T_n, if we decompose along the spine that connects the ROOT to leaf 1, we will find the leaf numbers of subtrees connected to the spine is a Chinese restaurant partition of $\{2, \ldots, n\}$ with parameters $(\alpha, 1-\alpha)$. Applying Lemma 2.7, we get following result.

Proposition 2.23. *Let $(T_n, n \geq 1)$ be alpha-gamma trees. Denote by $(P_1, P_2, \ldots)$ the limiting frequencies of the leaf numbers of each subtree of the spine connecting the* ROOT *to leaf 1 in the order of appearance. These can be represented as*

$$(P_1, P_2, \ldots) = (W_1, \overline{W}_1 W_2, \overline{W}_1 \overline{W}_2 W_3, \ldots)$$

where the W_i *are independent,* W_i *has* $\mathrm{beta}(1-\alpha, 1+(i-1)\alpha)$ *distribution, and* $\overline{W}_i = 1 - W_i$.

Observe that this result does not depend on γ. This observation also follows from Proposition 2.6, because colouring $(\mathrm{iv})^{\mathrm{col}}$ and crushing (cr) do not affect the partition of leaf labels according to subtrees of the spine.

Chapter 3

Continuum tree asymptotics of partly exchangeable fragmentation trees

Abstract

We extend exchangeablity into part exchangeablility and construct partly exchangeable fragmentation processes and partly exchangeable fragmentation trees. We investigate the integral representation of dislocation measure and subordinator representation of tagged fragments. We also present a general procedure to embed the partly exchangeable fragmentation trees into CRTs. We obtain the scaling limit of partly exchangeable fragmentation trees and apply the result to the alpha models and the alpha-gamma models. At the end of this chapter we construct a new three-parameter family of partly exchangeable fragmentation trees which contains the alpha-gamma model and another important two-parameter family based on Poisson-Dirichlet distributions.

Key words: Part exchangeability, fragmentation process, dislocation measure, continuum random tree, alpha model, alpha model, alpha-gamma model, Poisson-Dirichlet distributions

A version of this work will be submitted probably to the Annales de l'Institut Henri Poincaré, shortly.

3.1 Introduction

For a number of years there has been an increased interest in phylogenetic trees in mathematical literatures. These trees are described as connected acyclic graphs with no degree-2 vertex but some degree-1 vertices, one of which is distinguished as the root and the others as leaves. We denote by $\mathbb{T}_n^\circ$ the space of such trees with n leaves. Also, if labels $1, \ldots, n$ are given to the leaves, we call the new objects labelled trees and denote the space of them by $\mathbb{T}_n$. Obviously, we can assign probability measures to $\mathbb{T}_n^\circ$ or $\mathbb{T}_n$ and call an associated random variable a random tree. A key property for random trees is the *Markov branching* property, which can be specified as follows. Let $(P_n^\circ, n \geq 1)$ be a sequence of distributions on $(\mathbb{T}_n^\circ, n \geq 1)$. Given a random tree T_n° with n leaves and distribution P_n°, given it splits into $k \geq 2$ subtrees with $n_1 \geq \ldots \geq n_k$ leaves at the branch point adjacent the root, these subtrees are independent with distributions $P_{n_i}^\circ, i = 1, \ldots, k$. As random labelled trees one often considers unlabelled trees equipped with uniformly chosen labels among all possible labellings, which are called *exchangeable labels*.

Several models have been constructed such as the Yule, uniform and comb models, Aldous's beta-splitting models [4], models related to the stable trees [14], Ford's alpha models [21] and alpha-gamma models, see Chapter 2. All of these models are related to fragmentation processes introduced by Bertoin [8, 9]. Such fragmentation processes are right-continuous Markov processes and determined by a constant $c \geq 0$ called *erosion* and a measure ν on $\mathcal{S}^\downarrow := \{(s_1, \ldots) : s_1 \geq s_2 \geq \ldots \geq 0, \sum_{i \geq 1} s_i \leq 1\}$ called *dislocation measure*, which fulfils

$$\nu(\{(1, 0, \ldots)\}) = 0 \text{ and } \int_{\mathcal{S}^\downarrow} (1 - s_1)\nu(d\mathbf{s}) < \infty. \tag{3.1}$$

Specifically, the fragmentation processes associated with the alpha-gamma model have no erosion and a conservative dislocation measure i.e. $c = 0$ and $\nu(\{\mathbf{s} : \sum_{i \geq 1} s_i < 1\}) = 0$. This property makes the asymptotics of these models much easier since it has been shown by Haas and Miermont [27] that fragmentation processes with no erosion and conservative dislocation measures are in one-to-one correspondence with self-similar *continuum random trees* (CRT). Haas et al. [28] proved that properly scaled unlabelled random trees associated with a fragmentation process with no erosion and a conservative dislocation measure will converge to a CRT in probability if the dislocation measure ν satisfies the following regular variation condition

$$\nu(s_1 \leq 1-\varepsilon) = \varepsilon^{-\alpha}\ell(\frac{1}{\varepsilon}) \tag{3.2}$$

and the further regularity condition

$$\int_{\mathcal{S}^{\downarrow}} \sum_{i\geq 2} s_i |\ln(s_i)|^{\varrho}\nu(d\mathbf{s}) < \infty, \tag{3.3}$$

for some $\alpha \in (0,1)$ and $\rho > 0$, where the function $\ell(x)$ is slowly varying as $x \to \infty$.

All of these results are based on the feature that the leaf labels are exchangeable. This is natural in some models such as the beta splitting models and models related to the stable trees, but it does not seem to be so in some others like alpha models and alpha-gamma models. For the latter, they are constructed by inserting one leaf after another starting from the two leaf tree with some procedure called the growth rules. The natural labelling for them should be based on the order of appearance of the leaves. Unfortunately, such labels are not exchangeable. Some work has been done by Pitman and Winkel [40]. They introduced a family of binary trees called $(\alpha,\theta)-$trees which will be the alpha model when $\theta = 1-\alpha$, and got the asymptotics using ordered Chinese Restaurant processes introduced there as well.

In this chapter, we are interested in solving this non-exchangeable label problem by making some alternations in the associated fragmentation process. This allows us to cover multifurcating trees as well.

Denote by $[n]$ the set $\{1,\ldots,n\}$ for all $n \in \mathbb{N}$. Let $B \subseteq \mathbb{N}$, a *partition* of B is a countable collection $\pi = \{\pi_i, i \in \mathbb{N}\}$ of pairwise disjoint subsets of B such that $\cup_{i\in\mathbb{N}}\pi_i = B$. These disjoint subsets $\pi_i, i \in \mathbb{N}$ are called the *blocks* of π. We write $\mathcal{P}_B$ for the set of partitions of B. In the special case when $B = [n]$, we simply write $\mathcal{P}_n := \mathcal{P}_{[n]}$; also $\mathcal{P} := \mathcal{P}_{\mathbb{N}}$. Normally, we arrange the blocks $\pi_1, \pi_2, \ldots$ of π in the order of least element, i.e. $\min \pi_i \leq \min \pi_j$, for every $i \leq j$, with the convention that $\min \emptyset = \infty$. We also let $\pi_{(i)}$ be the block of π that contains the integer $i \in B$. If $\pi \in \mathcal{P}_B$, and $B' \subseteq \mathbb{N}$, we let $\pi|_{B'} = B' \cap \pi$ be the partition of $B' \cap B$ obtained by restricting π to the elements of $B' \cap B$. Let $\pi|_n := \pi|_{[n]}$ for every $n \geq 1$. By convention, we let $\mathbf{1}_B$ be the trivial partition $(B, \emptyset, \ldots)$ of B, and $\mathbf{0}_B = (\{i_1\}, \{i_2\}, \ldots)$ the partition of B into singletons, where $i_1 < i_2 < \ldots$ is the ranked list of elements of B. We say that a partition $\pi \in \mathcal{P}_B$ is *finer* than $\pi' \in \mathcal{P}_B$, and write $\pi \preceq \pi'$, if any block of π is included in some block of π'. This defines a partial order $\preceq$ on $\mathcal{P}_B$. A process or a sequence taking values on $\mathcal{P}_B$ is called *refining* if it is decreasing for this partial order. In the sequel, the set $\mathcal{P}$ will be endowed with the distance $\Delta(\pi, \pi') = 2^{-N(\pi,\pi')}$, where $N(\pi, \pi') = \sup\{n \geq 1 : \pi|_n = \pi'|_n\}$, and the associated Borel σ-algebra.

Given two partitions $\pi^{[m]}, \pi^{[n]}$ of $[m], [n]$, where $m < n$, $m \in \mathbb{N}$, and $n \in \mathbb{N}$, we say that $\pi^{[m]}$ and $\pi^{[n]}$ are *compatible* if $\pi^{[m]}$ coincides with the restriction of $\pi^{[n]}$ to $[m]$. In this terminology, we can cut $\mathcal{P}_n, n \in \mathbb{N}$, into two subsets as follows:

$$\begin{aligned} \mathcal{P}_n^1 &:= \{\pi \in \mathcal{P}_n : \pi \text{ is compatible with } \mathbf{1}_{[2]}\}; \\ \mathcal{P}_n^2 &:= \{\pi \in \mathcal{P}_n : \pi \text{ is compatible with } \mathbf{0}_{[2]}\}. \end{aligned}$$

Also, we define

$$\begin{aligned} \mathcal{P}^1 &:= \{\Gamma \in \mathcal{P} : \Gamma \text{ is compatible with } \mathbf{1}_{[2]}\}, \\ \mathcal{P}^2 &:= \{\Gamma \in \mathcal{P} : \Gamma \text{ is compatible with } \mathbf{0}_{[2]}\}. \end{aligned} \tag{3.4}$$

Definition 3.1. (i) Let $n \in \mathbb{N}$. For each partition $\pi_1 \in \mathcal{P}_n^1$ and $\pi_2 \in \mathcal{P}_n^2$ with same block sizes, a measure μ on $\mathcal{P}_n$ is called partly exchangeable if

$$\mu(\pi) = \begin{cases} \mu(\pi_1), & \textit{if } \pi \in \mathcal{P}_n^1, \\ \mu(\pi_2), & \textit{if } \pi \in \mathcal{P}_n^2, \end{cases} \tag{3.5}$$

for every partition π with same block sizes as π_1 and π_2.

(ii) Let μ be a measure on $\mathcal{P}$. We define μ_n a measure on $\mathcal{P}_n$ as follows

$$\mu_n(\pi) := \mu\{\Gamma \in \mathcal{P} : \Gamma|_n = \pi\}$$

for all $\pi \in \mathcal{P}_n$. If for every $n \in \mathbb{N}$, μ_n is partly exchangeable, we say that μ is partly exchangeable.

(iii) we denote the restrictions of a partly exchangeable measure μ to $\mathcal{P}^1$ and $\mathcal{P}^2$, μ^1 and μ^2 as partly exchangeable measure on $\mathcal{P}^1$ and $\mathcal{P}^2$.

(iv) A random partition with a partly exchangeable distribution is called partly exchangeable partition.

In the above notation, we can replace the exchangeability requirement in fragmentation processes by the part exchangeability and create *partly exchangeable fragmentation processes*:

Definition 3.2. Let $B \subset \mathbb{N}$, and consider a $\mathcal{P}_B$-valued Markov process $(\Pi(t), t \geq 0)$ with $\Pi(0) = \mathbf{1}_B$. We assume that for every t, $t' \geq 0$, the distribution of $\Pi(t+t')$ given $\Pi(t) = \pi$ is the same as that of the random partition whose blocks are given by

$$\sigma_{\Pi_i(t)}(\pi_j^{(i)}) \; i, j \geq 1,$$

where $(\pi^{(i)}, i = 1, \ldots)$ is an i.i.d. sequence of partly exchangeable partitions of $\mathbb{N}$ whose law only depends on t' and $\sigma_{\Pi_i(t)} : \mathbb{N} \to \Pi_i(t)$ is the map so that $\sigma_{\Pi_i(t)}(m)$ is the mth smallest element of $\Pi_i(t)$ for $m \leq \#\Pi_i(t)$ and the largest element for $m > \#\Pi_i(t)$. Then the process Π is called a homogeneous partly exchangeable fragmentation of B.

As for the homogeneous fragmentation process, such processes are in one-to-one correspondence with a σ-finite partly exchangeable measure κ on $\mathcal{P} \setminus \{\mathbf{1}_\mathbb{N}\}$ called *splitting rate*, which fulfils

$$\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_{[n]} \neq 1_{[n]}\}) < \infty \text{ for every } n \geq 2. \tag{3.6}$$

Such a measure has a unique integral representation as follows.

Theorem 3.1. *Let κ be a partly exchangeable measure on $\mathcal{P} \setminus \{\mathbf{1}_\mathbb{N}\}$ which fulfils* (3.6)*. Then there are unique constants $c_1, c_2, c_3 \geq 0$ and unique measures ν_1, ν_2 on $\mathcal{S}^\downarrow$, where ν_1 fulfils*

$$\nu_1(\{(0,0,\ldots)\}) = 0 \text{ and } \nu_1(\{(1,0,\ldots)\}) = 0 \text{ and } \int_{\mathcal{S}^\downarrow} \sum_{i=1}^\infty s_i^2(1-s_i)\nu_1(d\mathbf{s}) < \infty, \tag{3.7}$$

and ν_2 fulfils (3.1) *such that for every $n \in \mathbb{N}$,*

$$\begin{aligned} \kappa(\cdot) &= c_1\delta_{\omega^{(1,2)}}(\cdot) + c_2(\delta_{\epsilon^{(1)}}(\cdot) + \delta_{\epsilon^{(2)}}(\cdot)) + c_3 \sum_{i=3}^\infty \delta_{\epsilon^{(i)}}(\cdot) \\ &\quad + \int_{\mathcal{S}^\downarrow} \kappa_\mathbf{s}(\cdot \cap \mathcal{P}^1)\nu_1(d\mathbf{s}) + \int_{\mathcal{S}^\downarrow} \kappa_\mathbf{s}(\cdot \cap \mathcal{P}^2)\nu_2(d\mathbf{s}) \end{aligned}$$

where $\omega^{(1,2)} = (\{1,2\},\{3\},\{4\},\ldots)$, $\epsilon^{(i)} = (\{i\},\mathbb{N}\setminus\{i\})$, δ_π *stands for the Dirac point mass on* π *and* $\kappa_{\mathbf{s}}$ *is the law of Kingman's paintbox construction in Section* 3.2.1.

We call c_1 killing rate, c_2 and c_3 erosion rates and ν_1 and ν_2 dislocation measures.

As refining partition-valued processes, partly exchangeable fragmentation processes naturally correspond to random labelled trees. We call such trees *discrete partly exchangeable fragmentation trees.*

An important property for the family of random labelled trees is consistency. Given a family of labelled trees $(T_n, n \geq 1)$, we remove the leaf with label n, if the new tree is the same as T_{n-1}, we say that $(T_n, n \geq 1)$ is *consistent.* The most common version of consistency for the family of unlabelled trees is called *sampling consistency*, which can be expressed as follows. Given a family of random unlabelled trees $(T_n^\circ, n \geq 1)$, we choose uniformly at random a leaf from T_n° and delete it, if the new tree is distributed the same as T_{n-1}°, we say that $\{T_n^\circ, n \geq 1\}$ is *sampling consistent.* We show that labelled discrete partly exchangeable trees are consistent but the unlabelled discrete partly exchangeable trees are not necessarily sampling consistent.

Haas et al. [28] showed sampling consistent family of trees can be embedded into continuous random trees (CRT). In Section 3.3, we generate a procedure to embed partly exchangeable fragmentation into CRTs. This procedure works even if they are not sampling consistent As an application of this procedure, we show that properly scaled partly exchangeable trees will converge to CRT. A dislocation measure ν is called *conservative* if $\nu(\{\mathbf{s} \in \mathcal{S}^\downarrow : \sum_{i\geq 1} s_i < 1\}) = 0$.

Theorem 3.2. *Let* Π *be a partly exchangeable fragmentation process with* $c_1 = c_2 = c_3 = 0$ *and two conservative dislocation measures* ν_1 *and* ν_2, *where* ν_1 *fulfils* (3.7), (3.2), (3.3) *with* $\nu_1(s_1 \leq 1-\epsilon) = \epsilon^{-\alpha}\ell(1/\epsilon)$ *and* ν_2 *fulfils* (3.1). *If* $(T_n, n \geq 1)$ *is the associated sequence of discrete partly exchangeable fragmentation trees and* T_n° *is the unlabelled trees, then*

$$\frac{T_n^\circ}{n^\alpha \ell(n)\Gamma(1-\alpha)} \xrightarrow[n\to\infty]{(p)} \mathcal{T}_{(\alpha,\nu)}, \tag{3.8}$$

for the Gromov-Hausdorff metric, where

$$\nu(d\mathbf{s}) = \sum_{i=1}^{\infty} (s_i^2 \nu_1(d\mathbf{s}) + s_i(1-s_i)\nu_2(d\mathbf{s})).$$

In the final section, we show that the alpha model and alpha-gamma model are partly exchangeable fragmentation trees and get their convergence by applying Theorem 3.2. Also, we introduce another new family of Markov branching trees call *three-factor model* which is a subfamily of partly exchangeable fragmentation trees by setting the two dislocation measure as follows:

$$\begin{aligned} \nu_1(d\mathbf{s}) &= \lambda \mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s}), \\ \nu_2(d\mathbf{s}) &= (1-\lambda)\mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s}), \end{aligned}$$

for $\alpha \in (0,1), \theta \in [-2\alpha, \alpha]$ and $\lambda \in [0,1]$, where $\mathrm{PD}^*_{\alpha,\theta}$ is the Poisson-Dirichlet measure on $\mathcal{S}^{\downarrow}$. When $\lambda = \frac{\alpha+\theta}{2\alpha+\theta-1}$, the three-factor model is the alpha-gamma model; when $\lambda = 1/2$, it is Poisson-Dirichlet model [35]. The three-factor model is sampling consistent if and only if it is an alpha-gamma model or Poisson-Dirichlet model. As a family of partly exchangeable fragmentation trees, the properly scaled three-factor model will nevertheless converge to a CRT in probability.

3.2 Homogeneous partly exchangeable fragmentation processes

3.2.1 Exchangeable partitions and fragmentation processes

We say that a block π_i in a partition $\pi \in \mathcal{P}$ possesses an *asymptotic frequency* if and only if the limit

$$|\pi_i| := \lim_{n\to\infty} \frac{1}{n} \#(\pi_i \cap [n])$$

exists. If each block of some partition π has an asymptotic frequency, we say that π possesses asymptotic frequencies and write $|\pi| = (|\pi_1|, \ldots)$, and then write $|\pi|^{\downarrow}$ for the decreasing arrangement of the sequence $|\pi|$. Moreover, if $\sum_{i=1}^{\infty} |\pi_i| = 1$, we say that π has *proper* asymptotic frequencies.

Let Π be a random variable with values in $\mathcal{P}_B$, we say that Π is exchangeable if its law is invariant under the natural action of the permutations of B.

Definition 3.3. Let $B \subset \mathbb{N}$, and consider a $\mathcal{P}_B$-valued Markov process $(\Pi(t), t \geq 0)$. We assume that for every t, $t' \geq 0$, the distribution of $\Pi(t + t')$ given $\Pi(t) = \pi$ is the same as that of the random partition whose blocks are given by

$$\Pi_i(t) \cap \pi_j^{(i)}, i, j \geq 1,$$

where $\pi^{(i)}, i = 1, \ldots$ is i.i.d. sequence of exchangeable partitions of $\mathbb{N}$. Then the process Π is called a homogeneous fragmentation of B.

If a homogeneous fragmentation of B starts from the trivial partition $\mathbf{1}_B$ of B, we say that the process is *standard*. We also assume that the process is non-degenerate, namely that it is not constant a.s.. Then it is clear that homogeneous fragmentation processes are refining and their blocks all decrease to singletons.

As shown by Bertoin in [10][Proposition 3.2 and Proposition 3.3], the laws of standard homogeneous fragmentations of $\mathbb{N}$ are in one-to-one correspondence with σ-finite measures κ on $\mathcal{P} \setminus \{\mathbf{1}_\mathbb{N}\}$ fulfilling (3.6) and which correspond to the jump rates of the fragmentation process. As shown in [8], such measures have the following simple representations. For $\mathbf{s} \in \mathcal{S}^\downarrow$, we let $\kappa_\mathbf{s}$ be the distribution on $\mathcal{P}$ of the random variable Π obtained by Kingman's paintbox construction [31]: let $I_1, I_2, \ldots$ be i.i.d. with law $(s_j)_{j\geq 0}$, and let i, j be in the same block of Π if and only if $i = j$ or $I_i = I_j > 0$.

Lemma 3.3 (Theorem 3.1 of [10]). *Let κ be an exchangeable measure on $\mathcal{P}$ fulfilling* (3.6). *Then*

(i) *κ-almost every partition $\Gamma \in \mathcal{P}$ possesses asymptotic frequencies;*

(ii) *there exists a unique $c \geq 0$ and a unique measure ν on $\mathcal{S}^\downarrow$ which fulfils* (3.1) *such that*

$$\kappa(\cdot) = c\epsilon(\cdot) + \int_{\mathcal{S}^\downarrow} \kappa_\mathbf{s}(\cdot)\nu(d\mathbf{s}), \tag{3.9}$$

where $\epsilon := \sum_{i=1}^\infty \delta_{\epsilon^{(i)}}$, $\epsilon^{(i)}$ is the partition of $[n]$ into two blocks $\{i\}$ and $[n] \setminus \{i\}$;

(iii) *let $|\kappa|^\downarrow$ be the image measure of κ by the mapping $\Gamma \to |\Gamma|^\downarrow$, the restriction*

$$\nu(d\mathbf{s}) := \mathbf{1}_{\{\mathbf{s}\neq(1,0,\ldots)\}} |\kappa|^\downarrow(d\mathbf{s})$$

to $\mathcal{S}^\downarrow \setminus \{(1,0,\ldots)\}$ fulfils (3.1) *and*

$$\mathbf{1}_{\{|\cdot|^\downarrow \neq (1,0,\ldots)\}}\kappa(\cdot) = \int_{\mathcal{S}^\downarrow} \kappa_\mathbf{s}(\cdot)\nu(d\mathbf{s}).$$

(iv) *The restriction of κ to the subset of $\mathcal{P}$ with $|\Gamma|^{\downarrow} = (1, 0, \ldots)$ is proportional to c, i.e.*

$$\mathbf{1}_{\{|\cdot|^{\downarrow}=(1,0,\ldots)\}}\kappa(\cdot) = c\epsilon(\cdot).$$

3.2.2 Partially exchangeable random partitions and constrained exchangeable partitions

There have been other extensions of exchangeability in the literature , notably partial exchangeability and constrained exchangeability. *Partially exchangeable partitions* were introduced by Pitman [38]. A random partition Π_n of $\mathcal{P}_n$ is partially exchangeable if for every partition $\{B_1, \ldots, B_k\} \in \mathcal{P}_n$, where $B_1, \ldots, B_k$ are in order of least element,

$$\mathbb{P}\left(\Pi_n = \{B_1, \ldots, B_k\}\right) = f(\#B_1, \ldots, \#B_k)$$

for some function $f(n_1, \ldots, n_k)$ with $n_1, \ldots, n_k \geq 1$ and $\sum_{i=1}^{k} n_i = n$. A partially exchangeable partition is exchangeable if and only if f is a symmetric function. Partially exchangeable partitions are not partly exchangeable partitions in general and vice versa. Let us consider two partitions $\pi_1 = (\{1, 2\}, \{3, 4\})$ and $\pi_2 = (\{1, 3\}, \{2, 4\})$. π_1 and π_2 should have the same probability mass in partially exchangeable partitions but not necessarily in partly exchangeable partitions. On the other hand, for partitions $\pi_3 = (\{1, 2, 3\}, \{4, 5\})$ and $\pi_4 = (\{1, 2\}, \{3, 4, 5\})$, they should have the same probability mass in partly exchangeable partition but not necessarily in partly exchangeable partition. The intersection of two concepts is exchangeability.

Corollary 3.4. *A random partition Π_n of $\mathcal{P}_n$ is exchangeable if and only if it is both partially exchangeable and partly exchangeable.*

Proof. We only need to prove the sufficient condition. Suppose $\pi \in \mathcal{P}_n^1 \setminus \{\mathbf{1}_{[n]}\}$. Let σ_n be any permutation of $[n]$ and σ_n' be a permutation that only changes 2 and the smallest element in the second block of π. If $\sigma_n(\pi) \in \mathcal{P}_n^1$, then $\mathbb{P}(\Pi_n = \pi) = \mathbb{P}(\Pi_n = \sigma_n(\pi))$ by part exchangeability; if $\sigma_n(\pi) \in \mathcal{P}_n^2$, $\mathbb{P}(\Pi_n = \pi) = \mathbb{P}(\Pi_n = \sigma_n'(\pi)) = \mathbb{P}(\Pi_n = \sigma_n(\pi))$ the first equality comes from partial exchangeability as $\sigma_n'(\pi)$ has the same block sizes in order of least element as π. Hence, Π_n is exchangeable. □

The second one is constrained exchangeable partition. It was introduced by Gnedin in [22]. Let $\varsigma = (\varsigma_1, \ldots), \varsigma_k \geq 1$ be a fixed sequence. We call a partition $\Pi = (B_1, \ldots)$

is *constrained with respect to* ς if each block B_k contains the ς_k least elements of $\cup_{j\geq k}B_j$ for every k with $B_k \neq \emptyset$. We call a random partition Π *constrained exchangeable* if Π is a constrained partition with respect to ς and is invariant under all bijections $\sigma : \mathbb{N} \to \mathbb{N}$ that preserve this property. For instance, take $\varsigma = (1, 2, 1, \ldots)$. A constrained exchangeable partition assumes the values of $(\{1, 3\}, \{2, 4\}, \{5\})$ and $(\{1, 2\}, \{3, 4\}, \{5\})$ with the same probability. This example also shows that not all constrained exchangeable partitions are partly exchangeable. Also partly exchangeable partition is not constrained exchangeable. Due to the part exchangeability, the partitions in $\mathcal{P}^1$ will be constrained exchangeable if and only if we take $\varsigma = (2, 1, 1, \ldots)$, while the partitions in $\mathcal{P}^2$ will be constrained exchangeable if and only we take $\varsigma = (1, 1, \ldots)$. Hence we cannot find a ς such that the constrained exchangeability coincides with partial exchangeability.

3.2.3 Splitting rates and Poisson construction for partly exchangeable fragmentation processes

In this section, we give the formal construction and definition for partly exchangeable fragmentation process. To do this, we have to extend the term *fragmentation* defined by Bertoin [10] into the more general case.

Definition 3.4. Consider $B \subseteq \mathbb{N}$. Let π be a partition of B with $\#\pi = k$ non-empty blocks, and $\pi^{(\cdot)} = (\pi^{(i)}, i = 1, .., k)$ be a sequence in $\mathcal{P}$. For every integer i, we consider the partition of the i-th block π_i of π induced by the i-th term $\pi^{(i)}$ of the sequence $\pi^{(\cdot)}$, that is

$$\pi^{(i)}|_{\pi_i} = \left(\sigma_{\pi_i}(\pi_j^{(i)}|_{[\#\pi_i]}), j \in \mathbb{N}\right),$$

where $\sigma_{\pi_i} : [\#\pi_i] \to \pi_i$ satisfies that $\sigma_{\pi_i}(t)$ is t-th smallest element in π_i and $\#\pi_i$ is the size of π_i. As i varies in $[k]$, the collection $\{\sigma_{\pi_i}(\pi_j^{(i)}|_{[\#\pi_i]}) : i, j \in \mathbb{N}\}$ of the blocks of the these induced partitions forms a partition of B which we denote by $\mathrm{Frag}(\pi, \pi^{(\cdot)})$ and call fragmentation of π by $\pi^{(\cdot)}$.

As the above definition is just a slight extension of the work of Bertion, the following elementary continuity result still holds, the proof of which are the same of that argued by Bertoin [10, Lemma 3.1].

- For every sequence $\pi^{(\cdot)} = (\pi^{(i)}, i \in \mathbb{N})$ of partitions of $\mathbb{N}$ the map $\pi \mapsto \mathrm{Frag}(\pi, \pi^{(\cdot)})$, $\pi \in \mathcal{P}$ is Lipschitz-continuous.

- For each integer n, let $\pi^{(\cdot,n)} = (\pi^{(i,n)}, i \in \mathbb{N})$ be a sequence of partitions. Suppose that the limit $\lim_{n\to\infty} \pi^{(i,n)} := \pi^{(i)}$ exists for every $i \in \mathbb{N}$. Then for every partition $\pi \in \mathcal{P}$, it holds that $\lim_{n\to\infty} \mathrm{Frag}(\pi, \pi^{(\cdot,n)}) = \mathrm{Frag}(\pi, \pi^{(\cdot)})$.

We now present a general procedure for constructing functions: $\Pi : t \mapsto \Pi(t)$ with values in $\mathcal{P}$, such that $\Pi(t)$ gets finer as t increases. Typically, for every fixed $n \in \mathbb{N}$, the function $t \mapsto \Pi^{[n]}(t)$ is defined recursively by splitting exactly one block of the partition at certain discrete times. Then one checks that for every $t \geq 0$, the sequence $(\Pi^{[n]}, n \in \mathbb{N})$ is compatible, and thus can be identified as a unique partition $\Pi(t) \in \mathcal{P}$. The fundamental feature of this construction is that, even though each restricted function $t \mapsto \Pi^{[n]}(t) = \Pi|_n(t)$ has a discrete evolution, the set formed by its jump-times as n varies in $\mathbb{N}$ may be everywhere dense.

Specifically, call '*discrete point measure*' on $\mathbb{R}_+ \times \mathcal{P} \times \mathbb{N}$' any measure m which can be expressed in the form

$$m = \sum_{(t,\Gamma,k)\in\mathcal{D}}^{\infty} \delta_{(t,\Gamma,k)},$$

where $\mathcal{D}$ is a subset of $\mathbb{R}_+ \times \mathcal{P} \times \mathbb{N}$ such that the following two conditions hold. First, for every real number $t' \geq 0$ and integer $n \geq 1$, $\#\{(t,\Gamma,k) \in \mathcal{D} : t \leq t', \Gamma|_n \neq 1_{[n]}, k \leq n\} < \infty$. Second, m has at most one atom on each fiber $\{t\} \otimes \mathcal{P} \otimes \mathbb{N}$, that is

$$m(\{t\} \otimes \mathcal{P} \otimes \mathbb{N}) = 0 \text{ or } 1.$$

Starting from an arbitrary discrete point measure m on $\mathbb{R}_+ \times \mathcal{P} \times \mathbb{N}$, we first describe informally the construction of a family of nested partitions $(\Pi(t), t \geq 0)$. For every fixed $n \in \mathbb{N}$, the assumption that the point measure m is discrete enables us to define a càdlàg step-path $\Pi^{[n]} : t \to \Pi^{[n]}(t)$ with values in the space of partitions of $[n]$, which can only jump at times t at which the fiber $\{t\} \times \mathcal{P} \times \mathbb{N}$ carries an atom of m, say (t,Γ,k), such that $\Gamma|_n \neq 1_{[n]}$ and $k \leq n$. In that case, $\Pi^{[n]}(t)$ is the partition obtained by replacing the k-th block of $\Pi^{[n]}(t-)$, namely $\Pi^{[n]}_k(t-)$, by mapping the restriction of π to this block by $\sigma_{\Pi^{[n]}_k(t-)}$, and leaving other blocks unchanged.

To make this construction more formal, we introduce the following notation. For every $n \in \mathbb{N}$ and every pair $(\Gamma, k) \in \mathcal{P} \times \mathbb{N}$, we write $\Delta_n^{(\cdot)}(\Gamma, k)$ for the sequence of partitions in $\mathcal{P}_n$ given $i = 1, ..., n$ by

$$\Delta_n^{(i)}(\Gamma, k) = \begin{cases} 1_{[n]} & \text{if } i \neq k, \\ \Gamma|_{[n]} & \text{if } i = k. \end{cases} \tag{3.10}$$

Next, let $m^{[n]}$ be the point measure on $[0, \infty) \times (\mathcal{P}_n)^n$ whose atoms are images by the map $(t, \Gamma, k) \to (t, \Delta_n^{(\cdot)}(\Gamma, k))$ of the atoms (t, Γ, k) of m such that $\Gamma|_{[n]} \neq 1_{[n]}$ and $k \leq n$. We write $t_0 = 0$ and $\Big((t_1, \Delta_n^{(\cdot)}(1)), \ldots\Big)$ for the sequence of the atoms of $m^{[n]}$ ranked in increasing order of the first coordinate. We then set $\Pi^{[n]}(t) = 1_{[n]}$ for every $t \in [t_0, t_1)$ and define recursively

$$\Pi^{[n]}(t) = \mathrm{Frag}(\Pi^{[n]}(t_{i-1}), \Delta_n^{(\cdot)}(i)), \text{ for every } t \in [t_i, t_{i+1}).$$

It should be plain that this rigorous construction rephrases the informal one above.

The sequence of partitions $(\Pi^{[n]}(t), n \in \mathbb{N})$ is compatible and the proof is also that same as Lemma 3.3 [10] argued by Bertoin.

Lemma 3.5. *In the notation above, for every $t \geq 0$, the sequence $(\Pi^{[n]}(t), n \in \mathbb{N})$ is compatible, and thus there exists a unique partition $\Pi(t) \in \mathcal{P}$ such that $\Pi|_n(t) = \Pi^{[n]}(t)$ for every $n \in \mathbb{N}$. Moreover the function $\Pi : t \mapsto \Pi(t)$ is càdlàg.*

Definition 3.5. (i) Fix $n \in \mathbb{N}$, and let $\Pi = (\Pi(t), t \geq 0)$ be a Markov process with values in $\mathcal{P}_n$ with càdlàg sample paths. Π is called a homogeneous partly exchangeable fragmentation process if its semigroup can be described as follows. For every $t, t' \geq 0$, the conditional distribution of $\Pi(t + t')$ given $\Pi(t) = \pi$ is the law of $\mathrm{Frag}(\pi, \pi^{(\cdot)})$, where $\pi^{(\cdot)} = (\pi^{(1)}, ...)$ is an i.i.d sequence of partly exchangeable random partitions whose laws only depends on t'. Π is called standard if it starts from $\mathbf{1}_{[n]}$, the partition of $[n]$ into a single non-empty block, $([n], \emptyset, ...)$.

(ii) A Markov process $\Pi = (\Pi(t), t \geq 0)$ with values in $\mathcal{P}$ with càdlàg sample paths is called a homogeneous partly exchangeable fragmentation process if for every $n \in \mathbb{N}$, the restriction $\Pi|_n(\cdot)$ of $\Pi(\cdot)$ to $[n]$ is a homogeneous partly exchangeable fragmentation process.

The description of the semigroup implies that blocks in the partition $\Pi(t)$ split independently of each other, which should be viewed as the branching property in the setting of partition-valued process. The term *homogeneous* in Definition 3.5 refers to the fact that all the blocks of $\Pi(t)$ play the same role in the transition from $\Pi(t)$ to $\Pi(t+t')$ in the sense that they are split according to the same random procedure, independently of the size of these blocks.

We now turn to the dynamics of homogeneous partly exchangeable fragmentation processes. Let Π be a standard homogeneous partly exchangeable fragmentation process. Fix n, consider the following jump rates of the restrictions of Π,

$$q(\pi) := \lim_{t\to 0+} \frac{1}{t}\mathcal{P}(\Pi|_n(t) = \pi),\ \pi \in \mathcal{P}_n \setminus \{\mathbf{1}_{[n]}\}.$$

These jump rates entirely characterize the law of Π. More precisely, we have the following:

Lemma 3.6. *$q(\cdot)$ is a partly exchangeable measure on $\mathcal{P}_n$ and the family of jump rates $(q(\pi), \pi \in \mathcal{P}_n \setminus \{\mathbf{1}_{[n]}\}, n \in \mathbb{N})$ determines the law of Π.*

Proof. Note that the distribution of $\Pi|_n(t)$ is the law of $\mathrm{Frag}(\mathbf{1}_{[n]}, \pi^{(\cdot)})$. As $\pi^{(\cdot)}$ is a partly exchangeable partition, the part exchangeability of q is straightforward.

Consider some partition $\pi' \in \mathcal{P}_n$, and let $\pi'' \in \mathcal{P}_n$ be another partition with $\pi'' \neq \pi'$, which can be obtained from π' by the fragmentation of a single block. This means there is a partition $\Gamma \in \mathcal{P}$ and an index k which is smaller that the number of blocks of π such that $\pi' = \mathrm{Frag}(\pi, \Delta_n^{(\cdot)}(\Gamma, k))$, where $\Delta_n^{(\cdot)}(\Gamma, k)$ is defined in (3.10). As each block splits independently and identically with the same law, the jump rate from π' to π'' coincides with the jump rate $q(\pi)$ from $\mathbf{1}_{[\#\pi'_k]}$ to π, that is

$$\lim_{t\to 0+} \frac{1}{t}\mathbb{P}(\Pi|_n(t) = \pi''|\Pi|_n(0) = \pi') = q(\pi).$$

Finally we show that all other jump rates are zero. If π'' cannot be expressed in the form $\pi'' = \mathrm{Frag}(\pi', \pi^{(\cdot)})$ for any sequence $\pi^{(\cdot)}$, the jump rates will be zero. On the other hand, if π'' can be obtained from π' by the fragmentation of two or more of its blocks, the independence of each block implies that $\mathbb{P}(\Pi|_n(t) = \pi''|\Pi|_n(0) = \pi') = O(t^2)$ as $t \to 0$ and the jump rate from π' to π'' must be zero. Then the conclusion is obvious. □

The fundamental result about the collection of jump rates in Lemma 3.6 is that it can be described by a single measure on $\mathcal{P}$.

Proposition 3.7. *Let* $(q(\pi) : \pi \in \mathcal{P}_n \setminus \mathbf{1}_{[n]}, n \in \mathbb{N})$ *be the family of jump rates of some partly exchangeable homogeneous fragmentation* Π*. There exists a unique measure* κ *on* $\mathcal{P}$ *such that* $\kappa(\{1_{\mathbb{N}}\}) = 0$ *and*

$$\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n = \pi\}) = q(\pi) \tag{3.11}$$

for every $n \in \mathbb{N}$ *and every partition* $\pi \in \mathcal{P}_n \setminus \{\mathbf{1}_{[n]}\}$*. More precisely, the measure* κ *is partly exchangeable and assigns a finite mass to the sets* $\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]}\}$ *for all* $n \in \mathbb{N}$*.*

The proof of existence is exactly the same as the one of fragmentation process given by Bertoin [10, Proposition 3.2] and the part exchangeability is a direct result of Lemma 3.6.

Our purpose now is to construct a partly exchangeable homogeneous fragmentation Π with a given splitting rate κ. Consider a partly exchangeable measure κ on $\mathcal{P}$ fulfilling (3.6). Recall the construction of a family $(\Pi(t), t \geq 0)$ of nested partitions from a discrete point measure m. We now consider the situation when $m := M$ is random and distributed as a Poisson random measure on $[0, \infty) \times \mathcal{P} \times \mathbb{N}$ with intensity $dt \otimes \kappa(d\pi) \otimes \#$, where $\#$ stands the counting measure of $\mathbb{N}$ and κ is a splitting rate on $\mathcal{P}$. It is easy to see that M fulfils the requirements for discrete point measures. Followed by Lemma 3.6 and Proposition 3.7, the following result is straightforward.

Proposition 3.8. *In the notation above, the process* $\Pi = (\Pi(t), t \leq 0)$ *is a standard homogeneous fragmentation with splitting rate* κ*.*

We have seen that the distribution of a partly exchangeable homogeneous fragmentation is determined by its splitting rate, which is a partly exchangeable measure on $\mathcal{P}$ fulfilling (3.6). Conversely, the Poisson construction shows that any such measure can be viewed as the splitting rate of homogeneous partly exchangeable fragmentation.

3.2.4 Auxiliary σ-finite measures

Before we prove Theorem 3.1, let us express some other σ-finite measures. The exchangeable σ-finite measure on $\mathcal{P} \setminus \{\mathbf{1}_{\mathbb{N}}\}$ introduced by Bertoin fulfils (3.6). Parallel to the idea of Bertoin, we can define an exchangeable σ-finite measure κ on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}\}$ fulfilling

$$\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{0}_{[n]}\}) < \infty \text{ for every } n \geq 2. \tag{3.12}$$

Furthermore, we can consider another family of σ-finite exchangeable measures κ on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}, \mathbf{1}_{\mathbb{N}}\}$ as follows: for every $n \geq 3$

$$\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]} \text{ or } \mathbf{0}_{[n]}\}) < \infty. \tag{3.13}$$

For such measures, we have the following results.

Lemma 3.9. *Let κ be an exchangeable measure on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}, \mathbf{1}_{\mathbb{N}}\}$ which fulfills* (3.13). *Then there exist two unique constants $c_1, c_2 \geq 0$ and a unique measure ν on $\mathcal{S}^{\downarrow}$ that fulfills* (3.7) *such that*

$$\kappa(\cdot) = c_1\omega(\cdot) + c_2\epsilon(\cdot) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot)\nu(d\mathbf{s}), \tag{3.14}$$

where $\kappa_{\mathbf{s}}$ is the law of Kingman's paintbox construction, $\epsilon := \sum_{i=1}^{\infty} \delta_{\epsilon^{(i)}}$ with $\epsilon^{(i)} = (\{i\}, \mathbb{N} \setminus \{i\})$ and $\omega := \sum_{i \neq j} \delta_{\omega^{(i,j)}}$ with $\omega^{(i,j)} := (\{i,j\}, \{1\}, \{2\}, \ldots)$. Specially, following holds:

(i) *κ-almost every partition $\Gamma \in \mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}, \mathbf{1}_{\mathbb{N}}\}$ possesses asymptotic frequencies.*

(ii) *The restriction of κ to the subset of partitions Γ with $|\Gamma|^{\downarrow} \neq (0, \ldots)$ or $(1, 0, \ldots,)$ is a dislocation rate. More precisely, let $|\kappa|^{\downarrow}$ be the image measure of κ by the mapping $\Gamma \to |\Gamma|^{\downarrow}$. The restriction*

$$\nu(\mathrm{d}\mathbf{s}) := \mathbf{1}_{\{\mathbf{s} \neq (0,\ldots) \text{ or } (1,0,\ldots,)\}} |\kappa|^{\downarrow}(d\mathbf{s})$$

fulfills (3.7) *and*

$$\mathbf{1}_{\{|\Gamma|^{\downarrow} \neq (0,\ldots) \text{ or } (1,0,\ldots,)\}} \kappa(d\Gamma) = \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(d\Gamma)\nu(\mathbf{s}).$$

(iii) *The restriction of κ to the subset of partitions Γ with $|\Gamma|^{\downarrow} \in \{(0, \ldots), (1, 0, \ldots)\}$ is propositional to the erosion rate and killing rate, that is there are two real numbers $c_1, c_2 \geq 0$ such that*

$$\mathbf{1}_{\{|\Gamma|^{\downarrow} \in \{(0,\ldots),(1,0,\ldots)\}\}} \kappa(d\Gamma) = c_1\omega(d\Gamma) + c_2\epsilon(d\Gamma).$$

Proof. (i) For every integer n, write κ_n for the restriction of κ to $\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{0}_{[n]} \text{ or } \mathbf{1}_{[n]}\}$. Then κ_n is a finite measure on $\mathcal{P}$, define the n-shift $\vec{\Gamma}$ of a partition Γ by

$$i \overset{\vec{\Gamma}}{\sim} j \Leftrightarrow i+n \overset{\Gamma}{\sim} j+n,$$

and then write $\vec{\kappa}_n$ for the image of κ_n by this shift.

Then $\vec{\kappa}_n$ is an exchangeable finite measure on $\mathcal{P}$, and by Kingman's theorem [10, Theorem 2.1], $\vec{\kappa}_n$-almost every partition has asymptotic frequencies. More precisely,

$$\vec{\kappa}_n(d\Gamma) = \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(d\Gamma)\vec{\kappa}_n(|\Gamma|^\downarrow \in d\mathbf{s}) \tag{3.15}$$

is a regular disintegration of $\vec{\kappa}_n$. As shift does not affect asymptotic frequencies, κ_n-almost every partition has asymptotic frequencies. This establish the first claim.

(ii) By (3.15), for every $\mathbf{s} \in \mathcal{S}^\downarrow$

$$\begin{aligned}
& \kappa_n(\{\Gamma \in \mathcal{P} : \Gamma|_{\{n+1,n+2,n+3\}} = (\{n+1,n+2\},\{n+3\}))\\
= \ & \vec{\kappa}_n(\{\Gamma \in \mathcal{P} : \Gamma|_3 = (\{1,2\},\{3\})\})\\
= \ & \int_{\mathcal{S}^\downarrow} \kappa_{\mathbf{s}}(\{\Gamma|_3 = (\{1,2\},\{3\})\})\vec{\kappa}_n(|\Gamma|^\downarrow \in d\mathbf{s})\\
= \ & \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i^2(1-s_i)\vec{\kappa}_n(|\Gamma|^\downarrow \in d\mathbf{s})\\
= \ & \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i^2(1-s_i)\kappa_n(|\Gamma|^\downarrow \in d\mathbf{s}).
\end{aligned}$$

The last equality holds as $\kappa_n(|\Gamma|^\downarrow \in d\mathbf{s}) = \vec{\kappa}_n(|\Gamma|^\downarrow \in d\mathbf{s})$ by exchangeability. Then if we denote by $\nu_n(d\mathbf{s}) = \mathbf{1}_{\{\mathbf{s}\neq(0,\ldots) \text{ or } (1,0,\ldots)\}}|\kappa_n|^\downarrow(d\mathbf{s})$ the restriction to $\mathcal{S}^\downarrow \setminus \{(0,\ldots),(1,0,\ldots)\}$ of the image measure of κ_n by the map $\Gamma \mapsto |\Gamma|^\downarrow$, then

$$\infty > \kappa_n(\{\Gamma \in \mathcal{P} : \Gamma|_{\{n+1,n+2,n+3\}} = (\{n+1,n+2\},\{n+3\})\}) = \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i^2(1-s_i)\nu_n(d\mathbf{s}).$$

The finite measure ν_n increases as $n \to \infty$ to the measure ν defined in the statement, so

$$\lim_{n\to\infty} \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i^2(1-s_i)\nu_n(d\mathbf{s}) = \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i^2(1-s_i)\nu(d\mathbf{s}).$$

On the other hand

$$\begin{aligned}
&\int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i^2(1-s_i)\nu_n(d\mathbf{s}) \\
&= \kappa_n(\{\Gamma \in \mathcal{P} : \Gamma|_{\{n+1,n+2,n+3\}} = (\{n+1,n+2\},\{n+3\})\}) \\
&\leq \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_{\{n+1,n+2,n+3\}} = (\{n+1,n+2\},\{n+3\})\}) \\
&= \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_3 = (\{1,2\},\{3\})\}) \\
&< \infty.
\end{aligned}$$

We see that ν fulfills (3.7).

Finally, fix $k \in \mathbb{N}$ and pick a partition $\pi^k \neq \mathbf{0}_{[k]}$ of $[k]$. We have by monotone convergence

$$\begin{aligned}
&\kappa(\Gamma|_k = \pi^k, |\Gamma|^\downarrow \neq (0,\ldots) \text{ or } (1,0,\ldots,)) \\
&= \lim_{n\to\infty} \kappa(\Gamma|_k = \pi^k, \Gamma|_{\{k+1,\ldots,k+n\}} \neq \mathbf{0}_{\{k+1,\ldots,k+n\}} \text{ or } \mathbf{1}_{\{k+1,\ldots,k+n\}}, |\Gamma|^\downarrow \neq (0,\ldots) \text{ or } (1,0,\ldots)).
\end{aligned}$$

In the notation introduced in (i), we see from an obvious permutation that

$$\begin{aligned}
&\kappa(\Gamma|_k = \pi^k, \Gamma|_{\{k+1,\ldots,k+n\}} \neq \mathbf{0}_{\{k+1,\ldots,k+n\}} \text{ or } \mathbf{1}_{\{k+1,\ldots,k+n\}}, |\Gamma|^\downarrow \neq (0,\ldots) \text{ or } (1,0,\ldots)) \\
&= \vec{\kappa}_n(\Gamma|_k = \pi^k, |\Gamma|^\downarrow \neq (0,\ldots) \text{ or } (1,0,\ldots)). \qquad (3.16)
\end{aligned}$$

Applying (3.15) and then letting n tend to ∞, we conclude that

$$\kappa(\Gamma|_k = \pi^k, |\Gamma|^\downarrow \neq (0,\ldots) \text{ or } (1,0,\ldots)) = \int_{\mathcal{S}^\downarrow} \kappa_s(\Gamma|_k = \pi^k)\nu(d\mathbf{s}).$$

This establishes (ii) as k is arbitrary and the restriction $|\Gamma|^\downarrow \neq (0,\ldots)$ or $(1,0,\ldots)$ excludes the singleton partition $\mathbf{0}_{\mathbb{N}}$ and the trivial partition $\mathbf{1}_{\mathbb{N}}$.

(iii) Consider $\tilde{\kappa}$, the restriction of κ to the event $\{\Gamma|_3 = (\{1,2\},\{3\}), |\Gamma|^\downarrow = (0,\ldots)\}$, which has a finite mass. Its image by the 3-shift as defined in (i) is an exchangeable finite measure on $\mathcal{P}$ for which almost every partition has asymptotic frequencies $(0,\ldots)$, and hence it must be proportional to Dirac mass at the singleton partition. Let us denote by π the partition $(\{1,2\},\{3\},\{4\},\ldots)$, $\pi_{n,0}$ the partition $(\{1,2,n\},\{3\},\{4\},\ldots)$, $\pi_{0,n}$ the partition $(\{1,2\},\{3,n\},\{4\},\ldots)$ and $\pi_{m,n}$ the partition $(\{1,2,m\},\{3,n\},\{4,\},\ldots)$. We thus have that

$$\tilde{\kappa} = c_1\delta_\pi + \sum_{n=3}^{\infty}(c_n'\delta_{\pi_{n,0}} + c_n''\delta_{\pi_{0,n}}) + \sum_{\substack{m,n\geq 3\\ m\neq n}} c_n'''\delta_{\pi_{m,n}}$$

where δ stands for the Dirac point mass. By exchangeability, we see that the assumption $\kappa(\{\Gamma|_3 = (\{1,2\},\{3\})\}) < \infty$ forces $c'_n = c''_n = c'''_{m,n} = 0$. The fact that κ restricted to the event $\{|\Gamma|^{\downarrow} = (0,\ldots)\}$ coincides with $c_1\omega$ is now clear by exchangeability. Similarly we can get the fact that κ restricted to the event $\{|\Gamma|^{\downarrow} = (1,0,\ldots)\}$ coincides with $c_2\epsilon$. □

If we strengthen the σ-finite condition in (3.13) to that in (3.12), we have following corollary.

Corollary 3.10. *Let κ be an exchangeable measure on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}\}$ which fulfills* (3.12). *Then there exist unique $c_1 \geq 0$ and a unique measure ν on $\mathcal{S}^{\downarrow}$ that fulfills*

$$\nu(\{(0,\ldots)\} = 0 \text{ and } \int_{\mathcal{S}^{\downarrow}} s_1^2 \nu(d\mathbf{s}) < \infty \tag{3.17}$$

such that

$$\kappa(\cdot) = c_1\omega(\cdot) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot)\nu(\mathbf{s}). \tag{3.18}$$

Specifically, the following holds:

(i) *κ-almost every partition $\Gamma \in \mathcal{P}$ possesses asymptotic frequencies.*

(ii) *The restriction of κ to the subset of partitions Γ with $|\Gamma|^{\downarrow} \neq (0,\ldots)$ is a dislocation rate. More precisely, let $|\kappa|^{\downarrow}$ be the image measure of κ by the mapping $\Gamma \to |\Gamma|^{\downarrow}$. The restriction*

$$\nu(\mathrm{d}\mathbf{s}) := \mathbf{1}_{\{\mathbf{s}\neq(0,\ldots)\}}|\kappa|^{\downarrow}(d\mathbf{s})$$

fulfills (3.17) *and*

$$\mathbf{1}_{\{|\Gamma|^{\downarrow}\neq(0,\ldots)\}}\kappa(d\Gamma) = \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(d\Gamma)\nu(\mathbf{s}).$$

(iii) *The restriction of κ to the subset of partitions Γ with $|\Gamma|^{\downarrow} = (0,\ldots)$ is proportional to killing rate, that is there are two real numbers $c_1 \geq 0$ such that*

$$\mathbf{1}_{\{|\Gamma|^{\downarrow}=(0,\ldots)\}}\kappa(d\Gamma) = c_1\omega(d\Gamma).$$

Proof. The proof is nearly the same as that of Lemma 3.9, except for the condition (3.17).

By the same method in the proof of (ii) in Lemma 3.9, we have

$$\begin{aligned}\int_{\mathcal{S}^{\downarrow}} s_1^2 \nu(d\mathbf{s}) < \infty &= \lim_{n\to\infty} \int_{\mathcal{S}^{\downarrow}} s_1^2 \nu_n(d\mathbf{s}) \\ &\leq \kappa_n(\{\Gamma \in \mathcal{P} : \Gamma|_{\{n+1,n+2\}} = (\{n+1, n+2\})\}) \\ &= \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_2 = (\{1,2\})\}) \\ &< \infty.\end{aligned}$$

□

3.2.5 Decomposition of partly exchangeable measures

In this section, we start from finite partly exchangeable measures. In fact, a finite partly exchangeable measure is a mixture of a σ-finite measure on $\mathcal{P}\setminus\{\mathbf{1}_{\mathbb{N}}\}$ and a σ-finite measure on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}\}$.

Proposition 3.11. *Let κ be a partly exchangeable measure on $\mathcal{P}$, the follow holds:*

(i) *if κ is finite, there exist an exchangeable σ-finite measure κ_1 on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}\}$ fulfilling* (3.12) *and an exchangeable σ-finite measure κ_2 on $\mathcal{P} \setminus \{\mathbf{1}_{\mathbb{N}}\}$ fulfilling* (3.6) *such that*

$$\kappa = \kappa_1|_{\mathcal{P}^1} + \kappa_2|_{\mathcal{P}^2}, \tag{3.19}$$

where $\kappa_1|_{\mathcal{P}^1}, \kappa_2|_{\mathcal{P}^2}$ are the restrictions of κ_1, κ_2 on $\mathcal{P}^1, \mathcal{P}^2$ respectively;

(ii) *if κ is σ-finite on $\mathcal{P} \setminus \{\mathbf{1}_{\mathbb{N}}\}$ fulfilling* (3.6)*,* (3.19) *holds with κ_1 fulfilling* (3.13) *and κ_2 fulfilling* (3.6)*;*

(iii) *if κ is σ-finite on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}\}$ fulfilling* (3.12)*,* (3.19) *holds with κ_1 fulfilling* (3.12) *and κ_2 fulfilling* (3.13)*;*

(iv) *if κ is σ-finite on $\mathcal{P} \setminus \{\mathbf{0}_{\mathbb{N}}, \mathbf{1}_{\mathbb{N}}\}$ fulfilling* (3.13)*,* (3.19) *holds with both of κ_1 and κ_2 fulfilling* (3.13)*.*

Proof. (i) Note that for every $n \geq 3$ and every $\pi \in \mathcal{P}_n^2 \setminus \{\mathbf{0}_{[n]}\}$, there exists a $\pi' \in \mathcal{P}_n^1$ such that π' has the same block sizes as π. Due to the part exchangeability of κ, we define a function κ_1 on $\mathcal{P}_n \setminus \{\mathbf{0}_{[n]}\}$ as follows: for every $n \geq 3$

$$\kappa_1(\pi) = \begin{cases} \kappa(\{\Gamma \in \mathcal{P}^1 : \Gamma|_n = \pi\}), & \pi \in \mathcal{P}_n^1 \\ \kappa(\{\Gamma \in \mathcal{P}^1 : \Gamma|_n = \pi'\}), & \pi \in \mathcal{P}_n^2 \end{cases}$$

where π' is any partition in $\mathcal{P}_n^1$ that has the same block sizes as π. For any $\pi \in \mathcal{P}_n^1$, we have

$$\begin{aligned}
\kappa_1(\pi) &= \kappa(\{\Gamma \in \mathcal{P}^1 : \Gamma|_n = \pi\}) \\
&= \sum_{\{\pi_{n+1} \in \mathcal{P}_{n+1}^1 : \pi_{n+1}|_n = \pi\}} \kappa(\{\Gamma \in \mathcal{P}^1 : \Gamma|_n = \pi_{n+1}\}) \\
&= \sum_{\{\pi_{n+1} \in \mathcal{P}_{n+1}^1 : \pi_{n+1}|_n = \pi\}} \kappa_1(\pi_{n+1}). \qquad (3.20)
\end{aligned}$$

On the other hand, for any $\pi \in \mathcal{P}_n^2$, let σ_π be a permutation on $[n]$ such that $\sigma_\pi(\pi) \in \mathcal{P}_n^1$ and has the same block sizes as π. Note that σ_π induces a bijection from $\{\pi_{n+1} \in \mathcal{P}_n^2 : \pi_{n+1}|_n = \pi\}$ to $\{\pi'_{n+1} \in \mathcal{P}_n^1 : \pi'_{n+1}|_n = \sigma_\pi(\pi)\}$. According to the definition of κ_1, for any $\pi \in \mathcal{P}_n^2 \setminus \{\mathbf{0}_{[n]}\}$,

$$\begin{aligned}
\kappa_1(\pi) &= \kappa(\{\Gamma \in \mathcal{P}^1 : \Gamma|_n = \sigma_\pi(\pi)\}) \\
&= \sum_{\{\pi'_{n+1} \in \mathcal{P}_{n+1}^1 : \pi'_{n+1}|_n = \sigma_\pi(\pi)\}} \kappa(\{\Gamma \in \mathcal{P}^1 : \Gamma|_n = \pi'_{n+1}\}) \\
&= \sum_{\{\pi'_{n+1} \in \mathcal{P}_{n+1}^1 : \pi'_{n+1}|_n = \sigma_\pi(\pi)\}} \kappa_1(\sigma_\pi^{-1}(\pi'_{n+1})) \\
&= \sum_{\{\pi_{n+1} \in \mathcal{P}_{n+1}^2 : \pi_{n+1}|_n = \pi\}} \kappa_1(\pi_{n+1}) \qquad (3.21)
\end{aligned}$$

By combining (3.20) and (3.21), we can define consistently a measure on $\mathcal{P}_{\mathbb{N}} \setminus \{\mathbf{0}_{\mathbb{N}}\}$ and still denote it by κ_1. According to the definition of κ_1, κ_1 is an exchangeable measure and fulfills (3.12) as κ is finite. Now we have for $\Gamma \in \mathcal{P}^1$, $\kappa|_{\mathcal{P}^1}(d\Gamma) = \kappa_1|_{\mathcal{P}^1}(d\Gamma)$.

By the same method, we can define an exchangeable σ-finite measure κ_2 on $\mathcal{P}_{\mathbb{N}} \setminus \mathbf{1}_{\mathbb{N}}$, such that $\kappa_2|_{\mathcal{P}^2} = \kappa|_{\mathcal{P}^2}$ which complete the proof of (i).

Similarly (ii)(iii)(iv) is clear and the different σ-finite condition of κ_1, κ_2 in each statement is directly the result of different σ-finite condition of κ.

□

Combining Corollary 3.10 and Proposition 3.11, the following results hold.

Corollary 3.12. *Let κ be a finite partly exchangeable measure on $\mathcal{P}$. Then there exist two real numbers $c_1, c_2 \geq 0$, a dislocation measure ν_1 on $\mathcal{S}^\downarrow$ fulfilling* (3.17) *and a dislocation* *measure ν_2 on $\mathcal{S}^{\downarrow}$ fulfilling* (3.1)*, such that*

$$\kappa(\cdot) = c_1\delta_{\omega^{(1,2)}}(\cdot) + c_2(\delta_{\epsilon^{(1)}}(\cdot) + \delta_{\epsilon^{(2)}}(\cdot)) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^1)\nu_1(d\mathbf{s}) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^2)\nu_2(d\mathbf{s})$$

where δ is Dirac mass.

Proof. Applying Lemma 3.10, we have for $\Gamma \in \mathcal{P}^1$,

$$\kappa(d\Gamma) = \kappa_1(d\Gamma) = c_1\delta_{\omega^{(1,2)}}(d\Gamma) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(d\Gamma)\nu_1(d\mathbf{s}).$$

Applying Lemma 3.3, we have for every $\Gamma \in \mathcal{P}^2$,

$$\kappa(d\Gamma) = \kappa_2(d\Gamma) = c_2(\delta_{\epsilon^{(1)}}(\mathrm{d}\Gamma) + \delta_{\epsilon^{(2)}}(\mathrm{d}\Gamma)) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(d\Gamma)\nu_2(d\mathbf{s}).$$

□

Proof of Theorem 3.1. According to (ii) in Proposition 3.11, as κ is a σ-finite partly exchangeable measure on $\mathcal{P}\setminus\{\mathbf{1}_{\mathbb{N}}\}$ fulfilling (3.6), (3.19) holds with κ_1 fulfilling (3.13) and κ_2 fulfilling (3.6). Applying Lemma 3.9 to κ_1 and applying Lemma 3.3 to κ_2, the condition of ν_1, ν_2 is clear and we have

$$\begin{aligned}
\kappa(\cdot) &= c_1\omega(\cdot \cap \mathcal{P}^1) + c_3\epsilon(\cdot \cap \mathcal{P}^1) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^1)\nu_1(d\mathbf{s}) \\
&\quad + c_2\epsilon(\cdot \cap \mathcal{P}^2) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^2)\nu_2(d\mathbf{s}) \\
&= c_1\delta_{\omega^{(1,2)}}(\cdot) + c_2(\delta_{\epsilon^{(1)}}(\cdot) + \delta_{\epsilon^{(2)}}(\cdot)) + c_3\sum_{i=3}^{\infty}\delta_{\epsilon^{(i)}}(\cdot) \\
&\quad + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^1)\nu_1(d\mathbf{s}) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\cdot \cap \mathcal{P}^2)\nu_2(d\mathbf{s}),
\end{aligned}$$

which complete the proof. □

There is some elementary but useful consequences of Theorem 3.1.

Corollary 3.13. *Let Π be a partly exchangeable random partition of $\mathbb{N}$, $\Pi_{(i)}$ be the block containing i and $|\Pi|^{\downarrow}$ be the sequence of the asymptotic frequencies of its blocks in decreasing order. Then*

(i)

$$\mathbb{P}\big(\Pi|_{\{1,3,\dots,n\}} = \mathbf{1}_{\{1,3,\dots,n\}}\big||\Pi|^{\downarrow}, |\Pi_{(1)}|\big) = |\Pi_{(1)}|^{n-2} \tag{3.22}$$

and therefore

$$\mathbb{E}\big(|\Pi_{(1)}|^n\big) = \mathbb{P}(\Pi|_{\{1,3,\dots,n+2\}} = \mathbf{1}_{\{1,3,\dots,n+2\}}).$$

(ii)

$$\mathbb{P}\big(\Pi|_{\{2,3,\dots,n\}} = \mathbf{1}_{\{2,3,\dots,n\}}\big||\Pi|^{\downarrow}, |\Pi_{(2)}|\big) = |\Pi_{(2)}|^{n-2} \tag{3.23}$$

and therefore

$$\mathbb{E}\big(|\Pi_{(2)}|^n\big) = \mathbb{P}(\Pi|_{\{2,3,\dots,n+2\}} = \mathbf{1}_{\{2,3,\dots,n+2\}}).$$

(iii) *For* $i \geq 3$*,*

$$\mathbb{P}\big(\Pi|_{\{i,\dots,i+n\}} = \mathbf{1}_{\{i,\dots,i+n\}}\big||\Pi|^{\downarrow}, |\Pi_{(i)}|\big) = |\Pi_{(i)}|^{n} \tag{3.24}$$

and therefore

$$\mathbb{E}\big(|\Pi_{(i)}|^n\big) = \mathbb{P}(\Pi|_{\{i,\dots,i+n\}} = \mathbf{1}_{\{i,\dots,i+n\}}).$$

Proof. Define a measure ν on $\mathcal{S}^{\downarrow}$ as follows

$$\nu(d\mathbf{s}) = \sum_{i=1}^{\infty} s_i^2 \nu_1(d\mathbf{s}) + (s_0 + \sum_{i=1}^{\infty} s_i(1-s_i))\nu_2(d\mathbf{s}).$$

Obviously ν_1, ν_2 are absolute continuous relative to ν. We refer to $\frac{d\nu_1}{d\nu}$ and $\frac{d\nu_2}{d\nu}$ as their Radon-Nikodym derivatives. Conditionally on $|\Pi|^{\downarrow} = \mathbf{s} \in \mathcal{S}^{\downarrow}$, the integral representation in Theorem 3.1 implies that,

$$\mathbb{P}\big(\Pi|_{\{1,3,\dots,n\}} = \mathbf{1}_{\{1,3,\dots,n\}}, |\Pi_{(1)}| = s_i\big||\Pi|^{\downarrow} = \mathbf{s}\big) = s_i^{n+1}\frac{d\nu_1}{d\nu}(\mathbf{s}) + s_i^n(1-s_i)\frac{d\nu_2}{d\nu}(\mathbf{s}).$$

On the other hand,

$$\mathbb{P}\big(|\Pi_{(1)}| = s_i\big||\Pi|^{\downarrow} = \mathbf{s}\big) = s_i^2\frac{d\nu_1}{d\nu}(\mathbf{s}) + s_i(1-s_i)\frac{d\nu_2}{d\nu}(\mathbf{s}).$$

Therefore, (3.22) is straightforward. As 1 and 2 are exchangeable, (3.23) is just a copy of (3.22) for $\Pi_{(2)}$. Due to that $3, \dots,$ are exchangeable, (3.24) can be obtained similarly. □

3.2.6 Subordinator representation of the tagged partly exchangeable fragment

Now we turn our attention to the process of the asymptotic frequency of the block $|\Pi_{(1)}(\cdot)|$ containing 1 in a homogeneous partly exchangeable fragmentation process. As each random variable $\Pi_{(1)}(t)$ takes values in $[0,1]$ and is decreasing by t increases, we define a right-continuous increasing process $(\xi^{(1)}(t), t \geq 0)$ with values in $[0,\infty]$ by

$$\xi^{(1)}(t) = -\log|\Pi_{(1)}(t)|, \ t \geq 0.$$

We show that it is a subordinator.

Proposition 3.14. *Let Π be a homogeneous partly exchangeable fragmentation process with killing rate c_1, erosion rates c_2, c_3 and dislocation measures ν_1, ν_2. The process $\xi^{(1)} = (\xi^{(1)}(t), t > 0)$ defined by $\xi^{(1)}(t) = -\log|\Pi_{(1)}(t)|$, is an $(\mathcal{F}_t)$-subordinator, where $\mathcal{F}_t$ is the filtration of $\xi^{(1)}(t)$. The Laplace exponent of $\xi^{(1)}$, $\Phi(q) = -\log \mathbb{E}(|\Pi_{(1)}(1)|^q)$, $q > 0$ is given by*

$$\Phi(q) = c_1 + c_2 + c_3 q + \int_{\mathcal{S}^\downarrow} \sum_{i\geq 1} s_i^2(1-s_i^q)\nu_1(d\mathbf{s}) + \big(s_0 + \sum_{i\geq 1} s_i(1-s_i)(1-s_i^q)\big)\nu_2(d\mathbf{s}). \quad (3.25)$$

Alternatively, the drift coefficient d *coincides with the erosion rate c_3, the killing rate is given by*

$$k = c_1 + c_2 + \int_{\mathcal{S}^\downarrow} s_0 \nu_2(d\mathbf{s}),$$

and the Lévy measure by

$$\Lambda(dx) = \mathrm{e}^{-2x} \sum_{i=1}^{\infty} \nu_1(\{\mathbf{s} \in \mathcal{S}^\downarrow : -\log s_i \in dx\}) + \mathrm{e}^{-x}(1-\mathrm{e}^{-x}) \sum_{i=1}^{\infty} \nu_2(\{\mathbf{s} \in \mathcal{S}^\downarrow : -\log s_i \in dx\}).$$

Proof. The construction of partly exchangeable fragmentation processes Π in Section 3.2.3 implies that $\xi^{(1)}$ is increasing and right-continuous and that $\Pi_{(1)}(t+t') = \Pi_{(1)}(t) \cap \sigma_{\Pi_{(1)}(t)}(\Pi'_{(1)}(t'))$, where $\Pi'(t')$ is a partly exchangeable partition independent of $\mathcal{F}_t$ and has the same law as $\Pi(t')$. Therefore $|\Pi_{(1)}(t+t')| = |\Pi_{(1)}(t)||\sigma_{\Pi_{(1)}(t)}(\Pi'_{(1)}(t'))|$, which give us the independent and stationary increments. Thus $\xi^{(1)}$ is a subordinator.

Define $\tilde{\Phi}(q)$ by the righthand-side of (3.25). Recall the construction of $\Pi(\cdot)$ in Section 3.2.3 in terms of the Poisson random measure M on $[0,\infty) \times \mathcal{P} \times \mathbb{N}$ with intensity $dt \otimes \kappa \otimes \#$, where κ stands for the splitting rate which is a σ-finite partly exchangeable measure on $\mathcal{P} \setminus \{\mathbf{1}_\mathbb{N}\}$. Let M_1 be the random measure on $[0,\infty) \times \mathcal{P}$ derived from M by retaining only the atoms on fibre $[0,\infty) \times \mathcal{P} \times \{1\}$. Then M_1 is a Poisson random measure with intensity $dt \otimes \kappa$. The event $\{\Pi(t)|_{\{1,3,\ldots,k+2\}} = \mathbf{1}_{\{1,3,\ldots,k+2\}}\}$ holds if only if

$$M_1([0,t] \times \{\Gamma \in \mathcal{P} : \Gamma|_{\{1,3,\ldots,k+2\}} \neq \mathbf{1}_{\{1,3,\ldots,k+2\}}\}) = 0.$$

Since M_1 is a Poisson random measure, we obtain that

$$\mathbb{P}_{\mathcal{F}_t}(\Pi(t)|_{\{1,3,\ldots,k+2\}} = \mathbf{1}_{\{1,3,\ldots,k+2\}}) = \exp(-t\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_{\{1,3,\ldots,k+2\}} \neq \mathbf{1}_{\{1,3,\ldots,k+2\}}\})).$$

Applying Theorem 3.1, we obtain

$$\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_{\{1,3,\ldots,k+2\}} \neq \mathbf{1}_{\{1,3,\ldots,k+2\}}\}) = \tilde{\Phi}(k).$$

Hence, $\mathbb{P}_{\mathcal{F}_t}(\Pi(t)|_{\{1,3,\dots,k+2\}} = \mathbf{1}_{\{1,3,\dots,k+2\}}) = \exp(-t\tilde{\Phi}(k))$. According to Corollary 3.13,

$$\mathbb{E}_{\mathcal{F}_t}(|\Pi_{(1)}(t)|^k) = \mathbb{P}_{\mathcal{F}_t}(\Pi(t)|_{\{1,3,\dots,k+2\}} = \mathbf{1}_{\{1,3,\dots,k+2\}}) = \exp(-t\tilde{\Phi}(k)),$$

for every $t > 0$ and $k \in \mathbb{N}$. Therefore $\Phi(q) = \tilde{\Phi}(q)$ for every $q > 0$. □

Corollary 3.15. *Let Π be a homogeneous partly exchangeable fragmentation process with erosion rates c_1, c_2, c_3 and dislocation measures ν_1, ν_2. Denote by $F^{(1)}(t) \in \mathcal{P}$ the partition such that* Frag $\left(\Pi_{(1)}(t-), F^{(1)}(t)\right) = \Pi_{(1)}(t)$. *$\xi^{(1)}$ is the associated surbordinator specified in Proposition* 3.14. *Let $(\Delta\xi_t^{(1)}, t \geq 0)$ be all the jumps of $\xi^{(1)}$. Then*

$$\Delta\xi_t^{(1)} \stackrel{d}{=} \Delta\xi_t^{(1),1} + \Delta\xi_t^{(1),2},$$

where

$$\Delta\xi_t^{(1),1} := -\Delta \log |\Pi_{(1)}(t)| \mathbf{1}_{\{F^{(1)}(t) \in \mathcal{P}^1\}},$$

$$\Delta\xi_t^{(1),2} := -\Delta \log |\Pi_{(1)}(t)| \mathbf{1}_{\{F^{(1)}(t) \in \mathcal{P}^2\}}.$$

$\xi^{(1),1}, \xi^{(1),2}$ are independent subordinators with Laplace exponents

$$\begin{aligned} \Phi_1(q) &= c_1 + c_3 q + \int_{\mathcal{S}^\downarrow} \sum_{i \geq 1} s_i^2 (1 - s_i^q) \nu_1(d\mathbf{s}), \\ \Phi_2(q) &= c_2 + \int_{\mathcal{S}^\downarrow} \left(s_0 + \sum_{i \geq 1} s_i (1 - s_i)(1 - s_i^q)\right) \nu_2(d\mathbf{s}), \end{aligned} \tag{3.26}$$

respectively. Their corresponding Lévy measures are

$$\begin{aligned} \Lambda_1(dx) &= \mathrm{e}^{-2x} \sum_{i=1}^{\infty} \nu_1(\{\mathbf{s} \in \mathcal{S}^\downarrow : -\log s_i \in dx\}), \\ \Lambda_2(dx) &= \mathrm{e}^{-x}(1 - \mathrm{e}^{-x}) \sum_{i=1}^{\infty} \nu_2(\{\mathbf{s} \in \mathcal{S}^\downarrow : -\log s_i \in dx\}), \end{aligned}$$

and Λ_2 is a finite measure.

Proof. According to Proposition 3.11 (ii), $\kappa(\cdot) = \kappa_1(\cdot \cap \mathcal{P}^1) + \kappa_2(\cdot \cap \mathcal{P}^2)$. Then the intensity of the Poisson random measure M_1 is

$$dt \otimes \kappa_1(\cdot \cap \mathcal{P}^1) + dt \otimes \kappa_2(\cdot \cap \mathcal{P}^2).$$

Hence the Poisson point process associated with random measure M_1 can be decomposed to two independent Poisson point process associated with random measures $M_1^1 := dt \otimes$

$\kappa_1(\cdot\cap\mathcal{P}^1)$ and $M_1^2 := dt\otimes\kappa_2(\cdot\cap\mathcal{P}^2)$, respectively. In fact, M_1^1, M_1^2 are the Poisson random measure associated with $\xi^{(1),1}, \xi^{(1),2}$, respectively. Therefore, we only need to check the Laplace exponents to complete the proof. Let $\mathcal{F}_t^1$ be the filtration of M_1^1. Due to Corollary 3.13 (i)

$$\begin{aligned}
\mathbb{E}_{\mathcal{F}_t^1}\left(|\Pi_{(1)}(t)|^k\right) &= \mathbb{E}_{\mathcal{F}_t^1}\left(|\Pi_{(1)}(t)|^k \mathbf{1}_{\{F^{(1)}(u)\in\mathcal{P}^1, u\le t\}}\right) \\
&= \mathbb{P}_{\mathcal{F}_t}\left(\{\Pi(t)|_{\{1,3,\ldots k+2\}} = \mathbf{1}_{\{1,3,\ldots k+2\}}\}\mathbf{1}_{\{F^{(1)}(u)\in\mathcal{P}^1, u\le t\}}\right) \\
&= \mathbb{P}_{\mathcal{F}_t^1}\left(\Pi(t)|_{\{1,3,\ldots k+2\}} = \mathbf{1}_{\{1,3,\ldots k+2\}}\right) \\
&= \mathbb{P}_{\mathcal{F}_t^1}\left(\Pi(t)|_{k+2} = \mathbf{1}_{[k+2]}\right)
\end{aligned}$$

and under $\mathcal{F}_t^1$, the event $\{\Pi(t)|_{k+2} = \mathbf{1}_{[k+2]}\}$ holds if and only if

$$M_1^1([0,t]\times\{\Gamma\in\mathcal{P}^1 : \Gamma|_{k+2}\neq\mathbf{1}_{[k+2]}\}) = 0.$$

Similar to the discussion in (i), we can obtain that

$$\begin{aligned}
\mathbb{E}_{\mathcal{F}_t^1}\left(|\Pi_{(1)}(t)|^k\right) &= \mathbb{P}_{\mathcal{F}_t^1}\left(\Pi(t)|_{k+2} = \mathbf{1}_{[k+2]}\right) \\
&= \exp\left(-t\kappa(\{\Gamma\in\mathcal{P}^1 : \Gamma|_{k+2}\neq\mathbf{1}_{[k+2]}\}\right) \\
&= \exp(-t\Phi_1(k))
\end{aligned}$$

for all $k\in\mathbb{N}$. Consequently, the Laplace exponent and Lévy measure of $\xi^{(1),1}$ is specified and due to $\Phi_2(k) = \Phi(k) - \Phi_1(k)$, the Laplace exponent and Lévy measure of $\xi^{(1),2}$ is specified too.

To finish the proof, let us check that Λ_2 is finite.

$$\begin{aligned}
\Lambda_2([0,\infty)) &= \int_0^\infty \mathrm{e}^{-x}(1-\mathrm{e}^{-x})\sum_{i\ge1}\nu_2(\{\mathbf{s}\in\mathcal{S}^\downarrow : -\log s_i\in dx\}) \\
&= \int_{S^\downarrow}\sum_{i=1}^\infty s_i(1-s_i)\nu_2(d\mathbf{s}) \\
&= \int_{S^\downarrow}\left(s_1(1-s_1)+\sum_{i=2}^\infty s_i(1-s_i)\right)\nu_2(d\mathbf{s}) \\
&\le \int_{S^\downarrow}\left(s_1(1-s_1)+\sum_{i=2}^\infty s_i\right)\nu_2(d\mathbf{s})
\end{aligned}$$

$$
\begin{aligned}
&\leq \int_{S^{\downarrow}} (s_1(1-s_1)+(1-s_1))\,\nu_2(d\mathbf{s}) \\
&\leq 2\int_{S^{\downarrow}} (1-s_1)\nu_2(d\mathbf{s}) \\
&< \infty.
\end{aligned}
$$

□

Now let us consider the tagged fragment containing i, $i \geq 2$. For $|\Pi_{(2)}|$, we obtain the same distribution as $|\Pi_{(1)}|$ due to the partly exchangeability, while for $|\Pi_{(n)}|, n = 3, 4, \ldots,$ the situation is more complicated.

Proposition 3.16. *For $n \geq 3$, let $\Pi_{(n)}$ be the fragment containing n in a partly exchangeable fragmentation process Π and define*

$$\xi_t^{(n)} := -\log|\Pi_{(n)}(t)|.$$

Denote by $\mathcal{F}^{\Pi}$ the filtration associated with Π. Let

$$\tau_{[n]} := \inf\left\{t > 0 : \Pi|_n(t) \neq \mathbf{1}_{[n]}\right\},$$

where $\Pi_{(n)}|_n$ be the fragment $\Pi_{(n)}$ restricted on $[n]$. Then the following hold:

(i) $\tau_{[n]} < \infty$, a.s. *and $(\xi_t^{(n)}, 0 < t < \tau_{[n]})$ is a.s. a killed subordinator with Laplace exponent Φ_*. Its killing rate is given by*

$$k^{(n)} = c_1 + 2c_2 + (n-2)c_3 + \int_{S^{\downarrow}} \Big(\sum_{i\geq 1}(s_i^2 - s_i^n)\Big)\nu_1(d\mathbf{s}) + \Big(s_0 + \sum_{i\geq 1} s_i(1-s_i)\Big)\nu_2(d\mathbf{s})$$

and Lévy measure by

$$\Lambda^{(n)}(dx) = \mathrm{e}^{-(n-2)x}\Lambda_1(dx).$$

(ii) *Conditionally on $\tau_{[n]}$, the probability that $\Pi_{(n)}(\tau_{[n]})$ is killed i.e. it is a singleton*

$$\mathbb{P}\left(\Pi_{(n)}(\tau_{[n]}) = (\{n\})\Big|\tau_{[n]}\right) = \frac{c_1 + c_3 + \int_{\mathcal{S}^{\downarrow}} s_0\nu(d\mathbf{s})}{k^{(n)}},$$

where ν specified in (3.37)*; conditionally on τ^n and $\Pi_{(n)}(\tau_{[n]})$ is not killed, the joint law of $\left(\#\Pi_{(n)}|_n(\tau_{[n]}), \Delta\xi^{(n)}_{\tau_{[n]}}\right)$ is*

$$
\begin{aligned}
&\mathbb{P}\left(\#\Pi_{(n)}|_n(\tau_{[n]}) = m, \Delta\xi^{(n)}_{\tau_{[n]}} \in dx\Big|\tau_{[n]}, \Pi_{(n)}(\tau_{[n]}) \neq (\{n\})\right) \\
= \;& (2c_2 + (n-3)c_3)\,\mathbf{1}_{\{m=n-1\}} + \binom{n-3}{m-3}\mathrm{e}^{-(m-2)x}(1-\mathrm{e}^{x})^{n-m}\Lambda_1(dx) \\
&+2\binom{n-3}{m-2}\mathrm{e}^{-(m-1)x}(1-\mathrm{e}^{x})^{n-m-1}\Lambda_2(dx) + \binom{n-3}{m-1}\mathrm{e}^{-mx}(1-\mathrm{e}^{x})^{n-m-2}\Lambda_*(dx),
\end{aligned}
$$

where $\Lambda_*(dx) = \mathrm{e}^{-x} \sum_{i=1}^{\infty} \nu(\{\mathbf{s} \in \mathcal{S}^{\downarrow} : -\log s_i \in dx\})$ *with*

(iii) *Conditionally given* $\tau_{[n]}$ *and that* $\Pi_{(n)}(\tau_{[n]})$ *is not killed with* $\#\Pi_{(n)}|_n(\tau_{[n]}) = m$,

$$\xi^{(n)}(\tau_{[n]} + t) - \xi^{(n)}(\tau_{[n]}) \stackrel{d}{=} \xi^{(m)}(t).$$

Proof. (i) According to the definition of $\tau_{[n]}$, we have $\Pi_{(1)}(t) = \Pi_{(n)}(t), t < \tau_{[n]}$. By the analysis in the proof of Proposition 3.14, the random measure M_1 on $[0,\infty) \times \mathcal{P}$ by retaining only the atoms on fibre $[0,\infty) \times \mathcal{P} \times \{1\}$ is a Poisson random measure with intensity $dt \times \kappa$. Let $\big((t, \Delta^{(1)}(\Gamma, 1)), t < \tau_{[n]}\big)$ be the collection of atoms on $[0,\infty) \times \mathcal{P}$ defined in (3.10) associated with Π. Note that the event $\tau_{[n]} > t$ is equivalent to $\{\Pi(t)|_n = \mathbf{1}_{[n]}\}$ which hold if and only if $(u, \Delta^{(1)}(\Gamma, 1)) \in [0,t] \times \{\Gamma \in \mathcal{P} : \Gamma|_n = \mathbf{1}_{[n]}\}, u < t$. Hence $\big((t, \Delta^{(1)}(\Gamma, 1)), t < \tau_{[n]}\big)$ can be described in terms of two independent Poisson point process on $\{\Gamma \in \mathcal{P} : \Gamma|_n = \mathbf{1}_{[n]}\}$ with intensity $\kappa(\cdot \cap \{\Gamma \in \mathcal{P} : \Gamma|_n = \mathbf{1}_{[n]}\})$ and $\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]}\}$ with intensity $\kappa(\cdot \cap \{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]}\})$ respectively. Hence we obtain that for any $A \in [0,\infty)$,

$$\Lambda^{(n)}(A) = \mathbb{P}(\Delta\xi_t^{(n)} \in A) = \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n = \mathbf{1}_{[n]}, |\Gamma_{(1)}| \in A\}.$$

Applying Theorem 3.1, the form of $\Lambda^{(n)}$ is specified.

As M_1 is Poisson random measure, we obtain

$$\mathbb{P}(\tau_{[n]} > t) = \exp(-t\kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]}\})) > 0,$$

by the σ-finite condition (3.6) of κ. Hence $\tau_{[n]} < \infty$ a.s. and $k^{(n)} = \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]}\})$ is the killing rate of $(\xi_t^{(n)}, 0 < t < \tau^n)$.

(ii) Conditional given τ^n, the law of $(\tau_{[n]}, \Delta^{(1)}(\Gamma, 1))$ is $\kappa(\cdot \cap \{\Gamma \in \mathcal{P} : \Gamma|_n \neq \mathbf{1}_{[n]}\})/k^{(n)}$. By Theorem 3.1, the result is straightforward.

(iii) According to the definition of partly exchangeable fragmentation process, conditionally on $\#\Pi_{(n)}|_n(\tau_{[n]}) = m$, the law of $\Pi_{(n)}(\tau_{[n]} + t)$ is the same as $\Pi_{(m)}|(t)$, which induce the result. □

3.2.7 Self-similar partly exchangeable fragmentations

Throughout this section, we consider a càdlàg process $\Pi = (\Pi(t), t \geq 0)$ with values in $\mathcal{P}$ and started from the trivial partition $\Pi(0) = 1_{\mathbb{N}}$ a.s. We further assume that the following requirements are fulfilled with probability 1:

$$\Pi(t) \text{ possesses asymptotic frequencies } |\Pi(t)| \text{ for all } t \geq 0, \tag{3.27}$$

and for every $i \in \mathbb{N}$, if we denote by $B_i(t)$ the block of $\Pi(t)$ which contains i, then

$$\text{the process } t \to B_i(t) \text{ has right-continuous paths.} \tag{3.28}$$

We shall write $(\mathcal{F}_t)_{t\geq 0}$ for the natural filtration of Π after the completion by null events.

Definition 3.6. Let Π be a càdlàg process satisfying (3.27) and (3.28). We call Π a self-similar partly exchangeable fragmentation process with index $\alpha \in \mathbb{R}$ if for every $t, t' \geq 0$, the conditional distribution of $\Pi(t + t')$ given $\mathcal{F}_t$ is that of the law of $\text{Frag}(\Pi(t), \Pi^{(\cdot)})$, where $\Pi^{(\cdot)} = (\Pi^{(i)}, i \in \mathbb{N})$ is a family of independent random partitions, such that for each $i \in \mathbb{N}$, $\Pi^{(i)}$ has the same distribution as $\Pi(t'|\Pi(t)_i|^{\alpha})$.

Similar to self-similar fragmentation process, the self-similar partly exchangeable fragmentation process can still be constructed by homogeneous partly exchangeable fragmentation process under appropriate time changing. We state our result in the following lemma.

Lemma 3.17. *Let Π be a self-similar partly exchangeable fragmentation process with index $\alpha \in \mathbb{R}$. $\Pi_{(i)}(t)$ denotes the block of $\Pi(t)$ that contains i. Let $\beta \in \mathbb{R}$, we define a sequence of time-changes*

$$\eta_{(i)}(t) = \inf\{u \geq 0 : \int_0^u |\Pi_{(i)}(r)|^{-\beta} dr > t\},\ t \geq 0, \imath \geq 1. \tag{3.29}$$

Let $\Pi^{\beta}(t) \in \mathcal{P}$ be the partition whose blocks are those of the partitions $\Pi_{(i)}(\eta_{(i)}(t)), i \geq 1$. Then the process $(\Pi^{\beta}(t), t \geq 0)$ is a self-similar partly exchangeable fragmentation process with index $\alpha + \beta$. Moreover, if $\beta = -\alpha$, Π^{β} will be a homogeneous partly exchangeable fragmentation process.

Proof. The part exchangeability and asymptotic frequency for $\Pi^{\beta}(t)$ is a direct result of that its blocks are $\Pi_{(i)}(\eta_{(i)}(t)), i \geq 1$. Also, the process $t \to \eta_{(i)}(t)$ is right-continuous for every i can be deduced for (3.29). Hence, the processes $t \to |\Pi^{\beta}(t)|$ are right-continuous a.s..

Now let us turn to show the self-similarity index of Π^β is $\alpha+\beta$ to complete the proof. Let $\mathcal{F}_t^\beta = \cup_{i\geq 1}\mathcal{F}_{\eta_{(i)}(t)}$ be the sigma-field generated by the process Π^β up to time t. Fix $t, t' \geq 0$ and pick $i \in \mathbb{N}$. Let $\Pi^\beta(t)_i$ be the block that contains i. For every $r \geq 0$, denote $C_i(r)$ be the block of $\sigma_{\Pi_{(i)}(\eta_i(t))}(\Pi^{(j)}(r))$ containing i. By construction of the partly exchangeable fragmentation operator, we have

$$\Pi_{(i)}(\eta_i(t)+r) = \Pi_{(i)}^\beta(t) \cap C_i(r).$$

Hence, we can deduce the identity

$$\eta_i(t+t') = \eta_i(t) + \lambda_{(i)}(t'),$$

where

$$\lambda_{(i)}(t') := \inf\{s \geq 0 : \int_0^s |\Pi_{(i)}^\beta(t) \cap C_i(r)|^{-\beta} dr > t'\}.$$

Now we denote $\bar{\Pi}_{(i)}^\beta(t)$ as the partition by deleting the smallest two elements of $\Pi_{(i)}^\beta(t)$ and obviously

$$|\Pi_{(i)}^\beta(\eta_i(t)) \cap C_i(r)| = |\bar{\Pi}_{(i)}^\beta(\eta_i(t)) \cap C_i(r)|.$$

According to part exchangeability, $\bar{\Pi}_{(i)}^\beta(\eta_i(t))$ is an exchangeable partition. Hence

$$|\Pi_{(i)}^\beta(\eta_i(t)) \cap C_i(r)| = |\bar{\Pi}_{(i)}^\beta(\eta_i(t)) \cap C_i(r)| = |\bar{\Pi}_{(i)}^\beta(\eta_i(t))||C_i(r)| = |\Pi_{(i)}^\beta(\eta_i(t))||C_i(r)|$$

by Corollary 2.5 [10]. So we deduce that

$$\lambda_{(i)}(t') := \inf\{s \geq 0 : \int_0^s |C_i(r)|^{-\beta} dr > t'|\Pi_{(i)}^\beta(t)|^\beta\}. \tag{3.30}$$

The conditional distribution of $\Pi^\beta(t+t')$ given $\mathcal{F}_t^\beta$ is that of the law of $\text{Frag}(\Pi^\beta(\text{t}), \Pi^{(\cdot)})$, where $\Pi^{(\cdot)} = (\Pi^{(j)}, j \in \mathbb{N})$ is a family of independent random partitions. As Π is a self-similar partly exchangeable fragmentation process with index α, $\Pi^{(j)}$ has the same distribution as $\Pi^\beta(t'|\pi_j|^{\alpha+\beta})$ given $\Pi^\beta(t) = (\pi_j, j \geq 1)$ according to (3.30). So Π^β is a self-similar partly exchangeable fragmentation process with index $\alpha+\beta$.

□

3.2.8 Discrete partly exchangeable fragmentation trees

There is natural relation between labelled trees and refining partition-valued processes. For a finite subset B of $\mathbb{N}$ with n elements, let $\mathbf{t}_B$ be a collection of subsets of B and also contains ROOT $\in \mathbf{t}$, such that

- $B \in \mathbf{t}$, we call B the common ancestor in $\mathbf{t}$;
- $\{i\} \in \mathbf{t}$ for all $i \in B$, the leaves of $\mathbf{t}$;
- for all $A, C \in \mathbf{t}$, either $A \cap C = \emptyset$, or $A \subseteq C$ or $C \subseteq A$.

We call t_B the tree equipped with the partition of B and denote by $\mathbb{T}_B$ the set of $\mathbf{t}_B$.

Definition 3.7. Let $(\pi(t), t \geq 0)$ be a refining process taking values on $\mathcal{P}_B$. If $\pi(0) = \mathbf{1}_B$ and $\pi(t) = \mathbf{0}_B$ for some finite $t > 0$. We define the associated tree $\mathbf{t}_\pi := \{\text{ROOT}\} \cup \{A \subseteq B : A \in \pi(t) \text{ for some } t \geq 0\}$.

Now consider a partly exchangeable fragmentation process Π, its restriction $\sigma_B(\Pi|_{\#B})$ on a finite subset B is of course a refining partition-valued process on finite set, where $\sigma_B : [\#B] \to B$ satisfies that $\sigma_B(u)$ is the u-th smallest element of B for $1 \leq u \leq \#B$. Hence, we can define a random labelled tree $T_{\sigma_B(\Pi|_{\#B})}$ associated with Π by Definition 3.7 and call the associated *discrete partly exchangeable fragmentation tree* with the partly exchangeable fragmentation process Π for all $B \subset \mathbb{N}$. Moreover, we define the *partly exchangeable splitting rule* on $\mathcal{P}_B \setminus \{\mathbf{1}_B\}$ associated with Π as the partly exchangeable probability measure on $\mathcal{P}_B$ induced by $\sigma_B(\Pi|_{\#B})$.

The reason that we use $\sigma_B(\Pi|_{\#B})$ as the restriction of Π on B instead of $\Pi|_B$ is as follows. Suppose there are n elements in B, we want that the fragmentation of B evolves under the same laws as that of $[n]$. For the first partition occurring in the partly exchangeable fragmentation $\Pi|_n$, we have two laws based on whether 1 and 2 are separated. For a finite set B, we should distinguish the two laws by whether the two smallest elements are separated. Hence we have to use σ_B mapping $[n]$ to B. However, if we use $\Pi|_B$, mismatch will occur unless B contains 1 and 2. For instance, $B = \{1, 3, 4\}$, and $\pi = (\{1, 2\}, \{3, 4, ...\}) \in \mathcal{P}^1$. $\pi|_3$=({1,2},{3}) while $\pi|_B = \{1\}, \{3, 4\}$, then $\pi|_B$ will have different probability mass from $\pi|_3$ as the smallest two elements 1 and 3 are separated in $\pi|_B$.

Now according to the definition of partly exchangeable fragmentation processes, every block splits independently and under the same law. Consequently, the discrete partly exchangeable fragmentation tree is Markov branching which means that every subtree is a small copy of the whole tree and evolves independently. This Markov branching property implies two things as follows. First, the probability of the whole labelled tree can be expressed in terms of the product of the splitting rules at all branch points. Second, the splitting rules at any branch point is the same as that at the first branch point. The discrete partly exchangeable fragmentation tree is determined by its splitting rule at the first branch point. We express the splitting rules in the following Lemma.

Lemma 3.18. *Let* $(\Pi(t), t \geq 0)$ *be a homogeneous partly exchangeable fragmentation process with splitting rate* κ*,* B *be a finite subset of* $\mathbb{N}$ *with* n *elements. Then*

(i)
$$p_B(\pi) = \frac{\kappa(\{\Gamma \in \mathcal{P} : \sigma_B(\Gamma|_n) = \pi\})}{\kappa(\{\Gamma \in \mathcal{P} : \sigma_B(\Gamma|_n) \neq \mathbf{1}_B\})}, \ \pi \in \mathcal{P}_B \setminus \{\mathbf{1}_B\}$$

defines a partly exchangeable partition probability function (PEPPF) for the discrete partly exchangeable fragmentation tree $T_{\sigma_B(\Pi|_{\#B})}$ *associated with* Π*.*

(ii) *More explicitly, for* $\pi \in \sigma_B(\mathcal{P}_n)$*,*

$$\begin{aligned} p_B(\pi) &= \frac{1}{Y_n}\Bigg(c_1 \delta_{\omega^{(1,2)}}(\sigma_B^{-1}(\pi)) \\ &\quad + c_2(\delta_{\epsilon^{(1)}}(\sigma_B^{-1}(\pi)) + \delta_{\epsilon^{(2)}}(\sigma_B^{-1}(\pi)) + c_3 \sum_{i=3}^{\infty} \delta_{\epsilon^{(i)}}(\sigma_B^{-1}(\pi)) \\ &\quad + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\sigma_B^{-1}(\pi) \cap \mathcal{P}^1)\nu_1(d\mathbf{s}) + \int_{\mathcal{S}^{\downarrow}} \kappa_{\mathbf{s}}(\sigma_B^{-1}(\pi) \cap \mathcal{P}^2)\nu_2(d\mathbf{s}) \Bigg) \end{aligned} \tag{3.31}$$

where Y_n *is the normalization constant and in the form of*

$$\begin{aligned} Y_n &= c_1 + 2c_2 + (n-2)c_3 \\ &\quad + \int_{S^{\downarrow}} \Big(\sum_{i \geq 1} (s_i^2 - s_i^n) \Big) \nu_1(d\mathbf{s}) + \Big(s_0 + \sum_{i \geq 1} s_i(1 - s_i) \Big) \nu_2(d\mathbf{s})). \end{aligned} \tag{3.32}$$

Proof. (i) Note that $\bar{q}_B$ defined above is the distribution of $\sigma_B(\Pi|_n(D_n))$, where $D_n = \inf\{t \geq 0 : \Pi|_n(t) \neq \mathbf{1}_n\}$. As D_n is the first time that $[n]$ is split into non-empty blocks, $\sigma_B(\Pi|_n(D_n))$ is the partition at the first branch point in $T_{\sigma_B(\Pi|_{\#B})}$. Hence, p_B defines PEPPF for $T_{\sigma_B(\Pi|_{\#B})}$. The partly exchangeability of $\bar{q}_B$ comes from that of κ in Theorem 3.1.

(ii) It suffices to show that the normalization constant is as specified. Note that if $\pi \in \sigma_B(\mathcal{P}_n^2)$ i.e. the two leaves with smallest labels are separated in the partition, then $\pi \neq \mathbf{1}_B$. According to Theorem 3.1,

$$\begin{aligned}
&\kappa(\{\Gamma \in \mathcal{P} : \sigma_B(\Gamma|_n) \in \sigma_B(\mathcal{P}_n^1) \setminus \{\mathbf{1}_B\}\}) \\
&= \kappa(\{\Gamma \in \mathcal{P} : \Gamma|_n) \in \mathcal{P}_n^1 \setminus \{\mathbf{1}_{[n]}\}\}) \\
&= c_1 + (n-2)c_3 + \sum_{\pi \in \mathcal{P}_n^1} \int_{S^\downarrow} \kappa_s(\pi)\nu_1(d\mathbf{s}) \\
&= c_1 + (n-2)c_3 + \int_{S^\downarrow} \left((\sum_{i\geq 1} s_i^2)(s_0 + \sum_{i\geq 1} s_i)^{n-2} - \sum_{i=1}^{\infty} s_i^n \right) \nu_1(d\mathbf{s}) \\
&= c_1 + (n-2)c_3 + \int_{S^\downarrow} \left(\sum_{i\geq 1} s_i^2 - s_i^n \right) \nu_1(d\mathbf{s})
\end{aligned}$$

Similarly

$$\kappa(\{\Gamma \in \mathcal{P} : \sigma_B(\Gamma|_n) \in \sigma_B(\mathcal{P}_n^2)\}) = \int_{S^\downarrow} \Big(s_0 + \sum_{i\geq 1} s_i(1-s_i) \Big) \nu_2(d\mathbf{s}).$$

Hence, the form of Y_n is as specified. □

If $B = [n]$ and block sizes of π in ranked order are $n_1, \ldots, n_k$, due to the partly exchangeability, (3.31) are actually two functions of the block sizes $n_1, ..., n_k$. We can define the following two quantities as follows

$$\begin{aligned}
p_n^1(n_1, \ldots, n_k) &:= p_n(\pi), \text{ for } \pi \in \mathcal{P}_n^1 \\
p_n^2(n_1, \ldots, n_k) &:= p_n(\pi), \text{ for } \pi \in \mathcal{P}_n^2.
\end{aligned}$$

So far, we are focusing on the labelled trees associated with partly exchangeable fragmentation process. If we remove all the labels, it is easier to see that the unlabelled trees are also Markov branching. Given $n \in \mathbb{N}$, we call $(n_1, ..., n_k), n_1 \geq \ldots n_k \geq 1, n_1 + \ldots + n_k = n$ a *composition* of n. For the unlabelled trees, the splitting rules are defined as the probability measures on the set of compositions for every n and denoted by $\{q_n, n \geq 1\}$. Loosely speaking, the splitting rule $q_n(n_1, ..., n_k)$ is the probability that an unlabelled tree with n leaves splits into k subtrees with $n_1, ..., n_k$ leaves respectively at the first splitting.

Corollary 3.19. *The splitting rules of the unlabelled discrete partly exchangeable trees* $\{T_n^\circ, n \geq 1\}$ *are of the following form: if* $(n_1, ..., n_k)$ *is a composition of* n *with* $k \geq 2$ *and* $n_1 \geq \ldots \geq n_k \geq 1$, *of which exactly* $m_i \geq 0$ *parts are equal to* i, $1 \leq i \leq n$,

$$\begin{aligned} & q_n(n_1, ..., n_k) \\ = & \frac{(n-2)!}{n_1! \ldots n_k! m_1! \ldots m_n!} \left((\sum_{i=1}^k n_i(n_i-1)) p_n^1(n_1, ..., n_k) + (\sum_{i=1}^k n_i(n-n_i)) p_n^2(n_1, ..., n_k) \right). \end{aligned}$$

Proof. This is a direct result from Lemma 3.18 because $q_n(n_1, \ldots, n_k)$ is just the sum of $p_n(\pi)$ for all π with block sizes $(n_1, \ldots, n_k)$. More explicitly, there are totally $\frac{(n-2)!}{n_1!\ldots n_k! m_1! \ldots m_n!}(\sum_{i=1}^k n_i(n_i-1))$ partitions of which 1 and 2 are in the same block and $\frac{(n-2)!}{n_1!\ldots n_k! m_1! \ldots m_n!}(\sum_{i=1}^k n_i(n-n_i))$ partitions of which 1 and 2 are separated. Hence, the result is straightforward. □

3.2.9 Consistency

For Markov branching trees, Haas, et. al. in [28] showed that consistency is equivalent to the following equation of splitting rules,

$$p_n(\pi) = p_{n+1}(\{\Gamma \in \mathcal{P}_{n+1} : \Gamma|_n = \pi\}) + p_{n+1}([n], \{n+1\}))p_n(\pi) \tag{3.33}$$

For labelled partly exchangeable fragmentation trees, p_n^1 and p_n^2 naturally fulfill (3.33) as they are the restrictions of measure κ on $\mathcal{P}_n$.

Corollary 3.20. *The family of labelled partly exchangeable fragmentation trees is consistent.*

Haas et al. [28] also showed that sampling consistency is equivalent to the following equation of splitting rules,

$$\begin{aligned} q_n(n_1, \ldots, n_k) &= \sum_{i=1}^k \frac{(n_i+1)(m_{n_i+1}+1)}{(n+1)m_{n_i}} q_{n+1}((n_1, \ldots, n_i+1, \ldots, n_k)^\downarrow) \\ & + \frac{m_1+1}{n+1} q_{n+1}(n_1, \ldots, n_k, 1) \\ & + \frac{1}{n+1} q_{n+1}(n, 1) q_n(n_1, \ldots, n_k), \end{aligned} \tag{3.34}$$

for Markov branching trees. For each splitting rule, we can define the associated exchangeable partition probability function EPPF as

$$p_n(n_1, \ldots, n_k) = \frac{n_1! \ldots n_k! m_1! \ldots m_n!}{n!} q_n(n_1, \ldots, n_k).$$

Then sampling consistency is equivalent to

$$p_n(n_1,\dots,n_k) = \frac{\sum_{i=1}^k p_{n+1}(n_1,\dots,n_i+1,\dots,n_k) + p_{n+1}(n_1,\dots,n_k,1)}{1-p_{n+1}(n,1)}. \tag{3.35}$$

The sampling consistent fragmentation trees are correspondent to fragmentation process can be expressed the sampling consistent splitting rules as an integral of the dislocation measure ν on $\mathcal{S}^{\downarrow}$ which satisfies (3.1).

Lemma 3.21 (Theorem 1,Proposition 3 of [28])**.** *Let Π^* be a homogeneous fragmentation process with splitting rate κ. Then the exchangeable partition probability function (EPPF) can be expressed as follows:*

$$p_B(\pi) = \frac{\kappa(\{\Gamma\in\mathcal{P} : \sigma_B(\Gamma|_n) = \pi\})}{\kappa(\{\Gamma\in\mathcal{P} : \sigma_B(\Gamma|_n) \neq \mathbf{1}_B\})},\ \pi\in\mathcal{P}_B\setminus\{\mathbf{1}_B\},$$

for all finite subset B of $\mathbb{N}$.

More precisely, sampling consistent splitting rules $(q_n, n\geq 2)$ are all of the following form: if $(n_1,\dots,n_k)$ is a composition of n with $k\geq 2$ parts, of which exactly $m_i\geq 0$ parts are equal to i, $1\leq i\leq n$,

$$q_n(n_1,\dots,n_k) = \frac{C_{n_1,\dots,n_k}}{Z_n}\left(nc\mathbf{1}_{\{k=2,n_2=1\}} + \int_{\mathcal{S}^{\downarrow}}\sum_{l=0}^{m_1}\binom{m_1}{l}\sum_{\substack{i_1,\dots,i_k\\ \text{distinct}}} s_0^l\prod_{j=1}^{k-l}s_{i_j}^{n_j}\nu(d\mathbf{s})\right), \tag{3.36}$$

where $Z_n = \int_{\mathcal{S}^{\downarrow}}(1-\sum_{i\geq1}s_i^n)\nu(d\mathbf{s})$ and $C_{n_1,\dots,n_k} = \binom{n}{n_1,\dots,n_k}/(m_1!\dots m_n!)$.

Corollary 3.22. *Let $(T_n^\circ, n\geq 1)$ be a family of unlabelled trees associated with a partly exchangeable fragmentation process Π with killing rate c_1, erosion rates $c_2, c_3\geq 0$ and two dislocation measures ν_1,ν_2. If $(T_n^\circ, n\geq 1)$ is sampling consistent then $c_1 = 0$, $c_2 = c_3$ and there exists a fragmentation process $\tilde{\Pi}$ associated with $(T_n^\circ, n\geq 1)$, whose erosion coefficient $c = c_2 = c_3$ and dislocation measure*

$$\nu(d\mathbf{s}) = \left(\sum_{i=1}^\infty s_i^2\right)\nu_1(d\mathbf{s}) + \left(s_0 + \sum_{i=1}^\infty s_i(1-s_i)\right)\nu_2(d\mathbf{s}). \tag{3.37}$$

Proof. For the simplicity, we use following notations. Denote

$$\begin{aligned}
N &= (n_1, \dots, n_k) \\
N_{k+1} &= (n_1, \dots, n_k, 1) \text{ and } N_i = (n_1, \dots, n_i + 1, \dots, n_k)^{\downarrow} \text{ for } 1 \le i \le k, \\
N_{k+1,k+1} &= (n_1, \dots, n_k, 2)^{\downarrow} \text{ and } N_{i,i} = (n_1, \dots, n_i + 2, \dots, n_k)^{\downarrow} \text{ for } 1 \le i \le k, \\
N_{k+1,k+2} &= (n_1, \dots, n_k, 1, 1) \text{ and } N_{i,k+1} = (n_1, \dots, n_i + 1, \dots, n_k, 1)^{\downarrow} \text{ for } 1 \le i \le k, \\
N_{i,j} &= (n_1, \dots, n_i + 1, \dots, n_j + 1, \dots, n_k)^{\downarrow} \text{ for } 1 \le i < j \le k \\
A_N &= \sum_{i=1}^{k} n_i(n_i - 1) \text{ and } B_N = \sum_{i=1}^{k} n_i(n - n_i)
\end{aligned}$$

Let p_n be the EPPF of T_n°. Applying Corollary 3.19 first and then applying the sampling consistency of T_n° twice gives us

$$p_n(N) \propto \sum_{i=1}^{k+1} \sum_{j=1}^{k+2} \left(A_N p^1_{n+2}(N_{i,j}) + B_N p^2_{n+2}(N_{i,j}) \right);$$

applying sampling consistency first, then applying Corollary 3.19 to $p_{n+1}(N_i)$ and finally applying sampling consistency again gives us

$$p_n(N) \propto \sum_{i=1}^{k+1} \sum_{j=1}^{k+2} \left((A_N + 2n_i) p^1_{n+2}(N_{i,j}) + (B_N + 2(n - n_i)) p^2_{n+2}(N_{i,j}) \right);$$

applying sampling consistency twice and then applying Corollary 3.19 to $p_{n+2}(N_{i,j})$ gives us

$$\begin{aligned}
p_n(N) \quad \propto \quad & \sum_{i=1}^{k+1} \sum_{j=1}^{k+2} \big((A_N + 2(n_i + n_j + \mathbf{1}_{\{i=j\}})) p^1_{n+2}(N_{i,j}) \\
& + (B_N + 2(2n - n_i - n_j + \mathbf{1}_{\{i \ne j\}})) p^2_{n+2}(N_{i,j}) \big).
\end{aligned}$$

Combining the above three formulas, we obtain

$$p_n(N) \propto \sum_{1 \le i \le k+1} p^1_{n+2}(N_{i,i}) + \sum_{1 \le i < j \le k+2} p^2_{n+2}(N_{i,j}). \tag{3.38}$$

Apply Lemma 3.21 to LHS of (3.38), we obtain $c_1 = 0$ and the specified erosion coefficient and dislocation measure. □

3.2.10 Building discrete trees with edge lengths

Given a labelled tree $\mathbf{t}$, we define $E(\mathbf{t})$ be the set of its edges. A *labelled tree with edge lengths* is a pair $\vartheta = (\mathbf{t}, \mathbf{e})$ for a labelled tree $\mathbf{t}$ and $\mathbf{e} = (e_i, i \in E(\mathbf{t})) \in (\mathbb{R}_+ \setminus \{0\})^{E(\mathbf{t})}$.

Call $\mathbf{t}$ the *skeleton* of ϑ and denote by $S(\vartheta)$. Such a tree is naturally equipped with a distance $d(v, w)$ on the set of its vertices, by adding the lengths of edges that appear in the unique path connecting v and w in the skeleton. The height of a vertex is its distance to the root. Let e_{root} be the length of the unique edge connected to the root, and for $e < e_{\text{root}}$ write $\vartheta - e$ for the tree with edge lengths that has same skeleton and same edge lengths as ϑ, but for the edge pointing outward the root which is assigned length $e_{\text{root}} - e$.

We also define an operation MERGE as follows. Let $n \geq 2$ and take $\vartheta_1, \vartheta_2, ..., \vartheta_n$ with leaves $(L_i^1, 1 \leq i \leq k_1), ..., (L_i^n, 1 \leq n \leq k_n)$ respectively. Let also $e > 0$. The tree with edge lengths MERGE$((\vartheta_1, ..., \vartheta_n); e)$ is defined by merging together the roots of $\vartheta_1, ..., \vartheta_n$ into a single vertex a, and by drawing a new edge root $\to a$ with length e.

For a labelled tree ϑ with edge-lengths and i vertices $v_1, ..., v_i$, define the subtree spanned by the root and $v_1, ..., v_i$ as follows. Its skeleton is $R(S(\vartheta), v_1, ..., v_i)$, where $S(\vartheta)$ is the skeleton of ϑ. Its edge lengths are given by the respective distances between this subset of vertices of the original tree.

Now we use the method introduced by Haas and Miermont [27] to construct the discrete tree with edge lengths. For $B \subset \mathbb{N}$, we define $R(B)$, a random variable taking values in $\mathbb{T}_{\#B}$ with leaves $L_1, ..., L_{\#B}$, as follows. Given a partly exchangeable fragmentation process $\Pi(t)$, let $D_i = \inf\{t \geq 0 : \{i\} \in \Pi(t)\}$ be the first time when $\{i\}$ disappears, i.e. is isolated in a singleton of $\Pi(t)$. For B a finite subset of $\mathbb{N}$ with at least two elements, let $D_B = \inf\{t \geq 0 : \#(B \cap \Pi(t)) \neq 1\}$ be the first time when the restriction of $\Pi(t)$ to B is non-trivial, i.e. has more than one block. By convention, $D_{\{i\}} := D_i$. For every $i \geq 1$, define $R(\{i\})$ as a single edge root $\to L_i$, and assign this edge the length D_i. For B with $\#B \geq 2$, let $B_1, ..., B_i$ be the non-empty blocks of $B \cap \Pi(D_B)$, arranged in increasing order of least element, and define a tree $R(B)$ recursively by

$$R(B) = \text{MERGE}((R(B_1) - D_B, ..., R(B_i) - D_B); D_B).$$

3.2.11 $\mathbb{R}$-trees and reduced trees

Let $B \subset \mathbb{N}$, a rooted tree with edge lengths is a pair $\vartheta = (\mathbf{t}, (\eta_e, e \in E(\mathbf{t})))$, where $\mathbf{t} \in \mathbb{T}_B$, $E(\mathbf{t})$ is the set of edges of $\mathbf{t}$, and $(\eta_e, e \in E(\mathbf{t})) \in (0, \infty)^{E(\mathbf{t})}$ are positive marks, that are interpreted as the lengths of the associated edges. The tree $\mathbf{t}$ is called the shape, and we let Θ_B be the set of trees with edge lengths whose shape is in $\mathbb{T}_B$.

We now introduce the $\mathbb{R}$-trees, following Evans et al. in [17]. A complete separable metric space (ι, d) is called $\mathbb{R}$-tree, if and only if it satisfies the following two conditions:

1. for all $x, y \in \iota$, there is an isometry $\varphi_{x,y} : [0, d(x, y)] \to \iota$ such that $\varphi_{x,y}(0) = x$ and $\varphi_{x,y}(1) = y$,
2. for every injective path $c : [0, 1] \to \iota$ with $c(0) = x, c(1) = y$, one has $c([0, 1]) = \varphi_{x,y}([0, d(x, y)])$.

We will consider rooted $\mathbb{R}$-trees i.e. they have a distinguished element denoted by ρ.

Two rooted $\mathbb{R}$-trees $(\iota, \rho, d), (\iota', \rho', d')$ are called equivalent if there is a bijective isometry between the two metric spaces that maps the root of one to the root of the other. We also denote by Θ the set of equivalence classes of compact rooted $\mathbb{R}$-trees. It is has been shown that in [17] that Θ is a Polish space when endowed with the so-called pointed Gromov-Hausdorff distance $\mathrm{d_{GH}}$, where by definition the distance $\mathrm{d_{GH}}((\iota, \rho), (\iota', \rho'))$ is equal to the infimium of the quantities

$$\delta(r, r') \vee \delta_{\mathcal{H}}(T, T'),$$

where $(T, r), (T', r')$ are isometric embeddings of $(\iota, \rho), (\iota', \rho')$ into a common metric space (M, δ), and where $\delta_{\mathcal{H}}$ is the Hausdorff distance between compact subsets of (M, δ). It is elementary that this distance does not depend on particular choices in the equivalent classes of (ι, ρ) and (ι', ρ'). We endow Θ with the associated Borel σ-algebra. In the sequel, by a slight abuse of notation, we will still call the elements of Θ rooted $\mathbb{R}$-trees, and we will denote them by ι, omitting the mention of the root and the distance d. Also, by a measure on an element $\iota \in \Theta$, we will mean an equivalence class of a 4-tuple (ι, ρ, d, μ), where we call (ι, ρ, d, μ) and $\iota', \rho', d', \mu')$ equivalent if there exists an isometry from (ι, ρ, d) to (ι', ρ', d') such that μ' is the push-forward of μ

We denote the range of $\varphi_{x,y}$ by $[[x, y]]$ and call the quantity $d(\rho, x)$ the height of x. We say that x is an ancestor of y whenever $x \in [[\rho, y]]$. We let $x \wedge y$ be the unique element in ι such that $[[\rho, x]] \cap [[\rho, y]] = [[\rho, x \wedge y]]$, and call it the highest common ancestor of x and y in ι. Denoted by ι_x, endowed with the distance of d, and rooted at x, the set of

$y \in \iota$ such that x is an ancestor of y, is in turn a rooted $\mathbb{R}$-tree called the fringe subtree of ι rooted at x.

We call an element $x \in \iota$, $x \neq \rho$, in a rooted $\mathbb{R}$-tree ι, a leaf if its removal does not disconnect ι, and let $\mathcal{L}(\iota)$ be the set of leaves of ι. On the other hand, we characterize an element of ι a branch point, if it has the form $x \wedge y$ where x is neither an ancestor of y nor vice-versa. Equivalently, we can define the branch point as the one disconnecting the $\mathbb{R}$-tree into three or more components when removed. We let $\mathcal{B}(\iota)$ be the set of branch points of ι.

For ι an $\mathbb{R}$-tree and $x_1, \ldots, x_k \in \iota$, we let $R(\iota, x_1, \ldots, x_k) = \cup_{i=1}^{k}[[\rho, x_i]]$ be the *reduced tree* associated with $\iota, x_1, \ldots, x_k$.

There is natural connection between the trees with edge lengths with shape in $\mathbb{T}_B$ and rooted $\mathbb{R}$-trees with n leaves with labels $L_1, \ldots, L_n$. If ι is a rooted $\mathbb{R}$-tree with $\rho \notin \mathcal{B}(\iota)$ and exactly n leaves labelled $L_1, \ldots, L_n$, then we consider the graph whose vertices are the set $V = \{\rho\} \cup \{\mathcal{L}(\iota)\} \cup \{\mathcal{B}(\iota)\}$, and such that two vertices x, y are connected by an edge if and only if $[[x, y]] \cap V = \{x, y\}$. Then resulting graph is a tree which naturally rooted at ρ, and the edge connecting x and y naturally inherits the length $\mathrm{d}(x, y) = |\mathrm{d}(\rho, x) - \mathrm{d}(\rho, y)|$. Also if $\mathbf{t}$ is an element of $\mathbb{T}_B$, one naturally puts edge lengths equal to 1 on each edge, and consider $\mathbf{t}$ as an $\mathbb{R}$-tree as well, the restriction of the distance of this $\mathbb{R}$-tree to the set of branch points, leaves and the root being the usual combinatorial distance on the vertices of $\mathbf{t}$.

3.2.12 Continuum random trees

Introduced by Aldous [2], a pair $(\mathcal{T}, \mu)$ is called a continuum tree if $\mathcal{T} \in \Theta$, and μ is a probability measure on $\mathcal{T}$ such that

1. μ is supported by the set $\mathcal{L}(\mathcal{T})$,
2. μ has no atom,
3. for every $x \in \mathcal{T} \backslash \mathcal{L}(\mathcal{T})$, $\mu(\mathcal{T}_x) > 0$.

A continuum random tree (CRT) is a random variable whose values are continuum trees, defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. To formalize this, we should endow the set of continuum trees with a σ-algebra. A natural way would be to use Evans' and Winter's separable and complete metric structure [18] on the space of 'weighted $\mathbb{R}$-trees', although we would have to incorporate the fact that our trees are rooted. For technical simplicity, we use the method of Aldous [3]. Let the space $\ell_1 = \ell_1(\mathbb{N})$ be the base space for defining CRTs. We endow the set of compact subsets of ℓ_1 with the Hausdorff metric, and the set of probability measures on ℓ_1 with any metric inducing the topology of weak convergence, so that the set of pairs (T, μ) where T is a rooted $\mathbb{R}$-tree embedded as a subset of ℓ_1 and μ is a measure on T, is endowed with the product σ-algebra.

A $\mathcal{P}$-valued fragmentation process $(\Pi(t), t \geq 0)$ is self-similar with index $\alpha \in \mathbb{R}$ if given $\Pi(t) = \pi$, the random variable $\Pi(t+s)$ has the same law as the random partition whose blocks are those of $\pi_i \cap \Pi^{(i)}(|\pi_i|^\alpha s), i \geq 1$, where $(\Pi^{(i)}, i \geq 1)$ is a sequence of i.i.d. copies of $(\Pi(t), t \geq 0)$. $(|\Pi(t)|^\downarrow, t \geq 0)$ is a self-similar fragmentation valued in $S^\downarrow$. Bertoin proved in [7] that the distribution of a $\mathcal{P}$-valued self-similar fragmentation is determined by a triple (α, c, ν), where ν is a dislocation measure on $S^\downarrow$. For this report, we are only interested in the case $c = 0$ and when ν is conservative and call (α, ν) characteristic pair.

A continuum random tree $(\mathcal{T}, \mu)$ is a self-similar tree with index $-\alpha < 0$ if for every $t \geq 0$, given $(\mu(\mathcal{T}_t^i), i \geq 1))$ where $\mathcal{T}_t^i, i \geq 1$ is the ranked order of connected components of the open set $\{x \in \tau : d(x, \rho(\tau)) > t\}$, the continuum random trees

$$\left(\mu(\mathcal{T}_t^1)^{-\alpha}\mathcal{T}_t^1, \frac{\mu(\cdot \cap \mathcal{T}_t^1)}{\mu(\mathcal{T}_t^1)}\right), \left(\mu(\mathcal{T}_t^2)^{-\alpha}\mathcal{T}_t^2, \frac{\mu(\cdot \cap \mathcal{T}_t^2)}{\mu(\mathcal{T}_t^2)}\right), \ldots$$

are i.i.d copies of $(\mathcal{T}, \mu)$, where $\mu(\mathcal{T}_t^i)^{-\alpha}\mathcal{T}_t^i$ is the tree that has the same shape as $\mathcal{T}_t^i$, but whose edge lengths are divided by $\mu(\mathcal{T}_t^i)^\alpha$. Haas and Miermont in [27] have shown us that there exists a self-similar continuum random tree characterized by such a pair $(-\alpha, \nu)$, which is also in one-to-one correspondence with self-similar fragmentation processes with characteristic pair $(-\alpha, \nu)$.

3.3 Embedding partly exchangeable fragmentation trees into CRTs

3.3.1 Mass-fragmentation and CRT

Let $(\Pi^\alpha(t), t \geq 0)$ be a self-similar exchangeable fragmentation process with index α. As $\Pi^\alpha(t)$ possesses asymptotic frequencies for all $t \geq 0$ a.s., Bertoin referred the process of ranked asymptotic frequencies $|\Pi^\alpha(\cdot)|^\downarrow$ as a self-similar *mass-fragmentation* [10]. From now on, we use the following simpler notation

$$X^\alpha(t) = (X_1^\alpha(t), \dots) := |\Pi^\alpha(t)|^\downarrow, \ t \geq 0,$$

for the mass-fragmentation process.

Let $\mathbf{s} \in \mathcal{S}^\downarrow$ and $(\mathbf{s}^{(i)}, i \in \mathbb{N})$ be a sequence in $\mathcal{S}^\downarrow$. We denote by $\mathrm{Frag}(\mathbf{s}, \mathbf{s}^{(\cdot)})$ and call the fragmentation of $\mathbf{s}$ by $\mathbf{s}^{(\cdot)}$, i.e. the mass-partition given by the decreasing rearrangement of the collection of real numbers $(s_i s_j^{(i)}, i, j \in \mathbb{N})$. Bertoin showed that the process $(X^\alpha(t), t \geq 0)$ is Markovian and its semigroup can be described as follows. For every $t, t' \geq 0$, the conditional distribution of $X^\alpha(t + t')$ given $X^\alpha(t) = \mathbf{s}$ is the law of $\mathrm{Frag}(\mathbf{s}, \mathbf{S}^{(\cdot)})$, where each $\mathbf{S}^{(i)}$ is distributed as $X^\alpha(t' s_i^\alpha)$ [10, Proposition 3.7].

Haas and Miermont [27, Theorem 1 and Proposition 1] showed that mass-fragmentation processes X^α of self-similar exchangeable fragmentation processes Π^α with index α, no erosion and an infinite conservative dislocation measure are in one-to-one correspondence with an self-similar CRTs $(\mathcal{T}, \mu)$. More precisely, writing $Y^\alpha(t)$ for the decreasing sequence of μ-masses of connected components of the open set $\{x \in \mathcal{T}, \mathrm{d}(\rho, x) > t\}$, the process Y^α has the same law as X^α.

3.3.2 A generic procedure to sample a leaf from a CRT

Before we show the sampling procedure, let us highlight some notations. Let Σ be a leaf in $(\mathcal{T}, \mu)$, we denote by $b^\alpha(u)$ the branch point on $[[\rho, \Sigma]]$ with $\mathrm{d}(\rho, b^\alpha(u)) = u$ and by $\mathcal{T}_\Sigma^\alpha(u) := \{x \in \mathcal{T}, \mathrm{d}(\rho, x \wedge \Sigma) > u\}$ the subtree of $\mathcal{T}$ containing Σ rooted at $b^\alpha(u)$. Let η_Σ be the self-similar time change with

$$\eta_\Sigma(t) = \inf\left\{u : \int_0^u \mu(\mathcal{T}_\Sigma^\alpha(v))^{-\alpha} dv > t\right\}. \tag{3.39}$$

Let $\mathcal{T}_\Sigma(t) := \mathcal{T}^\alpha_\Sigma(\eta_\Sigma(t))$, $b(t) = b^\alpha(\eta_\Sigma(t))$ and $X_\Sigma(t) := \mu(\mathcal{T}_\Sigma(t))$. Denote by $S^\Sigma(t) = \big(S^\Sigma(t)_1, \ldots\big) \in \mathcal{S}^\downarrow$ such that $X_\Sigma(t-)S^\Sigma(t)$ is the decreasing sequence of μ-masses of the connected components of $\{x \in \mathcal{T} : b_1(t) \in [[\rho, x]]\}$.

Lemma 3.23. *Let Σ^* be a leaf chosen according to μ from a CRT $(\mathcal{T}, \mu)$ with characteristic pair (α, ν). Then $\left(\widetilde{X}_{\Sigma^*}(t) := \left(\frac{X_{\Sigma^*}(t)}{X_{\Sigma^*}(t-)}, S^{\Sigma^*}(t)\right), t > 0\right)$ is a Poisson point process on $(0,1) \times \mathcal{S}^\downarrow$ with intensity measure $\tilde{\nu}$, where for any $A \subset (0,1)$ and $B \subset \mathcal{S}^\downarrow$,*

$$\tilde{\nu}^*(A \times B) = \int_B \sum_{i=1}^{\infty} s_i \mathbf{1}_{\{s_i \in A\}} \nu(d\mathbf{s}).$$

Proof. Let X^α be the self-similar mass-fragmentation process corresponding to CRT $(\mathcal{T}, \mu)$ and X be the homogeneous mass-fragmentation process of X^α through the self-similar time-change. Denote by Π a homogeneous exchangeable fragmentation process associated with X. Without loss of generality, we assume that $X_{\Sigma^*}(t) = |\Pi_{(1)}(t)|$ by the exchangeability. Let $\Pi^{(1)}(t)$ be the partition of $\mathbb{N}$ such that $\Pi_{(1)}(t) = \mathrm{Frag}(\Pi_{(1)}(t-), \Pi^{(1)}(t))$ and then $S^{\Sigma^*}(t) = |\Pi^{(1)}(t)|^\downarrow$. By the Poissonian construction of exchangeable fragmentation processes, $\Pi^{(1)}$ is a Poisson point process with intensity measure κ which is $\kappa(\cdot) = \int_{\mathcal{S}^\downarrow} \kappa_\mathbf{s}(\cdot)\nu(d\mathbf{s})$. Hence, X^{Σ^*} is a Poisson point process on $\mathcal{S}^\downarrow$ with intensity ν.

As Σ^* is chosen according to μ, the conditional probability of $X_{\Sigma^*}(t) = X^{\Sigma^*}(t)_i X_{\Sigma^*}(t-)$ given $X_{\Sigma^*}(t-)$ and $X^{\Sigma^*}(t)$ is $X^{\Sigma^*}(t)_i$. As X_{Σ^*} is a Poisson point process with intensity measure ν, $\widetilde{X}_{\Sigma^*}$ is a Poisson point process with intensity $\tilde{\nu}^*$. □

Definition 3.8. A function $P : (0,1) \times \mathcal{S}^\downarrow \to [0,1]$ that fulfills following two conditions

- $P(x, \mathbf{s}) = 0$ if $x \neq s_i$ for all $i \in \mathbb{N}$;
- $\sum_{i=1}^\infty P(s_i, \mathbf{s}) = 1$.

is called selection probability function (SPF). Conditionally given a mass-partition $\mathbf{s}$, we refer $m_i P(s_i, \mathbf{s})$ as the conditional probability of choosing a mass s_i, where m_i is the number of the same s_i in $\mathbf{s}$.

Remark 3.1. It is worth noting that the selection probability associated with a leaf chosen according to μ is $P(s_i, \mathbf{s}) = s_i$.

Let us first explain a generic procedure to sample a special leaf Σ based on the selection probability P from CRT $(\mathcal{T}, \mu, \rho, \mathrm{d})$, where ρ is the root and d is the distance metric.

Procedure 3.1. Let P be a selection probability fulfilling

$$\int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{\infty} (1 - s_i) P(s_i, \mathbf{s}) \nu(d\mathbf{s}) < \infty.$$

(i) We start from $(\mathcal{T}_1, \mu_1, \rho_1, \mathrm{d}_1) := (\mathcal{T}, \mu, \rho, \mathrm{d})$ and $i = 1$.

(ii) Sample a leaf $\Sigma^{(i)}$ from CRT $(\mathcal{T}_i, \mu_i, \rho_i, \mathrm{d}_i)$ according to the measure μ_i.

(iii) For $t > 0$, Let $b^{(i)}(t)$ be the branch point on $[[\rho_i, \Sigma^{(i)}]]$ with $\mathrm{d}_i(\rho, b^{(i)}(t)) = \eta_{i,\Sigma^{(i)}}^{-1}(t)$, where $\eta_{i,\Sigma^{(i)}}$ defined in (3.39) is the self-similar time-change associated with $\Sigma^{(i)}$ in $(\mathcal{T}_i, \mu_i, \rho_i, \mathrm{d}_i)$. Conditionally given $(\widetilde{X}^i_{\Sigma^{(i)}}(u), u \leq t)$, the process in $(\mathcal{T}_i, \mu_i, \rho_i, \mathrm{d}_i)$ defined in Lemma 3.23, we pick a subtree $\mathcal{T}_{b^{(i)}(t)}$ above $b^{(i)}(t)$ with probability

$$\mathbb{P}\left(\mathcal{T}_{i+1} = \mathcal{T}_{b^{(i)}(t)} \middle| \left(\widetilde{X}^i_{\Sigma^{(i)}}(u), u \leq t\right)\right) = P\left(\mu_i\left(\mathcal{T}_{b^{(i)}(t)}\right), S^{\Sigma^{(i)}}(t)\right) \prod_{u<t} P\left(\widetilde{X}_{\Sigma^{(i)}}(u)\right).$$

(iv) Let $\tau^{(i)} := \min\{u \geq 0 : \mathcal{T}_{i+1} \nsubseteq \mathcal{T}_{\Sigma^{(i)}}(u)\}$. Conditionally on we pick a subtree $\mathcal{T}_{b^{(i)}(\tau^{(i)})}$ above $b^{(i)}(\tau^{(i)})$ not containing Σ_i^*, we define

$$(\mathcal{T}_{i+1}, \mu_{i+1}, \rho_{i+1}, \mathrm{d}_{i+1}) = \left(\mathcal{T}_{b^{(i)}(\tau^{(i)})}, \frac{\mu_i|_{\mathcal{T}_{b^{(i)}(\tau^{(i)})}}}{\mu_i(\mathcal{T}_{b^{(i)}(\tau^{(i)})})}, b^{(i)}(\tau^{(i)}), \frac{\mathrm{d}_i|_{\mathcal{T}_{b^{(i)}(\tau^{(i)})}}}{\mu_i^{\alpha}(\mathcal{T}_{b^{(i)}(\tau^{(i)})})}\right).$$

(v) Increase i by 1 and go to (ii).

(vi) We obtain a sequence $(\Sigma^{(1)}, \Sigma^{(2)}, \ldots)$ in $(\mathcal{T}_1, \mathcal{T}_2, \ldots)$. As each $\mathcal{T}_{i+1}$ is a subtree of $\mathcal{T}_i$, $(\Sigma^{(1)}, \Sigma^{(2)}, \ldots)$ is also a sequence in $\mathcal{T}$. Note that $\mu(\cap_{i=1}^{\infty} \mathcal{T}_i) = \prod_{i=1}^{\infty} \mu_i(\mathcal{T}_i)$. As $(\mu_i(\mathcal{T}_i), i = 1, 2, \ldots)$ is a i.i.d sequence, $\mu(\cap_{i=1}^{\infty} \mathcal{T}_i) = 0$. As $\mathcal{T}$ is compact and $\cap_{i=1}^{\infty} \mathcal{T}_i \neq \emptyset$, there exits a leaf Σ such that $\cap_{i=1}^{\infty} \mathcal{T}_i = \{\Sigma\}$. Hence $\Sigma = \lim_{i\to\infty} \Sigma^{(i)}$.

Roughly speaking, this sampling procedure is that we climb up the spine $[[\rho, \Sigma_1^*]]$ and keep choosing subtrees until the first time we choose a subtree not containing Σ_1^*. We show that in the following proposition that there is a spinal subordinator associated with Σ.

Proposition 3.24. (i) $\left(\widetilde{X}_\Sigma(t) = \left(\frac{X_\Sigma(t)}{X_\Sigma(t-)}, S^\Sigma(t)\right), 0 \le t < \tau^{(1)}\right)$ *is a Poisson point process killed at time* τ_1^* *with intensity measure*

$$\tilde{\nu}^1(A \times B) = \int_B \sum_{i=1}^{\infty} s_i P(s_i, \mathbf{s}) \mathbf{1}_{\{s_i \in A\}} \nu(d\mathbf{s})$$

for any $A \subset (0,1)$ *and* $B \subset \mathcal{S}^\downarrow$ *and killing rate* $\int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty}(1 - s_i)P(s_i, \mathbf{s})\nu(d\mathbf{s})$. *Furthermore,* $\left(\widetilde{X}_\Sigma(t), t \ge 0\right)$ *is a Poisson point process with intensity*

$$\tilde{\nu}(A \times B) = \int_B \sum_{i=1}^{\infty} P(s_i, \mathbf{s}) \mathbf{1}_{\{s_i \in A\}} \nu(d\mathbf{s}).$$

(ii) *Let* $\xi_t^\Sigma := -\log X_\Sigma(t)$. *Then* ξ^Σ *is a pure jump subordinator with Laplace exponent*

$$\Phi_\Sigma(q) = \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} (1 - s_i^q) P(s_i, \mathbf{s}) \nu(d\mathbf{s})$$

and the associated Lévy measure is

$$\Lambda_\Sigma(\cdot) = \sum_{i=1}^{\infty} \int_{\{-\log s_i \in \cdot\}} P(s_i, \mathbf{s}) \nu(d\mathbf{s}). \tag{3.40}$$

Proof. (i) According to Procedure 3.1, $\widetilde{X}_\Sigma(t) = \widetilde{X}_{\Sigma^{(2)}}(t)$ for all $(0 \le t < \tau^{(1)})$. Note that by standard thinning properties of the Poisson point process of Lemma 3.23 the killed Poisson point process $\left(\widetilde{X}_{\Sigma^{(2)}}(t), 0 \le t < \tau^{(1)}\right)$ can be described in terms of two independent Poisson point processes of points $\Delta\widetilde{X}_{\Sigma^{(2)}}(t)$ when $\mathcal{T}_{\Sigma^{(2)}}(t) = \mathcal{T}_{\Sigma^{(1)}}(t)$ and when $\mathcal{T}_{\Sigma^{(2)}}(t) \ne \mathcal{T}_{\Sigma^{(1)}}(t)$. The intensity of the former is $\tilde{\nu}^1$ and that of the latter is $\tilde{\nu} - \tilde{\nu}^1$. Hence we get the result.

The intensity of the jump $\Delta\widetilde{X}_{\Sigma^{(2)}}(\tau^{(1)})$ at $\tau^{(1)}$ is $P(s_i, \mathbf{s})\nu(d\mathbf{s})$. Due to the self-similarity of $(\mathcal{T}, \mu)$, we get the total intensity of $\widetilde{X}_\Sigma$ is $P(s_i, \mathbf{s})\nu(d\mathbf{s})$.

(ii) Following the result in (i), it is obvious that ξ^Σ is a pure jump subordinator. The Laplace exponent and Lévy measure can be deduced from the intensity of $\widetilde{X}_\Sigma$.

□

Remark 3.2. In fact, Proposition 3.24 (ii) shows that for any Lévy measure Λ with the form in (3.40), we can find a leaf Σ from this generic sampling procedure with some selection probability P such that the Lévy measure of its spinal subordinator coincides with Λ.

3.3.3 Sample $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_k)$ from $(\mathcal{T}, \mu)$.

In this section, we express a special procedure to sample k leaves $\Sigma_1, \ldots, \Sigma_k$ from a CRT $(\mathcal{T}, \mu)$ with characteristic pair (α, ν), where $\nu(d\mathbf{s}) = \sum_{i=1}^{\infty} s_i^2 \nu_1(d\mathbf{s}) + \sum_{i=1}^{\infty} s_i(1-s_i)\nu_2(d\mathbf{s})$. Clearly ν_1, ν_2 are absolutely continuous with respect to ν. We denote their Radon-Nikodym derivatives by $d\nu_1/d\nu, d\nu_2/d\nu$ respectively.

Procedure 3.2. (i) Sample leaf Σ_1 from $(\mathcal{T}, \mu)$

We sample the leaf Σ_1 according to Procedure 3.1 by setting the selection probability as

$$P_{\Sigma_1}(s_i, \mathbf{s}) = s_i^2 \frac{d\nu_1}{d\nu}(\mathbf{s}) + s_i(1-s_i)\frac{d\nu_2}{d\nu}(\mathbf{s}).$$

(ii) Sample leaf Σ_2 given the CRT $(\mathcal{T}, \mu)$ and Σ_1.

(a) For $t > 0$, Let $b_1(t)$ be the point on $[[\rho, \Sigma_1]]$ with $\mathrm{d}(\rho, b_1(t)) = \eta_{\Sigma_1}(t)$. Conditionally given $(\widetilde{X}_{\Sigma_1}(u), u \leq t)$ and, we pick a leaf Σ_2

- from the subtree above $b_1(t)$ containing Σ_1 with probability

$$\mathbb{P}\left(\mathcal{T}_{\Sigma_2}(u) = \mathcal{T}_{\Sigma_1}(u), u \leq t \,\middle|\, \left(\widetilde{X}_{\Sigma_1}(u), u \leq t\right)\right) = \prod_{u \leq t} P^1_{\Sigma_2}\left(\widetilde{X}_{\Sigma_1}(u)\right);$$

- from a subtree $\mathcal{T}_{b_1(t)}$ above $b_1(t)$ not containing Σ_1 with probability

$$\begin{aligned} &\mathbb{P}\left(\mathcal{T}_{\Sigma_2}(u) = \mathcal{T}_{\Sigma_1}(u), u < t, \mathcal{T}_{\Sigma_2}(t) = \mathcal{T}_{b_1(t)} \neq \mathcal{T}_{\Sigma_1}(t) \,\middle|\, \left(\tilde{X}_{\Sigma_1}(u), u \leq t\right)\right) \\ = \;& P^2_{\Sigma_2}\left(\frac{\mu\left(\mathcal{T}_{b_1(t)}\right)}{X_{\Sigma_1}(t-)}, \widetilde{X}_{\Sigma_1}(t)\right) \prod_{u<t} P^1_{\Sigma_2}\left(\widetilde{X}_{\Sigma_1}(u)\right), \end{aligned}$$

where

$$P^1_{\Sigma_2}(s_i, \mathbf{s}) = \frac{s_i^2 d\nu_1/d\nu(\mathbf{s})}{P_{\Sigma_1}(s_i, \mathbf{s})} \text{ and } P^2_{\Sigma_2}(s_i, s_j, \mathbf{s}) = \frac{s_i s_j d\nu_2/d\nu(\mathbf{s})}{P_{\Sigma_1}(s_j, \mathbf{s})}.$$

(b) Let $\tau_1^1 := \min\{u \geq 0 : \mathcal{T}_{\Sigma_2}(u) \neq \mathcal{T}_{\Sigma_1}(u)\}$ the first time when Σ_2 is not picked from the same subtree as Σ_1. Conditionally given that Σ_2 is picked from a subtree $\mathcal{T}_{b_1(\tau_1^1)}$ above $b_1(\tau_1^1)$ not containing Σ_1, Σ_2 is picked in

$$\left(\mathcal{T}_{b_1(\tau_1^1)}, \mu(\cdot \cap \mathcal{T}_{b_1(\tau_1^1)})/\mu(\mathcal{T}_{b_1(\tau_1^1)})\right)$$

by the same procedure as in (i).

(iii) Sample one leaf Σ_{k+1} given CRT $(\mathcal{T},\mu)$ and $\Sigma_1,\dots,\Sigma_k$, for $k\geq 2$.

(a) We pick a leaf Σ^*_{k+1} in $(\mathcal{T},\mu)$ according to μ.

(b) Denote by b^*_{k+1} the branch point adjacent to Σ^*_{k+1} in the reduced tree $R(\mathcal{T},\Sigma_1,\dots,\Sigma_k,\Sigma^*_{k+1})$ and by $\mathcal{T}^{\Sigma^*_{k+1}}_{b^*_{k+1}}$ the subtree of CRT $(\mathcal{T},\mu)$ containing Σ^*_{k+1} and rooted at b^*_{k+1}.

- If b^*_{k+1} is a branch point of the reduced tree $R(\mathcal{T},\Sigma_1,\dots,\Sigma_k)$ or belongs to an inner edge of it, we pick the leaf Σ_{k+1} from $\mathcal{T}^{\Sigma^*_{k+1}}_{b^*_{k+1}}$ by the procedure in (i).
- If b^*_{k+1} belongs to a leaf edge adjacent to the leaf Σ_i of $R(\mathcal{T},\Sigma_1,\dots,\Sigma_k)$, we denote by $b^{i,k}$ the branch point in $R(\mathcal{T},\Sigma_1,\dots,\Sigma_k)$ adjacent to Σ_i and by $\mathcal{T}^{\Sigma_i}_{b^{i,k}}$ the subtree of CRT $(\mathcal{T},\mu)$ containing Σ_i and rooted at $b^{i,k}$. We pick the leaf Σ_{k+1} from $\mathcal{T}^{\Sigma_i}_{b^{i,k}}$ given leaf Σ_i by the procedure in (ii).

Corollary 3.25. (i) $\left(\widetilde{X}_{\Sigma_1}(t),t>0\right)$ *is a Poisson point process with intensity*

$$\tilde{\nu}_{\Sigma_1}(A\times B):=\int_B\sum_{i=1}^{\infty}\mathbf{1}_{\{s_i\in A\}}\left(s_i^2\nu_1(d\mathbf{s})+s_i(1-s_i)\nu_2(d\mathbf{s})\right).$$

Let $\left(\xi^{\Sigma_1}_t:=-\log X_{\Sigma_1}(t),t>0\right)$, *then* ξ^{Σ_1} *is a pure jump subordinator with Laplace exponent*

$$\Phi(q):=\int_{\mathcal{S}^{\downarrow}}\sum_{i=1}^{\infty}(1-s_i^q)\left(s_i^2\nu_1(d\mathbf{s})+s_i(1-s_i)\nu_2(d\mathbf{s})\right).$$

(ii) $\left(\widetilde{X}_{\Sigma_2}(t),0\leq t<\tau_1^1\right)$ *is a Poisson point process with intensity* $\int_B\sum_{i=1}^{\infty}s_i^2\mathbf{1}_{\{s_i\in A\}}\nu_1(d\mathbf{s})$ *killed at time* τ_1^1 *with killing rate* $\int_{\mathcal{S}^{\downarrow}}\sum_{i=1}^{\infty}s_i(1-s_i)\nu_2(d\mathbf{s})$.

Proof. (i) is a direct result of Proposition 3.24. For (ii), we only need to note that the intensity of $\left(\widetilde{X}_{\Sigma_2}(t),0\leq t<\tau_1^1\right)$ is

$$\int_B\sum_{i=1}^{\infty}\mathbf{1}_{\{s_i\in A\}}P^1_{\Sigma_2}(s_i,\mathbf{s})P_{\Sigma_1}(s_i,\mathbf{s})\nu(d\mathbf{s})=\int_B\sum_{i=1}^{\infty}\mathbf{1}_{\{s_i\in A\}}s_i^2\nu_1(d\mathbf{s})$$

and the intensity $\widetilde{X}_{\Sigma_2}(\tau_1^1)$ is

$$\int_B\sum_{i=1}^{\infty}\mathbf{1}_{\{s_i\in A\}}\sum_{j\neq i}P^2_{\Sigma_2}(s_j,s_i,\mathbf{s})P_{\Sigma_1}(s_i,\mathbf{s})\nu(d\mathbf{s})=\int_B\sum_{i=1}^{\infty}\mathbf{1}_{\{s_i\in A\}}s_i(1-s_i)\nu_2(d\mathbf{s}).$$

□

Proposition 3.26. *Let* Π *be a partly exchangeable fragmentation process with no erosion or killing rate and two conservative dislocation measures* ν_1 *and* ν_2*, where* ν_1 *fulfils* (3.7), (3.2) *with* $\nu_1(s_1 \leq 1-\epsilon) = \epsilon^{-\alpha}\ell(1/\epsilon)$ *and* ν_2 *fulfils* (3.1)*. Let* $(\mathcal{T},\mu)$ *be a CRT with characteristic pair* (α,ν)*, where* $\nu(d\mathbf{s}) = \sum_{i=1}^{\infty}\left(s_i^2\nu_1(d\mathbf{s}) + s_i(1-s_i)\nu_2(d\mathbf{s})\right)$*. For every integer* k*, let* $R([k])$ *be an* $\mathbb{R}$*-tree correspondent with the discrete tree with edge length built in Section* 3.2.10 *associated with* Π *and* $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ *be a reduced tree sampled from* $\mathcal{T}$ *according to Procedure 3.2. Then*

$$(R([k]), k\geq 1) \stackrel{d}{=} (R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k), k\geq 1)\,.$$

Proof. Given $R([k])$, $(R([j]), j<k)$ are uniquely determined; given $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$,

$$(R(\mathcal{T},\Sigma_1,\ldots,\Sigma_j), j<k)$$

are uniquely determined too. Hence we only need to show that for every $k\geq 1$,

$$R([k]) \stackrel{d}{=} R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k).$$

The self-similarity of CRT $(\mathcal{T},\mu)$ and the sampling procedure implies that $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ is a Markov branching tree with edge lengths. Hence $R([k])$ and $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ are identical in distribution if and only if (i) the distributions of the length of the edge adjacent to the root are the same and (ii) the PEPPF of the two are the same, provided these two are independent.

For (i), the length of the edge adjacent to root in $R([k])$ is the self-similar time of the first time that $\Pi|_k \neq \mathbf{1}_{[k]}$ under self-similar time-change $\eta_{(1)}^{-1}$. Corollary 3.25 ensures that the subordinator ξ^{Σ_1} associated with Σ_1 is distributed the same as $\xi^{(1)}$ specified in Proposition 3.14, which shows that $\eta_{(1)}^{-1}$ is the same as $\eta_{\Sigma_1}^{-1}$. Hence (i) is equivalent to the homogeneous time of the first split in $R([k])$ and $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ are the same. By the Poissonian construction of Π, the homogeneous first split time of $R([k])$ is an exponential random variable with parameter $\kappa(\{\Pi\in\mathcal{P} : \Pi|_k \neq \mathbf{1}_{[k]}\})$. Let $b_{[k]}$ be the branch point in $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ adjacent to ρ and $\tau_{[k]}$ be the associated homogeneous time. Note that $b_{[k]}$ is located in the spine $[[\rho, b_1^1]]$, where b_1^1 is the branch point where Σ_1 and Σ_2 are separated. For $k=2$, $\tau_1^1 = \tau_{[2]}$. So $\tau_{[2]}$ is an exponential random variable with parameter

$\int_{\mathcal{S}^\downarrow}\sum_{i=1}^\infty s_i(1-s_i)\nu_2(d\mathbf{s})=\kappa(\mathcal{P}^2)$ i.e. is the same as that of $R([2])$. For $k\geq 3$, note that restricted in the spine $[[\rho,b_1^1]]$, the selection probability of $\Sigma_3,\ldots,\Sigma_k$ are $P_{\Sigma_j}(s_i,\mathbf{s})=s_i$. By the thinning property of Poisson point process, $\tau_{[k]}$ is an exponential random variable with parameter

$$\int_{\mathcal{S}^\downarrow}\sum_{i=1}^\infty\left((s_i^2-s_i^k)\nu_1(d\mathbf{s})+s_i(1-s_i)\nu_2(d\mathbf{s})\right)=\kappa(\{\Pi\in\mathcal{P}:\Pi|_k\neq\mathbf{1}_{[k]}\}).$$

Hence (i) is straightforward

For (ii) PEPPF, we only need to examine the case $k\geq 3$. Note that $\tau_{[k]}\leq\tau_1^1$ and $\tau_{[k]}<\tau_1^1$ corresponds to the split that Σ_1,Σ_2 are in the same subtree above $b_{[k]}$ while $\tau_{[k]}=\tau_1^1$ corresponds to the split that Σ_1,Σ_2 are in different subtrees above $b_{[k]}$. By Corollary 3.25, the intensity and killing rate of $\left(\widetilde{X}_{\Sigma_2}(t),0\leq t<\tau_1^1\right)$ are $\int_B\sum_{i=1}^\infty s_i^2\mathbf{1}_{\{s_i\in A\}}\nu_1(d\mathbf{s})$ and $\int_{\mathcal{S}^\downarrow}\sum_{i=1}^\infty s_i(1-s_i)\nu_2(d\mathbf{s})$. Also the selection probability for $\Sigma_3,\ldots,\Sigma_k$ is $P(s_j,\mathbf{s})=s_j$ by Remark 1. Therefore the probability that the first split of $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ is $k_1,\ldots,k_r$ and Σ_1,Σ_2 are in the same subtree above $b_{[k]}$ with leaves k_1 is proportional to

$$\int_{\mathcal{S}^\downarrow}\sum_{\substack{i_1,\ldots,i_r\\ \text{distinct}}}s_{i_1}^{k_1-2}s_{i_2}^{k_2}\ldots s_{i_r}^{k_r}P_{\Sigma_2}^1(s_{i_1},\mathbf{s})P_{\Sigma_1}(s_{i_1},\mathbf{s})\nu(d\mathbf{s})=\int_{\mathcal{S}^\downarrow}\sum_{\substack{i_1,\ldots,i_r\\ \text{distinct}}}s_{i_1}^{k_1}s_{i_2}^{k_2}\ldots s_{i_r}^{k_r}\nu_1(d\mathbf{s});$$

the probability that the first split of $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ is $k_1,\ldots,k_r$ and Σ_1,Σ_2 are in the subtree above b_k with leaves k_1,k_2 respectively is proportional to

$$\int_{\mathcal{S}^\downarrow}\sum_{\substack{i_1,\ldots,i_r\\ \text{distinct}}}s_{i_1}^{k_1-1}s_{i_2}^{k_2-1}s_{i_3}^{k_3}\ldots s_{i_r}^{k_r}P_{\Sigma_2}^2(s_{i_2},s_{i_1},\mathbf{s})P_{\Sigma_1}(s_{i_1},\mathbf{s})\nu(d\mathbf{s})=\int_{\mathcal{S}^\downarrow}\sum_{\substack{i_1,\ldots,i_r\\ \text{distinct}}}s_{i_1}^{k_1}s_{i_2}^{k_2}\ldots s_{i_r}^{k_r}\nu_2(d\mathbf{s});$$

and both the two probabilities are independent from $b_{[k]}$. Hence we get that the PEPPF of $R(\mathcal{T},\Sigma_1,\ldots,\Sigma_k)$ is the same as that of $R([k])$. □

3.4 Some applications of embedding procedure

3.4.1 Special branch points estimation

Gnedin [22] introduced a constrained paintbox based on an random sequence. Let $1=G_0\geq G_1\geq\ldots\geq 0$ be an arbitrary nonincreasing random sequence and (I_n) be a sequence of i.i.d uniform random variables on $[0,1]$ independent of (G_k). Define a new sequence $\bar{I}_n$

as follows. Replace I_1 by G_1. Then replace the first entries which belong to $I_2, \ldots$ and hit $(0, G_1]$ by G_2. Inductively, when $G_1, \ldots, G_k$ get used, keep on screening uniforms until replacing the first points hitting $[0, G_k]$ by G_{k+1}. Eventually all G_k's will enter the resulting sequence. Let J_n be the number of intervals $(G_1, 1], (G_2, G_1], \ldots$ discovered by the n first terms of the sequence. Gnedin obtains the asymptotics of J_n when $G_k = Y_1 \ldots Y_k$, where for $k \geq 1$ the Y_k's are i.i.d on $[0,1]$ with finite logarithmic moments $\mathbb{E}(-\log Y_1)$ and $\mathrm{Var}(-\log Y_1)$. Here we loose the requirement of the finite logarithmic moments.

Lemma 3.27. *Let $N_t = \#\{n \geq 1 : X_1 + \ldots + X_n \leq t\}$ be the renewal process associated with i.i.d. $X_j > 0$. Then for all $p \in \mathbb{N}$*

$$\limsup_{t\to\infty} \mathbb{E}\left(\frac{N_t^p}{t^p}\right) < \infty.$$

Proof. The case $p = 1$ is the well-known Elementary Renewal Theorem for $\mathbb{E}(N_t)$. To prove the general case inductively, we define $q_j(t) = \mathbb{E}(N_t^j)$ and consider the induction hypothesis: for all $t \geq 0$,

$$q_j(t) \leq \sum_{k=1}^{j} a_{jk}(q_1(t))^k \qquad \text{for all } 1 \leq j \leq p-1 \text{ and some } a_{jk} \geq 0. \tag{3.41}$$

This is trivially true for $p = 1$ and $p = 2$. Let F be the distribution function of X_1 and U be the renewal function i.e.

$$U(t) = \begin{cases} 1 + q_1(t) & t \geq 0, \\ 0 & t < 0. \end{cases}$$

To show the induction step, we condition on the first renewal time and obtain the renewal equation

$$q_p = F + \sum_{j=1}^{p-1} \binom{p}{j} q_j * F + q_p * F$$

where $*$ denotes the convolution i.e. $V * W(t) := \int_0^t V(t-s)dW(s)$. Let $F^{*(m)}$ be the distribution function of $T_m := \sum_{i=1}^m X_i$. Note that

$$\begin{aligned} \mathbb{E}(N_t^p) &= \mathbb{E}\left((\mathbf{1}_{\{T_m<t\}})^p\right) = \mathbb{E}\left(\sum_{m_1=1}^{\infty} \cdots \sum_{m_p=1}^{\infty} \mathbf{1}_{\{T_{\max_{1\leq i\leq p} m_i} < t\}}\right) \\ &= \sum_{m_1=1}^{\infty} \cdots \sum_{m_p=1}^{\infty} \mathbb{P}\left(T_{\max_{1\leq i\leq p} m_i} < t\right) \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{m=1}^{\infty} m^p F^{*(m)}(t) \leq \sum_{n=0}^{\infty}\sum_{j=1}^{k}(nk+j)^p F^{*(nk+j)}(t) \\
&\leq \sum_{n=0}^{\infty} k((n+1)k)^p F^{*(nk)}(t) \leq \sum_{n=0}^{\infty} k((n+1)k)^p \left(F^{*(k)}(t)\right)^n .
\end{aligned}$$

Choose k large enough such that $F^{*(k)}(t) < 1$, then $q_p(t)$ is bounded for fix t. Therefore this renewal equation has a unique locally bounded solution

$$q_p = F * U + \sum_{j=1}^{p-1} \binom{p}{j} q_j * F * U,$$

and particularly $q_1 = F * U$. Then using the induction hypothesis and we obtain

$$\begin{aligned}
q_p &\leq F * U + \sum_{j=1}^{p-1} \binom{p}{j} \left(\sum_{k=1}^{j} a_{jk} q_1^k * F * U \right) \\
&\leq q_1 + \sum_{k=1}^{p} \left(\sum_{j=k}^{p-1} a_{jk} q_1^k * q_1 \right) .
\end{aligned}$$

The monotonicity of q_1 implies that for all $t \geq 0$,

$$q_1^k * q_1(t) = \int_0^t q_1^k(t-s)dq_1(s) \leq (q_1(t))^{k+1}.$$

Hence the induction proceeds.

Since the Elementary Renewal Theorem guarantees $\limsup_{t\to\infty} q_1(t)/t < \infty$, this completes the proof, since now

$$\limsup_{t\to\infty} \mathbb{E}\left(\frac{N_t^p}{t^p}\right) = \limsup_{t\to\infty} \frac{q_p(t)}{t^p} \leq \limsup_{t\to\infty} \sum_{k=1}^{p} a_{jk} \frac{(q_1(t))^k}{t^p} < \infty.$$

□

Lemma 3.28. *Let $G_k = Y_1 \ldots Y_k$, where for $k \geq 1$ the Y_k's are i.i.d. on $[0,1]$. Then*

$$\lim_{n\to\infty} \frac{J_n}{\log n} = \frac{1}{\mathbb{E}(-\log Y_1)}$$

when $\mathbb{E}(-\log Y_1) = \infty$, we set $1/\infty = 0$, and for every $p \geq 1$,

$$\limsup_{n\to\infty} \mathbb{E}\left(\left(\frac{J_n}{\log n}\right)^p\right) < \infty,$$

Proof. The case for $\mathbb{E}(-\log Y_1) < \infty$ has been implicitly shown in the proof of in Proposition 8 of [22]. We only need to handle the case when $\mathbb{E}(-\log Y_1) = \infty$. Define $J'_n := \max\{k : G_k > 1/n\} = \min\{k : \sum_{i=1}^{k}(-\log Y_i) \leq \log n\} - 1$. According to the renewal theory [16, Theorem 4.1, Chapter 3], $J'_n/\log n$ is asymptotic to 0 a.s. when $\mathbb{E}(-\log Y_1) = \infty$. Let $I_{1,n} < \ldots < I_{n,n}$ be the order statistics of $I_1, \ldots, I_n$. Define ζ_n by $I_{\zeta_n,n} < 1/n < I_{\zeta_{n+1},n}$. According to Gnedin's discussion, J'_n and ζ_n are independent, ζ_n is binomial$(n, 1/n)$ and $J_n \leq J'_n + \zeta_n$. By Markov inequality, we have for all $\epsilon > 0$,

$$\mathbb{P}(\zeta_n > \epsilon \log n) = \mathbb{P}\left(e^{2\zeta_n/\epsilon} > n^2\right) \geq \frac{\mathbb{E}\left(e^{2\zeta_n/\epsilon}\right)}{n^2} = \frac{1}{n^2}\left(1 + \frac{e^{2/\epsilon} - 1}{n}\right)^n.$$

Hence we have $\sum_{n=1}^{\infty} \mathbb{P}(\zeta_n > \epsilon \log n) < \infty$. Borel-Cantelli Lemma implies that

$$\lim_{n\to\infty} \zeta_n/\log n = 0 \text{ a.s..}$$

It gives us $\limsup_{n\to\infty} \frac{J_n}{\log n} = 0$ when $\mathbb{E}(-\log Y_1) = \infty$.

For every $p \geq 1$, note that

$$\mathbb{E}\left(\left(\frac{J_n}{\log n}\right)^p\right) \leq \mathbb{E}\left(\left(\frac{J'_n + \zeta_n}{\log n}\right)^p\right) \leq 2^{p-1}\left(\mathbb{E}\left(\left(\frac{J'_n}{\log n}\right)^p\right) + \mathbb{E}\left(\left(\frac{\zeta_n}{\log n}\right)^p\right)\right).$$

The first term is bounded due to Lemma 3.27 and the second term converge to 0 because the moments of of ζ_n are bounded. □

For all $1 \leq j \leq n$, denote by $V_j^1, \ldots, V_j^n$, the homogeneous times when each leaf $\Sigma_1, \ldots, \Sigma_n$ leaves the spine $[[\rho, \Sigma_j]]$. Obviously $V_j^j = \infty$ and $V_j^i = V_i^j$. Define

$$\tau_1^m = \max\{V_1^2, \ldots, , V_1^{m+1}\}, 1 \leq m \leq n-1,$$

and $\tau_1^0 = 0$, we call the branch point b_1^m associated with τ_1^m on the spine $[[\rho, \Sigma_1]]$ its *special branch point*. Denote by

$$L_n^{(1)} = \#\{\tau_1^m : 1 \leq m \leq n-1\}$$

the number of special branch points in $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_n)$ along $[[\rho, \Sigma_1]]$. Similarly we can define τ_2^m and $L_n^{(2)}$ the corresponding quantity for the spine $[[\rho, \Sigma_2]]$.

For $j \geq 3$, define

$$\tau_j^m = \begin{cases} \min_{1\leq i\leq m+1} \max\{V_j^1, \ldots, V_j^{i-1}, V_j^{i+1}, \ldots, V_1^{m+1}\}, & 1 \leq m \leq j-2, \\ \max\{V_j^1, \ldots, V_j^{j-1}, V_j^{j+1}, \ldots, V_1^{m+1}\}, & j-1 \leq m \leq n-1. \end{cases} \tag{3.42}$$

We call the branch point b_j^m associated with τ_j^m on the spine $[[\rho, \Sigma_j]]$ its *special branch point.* Denote by $L_n^{(j)} = \#\{\tau_j^m : 1 \leq m \leq n-1\}$ the number of special branch points in $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_n)$ along $[[\rho, \Sigma_j]]$.

Roughly speaking, the special branch point along $[[\rho, \Sigma_1]]$ and $[[\rho, \Sigma_2]]$ are the ones where the partition induced by splitting belongs to $\mathcal{P}^2$; the special branch points along $[[\rho, \Sigma_j]]$ are the ones where the leaf with smallest label or second smallest in all of subtrees above this branch point is not split into the same subtree as Σ_j.

Proposition 3.29. *Let ν_1 and ν_2 be two conservative dislocation measures , where ν_1 fulfils* (3.7), (3.2) *with $\nu_1(s_1 \leq 1-\epsilon) = \epsilon^{-\alpha}\ell(1/\epsilon)$ and ν_2 fulfils* (3.1). *Let $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_n)$ be an $\mathbb{R}$-tree sampled from a self-similar CRT $(\mathcal{T}, \mu)$ with index α and dislocation measure $\nu(d\mathbf{s}) := \sum_{i=1}^{\infty} s_i^2 \nu_1(d\mathbf{s}) + \sum_{i=1}^{\infty} s_i(1-s_i)\nu_2(d\mathbf{s})$ by the procedure specified in Section* 3.3.3. *Then*

(i) $L_n^{(j)}/(n^\alpha \ell(n)) \xrightarrow[n\to\infty]{a.s.} 0$;

(ii) *for every $p \geq 1$,*

$$\limsup_{n\to\infty} \mathbb{E}\left(\left(L_n^{(n)}/\log n\right)^p\right) < \infty.$$

(iii) *for every $p \geq 1$ and $x > 0$, there exist a constant C_p depending on p such that for all $1 \leq j \leq n$,*

$$\mathbb{P}\left(L_n^{(j)} > 2x\bar{\Lambda}(n^{-1})\right) < \frac{C_p}{x^p n^{\alpha p - 1}}.$$

Proof. (i) Let us consider $L_n^{(1)}$ first. If $\nu_2 = 0$, Σ_1 and Σ_2 are always in the same subtree in $\mathcal{T}$. Then for every n, $L_n^{(1)} = 1$ and the result is straightforward. In the following part, we consider $\nu_2 \neq 0$. Recall that $X_{\Sigma_1}(V_1^i), i \geq 2$ are the residual masses of the subtrees containing Σ_1 when Σ_i has left the spine $[[\rho, \Sigma_1]]$. Let $Y_i^{(1)}, i \geq 2$ be i.i.d. copies of $X_{\Sigma_1}(V_1^2)$. We claim that there exists a sequence of i.i.d. uniform random variables $I_i^{(1)}, i \geq 1$ on $[0,1]$ such that for every $n \geq 2$,

$$(X_{\Sigma_1}(V_1^2), \ldots, X_{\Sigma_1}(V_1^n)) \stackrel{d}{=} (\bar{I}_1^{(1)}, \ldots, \bar{I}_{n-1}^{(1)}), \tag{3.43}$$

where $G_k^{(1)} = Y_1^{(1)} \ldots Y_{k+1}^{(1)}$ for all $k = 1, \ldots$. The above formula holds for $n = 2$ as $G_1^{(1)} = Y_2^{(1)} \stackrel{d}{=} X_{\Sigma_1}(V_1^2)$. Suppose it holds for $n-1$. Let $J_{n-2}^{(1)}$ be the number of intervals $(G_1^{(1)}, 1], (G_2^{(1)}, G_1^{(1)}], \ldots$, each of which contains at least one in $\bar{I}_1^{(1)}, \ldots, \bar{I}_{n-2}^{(1)}$. Then according to the induction hypothesis,

$$G^{(1)}_{J^{(1)}_{n-2}} \stackrel{d}{=} \min\{X_{\Sigma_1}(V_1^2), \ldots, X_{\Sigma_1}(V_1^{n-1})\} = X_{\Sigma_1}(\tau_1^{n-2}).$$

Now conditionally on τ_1^{n-2}, Σ_n is sampled according to μ along the spine $[[\rho, b_1^{n-2}]]$ and by the same procedure as leaf Σ_2 in the subtree containing Σ_1 above b_1^{n-2} along the spine $[[b_1^{n-2}, \Sigma_1]]$. Hence conditionally on $X_{\Sigma_1}(V_1^n) \le X_{\Sigma_1}(\tau_1^{n-2})$, $X_{\Sigma_1}(V_1^n)$ is uniformly distributed on $(X_{\Sigma_1}(\tau_1^{n-2}), 1]$ and conditionally on $X_{\Sigma_1}(V_1^n) > X_{\Sigma_1}(\tau_1^{n-2})$, it is distributed the same as $X_{\Sigma_1}(\tau_1^{n-2})Y_n^{(1)} \stackrel{d}{=} G^{(1)}_{J^{(1)}_{n-2}} Y_n^{(1)}$. Therefor we prove (3.43). Hence we obtain

$$(J_1^{(1)}, \ldots, J_n^{(1)}) \stackrel{d}{=} (L_1^{(1)}, \ldots, L_n^{(1)}). \tag{3.44}$$

Lemma 3.28 ensures that the asymptotics of $L_n^{(1)}/\log n$ is bounded by $1/\mathbb{E}(-\log Y_2^{(1)})$. Therefore $L_n^{(1)}/(n^\alpha \ell(n))$ converge to 0 a.s. as $n \to \infty$.

For $L_n^{(j)}, j \ge 3$, note that Σ_j is the leaf with smallest or second smallest label in the subtree containing Σ_j above the branch point b_j^{j-1} and $\Sigma_1, \ldots, \Sigma_{j-2}$ leave the spine $[[\rho, \Sigma_j]]$ before b_j^{j-1}. The number of leaves in $\Sigma_1, \ldots, \Sigma_n$ that belong to the subtree containing Σ_j rooted at b_j^{j-1} is at most $n - j + 2$. Hence the number of special points on the spine $[[b_j^{j-1}, \Sigma_j]]$ of $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_j)$ will be no larger than $\widetilde{L}^{(1)}_{n-j+2}$ where $(\widetilde{L}^{(1)}_k, k \ge 1) \stackrel{d}{=} (L^{(1)}_k, k \ge 1)$. The number of special branch points on the spine $[[\rho, b_j^{j-1}]]$ is $L_j^{(j)}$. Therefore

$$L_n^{(j)} \le L_j^{(j)} + \widetilde{L}^{(1)}_{n-j+2} \le j + \widetilde{L}^{(1)}_{n-j+2}. \tag{3.45}$$

Hence the convergence for $L_n^{(j)}/(n^\alpha \ell(n))$ follows.

(ii) By Procedure 3.2, $X_{\Sigma_j}(\tau_j^1) = X_{\Sigma_j^*}(\tau_j^1)$ for all $j \ge 3$. Also, note that τ_j^1 is determined by $\Sigma_1, \Sigma_2, \Sigma_j^*$. As Σ_j^* is sampled according to μ in $\mathcal{T}$, we have

$$X_{\Sigma_j}(\tau_j^1) = X_{\Sigma_j^*}(\tau_j^1) \stackrel{d}{=} X_{\Sigma_3^*}(\tau_3^1) = X_{\Sigma_3}(\tau_3^1). \tag{3.46}$$

Let $Y_k^{(3)}, k \ge 1$ be i.i.d. copies of $X_{\Sigma_3}(\tau_3^1)$. and $G_k^{(3)} = Y_1^{(3)} \ldots Y_k^{(3)}, k \ge 1$. Let $I_k^{(3)}, k \ge 1$ be uniform random variables on $[0, 1]$ and $(\bar{I}_n^{(3)}, n \ge 1)$ be the random sequence defined before Lemma 3.28. Denote by $J_n^{(3)}$ the number of intervals $(G_1^{(3)}, 1], (G_2^{(3)}, G_1^{(3)}], \ldots$, each of which contains at least one in $\bar{I}_1^{(3)}, \ldots, \bar{I}_n^{(3)}$. We claim that for all $n \ge 3$, and every $x > 0$,

$$\mathbb{P}(L_n^{(n)} - 1 > x) \le \mathbb{P}(J^{(3)}_{n-2} > x). \tag{3.47}$$

This formula holds for $n = 3$ as $L_3^{(3)} - 1 \leq J_1^{(3)} = 1$. Suppose (3.47) holds for all $n \leq j-1$. For $n = j$, the first special branch point b_j^1 on the spine $[[\rho, \Sigma_j]]$ is also located on the spine $[[\rho, b_1^1]]$, where b_1^1 is the first special branch point on the spine $[[\rho, \Sigma_1]]$. For $i = 3, \dots, j-1$, let $\mathcal{T}_{\Sigma_i}(\rho, b_1^1)$ be the subtree of $\mathcal{T}$ containing Σ_i which is rooted on a branch point along the spine $[[\rho, b_1^1]]$ (including the branch point b_1^1). By Procedure 3.2 (iii), $\Sigma_i^* \in \mathcal{T}_{\Sigma_i}(\rho, b_1^1)$ otherwise there will be a branch point b on $[[\rho, b_1^1]]$ such that Σ_i is the leaf with smallest or second smallest label in the subtree containing b_1^1 rooted at b, which is impossible. Let N_j^1 be the number of leaves in $\Sigma_3, \dots, \Sigma_{j-1}$ belonging to the subtree containing Σ_j above branch point b_j^1. Then we have

$$N_j^1 = \#\{\Sigma_i \in \mathcal{T}_{\Sigma_j}(\tau_j^1) : i = 3, \dots, j-1\} = \#\{\Sigma_i^* \in \mathcal{T}_{\Sigma_j}(\tau_j^1) : i = 3, \dots, j-1\}.$$

As $\Sigma_3^*, \dots, \Sigma_{j-1}^*$ are sampled according to μ and $X_{\Sigma_j}(\tau_j^1) \stackrel{d}{=} X_{\Sigma_3}(\tau_3^1) \stackrel{d}{=} Y_1^{(3)}$ by (3.46),

$$\begin{aligned} \mathbb{P}(N_j^1 = m) &= \mathbb{E}\left(\binom{j-3}{m}(X_{\Sigma_j}(\tau_j^1))^m(1 - X_{\Sigma_j}(\tau_j^1))^{j-3-m}\right) \\ &= \mathbb{E}\left(\binom{j-3}{m}(Y_1^{(3)})^m(1 - Y_1^{(3)})^{j-3-m}\right), \end{aligned}$$

for all $0 \leq m \leq j-3$. On the other hand, let $\widetilde{N}_j^1$ be the number of $I_2^{(3)}, \dots, I_{j-2}^{(3)}$ hitting the interval $(0, G_1^{(3)}]$. Then

$$\mathbb{P}(\widetilde{N}_3^1 = m) = \mathbb{E}\left(\binom{j-3}{m}(Y_1^{(3)})^m(1 - Y_1^{(3)})^{j-3-m}\right) = \mathbb{P}(N_j^1 = m).$$

Let $L_j^{(j)}(\tau_j^1, \infty)$ be the number of special points along the spine $[[b_j^1, \Sigma_j]]$ in $R(\mathcal{T}, \Sigma_1, \dots, \Sigma_j)$ and $J_{j-2}^{(3)}(0, Y_1^{(3)}]$ be the number of intervals $(G_2^{(j)}, G_1^{(j)}], \dots$ hit by $\bar{I}_2^{(3)}, \dots, \bar{I}_{j-2}^{(3)}$. Now conditionally on $N_j^1 = m$, the subtree containing Σ_j above b_j^1 in the reduced tree $R(\mathcal{T}, \Sigma_1, \dots, \Sigma_j)$ has $m+2$ leaves if Σ_1 or Σ_2 is contained in the subtree $\mathcal{T}_{\Sigma_j}(b_j^1)$ of $\mathcal{T}$; or it will have $m+1$ leaves if neither of Σ_1 or Σ_2 is contained in this subtree. For the former, we obtain by induction hypothesis for

$$\begin{aligned} &\mathbb{P}\left(L_j^{(j)}(\tau_j^1, \infty) - 1 > x \middle| N_j^1 = m, \{\Sigma_1, \Sigma_2\} \cap \mathcal{T}_{\Sigma_j}(b_j^1) \neq \emptyset\right) \\ &= \mathbb{P}(L_{m+2}^{(m+2)} - 1 > x) \leq \mathbb{P}(J_m^{(3)} > x); \end{aligned}$$

for the latter,

$$\begin{aligned} &\mathbb{P}\left(L_j^{(j)}(\tau_j^1, \infty) - 1 > x \middle| N_j^1 = m, \{\Sigma_1, \Sigma_2\} \cap \mathcal{T}_{\Sigma_j}(b_j^1) = \emptyset\right) \\ &= \mathbb{P}(L_{m+1}^{(m+1)} - 1 > x) \leq \mathbb{P}(J_{m-1}^{(3)} > x) \leq \mathbb{P}(J_m^{(3)} > x). \end{aligned}$$

Then

$$\mathbb{P}\left(L_j^{(j)}(\tau_j^1,\infty)-1>x\middle|N_j^1=m\right)\le \mathbb{P}(J_m^{(3)}>x)=\mathbb{P}\left(J_{j-2}^{(3)}(0,Y_1^{(3)}]>x\middle|\widetilde{N}_3^1=m\right).$$

As $(N_j^1, L_j^{(j)}(\tau_j^1,\infty)-1)\stackrel{d}{=}(\widetilde{N}_3^1, J_{j-2}^{(3)}(0,Y_1^{(3)}])$, we obtain that

$$\begin{aligned}\mathbb{P}\left(L_j^{(j)}-1>x\right) &= \mathbb{E}\left(\mathbb{P}\left(L_j^{(j)}(\tau_j^1,\infty)-1>x-1\middle|N_j^1\right)\right)\\ &\le \mathbb{E}\left(\mathbb{P}\left(J_{j-2}^{(3)}(0,Y_1^{(3)}]>x-1\middle|\widetilde{N}_3^1\right)\right)\\ &= \mathbb{P}\left(J_{j-2}^{(3)}>x\right).\end{aligned}$$

Now (3.47) is clear and we can deduce that for every $p\ge 1$,

$$\mathbb{E}\left((L_n^{(n)}-1)^p\right)\le \mathbb{E}\left((J_{n-2}^{(3)})^p\right).$$

Applying Lemma 3.28, the result in (ii) is clear.

(iii) (3.45) implies that for every $x>0$

$$\begin{aligned}\mathbb{P}\left(L_n^{(j)}>2x\bar{\Lambda}(n^{-1})\right) &\le \mathbb{P}\left(L_j^{(j)}>x\bar{\Lambda}(n^{-1})\right)+\mathbb{P}\left(\widetilde{L}_{n-j+2}^{(1)}>x\bar{\Lambda}(n^{-1})\right)\\ &\le \frac{\mathbb{E}\left((L_j^{(j)})^p\right)}{x^p(\bar{\Lambda}(n^{-1}))^p}+\frac{\mathbb{E}\left((\widetilde{L}_{n-j+2}^{(1)})^p\right)}{x^p(\bar{\Lambda}(n^{-1}))^p}.\end{aligned}$$

The last line is obtained by Markov inequality. Lemma 3.28 and the result in (ii) give the upper bound for the first probability, while (3.44) together with Lemma 3.28 gives the upper bound of the second one. As $\bar{\Lambda}(y)\sim y^{-\alpha}\ell(y), y\downarrow 0$, the result in (iii) follows.

□

3.4.2 The convergence of reduced discrete partly exchangeable fragmentation tree

Let T_n be a discrete partly exchangeable tree with unit edge lengths. In this section, we shall show the properly scaled reduced tree $R(T_n,[k])$ with unit edge length will converge to $R(\mathcal{T},\Sigma_1,\dots,\Sigma_k)$.

To prove a discrete tree with edge length, ϑ_n, converges to another discrete tree with edge lengths, ϑ, it is sufficient to show that the shape of ϑ_n is eventually that of ϑ and the edge lengths converge pointwise. As the shape of $R(T_n,[k])$ is exactly that of $R([k])$. Proposition 3.26 guarantee the identity in the distribution of shapes of $R(T_n,[k])$ and

$R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_k)$. Hence, we only need to obtain the convergence of edge lengths in this section.

Lemma 3.30 ([23], Lemma 8 in [28]). *Let $\xi = (\xi_t, t \geq 0)$ be a pure jump subordinator with Lévy measure Λ satisfying $\Lambda([x, \infty)) = x^{-\alpha}\ell(1/x)$, $x \downarrow 0$. Let $V_1, V_2, \ldots$ be a sequence of non-negative random variables which conditionally given ξ are independent and identically distributed with*

$$\mathbb{P}(V_i > v|\xi) = \exp(-\xi_v), \ v \geq 0,$$

and for $s \geq 0$ let

$$K_n(s) := \#\{V_i : 1 \leq i \leq n, V_i \leq s\}.$$

Then

$$\lim_{n\to\infty} \sup_{0\leq s\leq\infty} \left| \frac{K_n(s)}{n^{-\alpha}\ell(n)} - \Gamma(1-\alpha) \int_0^s \exp(-\alpha\xi_v)dv \right| = 0 \text{ a.s.}$$

and hence for every random variable S with values in $[0, \infty]$

$$\lim_{n\to\infty} \frac{K_n(S)}{n^{-\alpha}\ell(n)} = \Gamma(1-\alpha) \int_0^S \exp(-\alpha\xi_v)dv. \text{ a.s..}$$

We simply denote $K_n = K_n(\infty)$.

Proposition 3.31. *Let Π be a partly exchangeable fragmentation process Π with zero killing rate or erosion rates and two conservative dislocation measures ν_1, ν_2, where ν_1 fulfills (3.7), (3.2) with $\nu_1(s_1 \leq 1-\epsilon) = \epsilon^{-\alpha}\ell(1/\epsilon)$ and ν_2 fulfills (3.1). $(T_n)_{n\geq 1}$ is the associated family of labelled discrete partly exchangeable fragmentation trees with unit edge length. Let $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_k)$ be an $\mathbb{R}$-tree sampled from self-similar CRT $(\mathcal{T}, \mu)$ with index α and dislocation measure $\nu(d\mathbf{s}) := \sum_{i=1}^{\infty} s_i^2 \nu_1(d\mathbf{s}) + \sum_{i=1}^{\infty} s_i(1-s_i)\nu_2(d\mathbf{s})$ by the procedure specified in Section 3.3.3. Then*

$$n^{-\alpha}\ell(n)^{-1} R(T_n, [k]) \xrightarrow[n\to\infty]{a.s.} \Gamma(1-\alpha)\mathcal{R}_k$$

in the Gromov-Hausdorff sense, where $(\mathcal{R}_k, k \geq 1) \stackrel{d}{=} (R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_k), k \geq 1)$.

Proof. By Proposition 3.26, the shape of T_n is distributed the same as that of $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_n)$, we shall consider T_n as an $\mathbb{R}$-tree by assigning unit edge lengths to $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_n)$. We start from the one-leaf trees $R(\mathcal{T}, \Sigma_1)$ and $R(T_n, \{1\})$. Denote by D_1^n the length of

$R(T_n, 1)$. Let B_1^m be the branch point in T_n corresponding to b_1^m and $D_1^{n,m}$ is the distance from B_1^{m-1} to B_1^m (B_1^0 is the root of T_n).

(i) If $\nu_2 = 0$, which means that Σ_1 and Σ_2 are always in the same subtree in $\mathcal{T}$, then $\tau_1^2 = \ldots = \tau_1^{n-1} = \infty$. Conditionally on the subordinator ξ^{Σ_1} associated with leaf Σ_1, the leaves $\Sigma_3, \ldots, \Sigma_n$ are sampled according to μ along the spine $[[\rho, \Sigma_1]]$. Hence applying Lemma 3.30, the convergence result is straightforward.

(ii) Let us consider the case that $\nu_2 \neq 0$. In this case $\tau_1^m < \infty$ for every $m \in \mathbb{N}$. Recall Procedure 3.2, for $i \geq 3$, we sample a leaf Σ_i^* according to μ. If $\Sigma_1 \wedge \Sigma_i^*$ is located on the spine $[[\rho, b_1^{i-2}]]$, then we sample Σ_i in the subtree containing Σ_i^* rooted at $\Sigma_1 \wedge \Sigma_i^*$; if $\Sigma_1 \wedge \Sigma_i^*$ is located on the path $[[b_1^{i-2}, \Sigma_1]]$, we sample Σ_i as the leaf with second smallest label in the subtree containing Σ_1 rooted at b_1^{i-2}. Hence, $\Sigma_1 \wedge \Sigma_i = \Sigma_1 \wedge \Sigma_i^*$ if $\Sigma_1 \wedge \Sigma_i^*$ is located on the spine $[[\rho, b_1^{i-2}]]$. Let $V_1^{i,*}$ be the homogeneous time when Σ_i^* leaves the spine $[[\rho, \Sigma_1]]$. Then we have following identity:

$$V_1^i = V_1^{i,*}\mathbf{1}_{\{V_1^{i,*} \leq \tau_1^{i-2}\}} + V_1^i \mathbf{1}_{\{V_1^{i,*} > \tau_1^{i-2}\}}.$$

Set $\tau_1^0 = 0$. The above formula implies that

$$\begin{aligned}
D_1^n - 1 &= \#\{V_1^i, 2 \leq i \leq n\} \\
&= \sum_{m=1}^{n-1} \#\{V_1^i : \tau_1^{m-1} < V_1^i \leq \tau_1^m, 2 \leq i \leq n\} \\
&= \sum_{m=1}^{n-1} \mathbf{1}_{\{\tau_1^{m-1} < \tau_1^m\}} + \sum_{m=1}^{n-2} \#\{V_1^i : \tau_1^{m-1} < V_1^i < \tau_1^m, m+2 \leq i \leq n\} \\
&= L_n^{(1)} + \sum_{m=1}^{n-2} \#\{V_1^{i,*} : \tau_1^{m-1} < V_1^{i,*} < \tau_1^m, m+2 \leq i \leq n\} \\
&\leq L_n^{(1)} + \sum_{m=1}^{n-2} \#\{V_1^{i,*} : \tau_1^{m-1} < V_1^{i,*} < \tau_1^m, 3 \leq i \leq n\} \\
&\leq L_n^{(1)} + \#\{V_1^{i,*} : 0 < V_1^{i,*} \leq \tau_1^{n-2}, 3 \leq i \leq n\} \\
&= L_n^{(1)} + K_{n-2}^{(1)}(\tau_1^{n-2}), \qquad (3.48)
\end{aligned}$$

where $K_{n-2}^{(1)}(s) := \#\{V_1^{i,*} : 0 < V_1^{i,*} \leq s, 3 \leq i \leq n\}$. As $\Sigma_i^*, i \geq 3$ is sampled according to μ, $V_1^{i,*}$ fulfils the condition in Lemma 3.30. Applying Proposition 3.29 and Lemma 3.30, we obtain that

$$\limsup_{n\to\infty} \frac{D_1^n}{n^\alpha \ell(n)} \leq \Gamma(1-\alpha) \int_0^\infty \exp(-\alpha \xi_t^{\Sigma_1}) dt \text{ a.s..} \qquad (3.49)$$

On the other hand, for any fixed $l \geq 3$ and for every $n \geq l+1$,

$$\begin{aligned} D_1^n - 1 &\geq \#\{V_1^i : 0 < V_1^i \leq \tau_1^l, 2 \leq i \leq n\} \\ &\geq \#\{V_1^i : 0 < V_1^i \leq \tau_1^l, l+1 \leq i \leq n\} \\ &= \#\{V_1^{i,*} : 0 < V_1^{i,*} \leq \tau_1^l, l+1 \leq i \leq n\}. \end{aligned}$$

As $\nu_2 \neq 0$, the leaf with second smallest label in any subtree containing Σ_1 will always leave the spine $[[\rho, \Sigma_1]]$ at some finite time. Hence $\lim_{n\to\infty} L_n^{(1)} = \infty$. Also, by Procedure 3.2, $\tau_1^l - \tau_1^{l-1}|_{\tau_1^l \neq \tau_1^{l-1}} \stackrel{d}{=} \tau_1^1$. Therefore $\lim_{l\to\infty} \tau_1^l = \infty$. Then we obtain

$$\begin{aligned} \liminf_{n\to\infty} \frac{D_1^n}{n^\alpha \ell(n)} &\geq \sup_{l\geq 3} \liminf_{n\to\infty} \frac{\#\{V_1^{i,*} : 0 < V_1^{i,*} \leq \tau_1^l, l+1 \leq i \leq n\}}{n^\alpha \ell(n)} \\ &= \Gamma(1-\alpha) \int_0^\infty \exp(-\alpha \xi_t^{\Sigma_1}) dt. \end{aligned} \tag{3.50}$$

Combining (3.49) and (3.50), the convergence for D_1^n follows.

Next, assume that the conclusion holds for $1, \ldots, k$, we show that it then holds for $k+1$. Consider the branch point $b_{[k+1]}$ adjacent to ρ in $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_{k+1})$ Let $D_{[k+1]} = \mathrm{d}(\rho, b_{[k+1]})$. Similarly, we denote by $\tau_{[k+1]}$ the homogeneous time for $D_{[k+1]}$ i.e.

$$D_{[k+1]} = \eta_{\Sigma_1}^{-1}(\tau_{[k+1]}) = \int_0^{\tau_{[k+1]}} \exp(-\alpha \xi_t^{\Sigma_1}) dt.$$

Let $D_{[k+1]}^n$ be the height of the branch point adjacent to the root in $R(T_n, [k+1])$. Then $D_{[k+1]}^n - 1$ is the number of distinct branch points of $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_n)$ belonging to the root edge of $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_{k+1})$. So

$$D_{[k+1]}^n - 1 = \#\{\Sigma_1 \wedge \Sigma_i, k+2 \leq i \leq n, d(\rho, \Sigma_1 \wedge \Sigma_i) < D_{[k+1]}\}.$$

By the same argument as for $k = 1$,

$$D_{[k+1]}^n - 1 = \#\{V_1^{i,*} : V_1^{i,*} \leq \tau_{[k+1]}, k+2 \leq i \leq n\}.$$

Now applying Lemma 3.30 yields

$$n^{-\alpha} \ell(n)^{-1} D_{[k+1]}^n \xrightarrow[n\to\infty]{a.s.} \Gamma(1-\alpha) \int_0^{\tau_{[k+1]}} \exp(-\alpha \xi_s^{\Sigma_1}) ds = \Gamma(1-\alpha) D_{[k+1]}. \tag{3.51}$$

So the renormalized length of the root-edge of $R(T_n, [k+1])$ converges to the length of the root-edge of $R(\mathcal{T}, \Sigma_1, \ldots, \Sigma_k)$ up to the factor $\Gamma(1-\alpha)$.

Suppose the split of $R(T_n, [k+1])$ at the first branch point $B_{[k+1]}$ is $(\pi_1, \dots, \pi_r)$ with $\#\pi_i = k_i$. For all $N_1(n) + \dots + N_r(n) \le n$ and $N_i(n) \ge k_i$, denote by $T^i(n)$ the subtree of T_n rooted at $B_{[k+1]}$ with n_i leaves, within which the first k_i leaves with labels in π_i.

Let $b_{[k+1]}$ be the branch point adjacent to ρ in $R(\mathcal{T}, \Sigma_1, \dots, \Sigma_{k+1})$. Let $\mathcal{T}^i_{b_{[k+1]}}$ be a subtree of $\mathcal{T}$ rooted at $b_{[k+1]}$ containing k_i leaves $(\Sigma_{i,1}, \dots, \Sigma_{i,k_i})$ within $\{\Sigma_1, \dots, \Sigma_{k+1}\}$. By the strong law of large numbers and Procedure 3.2,

$$\frac{N_i(n)}{n} \xrightarrow[n\to\infty]{a.s.} X_{\Sigma_{i,1}}(\tau_{[k+1]}) = \mu(\mathcal{T}^i_{b_{[k+1]}}).$$

According to the induction hypothesis, for $1 \le i \le r$

$$\begin{aligned} \frac{R(T^i, \pi_i)}{n^\alpha \ell(n)} &= \frac{N_i(n)^\alpha \ell(N_i(n))}{n^\alpha \ell(n)} \frac{R(T^i, \pi_i)}{N_i(n)^\alpha \ell(N_i(n))} \\ &\xrightarrow[n\to\infty]{a.s.} \Gamma(1-\alpha)(R(\mathcal{T}, \Sigma_{i,1}, \dots, \Sigma_{i,k_i}) - [[\rho, b_{[k+1]}]]). \end{aligned} \tag{3.52}$$

In the notation of Section 3.2.10,

$$R(T_n, [k+1]) = \text{MERGE}((R(T^1, \pi_1) - D^n_{[k+1]}, \dots, R(T^r, \pi_r) - D^n_{[k+1]}); D^n_{[k+1]}).$$

Now combining (3.51) and (3.52), the convergence for $R(T_n, [k+1])$ is straightforward. □

Now we establish the convergence of $R(T_n, [k])$ to the CRT $(\mathcal{T}, \mu)$.

Proposition 3.32. *In the setting of Proposition 3.31,*

$$\lim_{k\to\infty} \lim_{n\to\infty} n^{-\alpha} \ell(n)^{-1} \Gamma^{-1}(1-\alpha) \left(R(T_n, [k])\right)^\circ = \mathcal{T} \text{ a.s.}$$

in the Gromov-Hausdorff sense as unlabelled trees.

Proof. This proof is following the proof of Proposition 22 in [40]. For simplicity, we denote $R(\mathcal{T}, \Sigma_1, \dots, \Sigma_k)$ by $\mathcal{R}_k$. Due to Proposition 3.31, we only need to show that $R(\mathcal{T}, \Sigma_1, \dots, \Sigma_k)$ will converge to $\mathcal{T}$ in the Gromov-Hausdorff sense as k tends to infinity. In other words, we need to show that conditionally given $(\mathcal{T}, \mu)$, for all $\epsilon > 0$, there a.s. exists $k_0 \ge 1$ such that all connected components of $\mathcal{T} \setminus \mathcal{R}_k$ have diameter less than ϵ.

Consider the connected components of

$$\{b \in \mathcal{T} : \{b' \in \mathcal{T}_b : \mathrm{d}(b, b') \ge \epsilon\} = \emptyset\},$$

each completed by their root on the branches of $\mathcal{T}$. Since $\mathcal{T}$ is compact, at most finitely many components $\mathcal{T}_1, \ldots, \mathcal{T}_N$ actually attain height ϵ. Fix subtree $\mathcal{T}_j$ with root ρ_j, denote by $\mathcal{T}_j^k$ the subtree of $\mathcal{T}$ rooted at a point ρ_j^k of $\mathcal{R}_k$ and containing $\mathcal{T}_j$. For $b \in [[\rho_j^k, \rho_j]]$, let $\mathcal{T}_b^j$ be the subtree rooted at b containing $\mathcal{T}_j$, $X_b^j := \mu(\mathcal{T}_b^j)$ and $\widetilde{X}_b^j$ is the analogous quantity defined in Lemma 3.23. $\mathcal{T}_j \cap (\mathcal{R}_{k+1} \setminus \mathcal{R}_k) \neq \emptyset$ is equivalent to that leaf Σ_{k+1} is sample from one of the subtrees rooted at ρ_j. If ρ_j^k is in an inner edge of $\mathcal{R}_k$, according to Procedure 3.2, we have to sample Σ_{k+1} from $\mathcal{T}_j^b$ and along the spine $[[\rho_j^k, \rho_j]]$ by the the same procedure of sampling leaf Σ_1 because non of $\Sigma_1, \ldots, \Sigma_k$ is located in $\mathcal{T}_j$. Hence

$$\mathbb{P}\left(\mathcal{T}_j \cap (\mathcal{R}_{k+1} \setminus \mathcal{R}_k) \neq \emptyset \middle| \mathcal{R}_k, \rho_j^k \text{ is in an inner edge of } \mathcal{R}_k\right) = X^j_{\rho_j^k} \prod_{b \in [[\rho_j^k, \rho_j]]} P_{\Sigma_1}\left(\widetilde{X}_b^j\right).$$

While if ρ_j^k is in a leaf edge of $\mathcal{R}_k$, suppose this leaf is Σ_i. Denote by $b^{i,k}$ the branch point in $\mathcal{R}_k$ adjacent to ρ_j^k and $X^i_{\rho_j^k}, \widetilde{X}^i_{\rho_j^k}$ be the analogous quantity of subtree rooted at ρ_j^k containing Σ_i. We have to consider two situations, the first is that $\Sigma_i \notin \mathcal{T}_j$. In this situation, we have to sample Σ_{k+1} in $\mathcal{T}_b^j$ by the same procedure of sampling leaf Σ_2. Hence by Procedure 3.2, we have

$$\begin{aligned} &\mathbb{P}\left(\mathcal{T}_j \cap (\mathcal{R}_{k+1} \setminus \mathcal{R}_k) \neq \emptyset \middle| \mathcal{R}_k, \Sigma_i \notin \mathcal{T}_j\ \rho_j^k \text{ is in a leaf edge of } \mathcal{R}_k\right) \\ =\ & X^j_{b^{i,k}} \left(\prod_{b \in [[b^{i,k}, \rho_j^k]]} P^1_{\Sigma_2}\left(\widetilde{X}_b^j\right)\right) P^2_{\Sigma_2}\left(X^j_{\rho_j^k}, \widetilde{X}^i_{\rho_j^k}\right) \left(\prod_{b \in [[\rho_j^k, \rho_j]]} P_{\Sigma_1}\left(\widetilde{X}_b^j\right)\right). \end{aligned}$$

On the other hand if $\Sigma_i \in \mathcal{T}_j$, the above probability will be

$$\begin{aligned} &\mathbb{P}\left(\mathcal{T}_j \cap (\mathcal{R}_{k+1} \setminus \mathcal{R}_k) \neq \emptyset \middle| \mathcal{R}_k, \Sigma_i \notin \mathcal{T}_j\ \rho_j^k \text{ is in a leaf edge of } \mathcal{R}_k\right) \\ =\ & X^j_{b^{i,k}} \left(\prod_{b \in [[b^{i,k}, \rho_j^k]]} P^1_{\Sigma_2}\left(\widetilde{X}_b^j\right)\right). \end{aligned}$$

Hence, the probability $\mathbb{P}\left(\mathcal{T}_j \cap (\mathcal{R}_{k+1} \setminus \mathcal{R}_k) \neq \emptyset | \mathcal{R}_k\right)$ is bounded below uniformly in k. Therefore, the step when $\mathcal{T}_j \cap \mathcal{R}_k \neq \emptyset$ is bounded by a geometric random variable and no subtrees of height ϵ can persist outside $\mathcal{R}_k$ forever, so there a.s. exists $k_0 \geq 1$ such that $\mathcal{T} \setminus \mathcal{R}_{k_0}$ has no connected components of diameter exceeding ϵ, which completes the proof.

□

3.4.3 Height estimation for discrete fragmentation trees

In this section, we give a height estimation of discrete fragmentation trees T_n. The condition (3.2) and (3.3) satisfied by ν_1 implies that the tail $\bar{\Lambda}$ of the Lévy measure Λ defined in Proposition (3.14) i.e. $\bar{\Lambda}(x) = \int_x^\infty \Lambda(dy), x > 0$, satisfies $\bar{\Lambda}(x) \sim x^{-\alpha}\ell(1/x)$ as $x \to 0$ and $\int_0^\infty x^\varrho \Lambda(dx) < \infty$. Also, the tail $\bar{\Lambda}_*$ of Lévy measure Λ_* defined in Proposition 3.16 satisfied the above two formulas. According to Potter's theorem [11, Theorem 1.5.6] and the monotonicity of $\bar{\Lambda}, \bar{\Lambda}_*$, this in turn implies the existence of some finite constants C_Λ, C_{Λ_*} such that

$$\bar{\Lambda}(xy) \leq C_\Lambda \bar{\Lambda}(x) y^{-\varrho} \text{ and } \bar{\Lambda}_*(xy) \leq C_{\Lambda_*} \bar{\Lambda}_*(x) y^{-\varrho}, \tag{3.53}$$

for all $y \geq 1$, $0 < x \leq 1$.

Lemma 3.33. *Let ξ be a pure jump subordinator with Lévy measure Λ satisfying $\Lambda([x,\infty)) = x^{-\alpha}\ell(1/x)$, $x \downarrow 0$ and (3.53). Let (ε, τ) be any random variable on $\mathbb{R}_+^2$ and τ' be any random variable on $[0,\infty]$. Given (ε,τ), let $U_1, \dots,$ be any sequence of i.i.d. random variables on $[0,\infty)$, the distribution of which satisfies*

$$\mathbb{P}\left(U_i \leq \tau | (\varepsilon,\tau), \xi\right) = \exp(-\varepsilon) \text{ and } \mathbb{P}\left(U_i > \tau + t | (\varepsilon,\tau), \xi\right) = \exp(-\varepsilon - \xi_t).$$

Denote by

$$K_n(\varepsilon,\tau,\tau') := \#\{U_i : 1 \leq i \leq n, \tau < U_i < \tau + \tau'\}.$$

Then for all $x \geq 1$ and all integer n,

$$\mathbb{P}\left(K_n(\varepsilon,\tau,\tau') > (1+x) Y(\varepsilon,\tau,\tau') \bar{\Lambda}(n^{-1})\right) \leq \frac{C_p}{x^p n^{\alpha p - 1}}, \tag{3.54}$$

where $Y(\varepsilon,\tau,\tau') = 1 + (1+A_\alpha) C_\Lambda \exp(-\varrho\varepsilon) \left(\sum_{k=0}^{[\tau']+1} \exp(-\varrho \xi_k)\right)$ with $A_\alpha = 2\sum_{k=1}^\infty \frac{(k+1)^{\sqrt{\alpha}}}{k(k+1)}$.

Proof. This Lemma is an extension of Lemma 12 in [28]. Let $N_y(t_1,t_2)$ denote the number of jumps of ξ of size at least y in the time interval $[t_1,t_2]$, $\widetilde{N}_y^{\varepsilon,\tau}(t_1,t_2)$ denote the number of jumps of $\exp(-\varepsilon)(1-\exp(-\xi))$ with size at least y in the same time interval. Let $\mathcal{F}_t^{\varepsilon,\tau}$ be the σ-field generated by (ε,τ) and ξ until time t and $\mathcal{F}_\infty^{\varepsilon,\tau}$ be the one generated by (ε,τ) and ξ.

Step 1. Large deviations for $\widetilde{N}_y^{\varepsilon,\tau}(0,t)$. Remark that

$$\widetilde{N}_y^{\varepsilon,\tau}(0,t) \leq \sum_{i=0}^{[t]} \widetilde{N}_y^{\varepsilon,\tau}(i,i+1) \leq \sum_{i=0}^{[t]} N_{y\exp(\varepsilon+\xi_i)}(i,i+1).$$

Conditional on $\mathcal{F}_i^{\varepsilon,\tau}$, $N_{y\exp(\varepsilon+\xi_i)}(i,i+1)$ is a Poisson random variable with mean $\bar{\Lambda}(y\exp(\varepsilon+\xi_i))$. But for any Poisson random variables P with mean λ, one has

$$\mathbb{E}\left(\exp(tP-(1+x)t\lambda)\right)=\exp\left(\lambda(\exp(t)-1-(1+x)t)\right),\ \forall t\in\mathbb{R}.$$

In particular, when $t=\ln(1+x), \exp(t)-1-(1+x)t=-a_x<0$ and the expectation is smaller than 1. Hence for all $n\in\mathbb{N}$, using (3.53) for the first inequality, we obtain, for all $y\leq 1$,

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=0}^{[t]}N_{y\exp(\varepsilon+\xi_i)}(i,i+1)\geq(1+x)C_\Lambda\sum_{i=0}^{[t]}\exp(-\varrho(\varepsilon+\xi_i))\bar{\Lambda}(y)\right)\\
\leq\ &\mathbb{P}\left(\sum_{i=0}^{[t]}N_{y\exp(\varepsilon+\xi_i)}(i,i+1)\geq(1+x)\sum_{i=0}^{[t]}\bar{\Lambda}(y\exp(\varepsilon+\xi_i))\right)\\
\leq\ &\mathbb{E}\left(\exp\left(t\left(\sum_{i=0}^{[t]}\left(N_{y\exp(\varepsilon+\xi_i)}(i,i+1)-(1+x)\bar{\Lambda}(y\exp(\varepsilon+\xi_i))\right)\right)\right)\right)\\
\leq\ &\mathbb{E}\Bigg(\mathbb{E}\left(\exp\left(t\left(N_{y\exp(\varepsilon+\xi_{[t]})}([t],[t]+1)-(1+x)\bar{\Lambda}(y\exp(\varepsilon+\xi_{[t]}))\right)\right)|\mathcal{F}_{[t]}^{\varepsilon,\tau}\right)\\
&\quad\times\exp\left(t\left(\sum_{i=0}^{[t]-1}\dots\right)\right)\Bigg)\\
\leq\ &\dots\leq\exp(-a_x\bar{\Lambda}(y)).
\end{aligned}\tag{3.55}$$

The last line being obtained by induction: at each step but the last we use the upper bound 1 for the conditional expectation and for the last step, we use the upper bound $\exp(-a_x\bar{\Lambda}(y))$ for the expectation $\mathbb{E}\left(\exp\left(t(N_y(0,1)-(1+x)\bar{\Lambda}(y))\right)\right)$. It remains for $t=\infty$ in the first probability involved in the above sequence of inequalities. We simply use that for any sequence of real-valued random variables $W_n\to W_\infty$ a.s., $\mathbb{P}(W_\infty>a)\leq\liminf_{n\to\infty}\mathbb{P}(Y_n>a)$ for all $a\in\mathbb{R}$ (by Fatou's Lemma).

Step 2. Large deviations for $\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)$. According to formula (4) of [23],

$$\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)=n\int_0^1(1-y)^{n-1}\widetilde{N}_y^{\varepsilon,\tau}(0,t)dy\leq\widetilde{N}_{1/n}^{\varepsilon,\tau}(0,t)+n\int_0^{1/n}\widetilde{N}_y^{\varepsilon,\tau}(0,t)dy.$$

Hence, setting $S := C_\Lambda \sum_{i=0}^{[t]+1} \exp(-\rho(\varepsilon+\xi_i))$,

$$\begin{aligned} &\mathbb{P}\left(\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right) > (1+x)(1+A_\alpha)S\bar\Lambda(n^{-1})\right) \\ \leq\ & \mathbb{P}\left(\widetilde{N}_{1/n}^{\varepsilon,\tau}(0,t) > (1+x)S\bar\Lambda(n^{-1})\right) + \mathbb{P}\left(n\int_0^{1/n} \widetilde{N}_y^{\varepsilon,\tau}(0,t)dy > (1+x)A_\alpha S\bar\Lambda(n^{-1})\right). \end{aligned}$$

The first probability in the right-hand side is smaller than $\exp(-a_x\bar\Lambda(n^{-1})$ by (3.55). To bound the second probability, we use $n\int_{1/(k+1)n}^{1/kn} \widetilde{N}_y^{\varepsilon,\tau}(0,t)dy \leq \frac{1}{k(k+1)}\widetilde{N}^{\varepsilon,\tau}_{\frac{1}{n(k+1)}}(0,t)$, which gives

$$\begin{aligned} &\mathbb{P}\left(n\int_0^{1/n} \widetilde{N}_y^{\varepsilon,\tau}(0,t)dy > (1+x)A_\alpha S\bar\Lambda(n^{-1})\right) \\ \leq\ & \sum_{k=1}^{\infty} \mathbb{P}\left(\widetilde{N}^{\varepsilon,\tau}_{\frac{1}{n(k+1)}}(0,t) > 2(k+1)^{\sqrt{\alpha}}(1+x)S\bar\Lambda(n^{-1})\right). \end{aligned}$$

Since $\bar\Lambda$ is regularly varying at 0 with index $-\alpha$ we have, provided that n is large enough, that $\bar\Lambda(n^{-1})(k+1)^{\alpha/2} \leq 2\bar\Lambda(1/n(k+1)) \leq 4\bar\Lambda(n^{-1})(k+1)^{\sqrt{\alpha}}$ for all $k \geq 1$. Combined with (3.55), this implies that the above sum of probabilities is smaller than

$$\sum_{k=1}^{\infty} \exp(-a_x\bar\Lambda(1/n(k+1))) \leq \sum_{k=1}^{\infty} \exp(-a_x\bar\Lambda(n^{-1})(k+1)^{\alpha/2}/2).$$

Last, the exponential in the latter sum can be split in two, using $(k+1)^{\alpha/2} \geq (k^{\alpha/2}+1)/2$, to get the upper bound

$$\exp(-a_x\bar\Lambda(n^{-1})/4)\sum_{k=1}^{\infty} \exp(-a_x\bar\Lambda(n^{-1})k^{\alpha.2}.4),$$

which is smaller than $B_\alpha \exp(-ax\bar\Lambda(n^{-1}))$ for n is large enough where $a = (2\ln 2 - 1)/4$ and $B_\alpha = \sum_{k=1}^{\infty} \exp(-ak^{\alpha/2})$. Hence, we obtain that for n is large enough,

$$\begin{aligned} &\mathbb{P}\left(\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right) > (1+x)(1+A_\alpha)\left(C_\Lambda \sum_{i=0}^{[t]+1} \exp(-\rho(\varepsilon+\xi_i))\right)\bar\Lambda(n^{-1})\right) \\ \leq\ & (1+B_\alpha)\exp(-ax\bar\Lambda(n^{-1})). \end{aligned} \tag{3.56}$$

Step 3. Proof of inequality (3.54). Fix $x \geq 1$, note that

$$\begin{aligned} &\mathbb{P}\left(K_n(\varepsilon,\tau,t) > (1+x)Y(\varepsilon,\tau,\tau')\bar\Lambda(n^{-1})\right) \\ \leq\ & \mathbb{P}\left(\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right) > (1+x)\left(Y(\varepsilon,\tau,t)-1\right)\bar\Lambda(n^{-1})\right) \\ &+\mathbb{P}\left(K_n(\varepsilon,\tau,t) - \mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right) > (1+x)\bar\Lambda(n^{-1})\right). \end{aligned}$$

(3.56) gives an upper bound for the first probability when n is large enough. The result on urn models (Section 6, [13]) ensures that

$$\mathbb{P}\left(K_n(\varepsilon,\tau,t)-\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)>y\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)\leq\exp\left(-\frac{y^2}{2\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)+2y/3}\right).$$

This implies that there exits some constant C'_p depending only on p such that

$$\begin{aligned}
&\mathbb{P}\left(K_n(\varepsilon,\tau,t)-\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)>(1+x)\bar{\Lambda}(n^{-1})\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)\\
\leq\ & C'_p\left(\frac{\mathbb{E}\left(K_n(\varepsilon,\tau,t)\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)+(1+x)\bar{\Lambda}(n^{-1})}{\left((1+x)\bar{\Lambda}(n^{-1})\right)^2}\right)^p\\
\leq\ & 2^{p-1}C'_p\frac{\mathbb{E}\left(K_n(\varepsilon,\tau,t)^p\middle|\mathcal{F}_t^{\varepsilon,\tau}\right)+\left((1+x)\bar{\Lambda}(n^{-1})\right)^p}{\left((1+x)\bar{\Lambda}(n^{-1})\right)^{2p}}\\
\leq\ & C_{p,\Lambda}\left((1+x)\bar{\Lambda}(n^{-1})\right)^{-p}, \qquad (3.57)
\end{aligned}$$

where $C_{p,\Lambda}$ only depends on p and Λ. The last line is obtained by

$$\mathbb{E}\left(K_n(\varepsilon,\tau,t)^p\middle|\varepsilon,\tau\right)\leq\mathbb{E}\left(K_n(0,0,t)^p\right)\sim\left(\bar{\Lambda}(n^{-1})\right)^p$$

(up to a constant) which is ensured by Theorem 6.3 of [23].

Then recall the upper bound given by (3.56) for the first probability involved in the right-hand side of (3.57). Together with the upper bound (3.57), it leads to the existence of $C'_{p,\Lambda}$ such that

$$\mathbb{P}\left(K_n(\varepsilon,\tau,t)>(1+x)Y(\varepsilon,\tau,\tau')\bar{\Lambda}(n^{-1})\right)\leq C'_{p,\Lambda}x^{-p}\left(\bar{\Lambda}(n^{-1})\right)^{-p},$$

for all $x\geq1$ and n large enough, say $n\geq n_0$. Since $\bar{\Lambda}(n^{-1})\sim n^{\alpha}\ell(n)$ when $n\to\infty$, this upper bound is in turn bounded from above by $x^{-p}n^{1-\alpha p}$ up to some constant. Last, inequality (3.54) is also true when $n\leq n_0$, since $K_n(\varepsilon,\tau,t)\leq n\leq n_0$ and $Y>1$, therefore the probability $\mathbb{P}\left(K_n(\varepsilon,\tau,t)>(1+x)Y(\varepsilon,\tau,\tau')\bar{\Lambda}(n^{-1})\right)$ is null whenever $1+x\geq n_0\left(\bar{\Lambda}(n^{-1})\right)^{-1}$. □

Remark 3.3. This Lemma is an extension from Lemma 12 in [28]. Setting $\tau=\epsilon=0$ and $\tau'=\infty$ gives their result.

Let Λ be the Lévy measure associated with the subordinator ξ^{Σ_1} and Λ_* the Lévy measure associated with a subordinator ξ^{Σ^*} generated by a leaf Σ^* sampled according to μ.

Lemma 3.34. *For every $p \geq 0$, $x \geq 1$ and all integers n, there exists a constant C'_p and a random variable*

$$Z_j := 4 + (1 + A_\alpha)C_\Lambda \left(\sum_{k=0}^{\infty} \left(X_{\Sigma_j}(k)\right)^\rho \right)$$

with finite positive moments of all orders such that for all D_j^n,

$$\mathbb{P}\left(D_j^n > 2(1+x)(2+Z_j)\max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}\right) \leq \frac{C'_p}{x^p n^{\alpha p - 1}}.$$

Proof. (i) Case D_1^n. Note that $D_1^n \leq 2(D_1^n - 1)$. According to (3.48), Lemma 3.33 by setting $(\varepsilon, \tau) = (0, 0)$ and Proposition 3.29,

$$\begin{aligned}
&\mathbb{P}\left(D_1^n > 2(1+x)(2+Z_1)\bar{\Lambda}(n^{-1})\right) \\
\leq\ &\mathbb{P}\left(D_1^n - 1 > (1+x)(2+Z_1)\bar{\Lambda}(n^{-1})\right) \\
\leq\ &\mathbb{P}\left(K_{n-2}^{(1)} + L_n^{(1)} > (1+x)(2+Z_1)\bar{\Lambda}(n^{-1})\right) \\
\leq\ &\mathbb{P}\left(K_{n-2}^{(1)} + L_n^{(1)} > (1+x)(2+Z_1)\bar{\Lambda}(n^{-1})\right) \\
\leq\ &\mathbb{P}\left(K_{n-2}^{(1)} > (1+x)Z_1\bar{\Lambda}(n^{-1})\right) + \mathbb{P}\left(L_n^{(1)} > 2(1+x)\bar{\Lambda}(n^{-1})\right) \\
\leq\ &\frac{C''_p}{x^p n^{\alpha p - 1}}.
\end{aligned} \tag{3.58}$$

Since 1 and 2 are exchangeable, (D_2^n, Z_2) is distributed the same as (D_1^n, Z_1) and hence the above formula also holds for D_2^n.

(ii) Case D_j^n. $j \geq 3$. Recall Procedure 3.2 where Σ_j^* is sampled according to μ. Let $V_{j,*}^i, i = 1, \ldots, j-1$ be the homogeneous time when leaf Σ_i leaves the spine $[[\rho, \Sigma_j^*]]$. Let $b_{j,*}^m, m = 1, \ldots, j-2$ be the special branch point along the spine $[[\rho, \Sigma_j^*]]$ and $\tau_{j,*}^m$ the associated homogeneous time in analogy with (3.42). By the Procedure 3.2, the leaves Σ_j and Σ_j^* will only be separated in the subtree containing both of them rooted at the special branch point $b_{j,*}^{j-2}$. Hence, $\tau_j^m = \tau_{j,*}^m$ and $b_j^m = b_{j,*}^m$ for $m = 1, \ldots, j-2$. Also,

$$V_j^i \mathbf{1}_{\{V_j^i \leq \tau_j^{j-2}\}} = V_{j,*}^i \mathbf{1}_{\{V_{j,*}^i \leq \tau_{j,*}^{j-2}\}}$$

for $i = 1, 2, \ldots, j-1, j+1, \ldots, n$. Now let $V_{j,*}^{i,*}$ be the homogeneous time when leaf Σ_i^* leaves the spine $[[\rho, \Sigma_j^*]]$ for $i = 3, \ldots, j-1, j+1, \ldots, n$. For $3 \leq i \leq j-1$, if $V_{j,*}^{i,*} \leq \tau_{j,*}^{i-2}$

i.e if Σ_i^* is not the leaf with smallest or second smallest label in the subtree containing itself of $R(\Sigma_1, \Sigma_2, \dots, \Sigma_{i-1}, \Sigma_i^*, \Sigma_j^*)$ rooted at the next branch point below $\Sigma_i^* \wedge \Sigma_j^*$, then Σ_i is sampled in the subtree containing Σ_i^* of $\mathcal{T}$ rooted at $\Sigma_i^* \wedge \Sigma_j^*$; for $j+1 \leq i \leq n$, if $V_{j,*}^{i,*} \leq \tau_{j,*}^{j-2}$, then Σ_i is sampled in the subtree containing Σ_i^* of $\mathcal{T}$ rooted at $\Sigma_i^* \wedge \Sigma_j^*$. Hence we obtain for $i = 3, \dots, j-1$

$$V_j^i \mathbf{1}_{\{V_j^i \leq \tau_j^{i-2}\}} = V_{j,*}^i \mathbf{1}_{\{V_{j,*}^i \leq \tau_{j,*}^{i-2}\}} = V_{j,*}^{i,*} \mathbf{1}_{\{V_{j,*}^{i,*} \leq \tau_{j,*}^{i-2}\}};$$

for $i = j+1, \dots, n$,

$$V_j^i \mathbf{1}_{\{V_j^i \leq \tau_j^{j-2}\}} = V_{j,*}^i \mathbf{1}_{\{V_{j,*}^i \leq \tau_{j,*}^{j-2}\}} = V_{j,*}^{i,*} \mathbf{1}_{\{V_{j,*}^{i,*} \leq \tau_{j,*}^{j-2}\}}.$$

By the same method, for the spine $[[b_j^{j-2}, \Sigma_j]]$, we have that for $i = j+1, \dots, n$,

$$V_j^i \mathbf{1}_{\{\tau_j^{j-2} < V_j^i \leq \tau_j^{i-2}\}} = V_j^{i,*} \mathbf{1}_{\{\tau_j^{j-2} < V_j^{i,*} \leq \tau_j^{i-2}\}},$$

where $V_j^{i,*}$ is the time when Σ_i^* leave the spine $[[\rho, \Sigma_j]]$.

Then we have

$$\begin{aligned}
D_j^n - 1 &= \#\{V_j^i : 3 \leq i \leq n, i \neq j\} \\
&= \sum_{m=1}^{n-2} \#\{V_j^i : \tau_j^{m-1} < V_j^i \leq \tau_j^m, m+2 \leq i \leq n, i \neq j\} \\
&= L_n^{(j)} + \sum_{m=1}^{j-2} \#\{V_j^i : \tau_j^{m-1} < V_j^i < \tau_j^m, m+2 \leq i \leq n, i \neq j\} \\
&\quad + \sum_{m=j-1}^{n-2} \#\{V_j^i : \tau_j^{m-1} < V_j^i < \tau_j^m, m+2 \leq i \leq n, i \neq j\} \\
&= L_n^{(j)} + \sum_{m=1}^{j-2} \#\{V_{j,*}^{i,*} : \tau_{j,*}^{m-1} < V_{j,*}^{i,*} < \tau_{j,*}^m, m+2 \leq i \leq n, i \neq j\} \\
&\quad + \sum_{m=j-1}^{n-2} \#\{V_j^{i,*} : \tau_j^{m-1} < V_j^{i,*} < \tau_j^m, m+2 \leq i \leq n, i \neq j\} \\
&\leq L_n^{(j)} + \#\{V_{j,*}^{i,*} : 0 < V_{j,*}^{i,*} < \tau_{j,*}^{j-2}, 3 \leq i \leq n, i \neq j\} \\
&\quad + \#\{V_j^{i,*} : V_j^{i,*} > \tau_j^{j-2}, j+1 \leq i \leq n, \} \\
&\leq L_n^{(j)} + \#\{V_{j,*}^{i,*} : 0 < V_{j,*}^{i,*} < \tau_{j,*}^{j-2}, 3 \leq i \leq n, i \neq j\} \\
&\quad + \#\{V_j^{i,*} : V_j^{i,*} > \tau_j^{j-2}, 3 \leq i \leq n, i \neq j\}.
\end{aligned}$$

Now by Lemma 3.33, setting $(\varepsilon, \tau) = (0, 0)$ and $\tau' = \tau_{j,*}^{j-2}$, we obtain that

$$\mathbb{P}\Bigg(\#\{V_{j,*}^{i,*} : 0 < V_{j,*}^{i,*} < \tau_{j,*}^{j-2}, 3 \le i \le n, i \ne j\} >$$
$$(1+x)\left(1 + (1+A_\alpha)C_\Lambda \left(\sum_{k=0}^{[\tau_j^{j-2}]+1} \left(X_{\Sigma_j^*}(k)\right)^\rho\right)\right)\bar{\Lambda}_*(n^{-1})\Bigg)$$
$$\le \frac{C_p^{(1)}}{x^p n^{\alpha p - 1}}, \tag{3.59}$$

as the Lévy measure associated with the process $-\log X_{\Sigma_j^*}$ is Λ_*.

Also, setting $(\varepsilon, \tau) = \left(-\log X_{\Sigma_j}(\tau_j^{j-2}), \tau_j^{j-2}\right)$ and $\tau' = \infty$, we obtain

$$\mathbb{P}\Bigg(\#\{V_j^{i,*} : V_j^{i,*} > \tau_j^{j-2}, 3 \le i \le n, i \ne j\} >$$
$$(1+x)\left(1 + (1+A_\alpha)C_\Lambda \left(\sum_{k=0}^{\infty} \left(X_{\Sigma_j}(k+\tau_j^{j-2})\right)^\rho\right)\right)\bar{\Lambda}(n^{-1})\Bigg)$$
$$\le \frac{C_p^{(2)}}{x^p n^{\alpha p - 1}}, \tag{3.60}$$

as the Lévy measure associated with the process $\left(-\log X_{\Sigma_j}(\tau_j^{j-2} + t), t \ge 0\right)$ is Λ. By Proposition 3.29 (iii),

$$\mathbb{P}\left(L_n^{(j)} > 2(1+x)\bar{\Lambda}(n^{-1})\right) \le \frac{C_p^{(3)}}{x^p n^{\alpha p - 1}}. \tag{3.61}$$

Note that $D_j^n \le 2(D_j^n - 1)$ for $n \ge 2$. Combining (3.59), (3.60)and (3.61), we obtain the desired result.

Last, let

$$Z_\infty := \sup_{j \ge 1} Z_j \le 4 + (1+A_\alpha)C_\Lambda \left(1 + \sup_{j\ge 1}\int_0^\infty \left(X_{\Sigma_j}(t)\right)^\rho dt\right) \le 4 + (1+A_\alpha)C_\Lambda\left(1 + H_{\mathcal{T}}^\rho\right),$$

where $H_{\mathcal{T}}^\rho$ is the height of the CRT $(\mathcal{T}, \mu)$ under self-similar time-change with parameter ρ. It was proved in Proposition 14 [26] that $H_{\mathcal{T}}^\rho$ has exponential moments. □

Corollary 3.35. *Let H_n be the height of T_n, that is $H_n := \max_{1\le j\le n} D_j^n$. For all $a > 0$ and $p \ge 2/\alpha$, there exists some constant $C_{p,a}$ such that for all $x \ge 1$ and all integers n,*

$$\mathbb{P}\left(H_n > ax\max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}\right) \le \frac{C_{p,a}}{x^p}.$$

Proof. Note that

$$\begin{aligned}&\mathbb{P}\left(H_n > (1+x)Z_\infty \max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}\right)\\ \le\ & \sum_{j=1}^{n} \mathbb{P}\left(D_j^n > (1+x)Z_j \max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}\right) \le \frac{C_p^{(4)}}{x^p}.\end{aligned}$$

Also,

$$\begin{aligned}&\mathbb{P}\left(H_n > ax \max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}\right)\\ \le\ & \mathbb{P}\left(H_n > ax \max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}, ax \ge (1+\sqrt{x})Z_\infty\right)\\ &+\mathbb{P}((1+\sqrt{x})Z_\infty > ax).\end{aligned}$$

The second probability is bounded by some $C'_{p,a}/x^p$ according to Chebysev's inequality as $\mathbb{E}(Z_\infty^{2p})$ is finite.

□

3.4.4 Tightness estimation and Proof of Theorem 3.2

Proposition 3.36. *Let $(T_n, n \ge 0)$ be a family of discrete partly exchangeable fragmentation trees associated with dislocation measures ν_1, ν_2. For $k \le n$, let*

$$\Delta(n,k) := \max_{1\le i\le n} \mathrm{d}_n(i, R(T_n, [k])),$$

d_n being the metric associated with T_n. Then for each $\eta > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\frac{\Delta(n,k)}{\max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}} > \eta\right) = 0.$$

Proof. For all $k \ge 1$, introduce $t_i^k := \inf\{t \ge 0 : \Pi_{(i)}(t) \cap [k] = \emptyset\}$, the first time at which the fragment containing i is disjoint from $[k]$. For all $t \ge 0$, the collection of blocks $(\Pi_{(i)}(t_i^k + t), i \ge k+1)$ induces a partly exchangeable fragmentation $\Pi(t^k + t)$ of $\mathbb{N} \setminus [k]$ and each $\Pi_j(t^k + t)$ admits asymptotic frequencies. We call $n_j^{k,n}$ the cardinality of $\Pi_j(t^k) \cap [n]$ and λ_j^k the a.s. limit of $n_j^{k,n}/n$ as $n \to \infty$. Clearly, $\lambda_{\max}^k := \max_{j\ge1} \lambda_j^k \to 0$ a.s. as $k \to \infty$. Let $\mathcal{G}(k)$ be the $\sigma-$field generated by $\Pi(t^k)$. According to the construction of partly exchangeable fragmentation process, $(\Pi(t^k + t), t \ge 0)$ is a partly exchangeable fragmentation process starting from $\Pi(t^k)$ given $\mathcal{G}(k)$. This implies that given $\mathcal{G}(k)$, the discrete fragmentation trees, with respectively $n_1^{k,n}, \ldots$ leaves, associated with the fragmentations of the blocks $\Pi_j(t^k), j \geq 1$, evolves independently as $n \to \infty$, with laws respectively distributed as $T_{n_j^{k,n}}, j \geq 1$. In particular, given $\mathcal{G}(k)$, the respective heights of those trees, $\bar{H}_{n_j^{k,n}}$ are independent and distributed as $H_{n_j^{k,n}}$.

Let $\eta > 0$. Applying the first dominated convergence for the limit in k and then Fatou's Lemma for the liminf in n, it is sufficient to show that

$$\lim_{k\to\infty} \liminf_{n\to\infty} \mathbb{P}(\Delta(n,k) \leq \eta\bar{\Lambda}(n^{-1})|\mathcal{G}(k)) \to 1 \text{ a.s.}$$

According to the discussion above

$$\mathbb{P}(\Delta(n,k) \leq \eta\bar{\Lambda}(n^{-1})|\mathcal{G}(k)) = \prod_{j\geq 1} \mathbb{P}\left(\bar{H}_{n_j^{k,n}} \leq \eta\bar{\Lambda}(n^{-1})|\mathcal{G}(k)\right),$$

and our goal turns into the proof of

$$\lim_{k\to\infty} \liminf_{n\to\infty} \sum_{j\geq 1} \ln\left(1 - \mathbb{P}\left(\bar{H}_{n_j^{k,n}} > \eta\bar{\Lambda}(n^{-1})|\mathcal{G}(k)\right)\right) = 0. \tag{3.62}$$

By the virtue of the proof of Proposition 9 of [28], (3.62) is a direct result of Corollary 3.35, which completes the proof of Proposition 3.36. □

Proof of Theorem 3.2. Fix $\epsilon, \eta > 0$ and choose n large enough such that

$$\mathbb{P}(\mathrm{d_{GH}}(R(T_n;[k])/\bar{\Lambda}(n^{-1}), \Gamma(1-\alpha)\mathcal{T}) > \eta) < \epsilon \tag{3.63}$$

according to Proposition 3.32. By Proposition 3.36, choose k large enough such that

$$\limsup_{n\to\infty} \mathbb{P}\left(\mathrm{d_{GH}}(R(T_n;[k]), T_n) > \max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\}\eta\right) < \epsilon.$$

Note that $\bar{\Lambda}_1(x) = x^{-\alpha}\ell(1/x), x \downarrow 0$ and Λ_2 is a finite measure. According to Proposition 3.14,

$$\bar{\Lambda}(x) = x^{-\alpha}\ell(1/x) + F(x),\ x \downarrow 0,$$

where $F(0) < \infty$. Similarly, we can obtain

$$\bar{\Lambda}_*(x) = x^{-\alpha}\ell(1/x) + F_*(x),\ x \downarrow 0,$$

where $F_*(0) < \infty$ from Proposition 3.16. Hence

$$\bar{\Lambda}(n^{-1}) \sim \max\{\bar{\Lambda}(n^{-1}), \bar{\Lambda}_*(n^{-1})\},\ n \to \infty,$$

which induce

$$\limsup_{n\to\infty} \mathbb{P}\left(\mathrm{d_{GH}}(R(T_n;[k]), T_n) > \bar{\Lambda}(n^{-1})\eta\right) < \epsilon. \tag{3.64}$$

Combining (3.63),(3.64), the result is straightforward. □

3.5 Three examples of partly exchangeable fragmentation trees

3.5.1 Ford's alpha models

Ford in [21] introduced a family of binary Markov branching model called the alpha model. There are two versions, labelled and unlabelled, and each can be described in different ways. The asymptotics of the unlabelled alpha tree have been completed by Haas, et. in [28]. The labelled alpha tree can be considered as an example of the (α, θ) tree introduced by Pitman and Winkel, whose convergence was established as well [40]. In this paper we focus on the labelled alpha tree and provide another method to establish the convergence.

The sequential construction of the labelled alpha tree starts with the unique binary labelled trees T_1 and T_2 with one and two leaves, respectively. Given the random tree T_n with n leaves and labels $\{1, 2, \ldots, n\}$ based on their appearance, a new leaf with label $n+1$ is added as follows. Choose an edge between two inner vertices with probability $\alpha/(n-\alpha)$ or an edge between a leaf and an inner vertex with probability $(1-\alpha)/(n-\alpha)$. Replace this edge by a new vertex and two edges linking to its two vertices to the new vertex. The resulting random tree with $n+1$ leaves is defined as T_{n+1}.

Corollary 3.37. *Let T_n be the labelled alpha tree. Then*

(i) *T_n is discrete partly exchangeable tree with PEPPF*

$$
\begin{aligned}
p_n^1(m, n-m) &= \alpha \frac{\Gamma_\alpha(m)\Gamma_\alpha(n-m)}{\Gamma_\alpha(n)}, \\
p_n^2(m, n-m) &= (1-\alpha) \frac{\Gamma_\alpha(m)\Gamma_\alpha(n-m)}{\Gamma_\alpha(n)},
\end{aligned}
$$

for $1 \leq m \leq n-1$, where $\Gamma_\alpha(n) = \Gamma(n-\alpha)/\Gamma(1-\alpha)$. The associated two dislocation measures are binary with

$$
\begin{aligned}
\nu_1(s_1 \in dx) &= \alpha(x(1-x))^{-\alpha-1}dx \\
\nu_2(s_1 \in dx) &= (1-\alpha)(x(1-x))^{-\alpha-1}dx,
\end{aligned}
$$

for all $x \in [\frac{1}{2}, 1)$.

(ii)

$$\frac{T_n}{n^\alpha} \xrightarrow{(p)} \mathcal{T}_{\alpha,\nu_{\mathrm{Ford}-\alpha}}$$

in the Gromov-Hausdorff sense, in probability as $n \to \infty$, where

$$\nu_{\mathrm{Ford}-\alpha}(s_1 \in dx) = \alpha(x(1-x))^{-\alpha-1} + (2-4\alpha)(x(1-x))^{-\alpha}dx$$

for all $x \in [\frac{1}{2}, 1)$.

Proof. It suffices to check that the splitting rules are as specified. The dislocation measures and convergence are just the application of Theorem 3.1 and 3.2. Now consider a first split at the first branch point b_1 of T_n into two subtrees with m and $n-m$ leaves and denoted by T^1 and T^2 respectively. Suppose both of the leaf 1 and 2 are in the subtree T^1. Let $k \geq 2$ be the smallest label in T^2. The first branch point b_1 is created when the leaf k is added to the root edge of T_{k-1} with probability $\alpha/(k-1-\alpha)$. Then other leaves are added to the two subtrees at b_1. Hence, the probability of such first split is

$$\frac{\alpha}{k-1-\alpha}\frac{(k-1-\alpha)\cdots(m-1-\alpha)(1-\alpha)\cdots(n-m-1-\alpha)}{(k-\alpha)\cdots(n-1-\alpha)} = \alpha\frac{\Gamma_\alpha(m)\Gamma_\alpha(n-m)}{\Gamma_\alpha(n)},$$

which is independent with k and is invariant under all possible labellings of T^1 and T^2 as long as 1 and 2 are in the same subtree.

On the other hand, if 1 and 2 are located in T^1 and T^2 respectively, the first branch point b_1 is created when 2 is added to the one leaf tree T_1. Hence, the probability of such kind of split is

$$(1-\alpha)\frac{\Gamma_\alpha(m)\Gamma_\alpha(n-m)}{\Gamma_\alpha(n)},$$

which is invariant under all possible labellings of T^1 and T^2 as long as 1 and 2 are in the different subtrees. Based on the above discussion, the alpha models are partly exchangeable with the splitting rules as specified. □

3.5.2 Alpha-gamma models

The alpha-gamma models introduced in Chapter 2 are extensions of alpha models into multifurcating case. As labelled alpha trees, the sequential construction of the labelled alpha-gamma trees starts with the unique labelled trees T_1 and T_2 with one and two leaves, respectively. Given the random tree T_n with n leaves and labels $\{1, 2, \dots, n\}$ based on their appearance, a new leaf with label $n + 1$ is added as follows.

- Choose an edge between two inner vertices with probability $\gamma/(n - \alpha)$ or an edge between a leaf and an inner vertex with probability $(1 - \alpha)/(n - \alpha)$. Replace this edge by a new vertex and two edges linking to its two vertices to the new vertex.

- Choose a branch point with k subtrees with probability $((k - 1)\alpha - \gamma)/(n - \alpha)$. Directly link the branch point a new vertex with a new edge and labelled this new vertex $n + 1$.

The resulting random tree with $n+1$ leaves is defined as T_{n+1}. We show that alpha-gamma trees are discrete partly exchangeable trees as well.

Corollary 3.38. *Let T_n be the labelled alpha-gamma tree. Then*

(i) *T_n is discrete partly exchangeable tree with PEPPF*

$$\begin{aligned} p_n^1(n_1, \dots, n_k) &= \gamma\alpha^{k-2}\Gamma_{1+\frac{\gamma}{\alpha}}(k)\frac{\Gamma_\alpha(n_1)\cdots\Gamma_\alpha(n_k)}{\Gamma_\alpha(n)}, \\ p_n^2(m, n - m) &= (1 - \alpha)\alpha^{k-2}\Gamma_{1+\frac{\gamma}{\alpha}}(k)\frac{\Gamma_\alpha(n_1)\cdots\Gamma_\alpha(n_k)}{\Gamma_\alpha(n)}, \end{aligned}$$

for $n_1 \geq \dots \geq n_k \geq 1, n_1 + \dots + n_k = n$. The associated two dislocation measures are

$$\begin{aligned} \nu_1(d\mathbf{s}) &= \gamma \mathrm{PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}), \\ \nu_2(d\mathbf{s}) &= (1 - \alpha)\mathrm{PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}), \end{aligned}$$

for all $\mathbf{s} \in \mathcal{S}^\downarrow$.

(ii)
$$\frac{T_n}{n^\gamma} \xrightarrow{(p)} \mathcal{T}_{\alpha,\nu_{\alpha,\gamma}}$$

in the Gromov-Hausdorff sense, in probability as $n \to \infty$, where

$$\nu_{\alpha,\gamma}(d\mathbf{s}) = \left(\gamma + (1 - \alpha - \gamma)\sum_{i \neq j} s_i s_j\right)\mathrm{PD}^*_{\alpha,-\alpha-\gamma}(d\mathbf{s}).$$

The $\mathrm{PD}^*_{\alpha,\theta}$ is Poisson-Dirichet type, where for $0<\alpha<1, \theta>-2\alpha$ we can express

$$\int_{\mathcal{S}^{\downarrow}} f(s)\mathrm{PD}^*_{\alpha,\theta}(ds) = \mathbb{E}\left(\sigma_1^{-\theta} f\left(\Delta\sigma_{[0,1]}/\sigma_1\right)\right),$$

for an α-stable subordinator σ with Laplace exponent $-\log(\mathbb{E}(e^{-\lambda\sigma_1})) = \lambda^\alpha$ and with ranked sequence of jumps $\Delta\sigma_{[0,1]} = (\Delta\sigma_t, t\in[0,1])^{\downarrow}$.

Proof. The proof is similar with the version for alpha trees. Consider a first split at the first branch point b_1 of T_n into k subtrees $T^1,\ldots,T^k$ with $n_1\geq\ldots\geq n_k\geq 1$ leaves respectively. If both of the leaf 1 and 2 are in the same subtree T^i, let j be the smallest labels in $T_n\setminus T^i$. Then b_1 is created when the leaf j is added to the root edge of T_{j-1} with probability $\gamma/(j-1-\alpha)$. Hence, the probability of such first split is

$$\gamma\alpha^{k-2}\Gamma_{1+\frac{\gamma}{\alpha}}(k)\frac{\Gamma_\alpha(n_1)\cdots\Gamma_\alpha(n_k)}{\Gamma_\alpha(n)},$$

which is independent with j and is invariant under all possible labellings of $T^1,\ldots,T^k$ as long as 1 and 2 are in the same subtree. Similarly we can get the result for the case that the leaf 1 and 2 are in different subtrees. Therefore, we obtain the part exchangeability and splitting rules of labelled alpha-gamma tree. □

3.5.3 The three-factor model

For $\alpha\in(0,1), \theta\in[-2\alpha,-\alpha]$ and $\lambda\in[0,1]$, we choose

$$\nu_1(\mathrm{d}\mathbf{s}) = \lambda\mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s}),\ \nu_2 = (1-\lambda)\mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s}). \tag{3.65}$$

ν_1 fulfils (3.7), (3.2), (3.3) with $\nu_1(s_1\leq 1-\epsilon) = \epsilon^{\alpha+\theta}\ell(1/\epsilon)$ and ν_2 fulfils (3.1).

Definition 3.9. We call a family of partly exchangeable trees three-factor model with parameters $\alpha\in(0,1), \theta\in[-2\alpha,-\alpha]$ and $\lambda\in[0,1]$, if it has no erosion rate or killing rate, and two dislocation measures specified in (3.65).

This three parameter model contains the alpha-gamma model and Poisson-Dirichlet model [35]. When $\theta=-\alpha-\gamma, \lambda=\frac{\gamma}{1-\alpha+\gamma}$, it is the alpha-gamma model; when $\lambda=\frac{1}{2}$, it is Poisson-Dirichlet model. However, unlike the alpha-gamma model and Poisson-Dirichlet model, the three-factor model is not necessarily sampling consistent.

Corollary 3.39. *Let $(T_n^\circ, n \geq 1)$ be unlabelled trees associated with the three-factor model with parameters $(\alpha, \theta, \lambda)$, then*

(i)
$$\frac{T_n^\circ}{n^{-\alpha-\theta}\Gamma(1-\alpha)} \xrightarrow{(p)} \mathcal{T}_{-\alpha-\theta,\nu}$$
in the Gromov-Hausdorff sense, in probability as $n \to \infty$, where
$$\nu(d\mathbf{s}) = \left(1 - \lambda + (1-2\lambda)\sum_{i=1}^{\infty} s_i^2\right) \mathrm{PD}^*_{\alpha,\theta}(d\mathbf{s});$$

(ii) *$(T_n^\circ, n \geq 1)$ is sampling consistent if and only if $\lambda = \frac{\alpha+\theta}{2\alpha+\theta-1}$ or $\lambda = \frac{1}{2}$ i.e. alpha-gamma model or Poisson-Dirichlet model.*

Proof. Applying Theorem 3.2, (i) is straightforward.

(ii) it is easy to deduce that the EPPF of $(T_n^\circ, n \geq 1)$ is
$$p_n(n_1, \ldots, n_k) = C_{n,\alpha,\theta}\left((1-\lambda)n^2 - \lambda n + (2\lambda - 1)\sum_{j=1}^{k} n_j^2\right)\prod_{i=1}^{k}\Gamma_\alpha(n_i), \tag{3.66}$$
for any $n_1 \geq \ldots \geq n_k, n_1 + \ldots + n_k = n$, where $C_{n,\alpha,\theta}$ is a constant determined by n, α, θ. The sampling consistency of $(T_n^\circ, n \geq 1)$ is equivalent to
$$p_n(n_1, \ldots, n_k) \propto \sum_{i=1}^{k} p_{n+1}(n_1, \ldots, n_i + 1, \ldots, n_k) + p_{n+1}(n_1, \ldots, n_k, 1). \tag{3.67}$$
According to (3.66),
$$\begin{aligned}
&p_{n+1}(n_1, \ldots, n_i + 1, \ldots, n_k)\\
&\quad\propto \left((1-\lambda)n^2 - \lambda n + (2\lambda - 1)\sum_{j=1}^{k} n_j^2 + 2\lambda n_i + 2(1-\lambda)(n - n_i)\right)\\
&\qquad\times(n_i - \alpha)\prod_{i=1}^{k}\Gamma_\alpha(n_i), \qquad (3.68)\\
&p_{n+1}(n_1, \ldots, n_k, 1)\\
&\quad\propto \left((1-\lambda)n^2 - \lambda n + (2\lambda - 1)\sum_{j=1}^{k} n_j^2 + 2(1-\lambda)n\right)\\
&\qquad\times(k\alpha + \theta)\prod_{i=1}^{k}\Gamma_\alpha(n_i). \qquad (3.69)
\end{aligned}$$

Taking (3.66), (3.68) and (3.69) into (3.67), we obtain that the sampling consistency of $(T_n^\circ, n \geq 1)$ is equivalent to

$$\left((1-\lambda)n^2 - \lambda n + (2\lambda - 1)\sum_{j=1}^{k} n_j^2\right) \propto \left((1-\lambda)n^2 - (\alpha(2\lambda-1) - \theta(1-\lambda))n + (2\lambda-1)\sum_{j=1}^{k} n_j^2\right) \tag{3.70}$$

holds for any $n_1 + \ldots + n_k = n$. If $\lambda = \frac{1}{2}$, (3.70) holds naturally. Otherwise, (3.70) is equivalent to $\lambda = \alpha(2\lambda - 1) - \theta(1-\lambda)$ i.e. $\lambda = \frac{\alpha+\theta}{2\alpha+\theta-1}$. Hence we can get the sampling consistency is equivalent to $\lambda = \frac{\alpha+\theta}{2\alpha+\theta-1}$ or $\lambda = \frac{1}{2}$, which completes the proof.

□

Bibliography

[1] D. Aldous. Exchangeability and related topics. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.

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

[3] D. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[4] D. Aldous. Probability distributions on cladograms. In *Random discrete structures (Minneapolis, MN, 1993)*, volume 76 of *IMA Vol. Math. Appl.*, pages 1–18. Springer, New York, 1996.

[5] D. J. Aldous. Mixing time for a Markov chain on cladograms. *Combin. Probab. Comput.*, 9(3):191–204, 2000.

[6] D. J. Aldous. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Statist. Sci.*, 16(1):23–34, 2001.

[7] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

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

[9] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

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

[11] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.

[12] B. Chen, D. Ford, and M. Winkel. A new family of Markov branching trees: the alpha-gamma model. *Electron. J. Probab.*, 14:no. 15, 400–430, 2009.

[13] L. Devroye. Laws of large numbers and tail inequalities for random tries and PATRICIA trees. *J. Comput. Appl. Math.*, 142(1):27–37, 2002. Probabilistic methods in combinatorics and combinatorial optimization.

[14] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.

[15] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.

[16] R. Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[17] S. N. Evans, J. Pitman, and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.

[18] S. N. Evans and A. Winter. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.*, 34(3):918–961, 2006.

[19] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[20] T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1:209–230, 1973.

[21] D. J. Ford. Probabilities on cladograms: introduction to the alpha model. 2005. Preprint, arXiv:math.PR/0511246.

[22] A. Gnedin. Constrained exchangeable partitions. *Discrete Mathematics and Theoretical Computer Science Proc. Ser. A*, G:391–398, 2006.

[23] A. Gnedin, J. Pitman, and M. Yor. Asymptotic laws for compositions derived from transformed subordinators. *Ann. Probab.*, 34(2):468–492, 2006.

[24] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields*, 145(1-2):285–322, 2009.

[25] R. C. Griffiths. Allele frequencies with genic selection. *J. Math. Biol.*, 17(1):1–10, 1983.

[26] B. Haas. Loss of mass in deterministic and random fragmentations. *Stochastic Process. Appl.*, 106(2):245–277, 2003.

[27] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab.*, 9:no. 4, 57–97 (electronic), 2004.

[28] 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.

[29] B. Haas, J. Pitman, and M. Winkel. Spinal partitions and invariance under re-rooting of continuum random trees. *Ann. Probab.*, 37(4):1381–1411, 2009.

[30] G. Kersting. On the height profile of a conditioned galton-watson tree. *Preprint,Unpublished manuscript*, 1998.

[31] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982.

[32] J. F. C. Kingman, S. J. Taylor, A. G. Hawkes, A. M. Walker, D. R. Cox, A. F. M. Smith, B. M. Hill, P. J. Burville, and T. Leonard. Random discrete distribution. *J. Roy. Statist. Soc. Ser. B*, 37:1–22, 1975. With a discussion by S. J. Taylor, A. G. Hawkes, A. M. Walker, D. R. Cox, A. F. M. Smith, B. M. Hill, P. J. Burville, T. Leonard and a reply by the author.

[33] D. E. Knuth. *The art of computer programming. Vol. 1: Fundamental algorithms.* Second printing. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont, 1969.

[34] P. Marchal. A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 489–499. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.

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

[36] G. Miermont. Self-similar fragmentations derived from the stable tree. I. Splitting at heights. *Probab. Theory Related Fields*, 127(3):423–454, 2003.

[37] G. Miermont. Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields*, 131(3):341–375, 2005.

[38] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields*, 102(2):145–158, 1995.

[39] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002.

[40] J. Pitman and M. Winkel. Regenerative tree growth: binary self-similar continuum random trees and Poisson-Dirichlet compositions. *Preprint, arXiv:0803.3098,* 2008*, to appear in Annals of Probability.*

[41] G. U. Yule. *A mathematical theory of evolution: based on the conclusions of Dr. J. C. Willis, F. R. S.* Harris and Sons, London, 1924.