Random Structured Phylogenies

Jean-Jil Duchamps

École doctorale de sciences mathématiques de Paris centre

Thèse de doctorat

Discipline : Mathématiques

présentée par

Jean-Jil Duchamps

Phylogénies aléatoires structurées

---

dirigée par Amaury Lambert

Soutenue le 2 décembre 2019 devant le jury composé de :

Igor Kortchemski	CNRS	rapporteur
Romain Abraham	Université d'Orléans	examinateur
Brigitte Chauvin	Université de Versailles Saint-Quentin-en-Yvelines	examinatrice
Thomas Duquesne	Sorbonne Université	examinateur
Amaury Lambert	Sorbonne Université	directeur

Laboratoire de Probabilités, Statistique et Modélisation. UMR 8001.
Boîte courrier 158
4 place Jussieu
75252 Paris Cedex 05

Sorbonne Université.
École doctorale de sciences mathématiques de Paris centre.
Boîte courrier 290
4 place Jussieu
75252 Paris Cedex 05

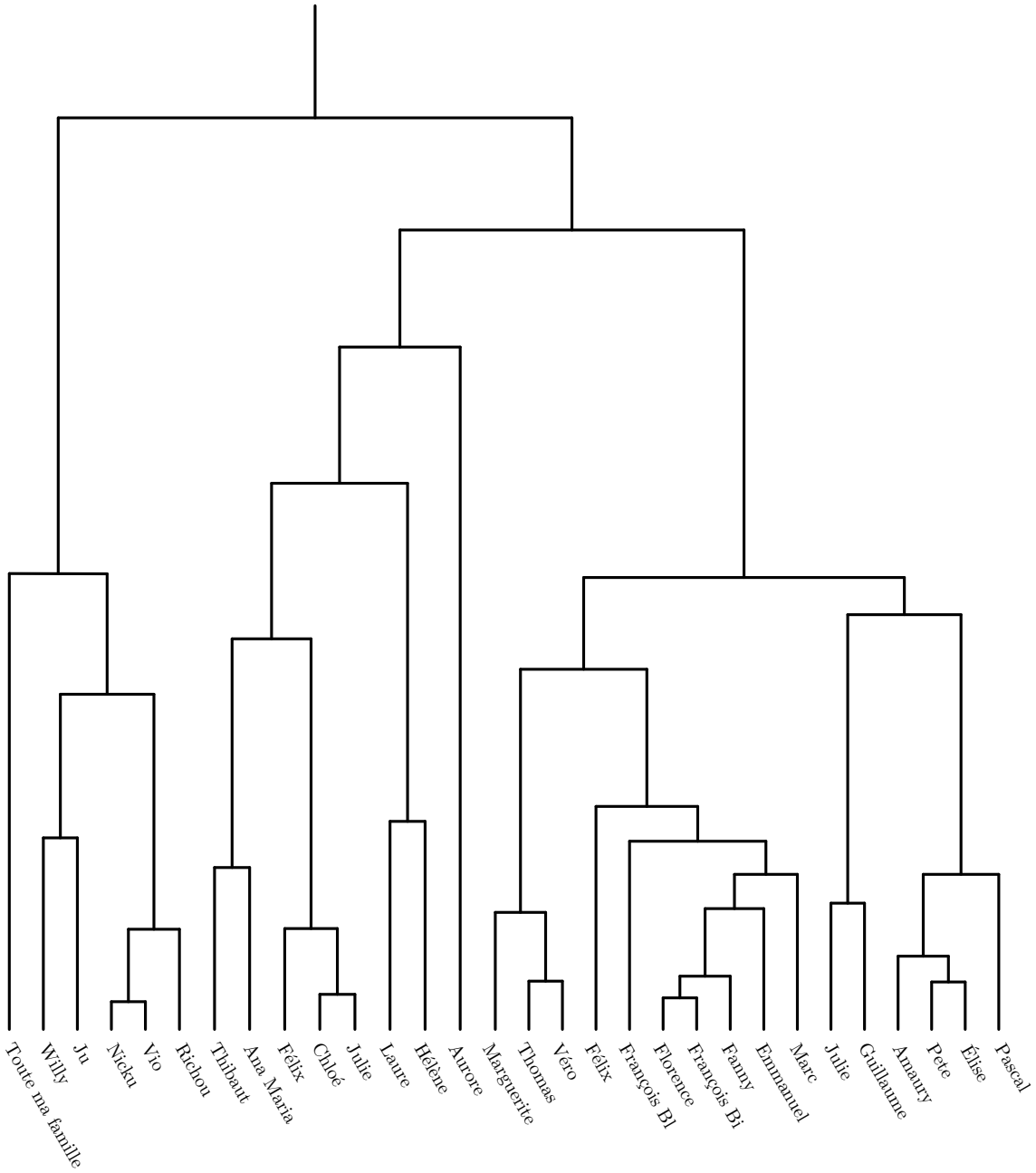
Résumé

Cette thèse est constituée de quatre chapitres indépendants, puisant leur origine en génétique des populations et en biologie évolutive, et liés à la théorie des processus de fragmentation ou de coalescence. Le chapitre 2 traite d'un arbre aléatoire binaire infini construit à partir d'un processus ponctuel coalescent équipé de mutations poissonniennes le long de ses branches et d'une mesure finie sur sa frontière. La partition allélique – partition de la frontière en parties qui portent la même combinaison de mutations – est définie pour cet arbre et sa mesure d'intensité est explicitée. Les chapitres 3 et 4 sont dédiés à l'étude de processus de coalescence et de fragmentation *emboîtés* – plus précisément à valeurs dans les couples de partitions emboîtées –, qui sont des analogues des Λ-coalescents et des fragmentations homogènes. Ces objets visent à modéliser un arbre de gènes niché dans un arbre d'espèces. Les coalescents emboîtés sont caractérisés par leurs coefficients de Kingman et leurs mesures de coagulation (éventuellement bivariées), tandis que les fragmentations emboîtées sont caractérisées par leurs coefficients d'érosion et leurs mesures de dislocation (éventuellement bivariées). Enfin le chapitre 5 pose la construction de processus de fragmentation à vitesses aléatoires, qui sont des processus de fragmentation où chaque fragment possède une marque qui accélère ou ralentit son taux de fragmentation, et où les marques de vitesse évoluent comme des processus de Markov positifs auto-similaires. Une caractérisation de type Lévy-Khintchine de ces processus de fragmentation généralisés est donné, ainsi que des conditions suffisantes pour l'absorption dans un état gelé, et pour que la généalogie du processus ait une longueur totale finie.

Abstract

This thesis consists of four self-contained chapters whose motivations stem from population genetics and evolutionary biology, and related to the theory of fragmentation or coalescent processes. Chapter 2 introduces an infinite random binary tree built from a so-called coalescent point process equipped with Poissonian mutations along its branches and with a finite measure on its boundary. The allelic partition – partition of the boundary into groups carrying the same combination of mutations – is defined for this tree and its intensity measure is described. Chapters 3 and 4 are devoted to the study of *nested* – i.e. taking values in the space of nested pairs of partitions – coalescent and fragmentation processes, respectively. These Markov processes are analogs of Λ-coalescents and homogeneous fragmentations in a nested setting – modeling a gene tree nested within a species tree. Nested coalescents are characterized in terms of Kingman coefficients and (possibly bivariate) coagulation measures, while nested fragmentations are similarly characterized in terms of erosion coefficients and (possibly bivariate) dislocation measures. Finally Chapter 5 gives a construction of fragmentation processes with speed marks, which are fragmentation processes where each fragment is given a mark that speeds up or slows down its rate of fragmentation, and where the marks evolve as positive self-similar Markov processes. A Lévy-Khinchin representation of these generalized fragmentation processes is given, as well as sufficient conditions for their absorption in finite time to a frozen state, and for the genealogical tree of the process to have finite total length.

Phylogénie des remerciements



N. B. : Cet arbre a été reconstitué sans données génétiques, ni véritable modèle d'ailleurs (l'aléatoire a en tout cas joué un grand rôle). Merci aussi à celles et ceux que j'ai pu oublier, et à celles et ceux qui lisent cette thèse.

Contents

Chapter 1

Introduction

Stochastic approaches to modeling genealogical or phylogenetic trees play a central role in population genetics and modern evolutionary biology. In this thesis, I study new models of random trees inspired by classic results and ideas in mathematical biology. Before going into detail, I briefly present two problems that motivated my research over the last years.

1.1 Kingman coalescent and Ewens sampling formula

1.1.1 Kingman coalescent: a universal random genealogy

First, consider the classical Moran model, which describes the evolution of a population of n individuals, where $n \in \mathbb{N}$ is constant in time. Its dynamics are as follows:

1. At any time $t \in \mathbb{R}$, n distinct individuals are alive and labeled with $[n] := \{1, \dots, n\}$.
2. From any point in time, at rate $\binom{n}{2}$ an individual j is chosen uniformly among the population; j is killed and replaced by an offspring of another individual i chosen among the remaining population.

This simple model can be entirely described by a random set of points $(t, i, j) \in \mathcal{M} \subset \mathbb{R} \times [n]^2$, where $\mathcal{M}$ is a Poisson point process with intensity $\binom{n}{2} \mathrm{d}t \otimes \mu$, where $\mathrm{d}t$ denotes the Lebesgue measure and μ is the uniform probability on ordered pairs of distinct elements of $[n]$. For two times $t < s$, the genealogy of individuals between those times is easily recovered from the point process $\mathcal{M}$ by following lineages backward in time (see Figure 1.1).

As the point process $\mathcal{M}$ is invariant under time-shifts, let us focus on the population at time 0. Note that going back in time, there exists a first time $T_n > 0$, which is almost surely finite, such that all n individuals alive at time 0 share the same ancestor at time $-T_n$. It is not hard to see that T_n is distributed as

$$T_n \overset{(d)}{=} \sum_{k=2}^{n} X_k,$$

where $(X_k, k \geq 1)$ are independent exponential random variables with parameter $\binom{k}{2}$, and that the genealogy of individuals at this time is a rooted tree with n leaves defined by the following backward-in-time procedure:

1. Start with n lineages labeled with $[n]$.

2. Let each pair of lineages coalesce (i.e. merge into a single lineage) at rate 1.

This random tree is known as the Kingman n-coalescent [58]. Other than being the genealogy of the Moran model, which is arguably the simplest toy model of population genetics one could think of, the n-coalescent is in fact a universal object. Indeed, it appears as the limiting genealogy of many population models, notably the Wright-Fisher model and more generally any Cannings model under suitable assumptions (these models are discrete population models that I do not describe here for conciseness – see e.g. [61] for an overview of these models). Two key properties of the Kingman coalescent make it an interesting mathematical object of study:

- it is sampling consistent, i.e. looking at the n first lineages in an m-coalescent for $m > n$ yields the n-coalescent.

- it comes down from infinity, i.e. it makes sense to define a coalescent started from an infinite number of lineages, which coalesce into finitely many lineages for any positive time.

We have identified a natural genealogy for a sample of individuals in a population. Now in order to study the effect of evolution, we need an additional process modeling it, superimposed on our random tree. This is the object of the next section.

1.1.2 Ewens sampling formula

Assume that individuals carry some genetic code (think of a long sequence of nucleotides) that is entirely duplicated when a parent produces an offspring (clonal reproduction). However, rare mutations, i.e. punctual changes in the sequence of an individual, may occur at random times, so at a fixed time individuals may be partitioned into classes with respect to the genetic code they share. A particular genetic sequence is called an allele so this partition into (allelic) classes is called the allelic partition. Let us encode the allelic

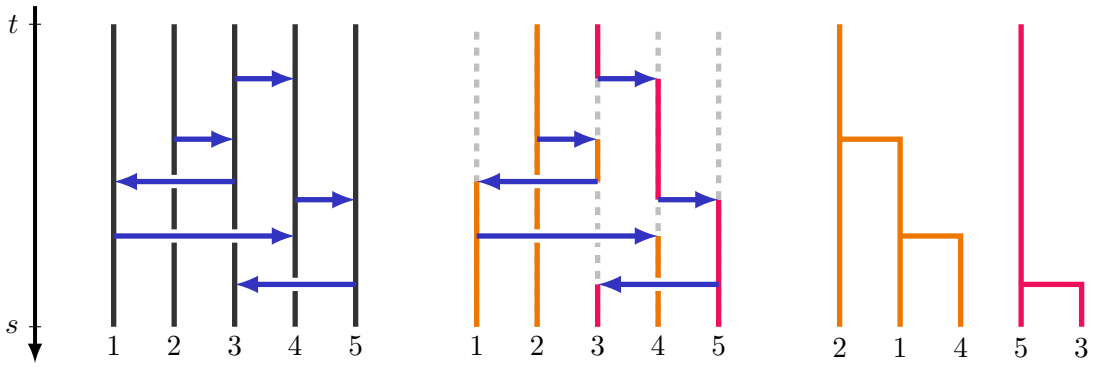


Figure 1.1 – Moran model for $n = 5$, between times $t < s$. Time flows from top to bottom, and events of replacement (u, i, j) are represented by horizontal blue arrows at time-coordinate u, pointing from the vertical line indexed by i towards the vertical line indexed by j. In this example, the ancestors at time t of the population at time s are individuals 2 and 3, respectively with descendants $\{1, 2, 4\}$ (orange lineages) and $\{3, 5\}$ (red lineages).

partition by the sequence $(A_1, A_2, \dots A_n)$, where A_k is the number of alleles shared by exactly k co-existing individuals in the population.

We take for granted the following simplifying assumptions:

- the infinite allele assumption, stating that any mutation gives rise to an allele that was never observed before.
- mutations are neutral, i.e. alleles do not affect the dynamics of the population in the underlying model.

A natural mathematical model for the allelic partition of a population of size n undergoing neutral selection is for instance the n-coalescent endowed with a point process on its branches recording where mutations have occurred. More precisely, given T_n the n-coalescent and a real number $\theta > 0$, assume that mutations arise at rate $\theta/2$ along the branches of T_n. Then the distribution of the allelic partition $(A_k, 1 \leq k \leq n)$ is given by the following so-called Ewens sampling formula.

$$\mathbb{P}(A_1 = a_1, \dots, A_n = a_n) = \frac{n!}{\theta(\theta+1)\cdots(\theta+n-1)} \prod_{k=1}^{n} \frac{\left(\frac{\theta}{k}\right)^{a_k}}{a_k!},$$

for any sequence $(a_1, \dots, a_n)$ of nonnegative integers such that $\sum_k k a_k = n$. This formula is named after Ewens [42] who discovered it in the study of the Wright-Fisher model. This sampling formula inspired Chapter 2 of this thesis, which aims at describing an analog of the allelic partition $(A_k, 1 \leq k \leq n)$ for some infinite trees – the so-called coalescent point processes, which can also be seen as limits of supercritical birth-death trees endowed with the uniform measure on leaves.

The other chapters of this thesis are connected to (or at least draw some inspiration from) the second idea I present.

1.2 Markov branching trees and fragmentation processes

Markov branching trees (MBT) were introduced by Aldous [3] in an attempt to identify some basic properties that models for phylogenetic trees should satisfy. Here, trees are rooted discrete trees with labeled leaves and for simplicity let us consider only binary trees, although the same ideas were developed for trees with any kind of degree distribution.

Definition 1.1. Let $q = (q_n, n \geq 2)$ be a family of probability measures, such that q_n is supported on $[n-1]$ and for all $i \in [n-1]$, $q_n(i) = q_n(n-i)$. The law of the Markov branching tree with n leaves associated with $(q_n, n \geq 2)$ is denoted by $\mathrm{MBT}(q, n)$ and defined inductively by:

1. Let $K \sim q_n$, and conditional on K, let $T' \sim \mathrm{MBT}(q, K)$ and $T'' \sim \mathrm{MBT}(q, n-K)$ be independent.
2. Define T_n as the grafting of T' and T'' on a root vertex, with uniformly relabeled leaves. Then T_n is distributed as $\mathrm{MBT}(q, n)$.

Note that in the last definition, q_n is the distribution of the number of leaves to the left of the root in the MBT with n leaves. It is immediate that these trees are exchangeable (i.e. invariant under permutations of the labels of the leaves). Some MBTs, including natural examples such as the uniform binary tree on n leaves or the Yule tree (where q_n is uniform on $[n-1]$), are sampling consistent, meaning that if $T_{n+1} \sim \mathrm{MBT}(q, n+1)$, then the tree obtained by erasing the leaf labeled $n+1$ in T_{n+1} has law $\mathrm{MBT}(q, n)$.

The distributions $(q_n, n \geq 2)$ associated with sampling consistent MBTs have been fully characterized in the 2000's, and in terms of binary Markov branching trees this translates to the following result.

Theorem 1.2 (Haas et al. [53, Proposition 3])**.** *Let $q = (q_n, n \geq 2)$ be the branching laws of a sampling consistent MBT. Then there is a measure μ on $[0,1]$ that is symmetric (i.e. invariant under $x \mapsto 1-x$) and satisfies*

$$\int_{[0,1]} x(1-x)\,\mu(\mathrm{d}x) \;+\; \mu(\{0,1\}) < \infty,$$

such that for all $n \geq 2$ and $1 \leq k < n$,

$$q_n(k) = \frac{1}{\alpha_n}\binom{n}{k}\left(\int_{(0,1)} x^k(1-x)^{n-k}\mu(\mathrm{d}x) + \mu(\{0\})\mathbb{1}_{k=1} + \mu(\{1\})\mathbb{1}_{k=n-1}\right), \tag{1.1}$$

with

$$\alpha_n = \int_{(0,1)} 1 - x^n - (1-x)^n \mu(\mathrm{d}x) + n\mu(\{0,1\}).$$

Furthermore, μ is unique up to a multiplicative constant.

Sampling consistent MBTs are in fact connected to the so-called exchangeable fragmentation processes with values in the partitions of $\mathbb{N}$ (see [10] for a comprehensive description of this framework). Informally, they can be described as branching processes recording a genealogy of fragments, which split independently of the others and in the same way as the original fragment. This genealogy is always a Markov branching tree, and the previous theorem establishes the converse. To be more specific, if μ satisfies the conditions of the theorem and if one defines

$$\nu := \mu_{|(0,1)} \circ \varphi^{-1},$$

where $\varphi : (0,1) \to \mathscr{S}^{\downarrow} := \{(s_1, s_2, \ldots) \in [0,1]^{\mathbb{N}},\ s_1 \geq s_2 \geq \ldots \text{ and } \sum_i s_i \leq 1\}$ is the map defined by

$$\varphi(x) = \begin{cases} (x, 1-x, 0, \ldots) & \text{if } x \geq 1/2, \\ (1-x, x, 0, \ldots) & \text{if } x < 1/2, \end{cases}$$

then the genealogy of a fragmentation process with erosion coefficient $\mu(\{0,1\})$ and dislocation measure ν (see Chapter 3 in [10]) is a binary MBT with distribution given by (1.1).

Viewing exchangeable partition-valued fragmentation processes as natural candidates for models of phylogenetic trees, I tried to study some of their generalizations in directions that were inspired, to some extent, by evolutionary biology.

1.3 Outline of the thesis

Chapter 2 is joint work with Amaury Lambert, and published in *The Annals of Applied Probability* [37]. We study a coalescent point process with mutations, that is a random infinite binary ultrametric real tree equipped with a point process of mutations on its branches and with a finite measure ℓ on its boundary (i.e. its leaves). As I have stated at the end of Section 1.1, this model of random tree can be seen as the limit as $t \to \infty$ of a supercritical birth-death tree conditioned on non-extinction at time t and endowed with a (rescaled) uniform measure on leaves. The allelic partition of this tree is represented by a point measure $\sum_i \delta_{\ell(A_i)}$ recording the mass $\ell(A_i)$ of each allelic class A_i.

We first study the clonal part of this tree, i.e. the subtree carrying the same allele as the root. We show that the clonal *boundary* can be viewed as a random regenerative subset of $\mathbb{R}$ whose distribution we characterize. From this, the intensity of the allelic partition is deduced. Finally, we study a natural coupling of clonal subtrees for varying mutation rates.

The rest of this thesis is set in the framework of exchangeable partition-valued fragmentation and coalescent processes (generalizations of the Kingman coalescent; I refer again to [10] for the theory).

More precisely, Chapter 3 and 4 are two sides (respectively, the coalescent side and the fragmentation side) of the same coin, which is the modeling of nested trees – a *gene* tree nested into a *species* tree, with terminology inspired by phylogenetics, and which meaning will be made clear in the aforementioned chapters. In each of these chapters, we define in a general way nested partition-valued processes $(\Pi^{\mathrm{g}}(t), \Pi^{\mathrm{s}}(t),\, t \geq 0)$, with $\Pi^{\mathrm{g}}(t)$ finer than $\Pi^{\mathrm{s}}(t)$ for all $t \geq 0$. The idea is the following: the genealogy of Π^{g} models a gene tree, which is nested into the genealogy of Π^{s} modeling the species tree. Using exchangeability, we characterize the possible distributions of these nested partition-valued processes, with results similar to Theorem 1.2.

Chapter 3 is joint work with Airam Blancas, Amaury Lambert and Arno Siri-Jégousse, and published in *Electronic Journal of Probability* [18]. In it, we only consider *simple* nested coalescents, which are a nested version of Λ-coalescents. More precisely, these simple processes are such that each partition only undergoes one coalescence event at a time (but possibly the same time). We characterize the law of these nested coalescents as follows. In the absence of gene coalescences, species blocks undergo Λ-coalescent type events and in the absence of species coalescences, gene blocks lying in the same species block undergo i.i.d. Λ-coalescents. Simultaneous coalescence of the gene and species partitions are governed by an intensity measure ν_s on $(0,1] \times \mathcal{M}_1([0,1])$ providing the frequency of species merging and the law in which are drawn (independently) the frequencies of genes merging in each coalescing species block. As an application, we also study the conditions under which a simple nested coalescent comes down from infinity.

In Chapter 4 (in press in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* [36]), I characterize the possible distributions of nested fragmentation processes in terms of erosion coefficients and dislocation measures. Three forms of erosion and two forms of dislocation are identified – one being specific to the nested setting and relating

to a bivariate paintbox process.

Chapter 5 (submitted to *Advances in Applied Probability*) aims at studying fragmentation processes with speed marks. More precisely, each fragment is given a mark that speeds up or slows down its rate of fragmentation, and the marks evolve as positive self-similar Markov processes. These processes are a natural generalization of self-similar fragmentations [11] and are comparable to the so-called self-similar growth-fragmentations of [9]. The main difference is that fragmentations with marks are still partition-valued processes whereas growth-fragmentation processes only describe the branching process of the marks, and need restrictive assumptions on their dynamics – notably binary, conservative splits.

Much like in previous work on fragmentation, I first give a Lévy-Khinchin representation of these generalized fragmentation processes using techniques from positive self-similar Markov processes and from classical fragmentation processes. Then I derive sufficient conditions for their absorption in finite time to a frozen state, and for the genealogical tree of the process to have finite total length.

References for the introduction

[3] D. Aldous. Probability Distributions on Cladograms. *Random Discrete Structures*. The IMA Volumes in Mathematics and Its Applications 76. Springer New York, 1996, pp. 1–18. DOI: 10.1007/978-1-4612-0719-1_1 (see pp. 9, 90).

[9] J. Bertoin. Markovian Growth-Fragmentation Processes. *Bernoulli*, 23.2 (May 2017), pp. 1082–1101. DOI: 10.3150/15-BEJ770 (see pp. 12, 126).

[10] J. Bertoin. *Random Fragmentation and Coagulation Processes*. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511617768 (see pp. 10, 11, 60, 61, 64, 65, 79, 81, 90, 91, 94, 95, 99, 117, 126–128, 135, 153).

[11] J. Bertoin. Self-Similar Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 38.3 (2002), pp. 319–340. DOI: 10.1016/S0246-0203(00)01073-6 (see pp. 12, 126, 134, 135).

[18] A. Blancas, J.-J. Duchamps, A. Lambert, and A. Siri-Jégousse. Trees within Trees: Simple Nested Coalescents. *Electron. J. Probab.*, 23.0 (2018). DOI: 10.1214/18-EJP219 (see pp. 11, 58, 133).

[36] J.-J. Duchamps. Trees within Trees II: Nested Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat. (to appear)*, (2019+). arXiv: 1807.05951 (see pp. 11, 89, 133, 153).

[37] J.-J. Duchamps and A. Lambert. Mutations on a Random Binary Tree with Measured Boundary. *Ann. Appl. Probab.*, 28.4 (Aug. 2018), pp. 2141–2187. DOI: 10.1214/17-AAP1353 (see pp. 11, 14).

[42] W. J. Ewens. The Sampling Theory of Selectively Neutral Alleles. *Theor. Popul. Biol.*, 3.1 (Mar. 1, 1972), pp. 87–112. DOI: 10.1016/0040-5809(72)90035-4 (see pp. 9, 15).

[53] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum Tree Asymptotics of Discrete Fragmentations and Applications to Phylogenetic Models. *Ann. Probab.*, 36.5 (Sept. 2008), pp. 1790–1837. DOI: 10.1214/07-AOP377 (see pp. 10, 90, 91, 126).

[58] J. Kingman. The Coalescent. *Stochastic Process. Appl.*, 13.3 (1982), pp. 235–248. DOI: 10.1016/0304-4149(82)90011-4 (see pp. 8, 15, 22, 59, 90, 95).

[61] A. Lambert. Population Dynamics and Random Genealogies. *Stoch. Models*, 24 (sup1 2008), pp. 45–163. DOI: 10.1080/15326340802437728 (see pp. 8, 58, 90).

Chapter 2

Mutations on a random binary tree with measured boundary

Joint work with Amaury Lambert. This chapter is published in *The Annals of Applied Probability* [37].

Contents

2.1 Introduction

In this paper, we give a new flavor of an old problem of mathematical population genetics which is to characterize the so-called *allelic partition* of a population. To address this problem, one needs to specify a model for the genealogy (i.e., a random tree) and a model for the mutational events (i.e., a point process on the tree). Two typical assumptions that we will adopt here are: the *infinite-allele assumption*, where each mutation event confers a new type, called *allele*, to its carrier; and the *neutrality of mutations*, in the sense that co-existing individuals are exchangeable, regardless of the alleles they carry. Here, our goal is to study the allelic partition of the boundary of some random real trees that can be seen as the limits of properly rescaled binary branching processes.

In a discrete tree, a natural object describing the allelic partition without labeling alleles is the *allele frequency spectrum* $(A_k)_{k\geq 1}$, where A_k is the number of alleles carried by exactly k co-existing individuals in the population. In the present paper, we start from a time-inhomogeneous, supercritical binary branching process with finite population $N(t)$ at any time t, and we are interested in the allelic partition of individuals 'co-existing at infinity' ($t \to \infty$), that is the allelic partition at the *tree boundary*. To define the analogue of the frequency spectrum, we need to equip the tree boundary with a measure ℓ, which we do as follows. Roughly speaking, if $N_u(t)$ is the number of individuals co-existing at time t in the subtree $\mathcal{T}_u$ consisting of descendants of the same fixed individual u, the measure $\ell(\mathcal{T}_u)$ is proportional to $\lim_{t\uparrow\infty} N_u(t)/N(t)$. It is shown in Section 2.5 that the tree boundary of any supercritical branching process endowed with the (properly rescaled) tree metric and the measure ℓ has the same law as a random real tree, called *coalescent point process* (CPP) generated from a Poisson point process, equipped with the so-called comb metric [60] and the Lebesgue measure. Taking this result for granted, we will focus in Sections 2.2, 2.3 and 2.4 on coalescent point processes with mutations.

In the literature, various models of random trees and their associated allelic partitions have been considered. The most renowned result in this context is *Ewens' Sampling Formula* [42], a formula that describes explicitly the distribution of the allele frequency spectrum in a sample of n co-existing individuals taken from a stationary population with genealogy given by the Wright-Fisher model with population size N and mutations occurring at birth with probability θ/N. When time is rescaled by N and $N \to \infty$, this model converges to the Kingman coalescent [58] with Poissonian mutations occurring at rate θ along the branches of the coalescent tree. In the same vein, a wealth of recent papers has dealt with the allelic partition of a sample taken from a Λ-coalescent or a Ξ-coalescent with Poissonian mutations, e.g., [5, 7, 47, 48].

In parallel, several authors have studied the allelic partition in the context of branching processes, starting with [51] and the monograph [89], see [24] and the references therein. In a more recent series of papers [22, 23, 33, 64], the second author and his co-authors have studied the allelic partition at a fixed time of so-called 'splitting trees', which are discrete branching trees where individuals live *i.i.d* lifetimes and give birth at constant rate. In particular, they obtained the almost sure convergence of the normalized frequency spectrum $(A_k(t)/N(t))_{k\geq 1}$ as $t \to \infty$ [22] as well as the convergence in distribution of the (properly rescaled) sizes of the most abundant alleles [23]. The limiting spectrum of these

trees is to be contrasted with the spectrum of their limit, which is the subject of the present study, as explained earlier.

Another subject of interest is the allelic partition of the entire progeny of a (sub)critical branching process, as studied in particular in [14]. The scaling limit of critical branching trees with mutations is a Brownian tree with Poissonian mutations on its skeleton. Cutting such a tree at the mutation points gives rise to a forest of trees whose distribution is investigated in the last section of [14], and relates to cuts of Aldous' CRT in [4] or the Poisson snake process [2]. The couple of previously cited works not only deal with the limits of allelic partitions for the whole discrete tree, but also tackle the limiting object directly. This is also the goal of the present work, but with quite different aims.

First, we construct in Section 2.2 an ultrametric tree with boundary measured by a 'Lebesgue measure' ℓ, from a Poisson point process with infinite intensity ν, on which we superimpose Poissonian neutral mutations with intensity measure μ. Section 2.2 ends with Proposition 2.12, which states that the total number of mutations in any subtree is either finite a.s. or infinite a.s. according to an explicit criterion involving ν and μ.

The structure of the allelic partition at the boundary is studied in detail in Section 2.3. Theorem 2.15 ensures that the subset of the boundary carrying no mutations (or clonal set) is a (killed) regenerative set with explicit Laplace exponent in terms of ν and μ and measure given in Corollary 2.20. The mean intensity Λ of the allele frequency spectrum at the boundary is defined by $\Lambda(B) := \mathbb{E}\sum \mathbb{1}_{\ell(R)\in B}$, where the sum is taken over all allelic clusters at the boundary. It is explicitly expressed in Proposition 2.23. An a.s. convergence result as the radius of the tree goes to infinity is given in Proposition 2.26 for the properly rescaled number of alleles with measure larger than $q > 0$, which is the analogue of $\sum_{k\geq q} A_k$ in the discrete setting.

Section 2.4 is dedicated to the study of the dynamics of the clonal (mutation-free) subtree when mutations are added or removed through a natural coupling of mutations in the case when $\mu(\mathrm{d}x) = \theta \mathrm{d}x$. It is straightforward that this process is Markovian as mutations are added. As mutations are removed, the growth process of clonal trees also is Markovian, and its semigroup and generator are provided in Theorem 2.29.

Section 2.5 is devoted to the links between measured coalescent point processes and measured pure-birth trees which motivate the present study. Lemma 2.35 gives a representation of every CPP with measured boundary, in terms of a rescaled pure-birth process with boundary measured by the rescaled counting measures at fixed times. Conversely, Theorem 2.36 gives a representation of any such pure-birth process in terms of a CPP with intensity measure $\nu(dx) = \frac{dx}{x^2}$, as in the case of the Brownian tree.

2.2 Preliminaries and Construction

2.2.1 Discrete Trees, Real Trees

Let us recall some definitions of discrete and real trees, which will be used to define the tree given by a so-called coalescent point process.

In graph theory, a tree is an acyclic connected graph. We call discrete trees such graphs

that are labeled according to Ulam–Harris–Neveu's notation by labels in the set $\mathcal{U}$ of finite sequences of non-negative integers:

$$\mathcal{U} = \bigcup_{n\geq 0} \mathbb{Z}_+^n = \{u_1 u_2 \dots u_n,\ u_i \in \mathbb{Z}_+, n \geq 0\},$$

with the convention $\mathbb{Z}_+^0 = \{\varnothing\}$.

Definition 2.1. A **rooted discrete tree** is a subset $\mathcal{T}$ of $\mathcal{U}$ such that

- $\varnothing \in \mathcal{T}$ and is called the **root** of $\mathcal{T}$
- For $u = u_1 \dots u_n \in \mathcal{T}$ and $1 \leq k < n$, we have $u_1 \dots u_k \in \mathcal{T}$.
- For $u \in \mathcal{T}$ and $i \in \mathbb{Z}_+$ such that $ui \in \mathcal{T}$, for $0 \leq j \leq i$, we have $uj \in \mathcal{T}$ and uj is called a **child** of u.

For $n \geq 0$, the **restriction of $\mathcal{T}$ to the first n generations** is defined by:

$$\mathcal{T}_{|n} := \{u \in \mathcal{T},\ |u| \leq n\},$$

where $|u|$ denotes the length of a finite sequence. For $u, v \in \mathcal{T}$, if there is $w \in \mathcal{U}$ such that $v = uw$, then u is said to be an **ancestor** of v, noted $u \preceq v$. Generally, let $u \wedge v$ denote the most recent common ancestor of u and v, that is the longest word $u_0 \in \mathcal{T}$ such that $u_0 \preceq u$ and $u_0 \preceq v$. The edges of $\mathcal{T}$ as a graph join the parents u and their children ui.

For a discrete tree $\mathcal{T}$, we define the **boundary of $\mathcal{T}$** as

$$\partial\mathcal{T} := \{u \in \mathcal{T},\ u0 \notin \mathcal{T}\} \cup \{v \in \mathbb{Z}_+^{\mathbb{N}},\ \forall u \in \mathcal{U}, u \preceq v \Rightarrow u \in \mathcal{T}\},$$

and we equip $\partial\mathcal{T}$ with the σ-field generated by the family $(B_u)_{u\in\mathcal{T}}$, where

$$B_u := \{v \in \partial\mathcal{T},\ u \preceq v\}.$$

Remark 2.2. With a fixed discrete tree $\mathcal{T}$, a finite measure $\mathscr{L}$ on $\partial\mathcal{T}$ is characterized by the values $(\mathscr{L}(B_u))_{u\in\mathcal{T}}$. Reciprocally if the number of children of u is finite for each $u \in \mathcal{T}$, by Carathéodory's extension theorem, any *finitely additive* map $\mathscr{L} : \{B_u,\ u \in \mathcal{T}\} \to [0, \infty)$ extends uniquely into a finite measure $\mathscr{L}$ on $\partial\mathcal{T}$.

By assigning a positive length to every edge of a discrete tree, one gets a so-called real tree. Real trees are defined more generally as follows, see e.g. [41].

Definition 2.3. A metric space $(\mathbb{T}, d)$ is a **real tree** if for all $x, y \in \mathbb{T}$,

- There is a unique isometry $f_{x,y} : [0, d(x, y)] \to \mathbb{T}$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x, y)) = y$,
- All continuous injective paths from x to y have the same range, equal to $f_{x,y}([0, d(x, y)])$.

This unique path from x to y is written $[\![x, y]\!]$. The **degree** of a point $x \in \mathbb{T}$ is defined as the number of connected components of $\mathbb{T} \setminus \{x\}$, so that we may define:

- The **leaves** of $\mathbb{T}$ are the points with degree 1.
- The **internal nodes** of $\mathbb{T}$ are the points with degree 2.
- The **branching points** of $\mathbb{T}$ are the points with degree larger than 2.

One can root a real tree by distinguishing a point $\varrho \in \mathbb{T}$, called the **root**.

From this definition, one can see that for a rooted real tree $(\mathbb{T}, d, \varrho)$, for all $x, y \in \mathbb{T}$, there exists a unique point $a \in \mathbb{T}$ such that $[\![\varrho, x]\!] \cap [\![\varrho, y]\!] = [\![\varrho, a]\!]$. We call a the **most recent common ancestor** of x and y, noted $x \wedge y$. There is also an intrinsic order relation in a rooted tree: if $x \wedge y = x$, that is if $x \in [\![\varrho, y]\!]$, then x is called an ancestor of y, noted $x \preceq y$.

We will call a rooted real tree a **simple** tree if it can be defined from a discrete tree by assigning a length to each edge. From now on, we will restrict our attention to simple trees.

Definition 2.4. A **simple (real) tree** is given by $(\mathcal{T}, \alpha, \omega)$, where $\mathcal{T} \subset \mathcal{U}$ is a rooted discrete tree, and α and ω are maps from $\mathcal{T}$ to $\mathbb{R}$ satisfying

$$\zeta(u) := \omega(u) - \alpha(u) > 0,$$

$$\forall u \in \mathcal{T}, \forall i \in \mathbb{Z}_+, \quad ui \in \mathcal{T} \Longrightarrow \alpha(ui) = \omega(u).$$

Here $\alpha(u)$ and $\omega(u)$ are called the **birth time** and **death time** of u and $\zeta(u)$ is the **life length** of u.

We will sometimes consider simple trees $(\mathcal{T}, \alpha, \omega, \mathscr{L})$ equipped with $\mathscr{L}$ a **measure on their boundary** $\partial\mathcal{T}$.

We call a **reversed simple tree** a triple $(\mathcal{T}, \alpha, \omega)$ where $(\mathcal{T}, -\alpha, -\omega)$ is a simple tree. We may sometimes omit the term "reversed" when the context is clear enough.

The **restriction** of $A = (\mathcal{T}, \alpha, \omega)$ to the first n generations is the simple tree defined by

$$A_{|n} = (\mathcal{T}_{|n}, \alpha_{|\mathcal{T}_{|n}}, \omega_{|\mathcal{T}_{|n}}).$$

One can check that a simple tree $(\mathcal{T}, \alpha, \omega)$ defines a unique real rooted tree defined as the completion of $(\mathbb{T}, d, \varrho)$, with

$$\begin{aligned} \varrho &:= (\varnothing, \alpha(\varnothing)), \\ \mathbb{T} &:= \{\varrho\} \cup \bigcup_{u \in \mathcal{T}} \{u\} \times (\alpha(u), \omega(u)] \quad \subset \mathcal{U} \times \mathbb{R}, \\ d((u,x),(v,y)) &:= \begin{cases} |x - y| & \text{if } u \preceq v \text{ or } v \preceq u, \\ x + y - 2\omega(u \wedge v) & \text{otherwise.} \end{cases} \end{aligned} \tag{2.1}$$

In particular, we have $(u, x) \wedge (v, y) = (u \wedge v, \omega(u \wedge v))$.

In this paper, we construct random simple real trees with marks along their branches. We see these trees as genealogical/phylogenetic trees and the marks as mutations that appear in the course of evolution. We will assume that each new mutation confers a new

type, called **allele**, to its bearer (infinitely-many alleles model). Our goal is to study the properties of the **clonal subtree** (individuals who do not bear any mutations, black subtree in Figure 2.1) and of the **allelic partition** (the partition into bearers of distinct alleles of the population at some fixed time).

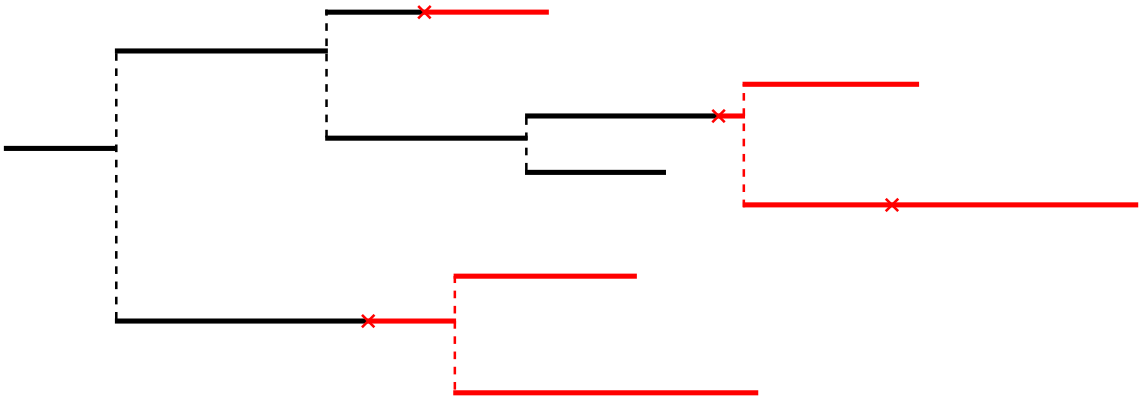


Figure 2.1 – Simple tree with mutations

2.2.2 Comb Function

Definition

We now introduce ultrametric trees, using a construction with comb functions following Lambert and Uribe Bravo [60].

Definition 2.5. Let $T > 0$ and $I = [0, T]$. Let also $f : I \to [0, \infty)$ such that

$$\#\{x \in I,\ f(x) > \varepsilon\} < \infty \qquad \varepsilon > 0.$$

The pair (f, I) will be called a **comb function**. For any real number $z > \max_I f$, we define the **ultrametric tree of height** z **associated with** (f, I) as the real rooted tree T_f which is the completion of $(\mathrm{Sk}, \varrho, d_f)$, where $\mathrm{Sk} \subset I \times [0, \infty)$ is the **skeleton** of the tree, and Sk, ϱ and d_f are defined by

$$\varrho := (0, z),$$
$$\mathrm{Sk} := \{0\} \times (0, z] \cup \{(t, y) \in I \times (0, z],\ f(t) > y\},$$
$$d_f : \begin{cases} \mathrm{Sk}^2 & \longrightarrow [0, \infty) \\ ((t, x), (s, y)) & \longmapsto \begin{cases} |\max_{(t,s]} f - x| + |\max_{(t,s]} f - y| & \text{if } t < s, \\ |x - y| & \text{if } t = s. \end{cases} \end{cases}$$

The set $\{0\} \times (0, z] \subset \mathrm{Sk}$ is called the **origin branch of the tree**.

For $t \in I$, $t > 0$, we call the **lineage of** t the subset of the tree $L_t \subset T_f$ defined as the closure of the set

$$\{(s, x) \in \mathrm{Sk},\ s \leq t,\ \forall s < u \leq t,\ f(u) \leq x\}.$$

For $t = 0$ one can define L_0 as the closure of the origin branch.

Remark 2.6. One can check that d_f is a distance which makes (Sk, d_f) a real tree, and so its completion (T_f, d_f) also is a real tree. Furthermore, the fact that $\{f > \varepsilon\}$ is finite for all $\varepsilon > 0$ ensures that it is a simple tree, since the branching points in Sk are the points $(t, f(t))$ with $f(t) > 0$. For a visual representation of the tree associated with a comb function, see Figure 2.2, where the skeleton is drawn in vertical segments and the dashed horizontal segments represent branching points.

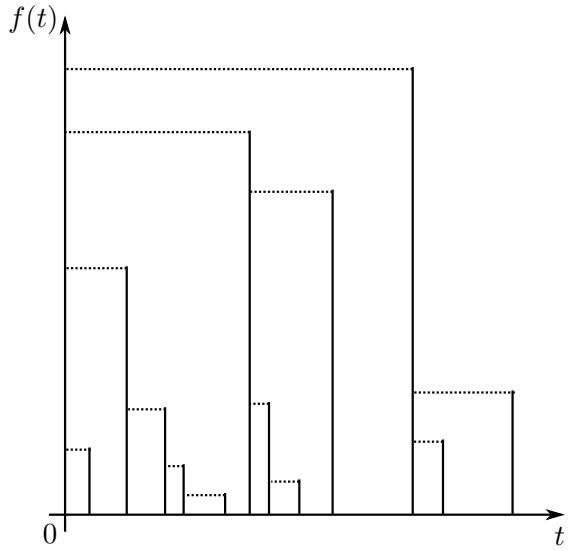


Figure 2.2 – Comb function and its associated tree.

Proposition 2.7. *With the same notation as in Definition 2.5, for a fixed comb function (f, I) and a real number $z > \max_I f$, writing T_f for the associated real tree, the following holds. For each $t \in I$, there is a unique leaf $\alpha_t \in T_f$ such that*

$$L_t = [\![\varrho, \alpha_t]\!].$$

Furthermore, the map $\alpha : t \mapsto \alpha_t$ is measurable with respect to the Borel sets of I and T_f.

Proof. For $t = 0$, L_0 is defined as the closure of the origin branch $\{0\} \times (0, z]$. Since $d_f((0, x), (0, y)) = |x - y|$, the map

$$\varphi_0 : \begin{cases} (0, z] & \longrightarrow \mathrm{Sk} \\ x & \longmapsto (0, x) \end{cases}$$

is an isometry, and since T_f is defined as the completion of the skeleton Sk, there is a unique isometry $\widetilde{\varphi}_0 : [0, z] \to T_f$ which extends φ_0. Therefore we define $\alpha_0 := \widetilde{\varphi}_0(0) \in T_f$, which satisfies $L_0 = [\![\varrho, \alpha_0]\!]$ since $\widetilde{\varphi}_0$ is an isometry. Also α_0 is a leaf of T_f because it is in $T_f \setminus \mathrm{Sk}$. Indeed, since T_f is the completion of Sk which is connected, $T_f \setminus \{\alpha_0\}$ is necessarily also connected, which means that α_0 has degree 1.

Now for a fixed $t \in I$, $t > 0$, write $(t_i, x_i)_{i \geq 0}$ for the (finite or infinite) sequence with values in

$$\{(0, z)\} \cup \{(s, x) \in I \times (0, \infty),\ f(s) = x\}$$

defined inductively (as long as they can be defined) by $(t_0, x_0) = (0, z)$ and

$$\forall i \geq 0, \quad x_{i+1} := \max_{(t_i, t]} f,$$
$$\text{and } t_{i+1} := \max\{s \in (t_i, t],\ f(s) = x_{i+1}\}.$$

- If the sequence $(t_i, x_i)_{i\geq 0}$ is well defined for all $i \geq 0$, then since f is a comb function, we necessarily have that $x_i \to 0$ as $i \to \infty$.
- On the other hand, the sequence $(t_i, x_i)_{0\leq i\leq n}$ is finite if and only if it is defined up to an index n such that either $t_n = t$ or f is zero on the interval $(t_n, t]$. In that case, we still define for convenience $x_{n+i} := 0$, $t_{n+i} := t_n$ for all $i \geq 1$.

Now it can be checked that we have

$$\bigcup_{i=0}^{\infty} [x_{i+1}, x_i) \setminus \{0\} = (0, z),$$

and that L_t is defined as the closure of the set

$$A_t := \bigcup_{i=0}^{\infty} \{t_i\} \times ([x_{i+1}, x_i) \setminus \{0\}) \subset \mathrm{Sk}.$$

Also, by definition of the sequence $(t_i, x_i)_{0\leq i}$, the distance d_f satisfies, for $(s, x), (u, y) \in A_t$,

$$d_f((s, x), (u, y)) = |x - y|.$$

Therefore the following map is an isometry (and it is well defined because $x_i \downarrow 0$).

$$\varphi_t : \begin{cases} (0, z) & \longrightarrow \mathrm{Sk} \\ x & \longmapsto (t_i, x) \quad \text{if } x \in [x_{i+1}, x_i) \text{ for an index } i \geq 0. \end{cases}$$

As in the case $t = 0$, this isometry can be extended to $\widetilde{\varphi}_t : [0, z] \to T_f$ and we define $\alpha_t := \widetilde{\varphi}_t(0)$. It is a leaf of T_f satisfying $L_t = [\![\varrho, \alpha_t]\!]$ for the same reasons as for 0.

It remains to prove that $\alpha : t \mapsto \alpha_t$ is measurable. It is enough to show that it is right-continuous, because in that case the pre-image of an open set is necessarily a countable union of right-open intervals, which is a Borel set. Now for $t < t' \in I$, by taking limits along the lineages L_t and $L_{t'}$, it is easily checked that the distance between α_t and $\alpha_{t'}$ can be written

$$d_f(\alpha_t, \alpha_{t'}) = 2 \max_{(t, t']} f,$$

and since f is a comb function, necessarily we have

$$\max_{(t, t']} f \xrightarrow[t' \downarrow t]{} 0.$$

Hence α is right-continuous, therefore measurable. □

It follows from Proposition 2.7 that the Lebesgue measure λ on the real interval I can be transported by the map α to a measure on the tree T_f, or more precisely on its boundary, that is the set of its leaves.

Definition 2.8. With the same notation as in Definition 2.5 and Proposition 2.7, for any fixed comb function (f, I) and $z > \max_I f$, writing T_f for the associated real tree, we define the **measure on the boundary of** T_f as the measure

$$\ell := \lambda \circ \alpha^{-1}$$

which concentrates on the leaves of the tree. From now on, we always consider the tree T_f associated with a comb function f as a rooted real tree equipped with the measure ℓ on its boundary.

The Coalescent Point Process

Here we will consider the measured tree associated to a random comb function. Let ν be a positive measure on $(0, \infty]$ such that for all $\varepsilon > 0$, we have

$$\overline{\nu}(\varepsilon) := \nu([\varepsilon, \infty]) < \infty,$$

and $\mathcal{N}$ be the support of the Poisson point process on $[0, \infty) \times (0, \infty]$ with intensity $\mathrm{d}t \otimes \nu$. Then we can define $f^{\mathcal{N}}$ as the function whose graph is $\mathcal{N}$.

$$f^{\mathcal{N}}(t) = \begin{cases} x & \text{if } (t, x) \in \mathcal{N}, \\ 0 & \text{if } \mathcal{N} \cap (\{t\} \times (0, \infty]) = \varnothing. \end{cases}$$

Now fix $z > 0$ such that $\overline{\nu}(z) > 0$ and set

$$T(z) := \inf\{t \geq 0, f^{\mathcal{N}}(t) \geq z\}.$$

Definition 2.9. The ultrametric random tree associated to $I = [0, T(z))$ and $f^{\mathcal{N}}_{|I}$ is called **coalescent point process (CPP)** of intensity ν and height z, denoted by $\mathrm{CPP}(\nu, z)$. It is equipped with the random measure ℓ, concentrated on the leaves, which is the push-forward of the Lebesgue measure on $[0, T(z))$ by the map α.

Note that a coalescent point process is not directly related to coalescent theory, a canonical example of which is Kingman's coalescent [58], although there exist links between the two: it is shown in [66] that a CPP appears as a scaling limit of the genealogy of individuals having a very recent common ancestor in the Kingman coalescent.

Formally, a CPP is a random variable valued in the space of finitely measured compact metric spaces endowed with the Gromov-Hausdorff-Prokhorov distance defined in [1] as an extension of the more classical Gromov-Hausdorff distance. Actually, it is easy to check that all the random quantities we handle are measurable, since we are dealing with a construction from a Poisson point process.

2.2.3 Mutations on a CPP

Here we set up how mutations appear on the random genealogy associated with a CPP of intensity ν. Let μ be a positive measure on $[0, \infty)$. We make the following assumptions:

$$\begin{gathered} \forall x > 0, \qquad 0 < \overline{\nu}(x) := \nu([x, \infty]) < \infty \ \text{ and } \ \underline{\mu}(x) := \mu([0, x]) < \infty, \\ \mu([0, \infty)) = \infty, \\ \nu \text{ and } \mu \text{ have no atom on } [0, \infty). \end{gathered} \tag{H}$$

We will now define the CPP of intensity ν and height $z > 0$ marked with rate μ.

Recall that the CPP is constructed from the support $\mathcal{N}$ of a Poisson point process with intensity $\mathrm{d}t \otimes \nu$ on $[0, \infty) \times [0, \infty]$ and has a root $\varrho = (0, z)$. Define independently for each point $N := (t, x)$ of $\mathcal{N} \cup \{\varrho\}$ the Poisson point process M_N of intensity μ on the interval $(0, x)$. Each atom $y \in [0, x]$ of M_N is a mark (t, y) on the branch $\{t\} \times (0, x) \subset \mathrm{Sk}$ at height y. The family $(M_N)_{N \in \mathcal{N}}$ therefore defines a point process M on the skeleton of the CPP tree:

$$M := \sum_{(t,x) \in \mathcal{N} \cup \{\varrho\}} \sum_{y \in M_{(t,x)}} \delta_{(t,y)}.$$

By definition, conditional on Sk, M is a Poisson point process on Sk whose intensity is such that for all non-negative real numbers t and $a < b$, we have:

$$\mathbb{E}\Big[M(\{t\} \times [a, b]) \;\Big|\; \{t\} \times [a, b] \subset \mathrm{Sk} \Big] = \mu([a, b]).$$

Definition 2.10. Let ν, μ be measures satisfying assumption (H). A **coalescent point process with intensity ν, mutation rate μ and height** z, denoted $\mathrm{CPP}(\nu, \mu, z)$, is defined as the random $\mathrm{CPP}(\nu, z)$ given by $\mathcal{N}$, equipped with the point process M on its skeleton.

(i) The **clonal subtree** of the rooted real tree $(\mathbb{T}, \varrho)$ equipped with mutations M is defined as the subset of $\mathbb{T}$ formed by the points :

$$\{x \in \mathbb{T}, M([\![\varrho, x]\!]) = 0\}.$$

Equipped with the distance induced by d, this is also a real tree.

(ii) Given the (ultrametric) rooted real tree $(\mathbb{T}, \varrho)$ equipped with mutations M and the application α from the real interval $I = [0, T(z))$ to $\mathbb{T}$ whose range is included in the leaves of $\mathbb{T}$, we can define the **clonal boundary** (or **clonal population**) $R = R(\mathbb{T}, M, \alpha) \subset I$:

$$R := \{t \in I, M([\![\varrho, \alpha_t]\!]) = 0)\}.$$

Remark 2.11. This set R is studied in a paper by Philippe Marchal [69] for a CPP with $\nu(dx) = \frac{\mathrm{d}x}{x^2}$ and mutations at branching points with probability $1 - \beta$. In that case the sets R_β have the same distribution as the range of a β-stable subordinator. In the present case of Poissonian mutations, R is not stable any longer but we will see in Section 2.3 that it remains a regenerative set.

Total number of mutations. Since μ is a locally finite measure on $[0, \infty)$, the number of mutations on a fixed lineage of the $\mathrm{CPP}(\nu, \mu, z)$ is a Poisson random variable with parameter $\mu([0, z]) < \infty$, and so is a.s. finite. However, it is possible that in a **clade** (here defined as the union of all lineages descending from a fixed point), there are infinitely many mutations with probability 1. For instance, if μ is the Lebesgue measure and if ν is such that

$$\int_0 x\nu(\mathrm{d}x) = \infty,$$

we know from the properties of Poisson point processes that the total length of any clade is a.s. infinite. In this case, the number of mutations in any clade is also a.s. infinite so that each point x in the skeleton of the tree has a.s. at least one descending lineage with infinitely many mutations. Such a lineage can be displayed by choosing iteratively at each branching point a sub-clade with infinitely many mutations.

One can ask under which conditions this phenomenon occurs. Conditional on the tree of height z, the total number of mutations follows a Poisson distribution with parameter

$$\Lambda := \underline{\mu}(z) + \sum_{(t,y)\in\mathcal{N}, t<T(z)} \underline{\mu}(y),$$

where $T(z)$ is the first time such that there is a point of $\mathcal{N}$ with height larger than z. Indeed, the origin branch is of height z and the heights of the other branches are the heights of points of $\mathcal{N}$. This number of mutations is finite *a.s.* on the event $A := \{\Lambda < \infty\}$ and infinite *a.s.* on its complement. But by the properties of Poisson point processes, two cases are distinguished: either A has probability 0 or it has probability 1.

Proposition 2.12. *There is the following dichotomy:*

$$\int_0 \underline{\mu}(x)\nu(\mathrm{d}x) < \infty \quad \Longrightarrow \quad \textit{the total number of mutations is finite a.s.}$$
$$\int_0 \underline{\mu}(x)\nu(\mathrm{d}x) = \infty \quad \Longrightarrow \quad \textit{the number of mutations in any clade is infinite a.s.}$$

In the former case, the total number of mutations has mean

$$\mathbb{E}[\Lambda] = \underline{\mu}(z) + \frac{1}{\overline{\nu}(z)} \int_{[0,z]} \underline{\mu}(x)\nu(\mathrm{d}x).$$

Proof. Conditional on $T(z)$, the set $\mathcal{N}' := \{(t,y) \in \mathcal{N}, t < T(z)\}$ is the support of a Poisson point process on $[0, T(z)] \times [0, z]$ with intensity $\mathrm{d}t \otimes \nu$. Therefore, from basic properties of Poisson point processes, conditional on $T(z)$, $\Lambda = \underline{\mu}(z) + \sum_{(t,y)\in\mathcal{N}'} \underline{\mu}(y)$ is finite *a.s.* if and only if

$$\int_0^{T(z)} \left(\int_{[0,z]} \left(\underline{\mu}(x) \wedge 1 \right) \nu(\mathrm{d}x) \right) \mathrm{d}t < \infty \qquad a.s.,$$

and since $T(z)$ is finite *a.s.* and $\underline{\mu}$ is increasing, this condition is equivalent to the condition of the proposition. Now let us write N_{tot} for the total number of mutations. The conditional distribution of N_{tot} given Λ is a Poisson distribution with mean Λ. Therefore we deduce

$$\begin{aligned}
\mathbb{E}[N_{\mathrm{tot}}] &= \mathbb{E}[\Lambda] \\
&= \underline{\mu}(z) + \mathbb{E}\left[\sum_{(t,y)\in\mathcal{N}'} \underline{\mu}(y) \right] \\
&= \underline{\mu}(z) + \mathbb{E}\left[T(z) \int_{[0,z]} \underline{\mu}(x)\nu(\mathrm{d}x) \right] \\
&= \underline{\mu}(z) + \frac{1}{\overline{\nu}(z)} \int_{[0,z]} \underline{\mu}(x)\nu(\mathrm{d}x),
\end{aligned}$$

which concludes the proof. □

2.3 Allelic Partition at the Boundary

In this section, we will identify the clonal boundary R in a mutation-equipped CPP, that is the set of leaves of the tree which do not carry mutations, and characterize the reduced subtree generated by this set.

2.3.1 Regenerative Set of the Clonal Lineages, Clonal CPP

Denote by $\mathbb{T}^z$ a CPP(ν, μ, z) where ν, μ satisfy assumptions (H). A leaf of $\mathbb{T}^z$ is said **clonal** if it carries the same allele as the root. Recall the canonical map α^z from the real interval $[0, T(z))$ to the leaves of $\mathbb{T}^z$ (see Proposition 2.7). The **clonal boundary** (see Definition 2.10) of $\mathbb{T}^z$ is then the set $R^z \subset [0, T(z))$ defined as the pre-image of the clonal leaves by the map α^z.

We define the event

$$O^z := \{M_\varrho([0, z]) = 0\}$$

that there is no mutation on the origin branch of $\mathbb{T}^z$. Note that this event has a positive probability equal to $\mathrm{e}^{-\underline{\mu}(z)}$. By definition, the point process of mutations on the origin branch M_ϱ is independent of $(M_N)_{N\in\mathcal{N}}$. Therefore conditioning on O^z amounts to considering the tree $\mathbb{T}^z$ equipped with the mutations on its skeleton which are given only by the point processes $(M_N)_{N\in\mathcal{N}}$. We now define a random set $\widetilde{R}$, whose distribution depends only on (ν, μ) and not on z, which will allow the characterization of the clonal boundaries R^z conditional on the event O^z.

Definition 2.13. Recall the notations $\mathcal{N}$ and $(M_N)_{N\in\mathcal{N}}$. For each fixed $t \in [0, \infty)$, let $(t_i, x_i)_{i\geq 1}$ be the (possibly finite) sequence of points of $\mathcal{N}$ such that

$$\begin{aligned} x_1 &= \sup\{x \in [0, \infty],\ \#\mathcal{N} \cap (0, t] \times [x, \infty] \geq 1\}, \\ t_1 &= \sup\{s \in [0, t],\ (s, x_1) \in \mathcal{N}\}, \\ x_{i+1} &= \sup\{x \in [0, x_i),\ \#\mathcal{N} \cap (t_i, t] \times [x, \infty] \geq 1\}, \\ t_{i+1} &= \sup\{s \in (t_i, t],\ (s, x_{i+1}) \in \mathcal{N}\}, \end{aligned}$$

with the convention $\sup \varnothing = 0$, and where the sequence is finite if there is a $n \geq 0$ such that $x_n = 0$. We define the following random point measure on $[0, \infty)$:

$$M_t := \sum_{i\geq 1,\ x_i>0} M_{(t_i,x_i)}(\,\cdot \cap [x_{i+1}, x_i]).$$

Now we define the random set $\widetilde{R}$ as:

$$\widetilde{R} := \{t \in [0, \infty),\ M_t([0, \infty)) = 0\}.$$

Remark 2.14. Recall that for a comb function (f, I) and a real number $t \in I$, in the proof of Proposition 2.7, we defined a sequence $(t_i, x_i)_{i\geq 0}$ in the same way as in the previous definition and we remarked that the lineage L_t of t is the closure of the set

$$\bigcup_{i\geq 0,\ x_i>0} \{t_i\} \times ([x_{i+1}, x_i) \setminus \{0\}) \subset \mathrm{Sk}.$$

It follows that in the case of the tree $\mathbb{T}^z$ equipped with the mutations M on its skeleton, we have the equality between events

$$O^z \cap \{M([\![\varrho, \alpha_t^z]\!]) = 0\} = O^z \cap \{M_t([0,\infty)) = 0\}.$$

Therefore, on the event O^z, the clonal boundary R^z of the tree $\mathbb{T}^z$ coincides with the restriction of $\widetilde{R}$ to the interval $[0, T(z))$, which explains why we study the set $\widetilde{R}$.

The subtree of $\mathbb{T}^z$ spanned by the clonal boundary R^z is called the **reduced clonal subtree** and defined as

$$\bigcup_{t\in R^z} [\![\varrho, \alpha_t^z]\!].$$

Note that it is a Borel subset of $\mathbb{T}^z$ because it is the closure of

$$\bigcap_{n\geq 1} \bigcup_{p\geq n} \bigcup_{x\in C_p} [\![\varrho, x]\!],$$

where C_p is the finite set $\{x \in \mathbb{T}^z,\ d(x,\varrho) = z(1-1/p),\ M([\![\varrho, x]\!]) = 0\}$. The set $\widetilde{R}$ is proven to be a regenerative set (see Appendix 2.A.3 for the results used in this paper and the references concerning subordinators and regenerative sets), and the reduced clonal subtree is shown to have the law of a CPP.

Theorem 2.15. *The law of $\widetilde{R}$ and of the associated reduced clonal subtree can be characterized as follows.*

(i) Under the assumptions (H) and with the preceding notation the random set $\widetilde{R}$ is regenerative. It can be described as the range of a subordinator whose Laplace exponent φ is given by:

$$\frac{1}{\varphi(\lambda)} = \int_{(0,\infty)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{\lambda + \overline{\nu}(x)} \mu(\mathrm{d}x).$$

(ii) The reduced clonal subtree, that is the subtree spanned by the set $\widetilde{R}$, has the distribution of a CPP with intensity ν^μ, where ν^μ is the positive measure on $\mathbb{R}_+ \cup \{\infty\}$ determined by the following equation. Letting $W(x) := (\overline{\nu}(x))^{-1}$ and $W^\mu(x) := (\overline{\nu^\mu}(x))^{-1}$, we have, for all $x > 0$,

$$W^\mu(x) = W(0) + \int_0^x \mathrm{e}^{-\underline{\mu}(z)} \mathrm{d}W(z).$$

Remark 2.16. The last formula of the theorem is an extension of Proposition 3.1 in [64], where the case when ν is a finite measure and $\mu(dx) = \theta\, \mathrm{d}x$ is treated. Here, we allow ν to have infinite mass and μ to take a more general form (provided (H) is satisfied).

Regenerative set. Here, we prove the first part of the theorem concerning $\widetilde{R}$.

Proof of Theorem 2.15, (i). Let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration of the marked CPP defined by:

$$\mathcal{F}_t = \sigma\left(\mathcal{N} \cap ([0,t] \times \mathbb{R}_+), M_{(s,x)}, s \leq t, x \geq 0\right).$$

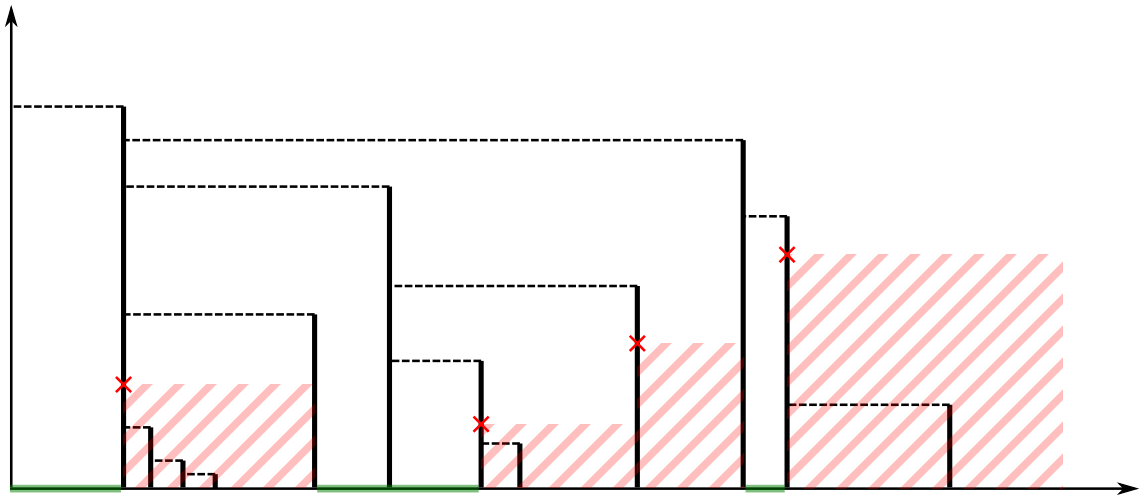


Figure 2.3 – Mutation-equipped CPP, regenerative set $\widetilde{R}$ shown in green

To show first that $\widetilde{R}$ is $(\mathcal{F}_t)$-progressively measurable, we show that for a fixed $t > 0$, the set

$$\{(s, \omega) \in [0, t] \times \Omega,\ s \in \widetilde{R}(\omega)\}$$

is in $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$. Basic properties of Poisson point processes ensure there exists an $\mathcal{F}_t$-measurable sequence of random variables giving the coordinates of the mutations in $\mathcal{N} \cap ([0, t] \times \mathbb{R}_+)$. Let $(U_i, X_i)_i$ be such a sequence, for instance ranked such that X_i is decreasing as in Figure 2.4. We also define the following $\mathcal{F}_t$-measurable random variables:

$$T_i := t \wedge \inf\{s \geq U_i, (s, x) \in \mathcal{N}, x \geq X_i\}.$$

Now we have

$$\widetilde{R} \cap [0, t] = \bigcap_i ([0, t] \setminus [U_i, T_i)),$$

which proves that the random set $\widetilde{R}$ is $(\mathcal{F}_t)$-progressively measurable, and almost-surely left-closed.

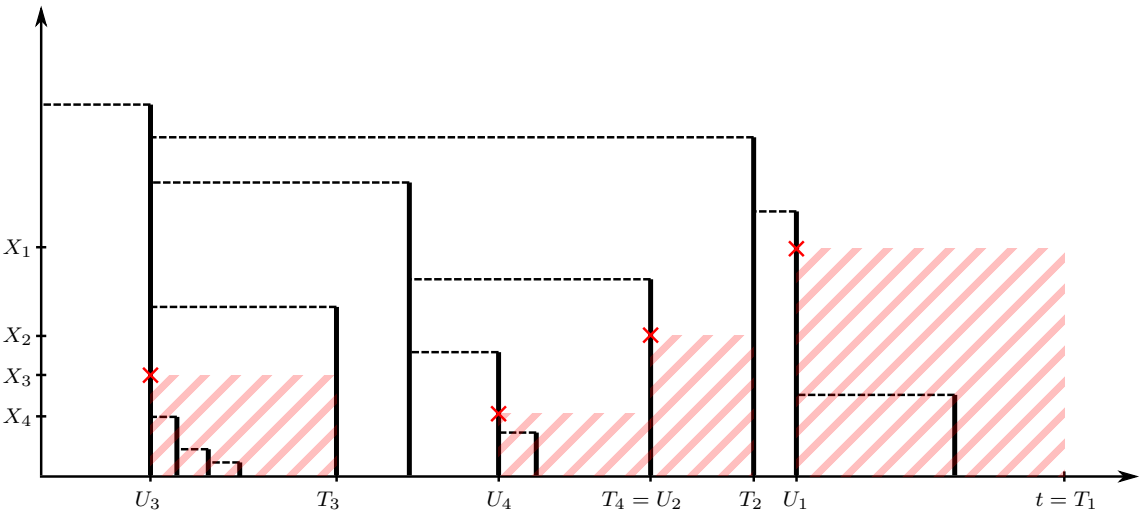


Figure 2.4 – Mutations localized by the variables (U_i, X_i, T_i)

Let us now show the regeneration property of $\widetilde{R}$. Define

$$H(s,t) := \max\{x \geq 0, (u,x) \in \mathcal{N}, s < u \leq t\},$$

the maximal height of atoms of $\mathcal{N}$ between s and t. We will note $H(t) := H(0,t)$ for simplicity. Remark that

$$\widetilde{R} = \{t \geq 0, M_t([0, H(t)]) = 0\}.$$

Let S be a $(\mathcal{F}_t)$-stopping time, and suppose that almost surely, $S < \infty$, and $S \in \widetilde{R}$ is not isolated to the right. From elementary properties of Poisson point processes and the fact that the random variables $(M_{(s,x)})_{s\geq 0, x\geq 0}$ are i.i.d, we know that the tree strictly to the right of S is independent of $\mathcal{F}_S$ and has the same distribution as the initial tree. Now since $S \in \widetilde{R}$ almost surely, we have, for all $t \geq S$,

$$M_t([0, H(t)]) = M_t([0, H(S,t)]),$$

because $M_t([H(S,t), H(0,t)]) = M_S([H(S,t), H(0,t)]) = 0$, in other words there are no mutations on the lineage of t that is also part of the lineage of S. As a consequence,

$$\widetilde{R} \cap [S, \infty) = \{t \geq S, M_t([0, H(S,t)]) = 0\},$$

which implies that $\widetilde{R} \cap [S,\infty) - S$ has the same distribution as $\widetilde{R}$ and is independent of $\mathcal{F}_S$.

Therefore it is proven that $\widetilde{R}$ has the regenerative property, so one can compute its Laplace exponent. Here we are in the simple case where $\widetilde{R}$ has a positive Lebesgue measure, and we have in particular, for all $t \in \mathbb{R}_+$,

$$\begin{aligned}
\mathbb{P}(t \in \widetilde{R}) &= \mathbb{E}\left[\mathrm{e}^{-\underline{\mu}(H_t)}\right] \\
&= \int_{[0,\infty]} \mathbb{P}(H_t \in \mathrm{d}x)\mathrm{e}^{-\underline{\mu}(x)} \\
&= \int_{(0,\infty)} \mathbb{P}(H_t \leq x)\mathrm{e}^{-\underline{\mu}(x)}\mu(\mathrm{d}x) \\
&= \int_{(0,\infty)} \mathrm{e}^{-t\overline{\nu}(x)-\underline{\mu}(x)}\mu(\mathrm{d}x).
\end{aligned}$$

The passage from the second to the third line is done integrating by parts thanks to the assumption that $\underline{\mu}$ is continuous and that μ has an infinite mass. The last displayed expression is therefore the density with respect to the Lebesgue measure of the renewal measure of $\widetilde{R}$ (see Remark 2.45). This is sufficient to characterize our regenerative set, and the expression given in the Proposition is found by computing the Laplace transform of this measure:

$$\begin{aligned}
\frac{1}{\varphi(\lambda)} &= \int_0^\infty \mathrm{e}^{-\lambda t}\left(\int_{(0,\infty)} \mathrm{e}^{-t\overline{\nu}(x)-\underline{\mu}(x)}\mu(\mathrm{d}x)\right)\mathrm{d}t \\
&= \int_{(0,\infty)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{\lambda + \overline{\nu}(x)}\mu(\mathrm{d}x),
\end{aligned}$$

which concludes the proof of (i). □

Remark 2.17. It is important to note that the particular case of a CPP with intensity $\nu(dx) = \frac{dx}{x^2}$ has the distribution of a (root-centered) sphere of the so-called Brownian CRT (Continuum Random Tree), the real tree whose contour is a Brownian excursion. This is shown for example by Popovic in [77] where the term 'Continuum genealogical point process' is used to denote what is called here a coalescent point process. The measure $\nu(dx) = \frac{dx}{x^2}$ is the push-forward of the Brownian excursion measure by the application which maps an excursion to its depth. In general, the sphere of radius say r of a totally ordered tree is an ultrametric space whose topology is characterized by the pairwise distances between 'consecutive' points at distance r from the root. When the order of the tree is the order associated to a contour process, these distances are the depths of the 'consecutive' excursions of the contour process away from r, see e.g. Lambert and Uribe Bravo [60].
If in addition to $\nu(dx) = \frac{dx}{x^2}$, we assume that $\mu(dt) = \theta\, \mathrm{d}t$, which amounts to letting Poissonian mutations at constant rate θ on the skeleton of the CRT, we have

$$\frac{1}{\varphi_\theta(\lambda)} = \int_0^\infty \frac{\theta \mathrm{e}^{-\theta x}}{\lambda + 1/x} \mathrm{d}x.$$

In particular, for all $\theta, c > 0$, we can compute:

$$\varphi_\theta(c\lambda) = c\varphi_{\theta/c}(\lambda).$$

This implies the equality in distribution $cR_\theta \overset{(d)}{=} R_{\theta/c}$. Nevertheless R_θ is not a so-called 'stable' regenerative set, contrary to the sets R_α in [69].

Reduced clonal subtree. To show that the reduced clonal subtree is a CPP, let us exhibit the Poisson point process that generates it. Let σ be the subordinator with drift 1 whose range is $\widetilde{R}$ and let $\mathcal{N}'$ be the following point process:

$$\mathcal{N}' := \{(t, x),\ t \in \mathbb{R}_+,\ x = H(\sigma_{t-}, \sigma_t) > 0\},$$

where $H(s,t) := \max\{x, (u,x) \in \mathcal{N}, s \le u \le t\}$. This point process generates the reduced clonal subtree, because $H(\sigma_{t-}, \sigma_t)$ is (up to a factor 1/2) the tree distance between the consecutive leaves σ_{t-} and σ_t in $\widetilde{R}$. To complete the proof of the theorem, it is sufficient to show that conditional on the death time ζ of the subordinator σ, $\mathcal{N}'$ is a Poisson point process on $[0, \zeta) \times \mathbb{R}_+$ with intensity $\mathrm{d}t \otimes \nu^\mu$.

Proof of Theorem 2.15, (ii). This is due to the regenerative property of the process. For fixed $t \ge 0$, σ_t is a $(\mathcal{F}_t)$-stopping time which is almost surely in $\widetilde{R}$ on the event $\{\sigma_t < \infty\} = \{\zeta > t\}$. This implies that conditional on $\{\sigma_t < \infty\}$, the marked CPP strictly to the right of σ_t is equal in distribution to the original marked CPP and is independent of $\mathcal{F}_{\sigma_t}$. In particular:

$$\Big(\{(s,x) \in \mathbb{R}_+^2, (\sigma_t + s, x) \in \mathcal{N}\}, \widetilde{R} \cap [\sigma_t, \infty) - \sigma_t\Big) \overset{(d)}{=} (\mathcal{N}, \widetilde{R}).$$

This implies that $\mathcal{N}' \cap ([t, \infty) \times \mathbb{R}_+) - (t, 0)$ has the same distribution as $\mathcal{N}'$ and is independent of $\mathcal{F}_{\sigma_t}$. For fixed $\varepsilon > 0$, let $(T_i, X_i)_{i \ge 1}$ be the sequence of atoms of $\mathcal{N}'$ such that $X_i > \varepsilon$, ranked with increasing T_i. Then T_i is a $(\mathcal{F}_{\sigma_t})$-stopping time and the sequence

$(T_i - T_{i-1}, X_i)_{i\geq 1}$ is i.i.d., with $T_0 := 0$ for convenience. It is sufficient to observe that T_1 is an exponential random variable to show that $\mathcal{N}'$ has an intensity of the form $\mathrm{d}t \otimes \nu^\mu$:

$$\begin{aligned}\mathbb{P}(T_1 > t+s \mid T_1 > t) &= \mathbb{P}(H(0,\sigma_{t+s}) \leq \varepsilon \mid H(0,\sigma_t) \leq \varepsilon) \\ &= \mathbb{P}(H(\sigma_t,\sigma_{t+s}) \leq \varepsilon \mid H(0,\sigma_t) \leq \varepsilon) \\ &= \mathbb{P}(H(0,\sigma_s) \leq \varepsilon) = \mathbb{P}(T_1 > s).\end{aligned}$$

It remains to characterize the measure ν^μ by computing $W^\mu(x)$. Note that the following computations are correct thanks to the assumption that ν has no atom, so that W is continuous. To simplify the notation, let $H_t := H(0,t) = \max\{x,\ (u,x) \in \mathcal{N},\ 0 \leq u \leq t\}$. Then we can compute:

$$\begin{aligned}W^\mu(x) &= \int_0^\infty \mathrm{e}^{-t\overline{\nu^\mu}(x)}\mathrm{d}t \\ &= \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{H_{\sigma_t}\leq x\}}\mathrm{d}t\right] \\ &= \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{H_u \leq x\}}\mathbb{1}_{\{u\in\widetilde{R}\}}\mathrm{d}u\right].\end{aligned} \tag{2.2}$$

Letting $F(y) := \mathbb{P}(H_u \leq y) = \mathrm{e}^{-u\overline{\nu}(y)}$, we have

$$\begin{aligned}\mathbb{P}(H_u \leq x, u \in \widetilde{R}) &= \mathbb{P}(H_u = 0) + \int_0^x \mathbb{P}(H_u \in \mathrm{d}y)\mathrm{e}^{-\underline{\mu}(y)} \\ &= F(0) + \int_0^x \mathrm{e}^{-\underline{\mu}(y)}\mathrm{d}F(y).\end{aligned}$$

Now $\mathrm{d}F(y) = u\mathrm{e}^{-u\overline{\nu}(y)}\nu(\mathrm{d}y)$, hence

$$\begin{aligned}W^\mu(x) &= \int_0^\infty \mathrm{e}^{-u\overline{\nu}(0)}\mathrm{d}u + \int_0^x \left(\int_0^\infty u\mathrm{e}^{-u\overline{\nu}(y)}\mathrm{d}u\right)\mathrm{e}^{-\underline{\mu}(y)}\nu(\mathrm{d}y) \\ &= \frac{1}{\overline{\nu}(0)} + \int_0^x \frac{1}{\overline{\nu}(y)^2}\mathrm{e}^{-\underline{\mu}(y)}\nu(\mathrm{d}y) \\ &= W(0) + \int_0^x \mathrm{e}^{-\underline{\mu}(y)}\mathrm{d}W(y),\end{aligned}$$

which concludes the proof. □

Remark 2.18. Equality (2.2) becomes, letting $x \to \infty$,

$$W^\mu(\infty) = \mathbb{E}[\lambda(\widetilde{R})].$$

Remark 2.19. In Remark 2.17, we explained that when the contour of a random tree is a strong Markov process as in the case of Brownian motion, the root-centered sphere of radius r of this tree is a CPP. In addition, the intensity measure of this CPP is the measure of the excursion depth under the excursion measure of the contour process (away from r). Let $\mathbf{n}_c$ denote the excursion measure of the process $(B_t^{(c)} - \inf_{s\leq t} B_s^{(c)})_{t\geq 0}$ away from 0, with $B^{(c)}$ a Brownian motion with drift c, and let h denote the depth of the excursion. In the case $\nu(\mathrm{d}x) = \frac{\mathrm{d}x}{x^2} = \mathbf{n}_0(h \in \mathrm{d}x)$ and $\mu(\mathrm{d}x) = \theta\,\mathrm{d}x$, we have

$$W^\theta(x) = \frac{1-\mathrm{e}^{-\theta x}}{\theta} = \mathbf{n}_{\theta/2}(h \in [x,\infty])^{-1}.$$

This is consistent with Proposition 4 in [2], which shows that putting Poissonian random cuts with rate θ along the branches of a standard Brownian CRT yields a tree whose contour process is $(e(s)-\theta s/2)_{s\geq 0}$ stopped at the first return at 0, where e is the normalized Brownian excursion.

2.3.2 Measure of the Clonal Population

Recall that for a CPP(ν, μ, z), conditional on O^z (no mutation on the origin branch), the Lebesgue measure $\lambda(\widetilde{R} \cap [0, T(z))$ is equal to the measure $\ell(R^z)$ of the set of clonal leaves .

Corollary 2.20. *Let ν, μ be two measures satisfying assumptions (H).*

(i) With the notation of Theorem 2.15, the random variable $\lambda(\widetilde{R})$ follows an exponential distribution with mean $W^\mu(\infty)$.

(ii) In a CPP(ν, μ, z), conditional on O^z, the measure $\ell(R^z)$ of the set of clonal leaves is an exponential random variable of mean $W^\mu(z)$.

Proof. Given a subordinator σ with drift 1 and range $\widetilde{R}$, it is known (a quick proof of this can be found in [12]) that

$$\lambda(\widetilde{R}) = \inf\{t > 0, \sigma_t = \infty\}.$$

Now the killing time of the subordinator σ is an exponential random variable of parameter $\varphi(0)$, where φ is the Laplace exponent of σ. We already know from Remark 2.18 the mean of that variable:

$$\varphi(0)^{-1} = \mathbb{E}\left[\lambda(\widetilde{R})\right] = W^\mu(\infty).$$

With a fixed height $z > 0$, one is interested in the law of $\lambda(\widetilde{R} \cap [0, T(z)))$. By the properties of Poisson point processes, stopping the CPP at $T(z)$ amounts to changing the intensity measure ν of the CPP for $\widehat{\nu}$, with

$$\widehat{\nu} = \nu(\,\cdot \cap [0, z]) + \overline{\nu}(z)\delta_\infty.$$

Then if $\widehat{W}(x) := \widehat{\nu}([x, \infty])^{-1}$, we have

$$\begin{aligned}
\widehat{W}(x) &= \left(\nu([x, \infty] \cap [0, z]) + \overline{\nu}(z)\right)^{-1} \\
&= \left(\nu([x \wedge z, z]) + \nu([z, \infty])\right)^{-1} \\
&= \left(\nu([x \wedge z, \infty])\right)^{-1} \\
&= W(x \wedge z),
\end{aligned}$$

and because of the characterization of W^μ given in Theorem 2.15, we also have $\left(\widehat{W}\right)^\mu(x) = W^\mu(x \wedge z)$. Therefore $\left(\widehat{W}\right)^\mu(\infty) = W^\mu(z)$, and we can conclude that $\lambda(\widetilde{R} \cap [0, T(z)])$ is an exponential random variable of mean $W^\mu(z)$. □

Probability of clonal leaves. Here, we consider a CPP(ν, μ, z) and aim at computing the probability of existence of clonal leaves in the tree.

Proposition 2.21. *In a CPP(ν, μ, z), under the assumptions (H) and with the notation of Theorem 2.15, there is a mutation-free lineage with probability*

$$\frac{W(z)\,\mathrm{e}^{-\underline{\mu}(z)}}{W^{\mu}(z)}.$$

Remark 2.22. Using a description of CPP trees in terms of birth-death trees (see Section 2.5), the previous result could alternatively be deduced from the expression of the survival probability of a birth-death tree up to a fixed time (see Proposition 2.37 in the appendix).

Proof. Suppose the CPP(ν, μ, z) is given by the usual construction with the Poisson point processes $\mathcal{N}$ and $(M_N)_{n\in\mathcal{N}}$. We use the regenerative property of the process with respect to the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ of the marked CPP defined by:

$$\mathcal{F}_t = \sigma\left(\mathcal{N} \cap ([0,t]\times\mathbb{R}_+), M_{(s,x)}, s \leq t, x \geq 0\right).$$

Let X be the first clone on the real half-line.

$$X := \inf\{x \in [0, T(z)),\ M(\llbracket \varrho, \alpha_x \rrbracket) = 0\},$$

with the convention $\inf \varnothing = \infty$ and with the usual notation. Then X is a $(\mathcal{F}_t)$-stopping time, and conditional on $\{X < \infty\}$, the law of the tree on the right of X is the same as that of the original tree conditioned on having no mutation on the origin branch. Let $C^z := \{X < \infty\}$ denote the event of existence of a mutation-free lineage. Recall that R^z denotes the set of clonal leaves and that O^z denotes the event that there is no mutation on the origin branch. Then we have

$$\begin{aligned}
\mathbb{E}\left[\ell(R^z)\right] &= \mathbb{P}(C^z)\mathbb{E}\left[\ell(R^z) \mid C^z\right] \\
&= \mathbb{P}(C^z)\mathbb{E}\left[\ell(R^z \cap [X,\infty) - X) \mid X < \infty\right] \\
&= \mathbb{P}(C^z)\mathbb{E}\left[\ell(R^z) \mid O^z\right] \\
&= \mathbb{P}(C^z)W^{\mu}(z),
\end{aligned}$$

where the last equality is due to Corollary 2.20 (ii). Furthermore,

$$\begin{aligned}
\mathbb{E}\left[\ell(R^z)\right] &= \mathbb{E}\int_0^{T(z)} \mathbb{1}_{\{t\in\widetilde{R}\}}\,\mathrm{d}t \\
&= \int_0^{\infty} \mathbb{P}(t \in \widetilde{R}, t < T(z))\,\mathrm{d}t \\
&= \int_0^{\infty} \mathrm{e}^{-t\overline{\nu}(z)}\mathrm{e}^{-\underline{\mu}(z)}\,\mathrm{d}t \\
&= \frac{\mathrm{e}^{-\underline{\mu}(z)}}{\overline{\nu}(z)} = W(z)\,\mathrm{e}^{-\underline{\mu}(z)}.
\end{aligned}$$

Therefore, the probability that there exists a clone of the origin in the present population is

$$\mathbb{P}(C^z) = \frac{W(z)\,\mathrm{e}^{-\underline{\mu}(z)}}{W^{\mu}(z)},$$

which concludes the proof. □

2.3.3 Application to the Allele Frequency Spectrum

Intensity of the Spectrum

From now on we fix two measures ν, μ satisfying assumptions (H), and we further assume for simplicity that $\overline{\nu}(z) \in (0, \infty)$ for all $z > 0$. We denote by $\mathbb{T}^z$ a CPP(ν, μ, z).

Under the infinitely-many alleles model, recall that each mutation gives rise to a new type called allele, so that the population on the boundary of the tree can be partitioned into carriers of the same allele, called *allelic partition*. The key idea of this section is that expressions obtained for the clonal population of the tree allow us to gain information on quantities related to the whole allelic partition. We call $m \in \mathbb{T}^z$ a mutation if $M(\{m\}) \neq 0$ and denote by $\mathbb{T}^z_m$ the subtree descending from m. If f is a functional of real trees (say simple, marked, equipped with a measure on the leaves), one might be interested in the quantity

$$\varphi(\mathbb{T}^z, f) := \sum_{\substack{m \in \mathbb{T}^z \\ \text{mutation}}} f(\mathbb{T}^z_m), \tag{2.3}$$

or in its expectation

$$\psi(z, f) := \mathbb{E}\left[\varphi(\mathbb{T}^z, f)\right].$$

For each mutation $m \in \mathbb{T}^z$, we define the set R^z_m of the leaves carrying m as their last mutation

$$R^z_m := \{t \in \mathbb{R}_+, \text{ the most recent mutation on the lineage of } \alpha^z_t \text{ is } m\}.$$

We define the random point measure putting mass on the measures of the different allelic clusters

$$\Phi_z := \sum_{\substack{m \in \mathbb{T}^z \\ \text{mutation}}} \mathbb{1}_{\{R^z_m \neq \varnothing\}}\, \delta_{\lambda(R^z_m)}.$$

The intensity of the allele frequency spectrum is the mean measure Λ_z of this point measure, that is the measure on $\mathbb{R}_+$ such that for every Borel set B of $\mathbb{R}_+$,

$$\Lambda_z(B) = \mathbb{E}[\Phi_z(B)].$$

The analog for this measure when the number of individuals in the population is finite is the mean measure $(\mathbb{E}A(k))_{k>0}$ of the number $A(k)$ of alleles carried by exactly k individuals (notation $A_\theta(k,t)$ in [64] and [22]). The goal here is then to identify Λ_z, by noticing that for a Borel set B,

$$\Phi_z(B) = \varphi(\mathbb{T}^z, f_B) \quad \text{and} \quad \Lambda_z(B) = \psi(z, f_B),$$

with $f_B(\mathbb{T}) := \mathbb{1}_{\ell(R)\in B}$, where $\mathbb{T}$ is an ultrametric tree with point mutations and measure ℓ supported by its leaves, and R denotes the set of its clonal leaves.

Proposition 2.23. *In a CPP(ν, μ, z), under the assumptions (H) and with the notation of Theorem 2.15, the intensity of the allele frequency spectrum has a density with respect to the Lebesgue measure:*

$$\frac{\Lambda_z(\mathrm{d}q)}{\mathrm{d}q} = W(z)\left(\frac{\mathrm{e}^{-\underline{\mu}(z)}}{W^\mu(z)^2}\mathrm{e}^{-q/W^\mu(z)} + \int_{[0,z)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)^2}\mathrm{e}^{-q/W^\mu(x)}\mu(\mathrm{d}x)\right).$$

Remark 2.24. This expression is to be compared with Corollary 4.3 in [22] (the term $(1-\frac{1}{W^\theta(x)})^{k-1}$ with discrete k becoming here $\mathrm{e}^{-q/W^\mu(x)}$ with continuous q).

Remark 2.25. Integrating this expression, we get the expectation of the number of different alleles in the population:

$$\Lambda_z(\mathbb{R}_+) = \mathbb{E}[\Phi_z(\mathbb{R}_+)] = W(z)\left(\frac{\mathrm{e}^{-\underline{\mu}(z)}}{W^\mu(z)} + \int_{[0,z)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)}\mu(\mathrm{d}x)\right).$$

Note that $W(z)$ is the expectation of the total mass of the measure ℓ in a CPP(ν,μ,z). It is then natural to normalize by this quantity and then let $z\to\infty$. In (H) we assumed that $\mu([0,\infty)) = \infty$, and since $W^\mu(z)$ is an increasing, positive function of z, we have clearly $\frac{\mathrm{e}^{-\underline{\mu}(z)}}{W^\mu(z)} \to 0$ when $z\to\infty$. Therefore we have

$$\lim_{z\to\infty} \frac{\mathbb{E}[\Phi_z(\mathbb{R}_+)]}{W(z)} = \int_{[0,\infty)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)}\mu(\mathrm{d}x).$$

This provides us with a limiting spectrum intensity, written simply Λ:

$$\frac{\Lambda(\mathrm{d}q)}{\mathrm{d}q} := \lim_{z\to\infty} \frac{1}{W(z)}\left(\frac{\Lambda_z(\mathrm{d}q)}{\mathrm{d}q}\right) = \int_{[0,\infty)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)^2}\mathrm{e}^{-q/W^\mu(x)}\mu(\mathrm{d}x). \tag{2.4}$$

Note that in the Brownian case $\nu = \mathrm{d}x/x^2$, we get a simple expression $\Lambda(\mathrm{d}q) = (\theta/q)\mathrm{e}^{-\theta q}\mathrm{d}q$.

Proof of Proposition 2.23. We aim at computing $\psi(z,f)$, for f a measurable non-negative function of a simple real tree $\mathbb{T}$ with point mutations equipped with a measure ℓ on its leaves. Suppose the mutations $(M_n)_{n\geq 1}$ on the tree $\mathbb{T}$ are numbered by increasing distances from the root. Here we use the fact that a CPP can be seen as the genealogy of a birth-death process (see Section 2.5 for the development of this argument), a Markovian branching process whose time parameter is the distance from the root. This description implies that, for all $n\geq 1$, conditional on the height H_n of mutation M_n, the subtree growing from M_n has the law of $\mathbb{T}^{H_n}$. Set

$$\widetilde{f}(x) := \mathbb{E}[f(\mathbb{T}^x)].$$

Denoting H_n^z the height of the n-th mutation M_n^z of $\mathbb{T}^z$, we get

$$\begin{aligned}
\psi(z,f) &= \mathbb{E}\left[\sum_n f(\{\text{subtree of } \mathbb{T}^z \text{ growing from } M_n^z\})\right] \\
&= \sum_n \mathbb{E}\left[f(\{\text{subtree of } \mathbb{T}^z \text{ growing from } M_n^z\})\right] \\
&= \sum_n \mathbb{E}\left[\widetilde{f}(H_n^z)\right] \\
&= \mathbb{E}\left[\sum_n \widetilde{f}(H_n^z)\right].
\end{aligned}$$

Now this expression is simple to compute knowing $\widetilde{f}$ and the intensity of the point process giving mutation heights. Indeed, by elementary properties of Poisson point processes

$$\begin{aligned}
\mathbb{E}\left[\sum_n \widetilde{f}(H_n^z)\right] &= \mathbb{E}\left[\widetilde{f}(z) + \sum_{y\in M_{(0,z)}} \widetilde{f}(y) + \sum_{(t,x)\in\mathcal{N},\, t\le T(z)} \left(\sum_{y\in M_{(t,x)}} \widetilde{f}(y)\right)\right] \\
&= \widetilde{f}(z) + \int_{[0,z)} \widetilde{f}(x)\mu(\mathrm{d}x) + \mathbb{E}\left[T(z)\int_{[0,z)} \nu(\mathrm{d}y)\int_{[0,y)} \widetilde{f}(x)\mu(\mathrm{d}x)\right] \\
&= \widetilde{f}(z) + \int_{[0,z)} \widetilde{f}(x)\mu(\mathrm{d}x) + \frac{1}{\overline{\nu}(z)}\int_{[0,z)} \widetilde{f}(x)(\overline{\nu}(x)-\overline{\nu}(z))\mu(\mathrm{d}x) \\
&= \widetilde{f}(z) + W(z)\int_{[0,z)} \frac{\widetilde{f}(x)}{W(x)}\mu(\mathrm{d}x).
\end{aligned}$$

Now consider, for a fixed $q > 0$, the function f given by $f(\mathbb{T}) := \mathbb{1}_{\ell(R)>q}$, where $\mathbb{T}$ is a generic ultrametric tree with point mutations and measure ℓ supported by its leaves, and R denotes the set of its clonal leaves. This allows us to compute the expectation $\Lambda_z((q,\infty))$ of the number of mutations carried by a population of leaves of measure greater than q. Since the law of the measure of clonal leaves is known for a CPP, (see Corollary 2.20), we deduce

$$\begin{aligned}
\widetilde{f}(z) &= \mathbb{P}(C^z)\,\mathbb{P}(\ell(R^z) > q \mid C^z) \\
&= \mathbb{P}(C^z)\,\mathbb{P}(\lambda(\widetilde{R}\cap[0,T(z))) > q) \\
&= \frac{W(z)\mathrm{e}^{-\underline{\mu}(z)}}{W^\mu(z)}\,\mathrm{e}^{-q/W^\mu(z)},
\end{aligned}$$

where C^z again denotes the event of existence of clonal leaves in $\mathbb{T}^z$ and $\widetilde{R}$ is the set defined in Definition 2.13. Thus we have

$$\begin{aligned}
\Lambda_z((q,\infty)) &= \mathbb{E}[\Phi_z((q,\infty))] \\
&= W(z)\left(\frac{\mathrm{e}^{-\underline{\mu}(z)}}{W^\mu(z)}\mathrm{e}^{-q/W^\mu(z)} + \int_{[0,z)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)}\mathrm{e}^{-q/W^\mu(x)}\mu(\mathrm{d}x)\right).
\end{aligned}$$

Differentiating the last quantity yields the expression in the Proposition. □

Convergence Results for Small Families

Recall the construction of a CPP from a Poisson point process $\mathcal{N}$ in Section 2.2.2, and the point processes of mutations $(M_N)_{N\in\mathcal{N}}$. Since a CPP(ν,μ,z) is given by the points of $\mathcal{N}$ with first component smaller than $T(z)$, this construction yields a coupling of $(\mathbb{T}^z)_{z>0}$, where for each $z > 0$, $\mathbb{T}^z$ is a CPP(ν,μ,z). Recall the notation Φ_z from the previous subsection. Then, similarly to Theorem 3.1 in [64], we have the following almost sure convergence.

Proposition 2.26. *Under the preceding assumptions, and further assuming $\nu(\{\infty\}) = 0$, for any $q > 0$, we have the convergence:*

$$\lim_{z\to\infty} \frac{\Phi_z((q,\infty))}{T(z)} = \int_{[0,\infty)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)}\mathrm{e}^{-q/W^\mu(x)}\mu(\mathrm{d}x) = \Lambda((q,\infty)) \qquad a.s.$$

Remark 2.27. Recall that $\Phi_z((q,\infty))$ is the number of alleles carried by a population of leaves of measure larger than q in the tree $\mathbb{T}^z$, and $T(z)$ is the total size of the population of $\mathbb{T}^z$. The result is a strong law a large numbers: it shows that the number of small families (with a fixed size) grows linearly with the total measure of the tree at a constant speed given by the measure Λ defined by (2.4) as the limiting allele frequency spectrum intensity.

Proof. We will use the law of large numbers several times. Let us first introduce some notation. For $z > 0$, define $(T_i(z))_{i\geq 1}$ as the increasing sequence of first components of the atoms of $\mathcal{N}$ with second component larger than z, that is $T_1(z) = T(z)$ and for any $i \geq 1$

$$T_{i+1}(z) = \inf\{t > T_i(z),\ \exists x > z,\ (t,x) \in \mathcal{N}\}.$$

For $z < z'$, let $N(z,z') := \#\{(t,x) \in \mathcal{N} : t \leq T(z'), x > z\}$, that is the unique number n such that

$$T_n(z) = T(z').$$

Notice that the assumptions $\overline{\nu}(z) \in (0,\infty)$ for all $z > 0$ and $\nu(\{\infty\}) = 0$ imply that $T(z') \to \infty$ and $N(z,z') \to \infty$ as $z' \to \infty$, for a fixed z. Because the times $(T_{i+1}(z) - T_i(z))_{i\geq 1}$ are *i.i.d.* exponential random variables with mean $W(z)$ and since we have

$$T(z') = T(z) + \sum_{i=2}^{N(z,z')} (T_{i+1}(z) - T_i(z)),$$

it is clear by the strong law of large numbers that

$$\frac{T(z')}{N(z,z')} \underset{z'\to\infty}{\longrightarrow} W(z) \qquad a.s.$$

Also, write $\mathbb{T}_1^z, \ldots, \mathbb{T}_{N(z,z')}^z$ for the sequence of subtrees of height z within $\mathbb{T}^{z'}$ that are separated by the branches higher than z. That is, $\mathbb{T}_i^z$ is the ultrametric tree generated by the points of $\mathcal{N}$ with first component between $T_{i-1}(z)$ and $T_i(z)$. From basic properties of Poisson point processes, they are *i.i.d.* and their distribution is that of $\mathbb{T}^z$.

Now, write $h(\mathbb{T})$ for the height of an ultrametric tree (i.e., the distance between the root and any of its leaves), and take any non-negative, measurable function f of simple trees, such that

$$f(\mathbb{T}) = 0 \text{ if } h(\mathbb{T}) > z. \tag{$*$}$$

Recall the definition of $\varphi(\mathbb{T}, f)$. Since f satisfies ($*$), we can write

$$\varphi(\mathbb{T}^{z'}, f) = \sum_{i=1}^{N(z,z')} \varphi(\mathbb{T}_i^z, f). \tag{2.5}$$

Therefore, again by the strong law of large numbers, we have the following convergence

$$\frac{\varphi(\mathbb{T}^{z'}, f)}{N(z,z')} \underset{z'\to\infty}{\longrightarrow} \mathbb{E}[\varphi(\mathbb{T}^z, f)] = \psi(z,f) \qquad a.s. \tag{2.6}$$

Combining the two convergence results, it follows that

$$\frac{\varphi(\mathbb{T}^{z'}, f)}{T(z')} \underset{z' \to \infty}{\longrightarrow} \frac{\psi(z, f)}{W(z)} \qquad a.s.$$

Let us apply this to the function $f(\mathbb{T}) = \mathbb{1}_{\ell(R)>q}$. This function f does not satisfy $(*)$ for any $z > 0$, so we cannot apply (2.6) directly because (2.5) does not hold. However, we can artificially truncate f by defining the restriction f^z:

$$f^z(\mathbb{T}) := f(\mathbb{T})\mathbb{1}_{h(\mathbb{T})<z},$$

which does satisfy $(*)$. Now since $f^z \leq f$, we have the inequality between random variables

$$\varphi(\mathbb{T}^{z'}, f^z) \leq \varphi(\mathbb{T}^{z'}, f),$$

and by taking limits,

$$\frac{\psi(z, f)}{W(z)} \leq \liminf_{z' \to \infty} \frac{\varphi(\mathbb{T}^{z'}, f)}{T(z')} \qquad a.s.$$

But we have $\psi(z, f) = \Lambda_z((q, \infty))$ and as a consequence of Proposition 2.23, we have

$$\begin{aligned}\frac{\Lambda_z((q,\infty))}{W(z)} &= \frac{\mathrm{e}^{-\underline{\mu}(z)}}{W^\mu(z)}\mathrm{e}^{-q/W^\mu(z)} + \int_{[0,z)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)}\mathrm{e}^{-q/W^\mu(x)}\mu(\mathrm{d}x) \\ &\underset{z\to\infty}{\longrightarrow} \int_{[0,\infty)} \frac{\mathrm{e}^{-\underline{\mu}(x)}}{W^\mu(x)}\mathrm{e}^{-q/W^\mu(x)}\mu(\mathrm{d}x),\end{aligned}$$

which is $\Lambda((q, \infty))$ by definition. Therefore, we now have the inequality

$$\Lambda((q, \infty)) \leq \liminf_{z' \to \infty} \frac{\varphi(\mathbb{T}^{z'}, f)}{T(z')} \qquad a.s.$$

The converse inequality stems from a simple remark. There are at most $N(z, z')$ mutations of height greater than z giving rise to an allele carried by some leaves of $\mathbb{T}^{z'}$. This is simply because a population of n individuals can exhibit at most n different alleles. Therefore, we have

$$\varphi(\mathbb{T}^{z'}, f) \leq \varphi(\mathbb{T}^{z'}, f^z) + N(z, z'),$$

which gives by taking limits

$$\limsup_{z' \to \infty} \frac{\varphi(\mathbb{T}^{z'}, f)}{T(z')} \leq \frac{\psi(z, f) + 1}{W(z)} \underset{z\to\infty}{\longrightarrow} \Lambda((q, \infty)) \quad a.s.$$

We can finally conclude

$$\frac{\varphi(\mathbb{T}^{z}, f)}{T(z)} \underset{z\to\infty}{\longrightarrow} \Lambda((q, \infty)) \qquad a.s.,$$

which is the announced result. □

2.4 The Clonal Tree Process

In this section we consider the clonal subtree A^z of a random tree $\mathbb{T}^z$ with distribution $\mathrm{CPP}(\nu, \mu, z)$, where ν, μ are measures satisfying assumptions (H) and $z > 0$. We further assume $\nu([0, \infty)) = \infty$, that is we ignore the case when $\mathbb{T}^z$ is a finite tree almost surely. We will focus on the case when $\mu(dx) = \theta\, dx$.

2.4.1 Clonal Tree Process

There is a natural coupling in θ of the Poisson processes of mutations, in such a way that the sets of mutations are increasing in θ for the inclusion. Let $\mathbb{M}$ denote a Poisson point process with Lebesgue intensity on $\mathbb{R}_+^2$, and define for $\theta \geq 0$,

$$\mathbb{M}^\theta := \mathbb{M}([0,\theta] \times \cdot).$$

Then $\mathbb{M}^\theta$ is a Poisson point process on $\mathbb{R}_+$ with intensity $\theta\, dx$, and the sequence of supports of $\mathbb{M}^\theta$ increases with θ. Let us use this idea to couple mutations with different intensities on the random tree $\mathbb{T}^z$. Recall the construction of a CPP with a Poisson point process $\mathcal{N}$ in Section 2.2. For each point $N = (t, x)$ of $\mathcal{N} \cup \{(0, z)\}$, let M_N be a Poisson point process on $\mathbb{R}_+ \times [0, x]$ with Lebesgue intensity. For fixed $\theta \geq 0$, we get the original construction with $\mu(dx) = \theta\, dx$ when considering

$$M_N^\theta := M_N([0,\theta] \times \cdot).$$

Therefore a natural coupling of mutations of different intensities $(M^\theta)_{\theta \in \mathbb{R}_+}$ is defined on the random tree $\mathbb{T}^z$. Denote A_θ^z the clonal subtree of height z at mutation level θ, that is the subtree of $\mathbb{T}^z$ defined by

$$A_\theta^z := \{x \in \mathbb{T}^z, M^\theta([\![\varrho, x]\!]) = 0\}.$$

It is natural to seek to describe the decreasing process of clonal subtrees $(A_\theta^z)_{\theta \in \mathbb{R}_+}$. As θ increases, it is clearly a Markov process since the distribution of $A_{\theta+\theta'}^z$ given A_θ^z is the law of the clonal tree obtained after adding mutations at a rate θ' along the branches of A_θ^z. We will now study the Markovian evolution of the time-reversed process, as θ decreases. Its transitions are relatively simple to describe using grafts of trees.

2.4.2 Grafts of Real Trees

Given two real rooted trees $(\mathbb{T}_1, d_1, \varrho_1)$, $(\mathbb{T}_2, d_2, \varrho_2)$, and a graft point $g \in \mathbb{T}_1$, one can define the real rooted tree that is the graft of the root of $\mathbb{T}_2$ on $\mathbb{T}_1$ at point g by

$$\mathbb{T}_1 \oplus_g \mathbb{T}_2 := (\mathbb{T}_1 \sqcup \mathbb{T}_2 \setminus \{\varrho_2\}, d, \varrho_1),$$

with the new distance d defined as follows. For any $x, y \in \mathbb{T}_1 \sqcup \mathbb{T}_2$,

$$d(x, y) := d_i(x, y) \quad \text{if} \quad x, y \in \mathbb{T}_i \text{ for } i \in \{1, 2\},$$

and

$$d(x, y) := d_1(x, g) + d_2(\varrho_2, y) \quad \text{if} \quad x \in \mathbb{T}_1, y \in \mathbb{T}_2.$$

For real simple trees, this graft has a nice representation when the graft point is a leaf of the first tree.

Definition 2.28. For a simple tree $A = (\mathcal{T}, \alpha, \omega)$, define the **buds** of A as the set $\mathcal{B}(A)$ of leaves of $\mathcal{T}$ that live a finite time

$$\mathcal{B}(A) := \{b \in \mathcal{T},\ b0 \notin \mathcal{T},\ \omega(b) < \infty\}.$$

For two simple trees $A_i = (\mathcal{T}_i, \alpha_i, \omega_i)$ with $i \in \{1,2\}$, and for $b \in \mathcal{B}(A_1)$, we define the **graft** of A_2 on A_1 on the bud b, denoted $A_1 \oplus_b A_2$ by:

$$\mathcal{T} := \mathcal{T}_1 \cup b\mathcal{T}_2,$$

$$\alpha(b) := \alpha_1(b), \quad \omega(b) := \omega_1(b) + \zeta_2(\varnothing),$$

$$\forall u \in \mathcal{T}_1 \setminus \{b\}, \quad \alpha(u) := \alpha_1(u), \quad \omega(u) := \omega_1(u),$$

$$\forall u \in \mathcal{T}_2 \setminus \{\varnothing\}, \quad \begin{cases} \alpha(bu) := \omega(b) + (\alpha_2(u) - \omega_2(\varnothing)), \\ \omega(bu) := \alpha(bu) + \zeta_2(u), \end{cases}$$

$$A_1 \oplus_b A_2 := (\mathcal{T}, (\alpha(u), \zeta(u), \omega(u))_{u \in \mathcal{T}}).$$

It is then clear that $\mathcal{B}(A_1 \oplus_b A_2) := \mathcal{B}(A_1) \setminus \{b\} \cup b\mathcal{B}(A_2)$. See Figure 2.5 for an example.

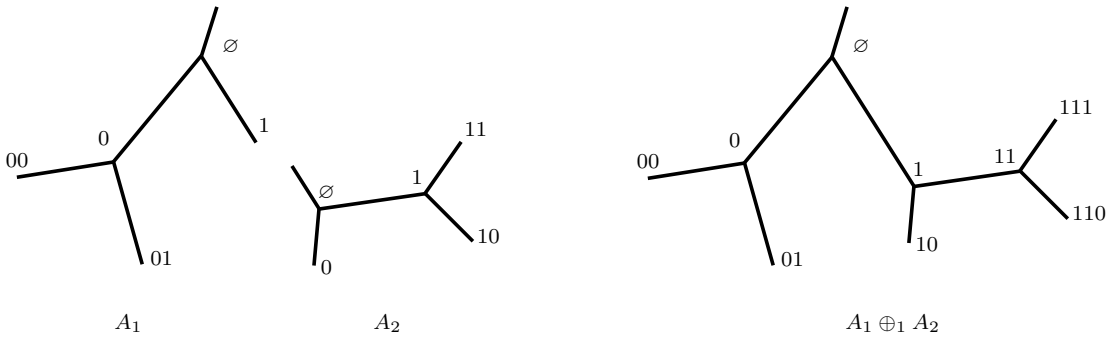


Figure 2.5 – Simple tree graft

2.4.3 Evolution of the Clonal Tree Process

We study the increasing clonal tree process as we remove mutations (decreasing θ). We therefore reverse time by denoting $\eta = -\ln \theta$, and defining $X_\eta^z := A_{\mathrm{e}^{-\eta}}^z$. Denote $\mathbb{Q}_\eta^z$ the distribution of X_η^z with values in the set of reversed (i.e., with time flowing from z to 0) simple binary trees. See Figure 2.6 for a sketch of the tree growth process. The increasing process $(X_\eta^z)_{\eta \in \mathbb{R}}$ is nicely described in terms of grafts.

Theorem 2.29.

(i) The process $(X_\eta^z)_{\eta \in \mathbb{R}}$ is a time-inhomogeneous Markov process, whose transitions conditional on X_η^z can be characterized as follows.

- *The buds of X_η^z are the leaves b of height $\omega(b)$. Independently of the others, each bud b is given an exponential clock T_b of parameter 1.*
- *At time $\eta' = \eta + T_b$, a tree is grafted on the bud b, following the distribution $\mathbb{Q}_{\eta'}^{\omega(b)}$, and each newly created bud b' is given an independent exponential clock $T_{b'}$ of parameter 1.*

(ii) The infinitesimal generator evaluated at a function φ of simple trees which depends only on a finite number of generations (i.e. such that the property $\exists\, n \geq 0,\ \varphi(\,\cdot\,) = \varphi(\,\cdot\,_{|n})$ holds) can be written as follows

$$\mathcal{L}_\eta \varphi(A) = \sum_{b \in \mathcal{B}(A)} \left(\mathbb{Q}_\eta^{\omega(b)}[\varphi(A \oplus_b Y)] - \varphi(A) \right),$$

where Y is the random tree drawn under the probability measure $\mathbb{Q}_\eta^{\omega(b)}$.

(iii) Write τ_z for the first time the clonal tree process reaches the boundary, that is the first time there is a leaf $x \in X_\eta^z$ with $d(\varrho, x) = z$, (where d is the distance in the real tree X_η^z):

$$\tau_z = \inf\{\eta \in \mathbb{R} : \exists x \in X_\eta^z,\, d(\varrho, x) = z\}.$$

Then the distribution of τ_z is given by

$$\mathbb{P}(\tau_z \leq \eta) = \frac{W(z)\,\mathrm{e}^{-\mathrm{e}^{-\eta}z}}{W_\eta(z)},$$

where as previously $W(z) = \overline{\nu}(z)^{-1}$, and

$$W_\eta(z) = W(0) + \int_{(0,z]} \mathrm{e}^{-\mathrm{e}^{-\eta}x}\mathrm{d}W(x),$$

that is $W_\eta = W^\mu$ with $\mu(dx) = \mathrm{e}^{-\eta}\,dx$.

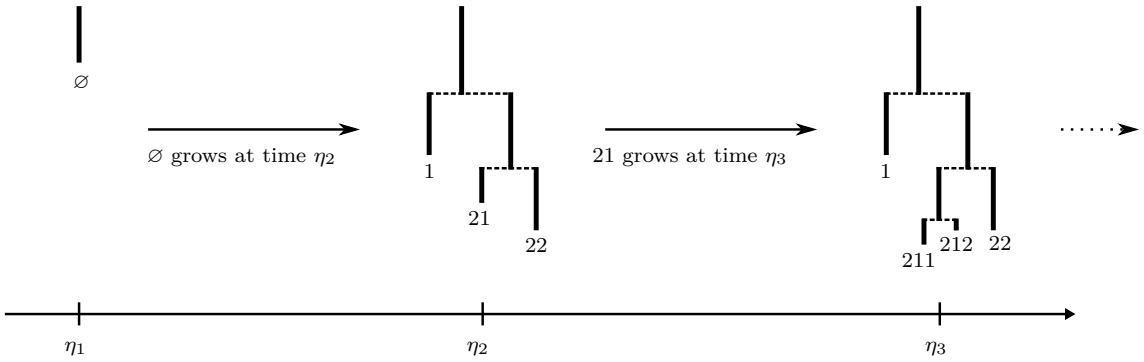


Figure 2.6 – Markovian evolution of an increasing tree process. In this example, the time $\eta_2 - \eta_1$ is an exponential time of parameter 1 and $\eta_3 - \eta_2$ is an exponential time of parameter 3.

We first state a result that is already interesting in itself, which ensures that CPP trees are reversed pure-birth trees (see next Section for details on birth-death trees and their links with CPPs). We refer the reader to Subsection 2.A.2, where a more general result is proved.

Lemma 2.30. *Let ν and μ be diffuse measures on $[0, \infty)$, satisfying assumptions (H) and $\nu([0, \infty)) = \infty$. Fix $z_0 \in [0, \infty)$ such that $\overline{\nu}(z_0) = 1$ and let $J = (0, z_0]$. Then for $z \in J$, a CPP(ν, z) is the genealogy of a reversed (i.e. with time flowing from z to 0) pure-birth process with birth intensity β defined as the Laplace-Stieltjes measure associated with the nondecreasing function $-\log \overline{\nu}$, started from z.*

Proof of Theorem 2.29. From Lemma 2.30, we can express the CPP in terms of a pure-birth tree, with time flowing from z to 0 (but measured from 0 to z) and birth intensity $\mathrm{d}\beta = \mathrm{d}(\log \circ W)$. Let $\mathcal{T} \subset \mathcal{U}$ denote the complete binary tree

$$\mathcal{T} := \bigcup_{n \geq 0} \{0, 1\}^n$$

Then we can define recursively $(\alpha(u), \omega(u))_{u\in\mathcal{T}}$ by setting $\alpha(\varnothing) = z$, and for $u = vi$, with $i \in \{0, 1\}$:

$$\alpha(u) = \omega(v) = \sup[0, \alpha(v)) \cap B_v,$$

with the convention $\sup \varnothing = 0$, and where $(B_v)_{v\in\mathcal{T}}$ are *i.i.d.* Poisson point processes on $[0, z]$ with intensity β. This defines the random reversed simple tree $(\mathcal{T}, \alpha, \omega)$ as the genealogy of a pure-birth process with birth intensity β, with time flowing from z to 0. In other words, by the definition of β, $(\mathcal{T}, \alpha, \omega)$ is the reversed simple tree with distribution $\mathrm{CPP}(\nu, z)$.

Now we define independently of $(\mathcal{T}, \alpha, \omega)$, a family $(M_u)_{u\in\mathcal{T}}$ of *i.i.d.* Poisson point processes on $[0, \infty) \times [0, z]$ with Lebesgue intensity. Writing for $\eta \in \mathbb{R}$ and $u \in \mathcal{T}$,

$$M_u^\eta = M_u([0, \mathrm{e}^{-\eta}] \times \cdot),$$

we define a coupling $((M_u^\eta)_{u\in\mathcal{T}})_{\eta\in\mathbb{R}}$ of point processes with intensity $\mathrm{e}^{-\eta}\mathrm{d}x$ on the branches of $(\mathcal{T}, \alpha, \omega)$.

Now let us define the process $(Y_\eta)_{\eta\in\mathbb{R}}$ by $Y_\eta = (\mathcal{T}_\eta, \alpha_\eta, \omega_\eta)$, with

$$\mathcal{T}_\eta := \{u \in \mathcal{T},\ \forall v \prec u,\ M_v^\eta([\alpha(v), \omega(v)]) = 0\},$$
$$\alpha_\eta(u) := \alpha(u) \quad \forall u \in \mathcal{T}_\eta,$$
$$\text{and } \ \omega_\eta(u) := \sup\left(\{\omega(u)\} \cup \{s < \alpha(u),\ M_u^\eta([s, \alpha(u)]) = 0\}\right) \quad \forall u \in \mathcal{T}_\eta,$$

By definition, one can check that Y_η is the clonal simple tree associated with the tree $(\mathcal{T}, \alpha, \omega)$ and the point process of mutations $(M_u^\eta)_{u\in\mathcal{T}}$. Therefore $(Y_\eta)_{\eta\in\mathbb{R}}$ has the same distribution as $(X_\eta^z)_{\eta\in\mathbb{R}}$. We define the filtration $(\mathcal{F}_\eta)_{\eta\in\mathbb{R}}$ as the natural filtration of the process $(Y_\eta)_{\eta\in\mathbb{R}}$, which we may rewrite:

$$\mathcal{F}_\eta := \sigma\left((\alpha_{\eta'})_{\eta'\le\eta}, (\omega_{\eta'})_{\eta'\le\eta}\right).$$

From our definitions, for $u \in \mathcal{T}$, we have:

$$\omega_\eta(u) = \inf\{s \in [0, \alpha(u)],\ M_u([0, \mathrm{e}^{-\eta}] \times [s, \alpha(u)]) = 0 \text{ and } B_u([s, \alpha(u)]) = 0\},$$

and since M_u and B_u are independent Poisson point processes, it is known that conditional on $\mathcal{F}_\eta$, we have: $M_u \cap [0, \infty) \times [0, \omega_\eta(u))$ and $B_u \cap [0, \omega_\eta(u))$ are independent Poisson point processes, with intensity Lebesgue for M_u and β for B_u, on their respective domains.

We can further notice that on the event $\{u \text{ is a bud of } Y_\eta\}$, conditional on $\mathcal{F}_\eta$, the families of point processes

$$(M_{uv} \cap [0, \infty) \times [0, \omega_\eta(u)))_{v\in\mathcal{T}} \text{ and } (B_{uv} \cap [0, \omega_\eta(u)))_{v\in\mathcal{T}}$$

are independent families of independent Poisson point process with intensity Lebesgue for M_{uv} and β for B_{uv}, on their respective domains.

Also, since M_u and B_u are independent and with diffuse intensities, we have the a.s. equalities between events

$$\begin{aligned}
&\{u \text{ is a bud of } Y_\eta\} \\
&\quad = \{\omega_\eta(u) = \inf\{s \in [0, \alpha(u)],\ M_u([0, \mathrm{e}^{-\eta}] \times [s, \alpha(u)]) = 0\}\} \\
&\quad = \{B_u(\{\omega_\eta(u)\}) = 0\}.
\end{aligned}$$

Moreover, since M_u is a Poisson point process with Lebesgue intensity on $[0,\infty)^2$, it is known that on this event, conditional on $\omega_\eta(u)$, the point process M_u restricted to $[0,\mathrm{e}^{-\eta}]\times[0,\omega_\eta(u)]$ has the conditional distribution of:

$$\delta_{(U,\,\omega_\eta(u))}+\widehat{M},$$

where U is a uniform random variable on $[0,\mathrm{e}^{-\eta}]$ and $\widehat{M}$ is an independent Poisson point process on $[0,\mathrm{e}^{-\eta}]\times[0,\omega_\eta(u))$ with Lebesgue intensity. Hence on the event $A:=\{u \text{ is a bud of } Y_\eta\}$, the distribution of

$$\begin{aligned}\widehat{\eta} &= \inf\{\eta'\geq\eta,\ M_u([0,\mathrm{e}^{-\eta'}]\times\{\omega_\eta(u)\})=0\}\\ &= \sup\{\eta'\geq\eta,\ M_u([0,\mathrm{e}^{-\eta'}]\times\{\omega_\eta(u)\})=1\}\end{aligned}$$

is given by

$$\begin{aligned}\mathbb{P}(\widehat{\eta}-\eta\geq t\mid A) &= \mathbb{P}(M_u([0,\mathrm{e}^{-(\eta+t)}]\times\{\omega_\eta(u)\})=1\mid A)\\ &= \mathbb{P}(U\in[0,\mathrm{e}^{-(\eta+t)}])\\ &= \mathrm{e}^{-t},\end{aligned}$$

And so if u is a bud of Y_η, the first time $\widehat{\eta}$ such that $\omega_{\widehat{\eta}}(u)$ is lower than $\omega_\eta(u)$ satisfies that $\widehat{\eta}-\eta$ has an exponential distribution with parameter 1.

We may now prove the first point *(i)* of the theorem. Fix $\eta\in\mathbb{R}$, and write $(b_1,b_2,\dots)$ for the distinct buds of Y_η. We define, for $i\geq 1$ and $\eta'\geq\eta$:

$$\begin{gathered}\widetilde{\mathcal{T}}^i_{\eta'}:=\{u,\ b_iu\in\mathcal{T}_{\eta'}\},\\ \widetilde{\alpha}^i_{\eta'}(\varnothing):=\omega_\eta(b_i)\text{ and for }u\in\mathcal{T}\setminus\{\varnothing\},\ \widetilde{\alpha}^i_{\eta'}(u):=\alpha_{\eta'}(b_iu),\\ \widetilde{\omega}^i_{\eta'}(u):=\omega_{\eta'}(b_iu),\\ \widetilde{Y}^i_{\eta'}:=\left(\widetilde{\mathcal{T}}^i_{\eta'},\widetilde{\alpha}^i_{\eta'},\widetilde{\omega}^i_{\eta'}\right).\end{gathered}$$

This definition formulates that for $\eta'\geq\eta$, $\widetilde{Y}^i_{\eta'}$ is the unique simple tree such that $Y_{\eta'}=A\oplus_{b_i}\widetilde{Y}^i_{\eta'}$ for another simple tree A in which b_i is a bud, with $\omega^A(b_i)=\omega_\eta(b_i)$. Note that when writing $Y_{\eta'}=A\oplus_{b_i}\widetilde{Y}^i_{\eta'}$, A may be different from Y_η, even for η' arbitrarily close to η, since other grafts may have occurred (possibly infinitely many grafts if Y_η has infinitely many buds).

Since $b_1,b_2,\dots$ are the buds of Y_η, the sets $b_1\mathcal{T},b_2\mathcal{T},\dots$ are disjoint. Thus, from our construction, the following families of random variables are independent conditional on $\mathcal{F}_\eta$:

$$(B_{b_1u})_{u\in\mathcal{T}},(B_{b_2u})_{u\in\mathcal{T}}\dots,(M_{b_1u})_{u\in\mathcal{T}},(M_{b_2u})_{u\in\mathcal{T}},\dots$$

Furthermore, we know how to describe their distributions conditional on $\mathcal{F}_\eta$ because of the previous observations. It follows that the trees $(\widetilde{Y}^i_{\eta'})_{i\geq1}$ are independent conditional on $\mathcal{F}_\eta$ and the distribution of $(\widetilde{Y}^i_{\eta'})_{\eta'\geq\eta}$ can be described by:

There is a random variable $\widehat{\eta}$ such that

- $\widehat{\eta}-\eta$ is exponentially distributed with parameter 1.

- For $\eta \leq \eta' < \widehat{\eta}$, we have $\omega_{\eta'}(b_i) = \omega_\eta(b_i)$ so $\widetilde{Y}^i_{\eta'}$ is the empty tree (or rather contains only one point, the root).
- Conditionally on $\widehat{\eta}$, the process $(\widetilde{Y}^i_{\eta'})_{\eta' \geq \widehat{\eta}}$ is distributed as our construction of the process $(Y_{\eta'})_{\eta' \geq \widehat{\eta}}$, with the initial condition $\alpha(\varnothing) = \omega_\eta(b_i)$.

This concludes the proof of *(i)*.

For *(ii)*, write $\mathfrak{T}$ for the set of simple binary trees and suppose we have a bounded measurable map $\varphi : \mathfrak{T} \to \mathbb{R}$ and a number $n \geq 0$ such that

$$\varphi(A) = \varphi(A_{|n}) \qquad A \in \mathfrak{T}.$$

Consider a fixed tree $A = (\mathcal{T}, \alpha, \omega) \in \mathfrak{T}$. There is a finite number of buds $b_1, \ldots, b_m$ in the first n generations $\mathcal{T}_{|n}$, therefore for a fixed $\eta \in \mathbb{R}$, conditional on $\{X^z_\eta = A\}$, the process $(\varphi(X^z_{\eta'}))_{\eta' \geq \eta}$ is a continuous time Markov chain. It follows from *(i)* that this Markov chain jumps after an exponential time with parameter m to a new state where one of the buds, uniformly chosen, grows into a new tree. That is, denoting $\mathcal{L}_\eta$ the infinitesimal generator of the process $(X^z_\eta)_{\eta \geq \eta_0}$,

$$\mathcal{L}_\eta \varphi(A) = \sum_{i=1}^{m} \left(\mathbb{Q}_\eta^{\omega(b_i)}[\varphi(A \oplus_{b_i} Y)] - \varphi(A) \right),$$

where Y is the random tree drawn under the probability measure $\mathbb{Q}_\eta^{\omega(b_i)}$.

For *(iii)*, note that the existence of a leaf in the clonal subtree at a distance z from the root coincides a.s. with the existence of a clonal leaf in $\mathbb{T}^z$, where $\mathbb{T}^z$ is the original CPP(ν, z) with mutation measure $\mu(\mathrm{d}x) = \mathrm{e}^{-\eta}\,\mathrm{d}x$. Then the formula in the proof follows from Proposition 2.21, which gives the probability that there is a clonal leaf in a CPP. □

The branching random walk of the buds. Forgetting the structure of the tree and considering only the height of the buds, the process becomes a rather simple branching random walk. Write $\chi^z_\eta := \sum_{b \in \mathcal{B}(X^z_\eta)} \delta_{\omega(b)}$ for the point measure on $\mathbb{R}_+$ giving the heights of the buds in X^z_η. Then $(\chi^z_\eta)_{\eta \geq \eta_0}$ is a branching Markov process where each particle stays at their height z' during their lifetime (an exponential time of parameter 1), then splits at their death time η according to the distribution of $\chi^{z'}_\eta$. Similarly to the preceding paragraph, one can describe the infinitesimal generator of this process as follows. For a map $f : \mathbb{R}_+ \to \mathbb{R}_+$ that is zero in a neighborhood of 0 and a Radon point measure Γ on $(0, \infty)$ (i.e. such that $\Gamma([x, \infty)) < \infty \; \forall x > 0$), write $\varphi^f(\Gamma)$ for the sum

$$\varphi^f(\Gamma) := \int f(z)\Gamma(\mathrm{d}z).$$

Then the infinitesimal generator $\mathcal{L}_\eta$ at time η of the time-inhomogeneous process $(\chi_\eta)_{\eta \in \mathbb{R}}$, evaluated at φ^f, is given by

$$\mathcal{L}_\eta \varphi^f(\Gamma) = \int \mathbb{Q}^z_\eta[\varphi^f(\chi)]\,\Gamma(\mathrm{d}z) - \varphi^f(\Gamma).$$

2.5 Link between CPP and Birth-Death Trees

2.5.1 Birth-Death Processes

An additional well-known example of random tree is given by the genealogy of a birth-death process, which will appear as an alternative description of our CPP trees. Here, a birth-death process is a time-inhomogeneous, time-continuous Markovian branching process living in $\mathbb{Z}_+$ with jumps in $\{-1, 1\}$. In a general context, we will define the genealogy of a birth-death process as a random simple tree, which we may equip with a canonical limiting measure on the set of its infinite lineages.

Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$. Suppose there are two measures on J, β and κ, respectively called the birth intensity measure and death intensity measure, or simply birth rate and death rate, which satisfy for all $t \in J$

$$\begin{aligned} \beta([t_0, t]) < \infty, \qquad & \kappa([t_0, t]) < \infty \\ \beta(\{t\}) = 0, \qquad & \kappa(\{t\}) = 0. \end{aligned} \tag{2.7}$$

In other words, β and κ are diffuse Radon measures on J.

Informally, the population starts with one individual at time t_0, and each individual alive at time $t \geq t_0$ may give birth to a new individual at rate $\beta(\mathrm{d}t)$, and die at rate $\kappa(\mathrm{d}t)$.

Definition 2.31. Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$, and β and κ measures on J satisfying (2.7). Independently for each $u \in \bigcup_n \{0, 1\}^n$, we define B_u and D_u two independent point processes, such that B_u (resp. D_u) is a Poisson point process on J with intensity β (resp. κ).

The **genealogy of a** (β, κ) **birth-death process** started from $t \in J$ is the random binary simple tree $(\mathcal{T}, \alpha, \omega)$ defined recursively by:

1. $\varnothing \in \mathcal{T}$, with $\alpha(\varnothing) = t$.
2. For each $u \in \mathcal{T}$, we set $T_B(u) := \inf B_u \cap (\alpha(u), t_\infty)$, and $T_D(u) := \inf D_u \cap (\alpha(u), t_\infty)$. Then there are three different possibilities:
   - if $T_B(u) < T_D(u)$, then we set $u0, u1 \in \mathcal{T}$, and $\alpha(u0) = \alpha(u1) = \omega(u) := T_B(u)$,
   - if $T_D(u) < T_B(u)$, then we set $\omega(u) = T_D(u)$, and $u0, u1 \notin \mathcal{T}$,
   - if $T_B(u) = T_D(u) = t_\infty$, then we set $\omega(u) = t_\infty$, and $u0, u1 \notin \mathcal{T}$.

Birth-death processes have been known for a long time. They have been studied thoroughly as early as 1948 [56]. In the case of pure-birth processes with infinite descendance, we introduce a canonical measure on the boundary of the tree.

Definition 2.32. Under the assumption $\kappa = 0$ and $\beta(J) = \infty$, the tree $(\mathcal{T}, \alpha, \omega)$ is said to be the genealogy of a **pure-birth process**. It may then be equipped with a **measure** $\mathscr{L}$ **on its boundary** $\partial\mathcal{T} = \{0, 1\}^{\mathbb{N}}$ defined by

$$\mathscr{L}(B_u) := \lim_{s \uparrow t_\infty} \frac{N_u(s)}{\mathrm{e}^{\beta([t_0, s])}} \qquad u \in \mathcal{T},$$

where $B_u = \{v \in \partial\mathcal{T}, u \prec v\}$ is defined as in Definition 2.4, and $N_u(s)$ is the number of descendants of u at time s:

$$N_u(s) := \#\{v \in \mathcal{T},\ u \preceq v,\ \alpha(v) < s \leq \omega(v)\}.$$

Remark 2.33. The limits in the definition are well-defined because for each $u \in \mathcal{T}$, conditional on $\alpha(u)$, the process $\left(\frac{N_u(s)}{e^{\beta([t_0,s])}}\right)_{s\geq\alpha(u)}$ is a non-negative martingale. Also, the fact that the map $u \mapsto N_u(s)$ is additive combined with Remark 2.2 justifies that the measure $\mathscr{L}$ is well defined.

Finally, let us introduce random mutations on a birth-death tree as a random discrete set of points.

Definition 2.34. Let μ be a diffuse Radon measure on J, and let $\#$ denote the counting measure on $\bigcup_n \{0,1\}^n$. A birth-death tree $(\mathcal{T}, \alpha, \omega)$ may be equipped with a set M of **neutral mutations** at rate μ by defining, independently of the preceding construction, a Poisson point process $\widetilde{M}$ on $(\bigcup_n \{0,1\}^n) \times J$ with intensity $\# \otimes \mu$, and then defining:

$$M := \{(u,s) \in \widetilde{M},\ \ u \in \mathcal{T}, \alpha(u) < s \leq \omega(u)\}.$$

This point process M is then a discrete subset of the skeleton of the real tree (defined as in (2.1)) associated with $(\mathcal{T}, \alpha, \omega)$.

Example. The Yule tree is the genealogy of a pure-birth process with $J = [0, \infty)$ and a birth rate β equal to the Lebesgue measure, which means that the branches separating two branching points are *i.i.d* exponential random variables with parameter 1. Every pure-birth tree with $\beta(J) = \infty$ can be time-changed into a Yule tree, with the time-change $\varphi : J \to [0, \infty)$, $t \mapsto \beta([t_0, t])$ (see Proposition 2.39).

2.5.2 Link between CPP and Supercritical Birth-Death Trees

We first provide a refined version of Lemma 2.30 which is proved in Subsection 2.A.2.

Lemma 2.35. *Under the assumptions of Lemma 2.30, the CPP(ν, z) with* ***boundary measured by*** ℓ *is the genealogy of a reversed pure-birth process with birth intensity* $\mathrm{d}\beta = -\mathrm{d}\log\overline{\nu}$ *started from* z*,* ***with boundary measured by*** $\mathscr{L}$*.*

Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$, and let β and κ be diffuse Radon measures on J, i.e. measures satisfying (2.7). Consider a birth-death process started from t_0 with birth rate β and death rate κ. Let us define

$$\mathcal{I}_t := \int_{[t,t_\infty)} e^{-\beta([t,s])+\kappa([t,s])}\, \beta(\mathrm{d}s)$$

$$\beta^*(\mathrm{d}t) := \frac{\beta(\mathrm{d}t)}{\mathcal{I}_t}$$

In a birth-death process with $\beta(J) = \infty$, we say that an individual i alive at time t has an *infinite progeny* if $N_i(s) > 0$ for any time $s > t$. It is known (see [56]) that the process is supercritical (i.e., the event $\{\liminf_{t\to t_\infty} N_\varnothing(t) > 0\}$ has positive probability) if

and only if $\mathcal{I}_{t_0} < \infty$, and that the probability of non-extinction for a process started at time $t \in J$ is then $\mathcal{I}_t^{-1}$. Also, if the birth-death process with rates (β, κ) is supercritical, then conditional on non-extinction, the subtree of individuals with infinite progeny is a pure-birth tree with birth rate β^*.

Now we assume Poissonian neutral mutations are set on the genealogy of a (β, κ) supercritical birth-death process, according to a rate μ, where μ is a diffuse Radon measure on J. We also assume $\beta^*(J) = \infty$ so that $\lim_{t \to t_\infty} N_\varnothing(t) = +\infty$ conditional on non-extinction. Conditional on non-extinction, the subtree of individuals with infinite progeny is a measured simple tree equipped with mutations $(\mathcal{T}, \alpha, \omega, \mathscr{L}, M)$, where:

- $(\mathcal{T}, \alpha, \omega, \mathscr{L})$ is a random simple binary tree constructed (see Definition 2.32) from a pure-birth process with birth rate β^*.
- With $\widehat{M}$ a Poisson point process on $(\bigcup_{n \geq 0} \{0,1\}^n) \times J$ with intensity $\# \otimes \mu$, the mutations on the branches of $\mathcal{T}$ are defined as the set

$$M = \{(i,t) \in \widehat{M},\, i \in \mathcal{T}, \alpha(i) < t \leq \omega(i)\}$$

One may study this measured tree with mutations as the limit in time of the genealogy of the birth-death process with neutral mutations. We show that this measured tree with mutations is in fact a time-changed CPP tree.

Theorem 2.36. *Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$, and let β and μ be diffuse Radon measures on J, with $\beta(J) = \infty$. Let $\mathbb{T} = (\mathcal{T}, \alpha, \omega, M, \mathscr{L})$ be a random measured simple tree representing the genealogy of a pure-birth process with rate β started from t_0, equipped with mutations at rate μ. Let $\varphi : J \to (0,1]$ be the time-change defined by*

$$\varphi : t \mapsto \mathrm{e}^{-\beta([t_0, t))}.$$

Then the time-changed tree $\varphi(\mathbb{T})$ (see Proposition 2.39) has the distribution of a $CPP\Big(\frac{\mathrm{d}x}{x^2}, \mu \circ \varphi^{-1}, 1\Big)$.

Proof. Thanks to Lemma 2.35, we only need to exhibit a correct time change to prove the Theorem. We know that a time-changed birth-death tree is still a birth-death tree: this is explicitly stated in Proposition 2.39 in the appendix. This implies here that the time-changed tree $\varphi(\mathbb{T})$ is a (reversed) pure-birth process with birth rate $\beta \circ \varphi^{-1}$, started from $\varphi(t_0) = 1$, and equipped with mutations with rate $\mu \circ \varphi^{-1}$. Let us first check that $\beta \circ \varphi^{-1}(\mathrm{d}x) = \mathrm{d}\log(x)$. Since β is diffuse, φ is continuous decreasing, so for all $x \in (0, z_0]$, we have $\varphi(\varphi^{-1}(x)) = x$, where φ^{-1} is the right-continuous inverse of φ. Therefore we have, for all $a < b \in (0,1]$:

$$\begin{aligned}
\beta \circ \varphi^{-1}([a,b]) &= \beta([\varphi^{-1}(b), \varphi^{-1}(a)]) \\
&= \log \varphi(\varphi^{-1}(b)) - \log \varphi(\varphi^{-1}(a)) \\
&= \log(b) - \log(a).
\end{aligned}$$

Now notice that for $x \in (0,1]$,

$$-\log\left(\int_x^\infty \frac{1}{y^2} \mathrm{d}y\right) = \log x,$$

so according to Lemma 2.35, a $\mathrm{CPP}\left(\frac{\mathrm{d}x}{x^2}, \mu \circ \varphi^{-1}, 1\right)$ is a pure-birth process with birth rate $\beta(\mathrm{d}x) = \mathrm{d}\log(x)$, started from 1 and equipped with mutations at rate $\mu \circ \varphi^{-1}$. Therefore its distribution is identical to the distribution of $\varphi(\mathbb{T})$. □

2.A Appendix

2.A.1 Birth-Death Processes

Proposition 2.37. *Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$, and β and κ diffuse Radon measures on J (i.e. satisfying (2.7)). Let $\mathbb{P}_t$ denote the distribution of the genealogy of a (β, κ) birth-death process started with one individual at time $t \in J$, and let N_T be the number of individuals alive at time $T \in J$. For $T > t$ and $\alpha \geq 0$, we have:*

$$\mathbb{E}_t(\mathrm{e}^{-\alpha N_T}) = 1 - \frac{(1 - \mathrm{e}^{-\alpha})}{\mathrm{e}^{\kappa([t,T]) - \beta([t,T])} + (1 - \mathrm{e}^{-\alpha}) \int_{[t,T]} \mathrm{e}^{\kappa([t,s]) - \beta([t,s])} \beta(\mathrm{d}s)},$$

and in particular,

$$\mathbb{P}_t(N_T > 0) = \left(\mathrm{e}^{\kappa([t,T]) - \beta([t,T])} + \int_{[t,T]} \mathrm{e}^{\kappa([t,s]) - \beta([t,s])} \beta(\mathrm{d}s) \right)^{-1}.$$

Remark 2.38. Note that the previous proposition shows that conditional on being non-zero, N_T is a geometric random variable, which is a known fact about birth-death processes (see for instance [56]). We still provide a proof in our case where the birth and death intensity measures are not necessarily absolutely continuous with respect to Lebesgue.

Proof. With a fixed time horizon $T \in J$ and a fixed real number $\alpha \geq 0$, write for $t < T$,

$$q(t) = \mathbb{E}_t(\mathrm{e}^{-\alpha N_T}).$$

We use a different description of the birth-death process than the one used in Section 2.5, and consider a population where individuals die at rate κ, and during their lifetime, produce a new individual at rate β. Notice that for any $s > t$, the number of individuals alive at time s has the same distribution in both models.

Thus we write D for the death time of the first individual, and B_i for the possible birth time of her i-th child. With our description, D has the distribution of the first atom of a Poisson point process on $[t, t_\infty)$ with intensity κ and conditional on D, the set $\{B_1, B_2, \ldots, B_N\}$ is a Poisson point process on $[t, D]$ with intensity β. Also, write $\widetilde{N}_T^i$ for the number of alive descendants of the i-th child at time T. Since we have $N_T = \mathbb{1}_{D>T} + \sum_i \widetilde{N}_T^i$, we have

$$q(t) = \mathbb{E}_t\left[\mathrm{e}^{-\alpha \mathbb{1}_{D>T}} \prod_i \mathrm{e}^{-\alpha \widetilde{N}_T^i}\right],$$

where we define by convention $\widetilde{N}_T^i = 0$ if $B_i > T$. Now conditional on D and (B_i), $(\widetilde{N}_T^i)$ are independent, with $\widetilde{N}_T^i$ equal to the distribution of N_T under $\mathbb{P}_{B_i}$. Hence

$$q(t) = \mathbb{E}_t\left[\mathrm{e}^{-\alpha \mathbb{1}_{D>T}} \prod_i q(B_i)\right],$$

where we use the convention $q(u) := 1$ if $u > T$. Now conditional on D, (B_i) are the atoms of a Poisson point process with intensity $\beta(\mathrm{d}s)$ on $[t, D]$, so we have

$$\begin{aligned} q(t) &= \mathbb{E}_t \left[\mathrm{e}^{-\alpha \mathbb{1}_{D>T}} \exp\left(- \int_{[t,D]} (1 - q(s)) \beta(\mathrm{d}s) \right) \right] \\ &= \int_{[t,\infty)} \kappa(\mathrm{d}u)\, \mathrm{e}^{-\kappa([t,u))} \mathrm{e}^{-\alpha \mathbb{1}_{u>T}} \exp\left(- \int_{[t,u]} (1 - q(s)) \beta(\mathrm{d}s) \right), \end{aligned}$$

which implies by differentiation

$$\mathrm{d}q(t) = -\kappa(\mathrm{d}t) + q(t) \left[\kappa(\mathrm{d}t) + (1 - q(t)) \beta(\mathrm{d}t) \right],$$

which in turn may be rewritten

$$\mathrm{d}\left(\frac{1}{1 - q(t)} \right) = -\beta(\mathrm{d}t) + \left(\frac{1}{1 - q(t)} \right) (\beta(\mathrm{d}t) - \kappa(\mathrm{d}t)).$$

Remark that with $F(t) := \mathrm{e}^{\beta([t,T]) - \kappa([t,T])}$, we have $\mathrm{d}F(t) = F(t)(\kappa(\mathrm{d}t) - \beta(\mathrm{d}t))$, so that

$$\mathrm{d}\left(\frac{F(t)}{1 - q(t)} \right) = -F(t) \beta(\mathrm{d}t),$$

and since $q(T) = \mathrm{e}^{-\alpha}$, we have by integration on $[t, T]$:

$$\frac{1}{1 - \mathrm{e}^{-\alpha}} - \frac{F(t)}{1 - q(t)} = - \int_{[t,T]} F(s) \beta(\mathrm{d}s),$$

that is

$$1 - q(t) = \frac{(1 - \mathrm{e}^{-\alpha})}{\mathrm{e}^{\kappa([t,T]) - \beta([t,T])} + (1 - \mathrm{e}^{-\alpha}) \int_{[t,T]} \mathrm{e}^{\kappa([t,s]) - \beta([t,s])} \beta(\mathrm{d}s)}.$$

This characterizes the distribution of N_T under $\mathbb{P}_t$ for all T. In particular, letting $\alpha \to \infty$, we get

$$\mathbb{P}_t(N_T > 0) = \left(\mathrm{e}^{\kappa([t,T]) - \beta([t,T])} + \int_{[t,T]} \mathrm{e}^{\kappa([t,s]) - \beta([t,s])} \beta(\mathrm{d}s) \right)^{-1},$$

which concludes the proof. □

Proposition 2.39 (Time-changed birth-death processes)**.** *Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$, and β, κ, and μ diffuse Radon measures on J (i.e. satisfying (2.7)). Let $\varphi : J \to \mathbb{R}$ be an increasing function, and define $t'_0 := \varphi(t_0)$, $t'_\infty := \lim_{t \uparrow t_\infty} \varphi(t)$ and $J' = [t'_0, t'_\infty)$. We assume that φ satisfies*

$$\forall t < t_\infty,\ \varphi(t) < t'_\infty.$$

Let $\mathbb{T} = (\mathcal{T}, \alpha, \omega, M)$ be the genealogy of a (β, κ) birth-death process, started at $t \in J$ and equipped with Poissonian mutations with rate μ, as in Definition 2.34. We define the time-changed simple tree:

$$\varphi(\mathbb{T}) := (\mathcal{T}, \varphi \circ \alpha, \varphi \circ \omega, \{(u, \varphi(s)),\ (u, s) \in M\}).$$

If $\beta\circ\varphi^{-1}$ and $\kappa\circ\varphi^{-1}$ (the push-forwards of β and κ by φ) still have no atoms, then $\varphi(\mathbb{T})$ has the distribution of the genealogy of a $(\beta\circ\varphi^{-1},\kappa\circ\varphi^{-1})$ birth-death process, started at $\varphi(t)\in J'$ and equipped with Poissonian mutations with rate $\mu\circ\varphi^{-1}$.

Also, if $\kappa=0$ and $\beta(J)=\infty$, then $\kappa\circ\varphi^{-1}=0$ and $\beta\circ\varphi^{-1}(J')=\infty$, and the measures $\mathscr{L}_{\mathbb{T}}$ and $\mathscr{L}_{\varphi(\mathbb{T})}$ on $\partial\mathcal{T}$, defined for $\mathbb{T}$ and for $\varphi(\mathbb{T})$, are the same.

Proof. Suppose $\mathbb{T}$ is constructed as in Definition 2.31 with independent Poisson point processes B_u and D_u with respective intensities β and κ, for each $u\in\bigcup_n\{0,1\}^n$. This implies that the random sets defined by

$$\varphi(B_u):=\{\varphi(s),\ s\in B_u\},$$
$$\varphi(D_u):=\{\varphi(s),\ s\in D_u\},$$

are independent Poisson point processes on the interval J' with respective intensities $\beta\circ\varphi^{-1}$ and $\kappa\circ\varphi^{-1}$. Remark that by assumption, for $\eta\in\{\beta,\kappa\}$, for all $t'\in J'$, we have $\eta\circ\varphi^{-1}(\{t'\})=0$, so we a.s. have $t'\notin\varphi(B_u)$ and $t'\notin\varphi(D_u)$. Now since $\alpha(u)$ is independent of B_u and D_u, we have also a.s.

$$\varphi\circ\alpha(u)\notin\varphi(B_u)\cup\varphi(D_u). \tag{2.8}$$

By definition, we have $\varnothing\in\mathcal{T}$ and $\alpha(\varnothing)=t$, so $\varphi\circ\alpha(\varnothing)=\varphi(t)$. Then, if $u\in\mathcal{T}$, with $T_B(u)=\inf B_u\cap(\alpha(u),t_\infty)$, and $T_D(u)=\inf D_u\cap(\alpha(u),t_\infty)$, the following assertions hold.

- Since we have (2.8), we know that a.s. for all $s\in B_u\cap(\alpha(u),t_\infty)$, we have $\varphi(\alpha(u))<\varphi(s)$. This ensures that $\varphi(T_B(u))=\inf\varphi(B_u)\cap(\varphi\circ\alpha(u),t'_\infty)$.
- For the same reason, we have $\varphi(T_D(u))=\inf\varphi(D_u)\cap(\varphi\circ\alpha(u),t'_\infty)$.
- Because $\varphi(B_u)$ is independent of $\varphi(D_u)$ and because $\beta\circ\varphi^{-1}$ and $\kappa\circ\varphi^{-1}$ are diffuse by assumption, we have $\varphi(B_u)\cap\varphi(D_u)=\varnothing$ almost surely. Therefore, we have:
  - $\varphi(T_B(u))<\varphi(T_D(u))\iff T_B(u)<T_D(u)$, which implies $u0,u1\in\mathcal{T}$, and $\varphi\circ\alpha(u0)=\varphi\circ\alpha(u1)=\varphi\circ\omega(u)=\varphi(T_B(u))$,
  - $\varphi(T_D(u))<\varphi(T_B(u))\iff T_D(u)<T_B(u)$, which implies $\varphi\circ\omega(u)=\varphi(T_D(u))$, and $u0,u1\notin\mathcal{T}$,
  - $\varphi(T_B(u))=\varphi(T_D(u))=t'_\infty\iff T_B(u)=T_D(u)=t_\infty$, which implies $\varphi\circ\omega(u)=t'_\infty$, and $u0,u1\notin\mathcal{T}$.

Thus $(\mathcal{T},\varphi\circ\alpha,\varphi\circ\omega)$ is defined as a $(\beta\circ\varphi^{-1},\kappa\circ\varphi^{-1})$ birth-death process, started at $\varphi(t)$.

For the neutral mutations, we assume there is, as in Definition 2.34, a Poisson point process $\widetilde{M}$ on $(\bigcup_n\{0,1\}^n)\times J$ with intensity $\#\otimes\mu$, and such that:

$$M=\{(u,s)\in\widetilde{M},\ u\in\mathcal{T},\alpha(u)<s\leq\omega(u)\}.$$

Now $\{(u,\varphi(s)),\ (u,s)\in\widetilde{M}\}$ is a Poisson point process on $(\bigcup_n\{0,1\}^n)\times J'$ with intensity $\#\otimes\mu\circ\varphi^{-1}$, so

$$\{(u,\varphi(s)),\ (u,s)\in M\}=\{(u,\varphi(s)),\ (u,s)\in\widetilde{M},\ u\in\mathcal{T},\alpha(u)<s\leq\omega(u)\}$$

is the definition of random neutral mutations at rate $\mu \circ \varphi^{-1}$ on the tree $(\mathcal{T}, \varphi \circ \alpha, \varphi \circ \omega)$.

It remains to prove that in the case $\kappa = 0$ and $\beta(J) = \infty$, the measures $\mathscr{L}_{\mathbb{T}}$ and $\mathscr{L}_{\varphi(\mathbb{T})}$ are the same. By definition, we have for $u \in \bigcup_n \{0,1\}^n$,

$$\begin{aligned}\mathscr{L}_{\varphi(\mathbb{T})}(B_u) &= \lim_{s' \uparrow t'_\infty} \frac{N'_u(s')}{\mathrm{e}^{\beta \circ \varphi^{-1}([t'_0, s'])}} \\ &= \lim_{s \uparrow t_\infty} \frac{N'_u(\varphi(s))}{\mathrm{e}^{\beta \circ \varphi^{-1}([t'_0, \varphi(s)])}},\end{aligned}$$

where $N'_u(s') := \#\{v \in \mathcal{T},\ u \preceq v,\ \varphi \circ \alpha(v) < s \leq \varphi \circ \omega(v)\}$ is the number of descendants of u in the time-changed tree at time s'. But we have a.s. for all $s \in J$, $N'_u(\varphi(s)) = N_u(s)$, and also $\beta \circ \varphi^{-1}([t'_0, \varphi(s)]) = \beta([t_0, s])$, so finally

$$\begin{aligned}\mathscr{L}_{\varphi(\mathbb{T})}(B_u) &= \lim_{s \uparrow t_\infty} \frac{N'_u(\varphi(s))}{\mathrm{e}^{\beta \circ \varphi^{-1}([t'_0, \varphi(s)])}}, \\ &= \lim_{s \uparrow t_\infty} \frac{N_u(s)}{\mathrm{e}^{\beta([t_0, s])}} \\ &= \mathscr{L}_{\mathbb{T}}(B_u),\end{aligned}$$

which ends the proof. □

Proposition 2.40 (Characterization of pure-birth processes)**.** *Let $J = [t_0, t_\infty)$ be a real interval, with $-\infty < t_0 < t_\infty \leq \infty$, and β a diffuse Radon measure on J, such that $\beta(J) = \infty$.*

There is a unique family $(\mathbb{P}_t)_{t \in J}$ of distributions on simple trees $(\mathcal{T}, \alpha, \omega, \mathscr{L})$ equipped with a measure $\mathscr{L}$ on $\partial\mathcal{T} := \{0,1\}^{\mathbb{N}}$, such that for all $t \in J$

(i) $\mathcal{T} = \bigcup_n \{0,1\}^n$ *and* $\alpha(\varnothing) = t$ $\quad \mathbb{P}_t$*-almost surely.*

(ii) $\mathbb{P}_t(\omega(\varnothing) > s) = \mathrm{e}^{-\beta([t,s))}$.

(iii) Under $\mathbb{P}_t$, $\mathscr{L}(\partial\mathcal{T})$ is an exponential r.v. with mean $\mathrm{e}^{-\beta([t_0,t))}$.

(iv) Under $\mathbb{P}_t$, define for $i \in \{0,1\}$, $\alpha_i(u) := \alpha(iu)$, $\omega_i(u) := \omega(iu)$, $\mathscr{L}_i$ the measure on $\partial\mathcal{T}$ such that $\mathscr{L}_i(B_u) = \mathscr{L}(B_{iu})$ for all $u \in \mathcal{T}$ and finally $\mathbb{T}_i := (\mathcal{T}, \alpha_i, \omega_i, \mathscr{L}_i)$. Then the conditional distribution of the pair of trees $(\mathbb{T}_0, \mathbb{T}_1)$ given $\omega(\varnothing)$ is $\mathbb{P}^{\otimes 2}_{\omega(\varnothing)}$, i.e. they are independent with the same distribution $\mathbb{P}_{\omega(\varnothing)}$.

Furthermore, for all $t \in J$, $\mathbb{P}_t$ is the distribution of the genealogy of a pure-birth process with birth rate β started with one individual at time $t \in J$, equipped with $\mathscr{L}$ the measure on $\partial\mathcal{T}$ introduced in Definition 2.32.

Proof. Let $\mathbb{Q}_t$ be the law of the genealogy of a β pure-birth process started from t. We will first show that the family $(\mathbb{Q}_t)_{t \in J}$ satisfies the assertions *(i)-(iv)* of the theorem.

(i) By definition $\alpha(\varnothing) = t$. Also, the fact that for all $t \in T$, $\beta([t, t_\infty)) = \infty$, implies that for each Poisson point process with intensity β on J, there are infinitely many points in $[t, t_\infty)$. This implies that each individual in the process will eventually split into two, so that $\mathcal{T} = \bigcup_n \{0,1\}^n$ $\quad \mathbb{P}_t$-almost surely.

(ii) Under $\mathbb{Q}_t$, $\omega(\varnothing)$ is distributed as the first point of a Poisson point process $B_\varnothing$ on $[t, t_\infty)$ with intensity β. Therefore,

$$\mathbb{Q}_t(\omega(\varnothing) > s) = \mathbb{Q}_t(\#B_\varnothing \cap [t, s) = 0) = \mathrm{e}^{-\beta([t,s))}.$$

(iii) By Proposition 2.37, writing $\mathbb{E}_t$ for the expectation under $\mathbb{Q}_t$, we have for $t < T < t_\infty$,

$$\mathbb{E}_t(\mathrm{e}^{-\alpha N_T}) = 1 - \frac{(1 - \mathrm{e}^{-\alpha})}{\mathrm{e}^{-\beta([t,T])} + (1 - \mathrm{e}^{-\alpha})(1 - \mathrm{e}^{-\beta([t,T])})}.$$

Replacing α by $\alpha \mathrm{e}^{-\beta([t_0,T])}$ and letting $T \to t_\infty$, we have by dominated convergence:

$$\mathbb{E}_t(\mathrm{e}^{-\alpha \mathscr{L}(\partial \mathcal{T})}) = \frac{1}{\alpha \mathrm{e}^{-\beta([t_0,t))} + 1,}$$

which implies that $\mathscr{L}(\partial \mathcal{T})$ is an exponential random variable with mean $\mathrm{e}^{-\beta([t_0,t))}$.

(iv) Let us define a family $(B_u)_{u \in \mathcal{T}}$ of independent Poisson point processes on J with intensity β. Let us write F for the deterministic function such that for all $t \in J$, $F(t, (B_u)_{u \in \mathcal{T}})$ is the simple tree $\mathbb{T} = (\mathcal{T}, \alpha, \omega, \mathscr{L})$ constructed as in Definition 2.31, which follows the distribution $\mathbb{Q}_t$. By assumption, the two families $(B_{0u})_{u \in \mathcal{T}}$ and $(B_{1u})_{u \in \mathcal{T}}$ are independent, and by construction, we have

$$\mathbb{T}_0 = F(\omega(\varnothing), (B_{0u})_{u \in \mathcal{T}}) \text{ and } \mathbb{T}_1 = F(\omega(\varnothing), (B_{1u})_{u \in \mathcal{T}}),$$

where $\mathbb{T}_0$ and $\mathbb{T}_1$ are defined as in the statement of the Proposition. Therefore, under $\mathbb{Q}_t$, the conditional distribution of $(\mathbb{T}_0, \mathbb{T}_1)$ given $\omega(\varnothing)$ is $\mathbb{P}^{\otimes 2}_{\omega(\varnothing)}$.

Now, let us show that if a family $(\mathbb{P}_t)_{t \in J}$ satisfies the assertions *(i)-(iv)* of the Proposition, it satisfies also the following one. Let $\mathcal{T}_n$ be the complete binary tree with n generations

$$\mathcal{T}_n := \bigcup_{k=0}^{n} \{0, 1\}^k,$$

and let $\mathbb{P}^n_t$ be the distribution of $(\alpha(u), \omega(u), \mathscr{L}(B_u))_{u \in \mathcal{T}_n}$, where $(\mathcal{T}, \alpha, \omega, \mathscr{L})$ has distribution $\mathbb{P}_t$. Now we view $\mathbb{P}^n_t$ as a probability measure on the space $(\mathbb{R}^3)^{\mathcal{T}_n} = \{(x(u), y(u), z(u)), u \in \mathcal{T}_n\}$. Then we have

1. $x(\varnothing) := t \quad \mathbb{P}^n_t$-almost surely.
2. For all $m \leq n$ and $u \in \mathcal{T}_m$, conditional on $x(u)$ and independently of the variables $(x(v), y(v))_{v \in \mathcal{T}_m \setminus \{u\}}$, the distribution of $y(u)$ is given by:

$$\mathbb{P}^n_t(y(u) > s) = \mathrm{e}^{-\beta([x(u),s))} \qquad s \geq x(u).$$

3. For all $u \in \{0, 1\}^n$, conditional on $x(u)$ and independently of the rest, $z(u)$ is defined as an exponential random variable with mean $\mathrm{e}^{-\beta([t_0,x(u)))}$.
4. For all $u \in \mathcal{T}_{n-1}$, $x(u0) = x(u1) := y(u)$.
5. For all $u \in \mathcal{T}_{n-1}$, $z(u) := z(u0) + z(u1)$.

Indeed, assertion 1 is directly deduced from *(i)*, 5 is trivial because $\mathscr{L}$ is additive, and 2, 3 and 4 are proved by induction on n using *(iv)*. One can check that 2 stems from *(ii)* and *(iv)*, 3 from *(iii)* and *(iv)*, and 4 from *(i)* and *(iv)*.

Now it is clear that these five assumptions define $\mathbb{P}_t^n$ uniquely for $n \geq 0$ and $t \in J$. Also, a measured simple tree $(\mathcal{T}, \alpha, \omega, \mathscr{L})$ for which $\mathcal{T} = \bigcup_n \{0,1\}^n$ is entirely described by $(\alpha(u), \omega(u), \mathscr{L}(B_u))_{u \in \mathcal{T}} \in (\mathbb{R}^3)^{\mathcal{T}}$. This implies that $\mathbb{P}_t$ is uniquely determined by its marginal distribution $(\mathbb{P}_t^n)_{n \geq 0}$.

Finally, we have shown that the family $(\mathbb{Q}_t)_{t \in J}$, where $\mathbb{Q}_t$ is the law of the genealogy of a β pure-birth process started from t, satisfies assertions *(i)-(iv)*. In addition, we have shown that there is at most one family $(\mathbb{P}_t)$ of simple tree distributions satisfying assertions *(i)-(iv)*. Therefore, such a family exists and is unique, which concludes the proof.

□

2.A.2 Proof of Lemmas 2.30 and 2.35

Let us write $\mathbb{P}_z$ for the distribution of a CPP(ν, z). Let $\mathcal{N}$ be a Poisson point process with intensity $\mathrm{d}t \otimes \nu$ as in our construction of CPP trees. Recall that $T(z) = \inf\{t \geq 0, (x,t) \in \mathcal{N}, x \geq z\}$ and define

$$\mathcal{N}_z := \mathcal{N} \cap ([0, T(z)) \times [0, z]).$$

Define also $\mathbb{T}^z$ as the comb function tree given by $\mathcal{N}_z$ with distribution denoted $\mathbb{P}_z$. Write $\mathcal{P}_z$ for the distribution of the pair $(\mathcal{N}_z, T(z))$.

In Proposition 2.40, we characterized the distributions of pure-birth processes. As a result, to conclude the present proof, it is sufficient to show that the family $(\mathbb{P}_z)_{z \in J}$ satisfies the following conditions:

(i) We have $\mathcal{T} = \bigcup_n \{0,1\}^n$ and $\alpha(\varnothing) = z \quad \mathbb{P}_z$-almost surely.

(ii) We have $\mathbb{P}_z(\omega(\varnothing) < x) = \mathrm{e}^{-\beta((x,z])}$.

(iii) Under $\mathbb{P}_z$, $\mathscr{L}(\partial\mathcal{T})$ is an exponential r.v. with mean $\mathrm{e}^{-\beta((z,z_0])}$.

(iv) Under $\mathbb{P}_z$, define for $i \in \{0,1\}$, $\alpha_i(u) := \alpha(iu)$, $\omega_i(u) := \omega(iu)$, $\mathscr{L}_i$ the measure on $\partial\mathcal{T}$ such that $\mathscr{L}_i(B_u) = \mathscr{L}(B_{iu})$ for all $u \in \mathcal{T}$ and finally $\mathbb{T}_i := (\mathcal{T}, \alpha_i, \omega_i, \mathscr{L}_i)$. Then the conditional distribution of the pair of trees $(\mathbb{T}_0, \mathbb{T}_1)$ given $\omega(\varnothing)$ is $\mathbb{P}_{\omega(\varnothing)}^{\otimes 2}$, i.e. they are independent with the same distribution $\mathbb{P}_{\omega(\varnothing)}$.

Let us now prove each assertion.

(i) Since $\nu([0,\infty)) = \infty$ we have a.s. for any $0 \leq a < b \leq T(z)$:

$$\#(\mathcal{N}_z \cap [a,b] \times [0,\infty)) = \infty.$$

Also, since ν is diffuse, we have a.s. for all $x > 0$ that $\#(\mathcal{N} \cap [0,\infty) \times \{x\}) \leq 1$ Those two conditions imply that $\mathbb{T}^z$ is a complete binary tree.

(ii) – (iii) The first branching point of the tree $\mathbb{T}^z$ is $\omega(\varnothing) = \max\{x > 0,\ (t,x) \in \mathcal{N}_z\}$. Also the total mass of the tree is $\mathscr{L}(\partial\mathcal{T}) = T(z)$, which is an exponential random variable with mean $(\overline{\nu}(z))^{-1} = \mathrm{e}^{-\beta((z,z_0])}$. We can easily compute the distribution of $\omega(\varnothing)$ under

$\mathcal{P}_z$, since conditional on $T(z)$, $\mathcal{N}_z$ is a Poisson point process on $[0, T(z)) \times [0, z]$ with intensity $\mathrm{d}t \otimes \nu$. Therefore, for $x \in (0, z]$:

$$\begin{aligned}
\mathcal{P}_z(\omega(\varnothing) < x) &= \int_0^\infty \mathbb{P}(T(z) \in \mathrm{d}t) \mathrm{e}^{-t\nu([x,z])} \\
&= \int_0^\infty \overline{\nu}(z) \mathrm{e}^{-\overline{\nu}(z)t} \mathrm{e}^{-t(\overline{\nu}(x)-\overline{\nu}(z))} \mathrm{d}t \\
&= \int_0^\infty \overline{\nu}(z) \mathrm{e}^{-\overline{\nu}(x)t} \mathrm{d}t \\
&= \frac{\overline{\nu}(z)}{\overline{\nu}(x)} = \mathrm{e}^{-\beta((x,z])}.
\end{aligned}$$

(iv) It remains to prove the branching property for the family $(\mathbb{P}_z)_{z \in (0, z_0]}$.

Under $\mathcal{P}_z$, conditional on $\omega(\varnothing)$, let $(\mathcal{N}_1, T_1)$ and $(\mathcal{N}_2, T_2)$ be independent random variables of identical distribution $\mathcal{P}_{\omega(\varnothing)}$. We concatenate $\mathcal{N}_1$ and $\mathcal{N}_2$, adding a point of height $\omega(\varnothing)$ between the two sets:

$$\widetilde{\mathcal{N}} = \mathcal{N}_1 \cup \{(T_1, \omega(\varnothing))\} \cup \{(T_1 + t, x), (t, x) \in \mathcal{N}_2\}.$$

We claim that the following equality in distribution holds:

$$(\widetilde{\mathcal{N}}, T_1 + T_2) \overset{(d)}{=} (\mathcal{N}_z, T(z)), \tag{2.9}$$

which formulates the branching property for the family $(\mathbb{P}_z)_{z \in (0, z_0]}$.

From basic properties of Poisson point processes, we know that conditional on $T(z)$, the highest atom of $\mathcal{N}_z$ is (U, Z), with U having a uniform distribution on $[0, T(z)]$ and $Z := \omega(\varnothing)$ independent of U, such that

$$\mathcal{P}^z(Z \leq x \mid T(z)) = \mathrm{e}^{-T(z)(\overline{\nu}(x)-\overline{\nu}(z))}.$$

The joint distribution of $(Z, T(z))$ is therefore given by:

$$\begin{aligned}
\mathbb{E}[f(T(z)) \mathbb{1}_{Z \leq x}] &= \int_0^\infty \overline{\nu}(z) \mathrm{e}^{-\overline{\nu}(z)t} \mathrm{e}^{-t(\overline{\nu}(x)-\overline{\nu}(z))} f(t) \mathrm{d}t \\
&= \int_0^\infty \overline{\nu}(z) \mathrm{e}^{-\overline{\nu}(x)t} f(t) \mathrm{d}t \\
&= \int_0^\infty \overline{\nu}(z) \int_{\overline{\nu}(x)}^\infty t \mathrm{e}^{-ut} \mathrm{d}u \; f(t) \mathrm{d}t \\
&= \int_{\overline{\nu}(x)}^\infty \frac{\overline{\nu}(z)}{u^2} \int_0^\infty t u^2 \mathrm{e}^{-ut} f(t) \mathrm{d}t \; \mathrm{d}u
\end{aligned}$$

In other words, the random variable $\overline{\nu}(Z)$ has a density $\frac{\overline{\nu}(z)}{u^2} \mathbb{1}_{u \geq \overline{\nu}(z)} \mathrm{d}u$, and conditional on $\overline{\nu}(Z)$, $T(z)$ follows a Gamma distribution with parameter $(\overline{\nu}(Z), 2)$. As $U/T(z)$ is uniform on $[0, 1]$ and independent of Z, one can check that $(Z, T(z), U)$ has the same distribution as $(Z, T_1 + T_2, T_1)$, where conditional on Z, the variables T_1 and T_2 are independent with the same exponential distribution with parameter $\overline{\nu}(Z)$. This concludes the proof of (2.9) since conditional on $(Z, T(z), U)$ (resp. $(Z, T_1 + T_2, T_1)$), $\mathcal{N}_z \setminus \{(U, Z)\}$ (resp. $\widetilde{\mathcal{N}} \setminus \{(T_1, Z)\}$) is a Poisson point process on $[0, T(z)) \times [0, Z]$ (resp. on $[0, T_1 + T_2) \times [0, Z]$) with intensity $\mathrm{d}t \otimes \nu$.

2.A.3 Subordinators and Regenerative Sets

We use some classical results about regenerative sets and subordinators, whose proofs can be found in the first two sections of Bertoin's Saint-Flour lecture notes [12].

Definition 2.41. A **subordinator** is a right-continuous, increasing Markov process $(\sigma_t)_{t\geq 0}$ started from 0 with values in $[0,\infty]$, where ∞ is an absorbing state, such that for all $s<t$, conditional on $\{\sigma_s<\infty\}$, we have

$$\sigma_t-\sigma_s \overset{(d)}{=} \sigma_{t-s}.$$

Theorem 2.42. *The distribution of a subordinator is characterized by its* ***Laplace exponent*** *defined as the increasing function $\varphi:[0,\infty)\to[0,\infty)$, such that for all $\lambda,t\geq 0$,*

$$\mathbb{E}[\mathrm{e}^{-\lambda\sigma_t}]=\mathrm{e}^{-t\varphi(\lambda)},$$

with the convention $\mathrm{e}^{-\lambda\infty}=0$ for all $\lambda\geq 0$. The Laplace exponent can be written under the form

$$\varphi(\lambda)=k+d\lambda+\int_{(0,\infty)}(1-\mathrm{e}^{-\lambda x})\pi(\mathrm{d}x),$$

where k is called the ***killing rate****, d the* ***drift coefficient*** *and π the* ***Lévy measure*** *of the subordinator. Necessarily, we have $k,d\geq 0$ and π satisfies*

$$\int_{(0,\infty)}(1\wedge x)\pi(\mathrm{d}x)<\infty.$$

Letting $\zeta:=\inf\{t\geq 0,\ \sigma_t=\infty\}$ be the lifetime of the subordinator, ζ follows an exponential distribution with parameter k (if $k=0$, then $\zeta\equiv\infty$). Also we have almost surely for all $t<\zeta$,

$$\sigma_t=dt+\sum_{s\leq t}\Delta\sigma_s,$$

and the set of jumps $\{(s,\Delta\sigma_s),\ \Delta\sigma_s>0\}$ is a Poisson point process with intensity $\mathrm{d}s\otimes\pi$.

The **renewal measure** of a subordinator is defined as the measure $U(\mathrm{d}x)$ on $[0,\infty)$ such that for any non-negative measurable function f

$$\int_{[0,\infty)}f(x)U(\mathrm{d}x)=\mathbb{E}\left[\int_0^{\zeta}f(\sigma_t)\mathrm{d}t\right].$$

This renewal measure characterizes the distribution of σ since its Laplace transform is the inverse of φ

$$\frac{1}{\varphi(\lambda)}=\int_{[0,\infty)}\mathrm{e}^{-\lambda x}U(\mathrm{d}x).$$

Remark also that setting $L_x:=\inf\{t\geq 0,\sigma_t>x\}$ the right-continuous inverse of σ, we have

$$U(x):=U([0,x])=\mathbb{E}\left[\int_0^{\infty}\mathbb{1}_{\sigma_t\leq x}\,\mathrm{d}t\right]=\mathbb{E}[L_x].$$

Definition 2.43. Given a probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with a complete, right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$, a **regenerative set** R is a random closed set containing 0 for which the following properties hold

- *Progressive measurability.* For all $t \geq 0$, the set $\{(s, \omega) \in [0, t] \times \Omega,\ s \in R(\omega)\}$ is in $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$.
- *Regeneration property.* For a $(\mathcal{F}_t)_{t\geq 0}$-stopping time T such that a.s. on $\{T < \infty\}$, $T \in R$ and T is not right-isolated in R, we have:

$$R \cap [T, \infty[- T \overset{(d)}{=} R,$$

where $R \cap [T, \infty[- T$ is defined formally as the set $\{t \geq 0,\ T + t \in R\}$.

We define the range of a subordinator σ as the closed set $\overline{\{\sigma_t,\ t \geq 0\}}$, and see that all regenerative sets can be expressed in this form.

Theorem 2.44. *The range of a subordinator is a regenerative set. Conversely, if R is a regenerative set without isolated points, there exists a subordinator σ whose range is R almost surely.*

Remark 2.45. In the case where $\lambda(R) > 0$ a.s., one can define such a subordinator as

$$\sigma_t := \inf\{x \geq 0,\ \lambda([0, x] \cap R) > t\}.$$

Then σ is the unique subordinator with drift 1 and range R, and its renewal measure is $U(\mathrm{d}x) = \mathbb{P}(x \in R)\,\mathrm{d}x$. Notice that $\lambda(R) = \inf\{t \geq 0,\ \sigma_t = \infty\} = \zeta$ by definition. Therefore $\lambda(R)$ is an exponential random variable with parameter k, the killing rate of σ.

References for Chapter 2

[1] R. Abraham, J.-F. Delmas, and P. Hoscheit. A Note on the Gromov-Hausdorff-Prokhorov Distance between (Locally) Compact Metric Measure Spaces. *Electron. J. Probab.*, 18 (2013). paper no. 14. DOI: 10.1214/EJP.v18-2116 (see p. 22).

[2] R. Abraham and L. Serlet. Poisson Snake and Fragmentation. *Electron. J. Probab.*, 7 (2002). paper no. 17. DOI: 10.1214/EJP.v7-116 (see pp. 16, 31).

[4] D. Aldous and J. Pitman. The Standard Additive Coalescent. *Ann. Probab.*, 26.4 (Oct. 1998), pp. 1703–1726. DOI: 10.1214/aop/1022855879 (see pp. 16, 126).

[5] A.-L. Basdevant and C. Goldschmidt. Asymptotics of the Allele Frequency Spectrum Associated with the Bolthausen-Sznitman Coalescent. *Electron. J. Probab.*, 13 (2008), pp. 486–512. DOI: 10.1214/EJP.v13-494 (see p. 15).

[7] J. Berestycki, N. Berestycki, and V. Limic. A Small-Time Coupling between Λ-Coalescents and Branching Processes. *Ann. Appl. Probab.*, 24.2 (Apr. 2014), pp. 449–475. DOI: 10.1214/12-AAP911 (see pp. 15, 82).

[12] J. Bertoin. Subordinators: Examples and Applications. *Lectures on Probability Theory and Statistics: École d'Été de Probabilités de Saint-Flour XXVII*. Springer, 1997, pp. 1–91. DOI: 10.1007/978-3-540-48115-7_1 (see pp. 31, 54).

[14] J. Bertoin. The Structure of the Allelic Partition of the Total Population for Galton–Watson Processes with Neutral Mutations. *Ann. Probab.*, 37.4 (July 2009), pp. 1502–1523. doi: 10.1214/08-AOP441 (see pp. 16, 59, 90).

[22] N. Champagnat and A. Lambert. Splitting Trees with Neutral Poissonian Mutations I: Small Families. *Stochastic Process. Appl.*, 122.3 (2012), pp. 1003–1033. doi: 10.1016/j.spa.2011.11.002 (see pp. 15, 33, 34).

[23] N. Champagnat and A. Lambert. Splitting Trees with Neutral Poissonian Mutations II: Largest and Oldest Families. *Stochastic Process. Appl.*, 123.4 (2013), pp. 1368–1414. doi: 10.1016/j.spa.2012.11.013 (see p. 15).

[24] N. Champagnat, A. Lambert, and M. Richard. *Birth and Death Processes with Neutral Mutations*. 2012. url: https://www.hindawi.com/journals/ijsa/2012/569081/ (visited on 09/05/2017) (see p. 15).

[33] C. Delaporte, G. Achaz, and A. Lambert. Mutational Pattern of a Sample from a Critical Branching Population. *J. Math. Biol.*, 73.3 (Sept. 1, 2016), pp. 627–664. doi: 10.1007/s00285-015-0964-2 (see p. 15).

[37] J.-J. Duchamps and A. Lambert. Mutations on a Random Binary Tree with Measured Boundary. *Ann. Appl. Probab.*, 28.4 (Aug. 2018), pp. 2141–2187. doi: 10.1214/17-AAP1353 (see pp. 11, 14).

[41] S. N. Evans. *Probability and Real Trees*. Red. by J. .-.-M. Morel, F. Takens, and B. Teissier. Vol. 1920. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. doi: 10.1007/978-3-540-74798-7 (see p. 17).

[42] W. J. Ewens. The Sampling Theory of Selectively Neutral Alleles. *Theor. Popul. Biol.*, 3.1 (Mar. 1, 1972), pp. 87–112. doi: 10.1016/0040-5809(72)90035-4 (see pp. 9, 15).

[47] F. Freund and M. Möhle. On the Number of Allelic Types for Samples Taken from Exchangeable Coalescents with Mutation. *Adv. in Appl. Probab.*, 41.04 (Dec. 2009), pp. 1082–1101. doi: 10.1017/S000186780000375X (see p. 15).

[48] F. Freund. Almost Sure Asymptotics for the Number of Types for Simple Ξ-Coalescents. *Electron. Commun. Probab.*, 17 (2012). doi: 10.1214/ECP.v17-1704 (see p. 15).

[51] R. C. Griffiths and A. G. Pakes. An Infinite-Alleles Version of the Simple Branching Process. *Adv. in Appl. Probab.*, 20.3 (Sept. 1988), p. 489. doi: 10.2307/1427033 (see p. 15).

[56] D. G. Kendall. On the Generalized "Birth-and-Death" Process. *Ann. Math. Statist.*, 19.1 (Mar. 1948), pp. 1–15. doi: 10.1214/aoms/1177730285 (see pp. 44, 45, 47).

[58] J. Kingman. The Coalescent. *Stochastic Process. Appl.*, 13.3 (1982), pp. 235–248. doi: 10.1016/0304-4149(82)90011-4 (see pp. 8, 15, 22, 59, 90, 95).

[60] A. Lambert and G. Uribe Bravo. The Comb Representation of Compact Ultrametric Spaces. *p-Adic Numbers Ultrametric Anal. Appl.*, 9.1 (Jan. 2017), pp. 22–38. doi: 10.1134/S2070046617010034 (see pp. 15, 19, 29).

[64] A. Lambert. The Allelic Partition for Coalescent Point Processes. *Markov Process. Related Fields*, 15 (2009), pp. 359–386. arXiv: 0804.2572 (see pp. 15, 26, 33, 35).

[66] A. Lambert and E. Schertzer. Recovering the Brownian Coalescent Point Process from the Kingman Coalescent by Conditional Sampling. *Bernoulli*, 25.1 (Feb. 2019), pp. 148–173. doi: 10.3150/17-BEJ971 (see p. 22).

[69] P. Marchal. Nested Regenerative Sets and Their Associated Fragmentation Process. *Mathematics and Computer Science III*. Trends in Mathematics. Birkhäuser Basel, 2004, pp. 461–470. doi: 10.1007/978-3-0348-7915-6_45 (see pp. 23, 29).

[77] L. Popovic. Asymptotic Genealogy of a Critical Branching Process. *Ann. Appl. Probab.*, 14.4 (Nov. 2004), pp. 2120–2148. doi: 10.1214/105051604000000486 (see p. 29).

[89] Z. Taïb. *Branching Processes and Neutral Evolution*. Red. by S. A. Levin. Vol. 93. Lecture Notes in Biomathematics. Springer Berlin Heidelberg, 1992. doi: 10.1007/978-3-642-51536-1 (see p. 15).

Chapter 3

Trees within trees: Simple nested coalescents

Joint work with Airam Blancas, Amaury Lambert and Arno Siri-Jégousse. This chapter is published in *Electronic Journal of Probability* [18].

Contents

3.1 Introduction

In the framework of population biology, one can see asexual organisms, but also DNA sequences or even species, as replicating particles. The genealogical ascendance of co-existing replicating particles can always be represented by a tree whose tips are labelled by the names of these particles [61, 63, 85]. Even if species are not strictly speaking replicating particles, ancestral relationships between species are also usually represented by a tree whose nodes are interpreted as *speciation* events, i.e., the emergence of two or more species from one single species. The inference of the so-called *gene tree* of contemporary DNA sequences from their comparison has a decade-long history. It is considered as a field in its own right, called *molecular phylogenetics* [43, 72], which relies heavily on the theory of Markov processes. (This can be misleading, but the *species tree*, much more often than the gene tree, is called a *phylogeny*.)

When one type of replicating particle is physically embedded in another type of particle, like a virus in its host, their common history can be depicted as a *tree within a tree* [35, 68, 74]: tree of dividing parasites inside the tree of dividing hosts, tree of paralogous genes (i.e. distinct DNA segments resulting from gene duplication and coding for similar functions) inside the gene family tree, gene tree inside the species tree. In many such cases, biologists are more interested in the coarser tree rather than in the finer tree. Typically, the finer tree is a gene tree and is inferred thanks to methods developed in molecular phylogenetics. One of the current methodological challenges in quantitative biology is to devise fast statistical algorithms able to also infer the coarser tree. When the genes are sampled from infecting pathogens of the same species (Influenza, HIV...), the coarser tree is the epidemic transmission process [50, 91]. When the genes are sampled from (any kind of) different species, the coarser tree is the *species tree* [54, 73, 88]. It is often required to use several gene trees nested in the same species tree to infer the latter.

In terms of stochastic modeling, the standard strategy is to define the two nested trees in a hierarchical model referred to as the *multispecies coalescent model* [32, 78] (see also [14, 40] for recent surveys on general coalescent theory and applications to population genetics). First, the species tree is fixed or drawn from some classic probability distribution (e.g., pure-birth process stopped at some fixed time, viewed as present time). Second, each gene sequence is assigned to the contemporary species it is (supposed to be) sampled from. Recall that each contemporary species is in correspondence with a tip of the species tree. Third, conditional on the species tree, each gene lineage can then be traced backwards in time inside the species tree, starting from the tip species harboring it and traveling through its ancestral species successively. In addition, gene lineages are assumed to coalesce according to the *censored Kingman coalescent* [58], i.e., each pair of lineages *lying in the same species* independently coalesces at constant rate.

In the case when the species tree is also distributed as a Kingman coalescent, the former two-type coalescent process is a Markov process as time runs backward, that we call the *nested Kingman coalescent* (or 'Kingman-in-Kingman') [19, 31, 65]. Our goal here is to display a much richer class of Markov models for trees within trees, called *simple nested exchangeable coalescent* (SNEC) processes, where multiple species lineages can merge into one single species lineage, and where simultaneously, within those merging species, multiple gene lineages can merge into one single gene lineage. To make this more precise, we show in the next display some valid and invalid coalescence events from an initial state where six genes, labeled from 1 to 6, are grouped by pairs in three species lineages. We represent this situation in the next display by a pair of partitions $\binom{\pi^s}{\pi^g}$, as in the left-hand side of the display. Event (A) is valid because the first two species merge and simultaneously, *within* these species, genes labeled 1, 2 and 3 coalesce. On the contrary, event (B) is not a valid transition because there are two distinct gene coalescences (1 with 2, and 3 with 4), which is proscribed, and event (C) is not valid because the gene coalescence (5 with 6)

is outside the species coalescence.

$$\left(\begin{array}{cccccc}\{1, & 2\}\{3, & 4\}\{5, & 6\}\\ \{1\}\{2\}\{3\}\{4\}\{5\}\{6\}\end{array}\right) \to \left(\begin{array}{cccccc}\{1, & 2, & 3, & 4\}\{5, & 6\}\\ \{1, & 2, & 3\}\{4\}\{5\}\{6\}\end{array}\right) \quad \text{(A)}$$

$$\not\to \left(\begin{array}{cccccc}\{1, & 2, & 3, & 4\}\{5, & 6\}\\ \{1, & 2\}\{3, & 4\}\{5\}\{6\}\end{array}\right) \quad \text{(B)}$$

$$\not\to \left(\begin{array}{cccccc}\{1, & 2, & 3, & 4\}\{5, & 6\}\\ \{1\}\{2\}\{3\}\{4\}\{5, & 6\}\end{array}\right) \quad \text{(C)}$$

In brief, SNEC processes are the generalization of Λ-*coalescents* to processes valued, not in partitions of $\mathbb{N}$, but in pairs of nested partitions of $\mathbb{N}$. The class of Λ-coalescents [76, 79], for which only one coalescence event can occur at a time, is a subclass of Markov, exchangeable processes with possibly non-binary nodes, called Ξ-*coalescents*, where several coalescence events can be simultaneous [10, 83].

Non-binary nodes in species trees can be interpreted as *unresolved nodes* (a sequence of binary nodes following each other too closely in time for their order to be inferred correctly) or *radiation* events (periods of frequent speciations due to the opening of new ecological opportunities that can be exploited by different, new species). In gene trees, non-binary nodes are increasingly recognized as a conspicuous sign of natural selection both by biologists [70, 90] and by mathematicians and physicists [8, 21, 34, 38, 71, 84] ; it is also well understood that non-binary nodes could be consequences of bottlenecks as well as large variance in offspring distributions [39, 82]. The class of SNEC processes includes all these features. They can distinguish unresolved nodes (sequence of stochastically close, binary coalescences) from radiations (multiple merger in the species tree). Under the interpretation of non-binary nodes as a result of natural selection, SNEC processes can model the appearance of alleles responsible for positive selection (multiple merger in the gene tree) or for divergent adaptation (multiple merger simultaneously in the gene tree and in the species tree).

From a mathematical point of view as well, SNEC processes open up the door to many possible new investigations. For example some of us are currently studying the speed of coming down from infinity of SNEC processes [19, 65] as well as similar extensions (see Chap. 4) to fragmentation processes [10]. It will be interesting to investigate how the nested trees generated by SNEC processes can be cast in the frameworks of multilevel measure-valued processes [31] and flows of bridges [15, 16] as well as of exchangeable combs [46, 63]. It would also be natural to study the extension of Ξ-coalescents to nested partitions.

Organization of the article. In Section 3.2, we introduce some notation, and give examples of nested coalescent processes whose distributions are characterized by four parameters. Section 3.3 formally defines our object of study, the SNEC processes. We prove our main result in Section 3.4, and show in Section 3.5 how SNEC processes can be constructed from a collection of Poisson point processes. Finally, Section 3.6 gives a necessary and sufficient condition under which SNEC processes come down from infinity.

3.2 Statement of results and notation

3.2.1 Statement of results and examples

An exchangeable partition is a random partition of $\mathbb{N}$ whose law is invariant by permutations of $\mathbb{N}$ (with finite support). A Λ-coalescent is a Markov process valued in the exchangeable partitions of $\mathbb{N}$ typically starting from the partition $\mathbf{0}_\infty$ of $\mathbb{N}$ into singletons, and such that only one coalescence event can occur at a time. The generator of a Λ-coalescent $\mathcal{R} = (\mathcal{R}(t), t \geq 0)$ is characterized by a σ-finite measure ν on $(0, 1]$ called the coagulation measure and a non-negative real number a called the Kingman coefficient. Then $\mathcal{R}$ can be constructed from a Poisson point process as follows.

For $x \in (0, 1]$, let P_x denote the law of a sequence of i.i.d. Bernoulli(x) r.v.'s and define

$$P := \int_{(0,1]} \nu(dx) P_x$$

Also define $\mathtt{K}_{i,i'}$ the (Dirac) law of the sequence with only zero entries except a 1 at positions i and i' and set

$$\mathtt{K} := \sum_{1 \leq i < i'} \mathtt{K}_{i,i'}$$

Finally, let M be a Poisson point process with intensity measure $dt \otimes (P + a\mathtt{K})$. Roughly speaking, at each atom $(t, (X_i, i \geq 1))$ of M, $\mathcal{R}(t)$ is obtained from $\mathcal{R}(t-)$ by merging exactly the i-th block of $\mathcal{R}(t-)$ together, for all i such that $X_i = 1$. The rigorous description is given through restrictions of $\mathcal{R}$ to $[n] := \{1, \ldots, n\}$ and by applying Kolmogorov extension theorem. See [10] for details. Note that for this description to apply (i.e., for restrictions of $\mathcal{R}$ to $[n]$ to have positive holding times), one needs the coagulation measure to satisfy

$$\int_{(0,1]} x^2 \, \nu(dx) < \infty. \tag{3.1}$$

The finite measure $x^2 \, \nu(dx)$ is usually denoted $\Lambda(dx)$, hence the name Λ-coalescent.

We can now draw the parallel with the results obtained in this paper. We want to define a Markov process $\mathcal{R} = ((\mathcal{R}^s(t), \mathcal{R}^g(t)),\ t \geq 0)$ valued in exchangeable bivariate, nested partitions of $\mathbb{N}$, in the sense that the *gene partition* $\mathcal{R}^g(t)$ is finer than the *species partition* $\mathcal{R}^s(t)$ for all t a.s.

We now have to allow for coalescences in both the gene partition and the species partition. To this aim, we will consider a doubly indexed array of 0's and 1's $\mathbf{Z} = (\mathbf{X}, (\mathbf{Y}_i,\ i \geq 1)) = (X_i, Y_{ij},\ i, j \geq 1)$. The goal is to give a characterization and a Poissonian construction of $\mathcal{R}$ under the assumptions that the semigroup of $\mathcal{R}$ is exchangeable and that both $\mathcal{R}^s$ and $\mathcal{R}^g$ undergo only one coalescence at a time (but possibly the same time), as detailed in forthcoming Definition 3.2. Roughly speaking, and similarly as previously, X_i will determine whether the i-th species block participates in the coalescence in the species partition $\mathcal{R}^s$, and Y_{ij} whether the j-th gene block of the i-th species block participates in the coalescence in the gene partition $\mathcal{R}^g$.

Let us start with the Kingman-type coalescences. Let $\mathtt{K}^s_{i,i'}$ be the (Dirac) law of the array $\mathbf{Z}$ with only zero entries except $X_i = X_{i'} = 1$ and let $\mathtt{K}^g_{i;j,j'}$ be the (Dirac) law of the array

$\mathbf{Z}$ with only zero entries except $X_i = Y_{ij} = Y_{ij'} = 1$. Finally, define

$$\mathtt{K}^s = \sum_{1\le i<i'} \mathtt{K}^s_{i,i'} \quad \text{and} \quad \mathtt{K}^g = \sum_{1\le i} \sum_{1\le j<j'} \mathtt{K}^g_{i;j,j'}$$

Let us carry on with multiple gene mergers without simultaneous species coalescences. Let $x \in (0,1]$ and $i \in \mathbb{N}$. Let $P^g_{i,x}$ be the distribution of the array $\mathbf{Z}$ with only zero entries except at row i, where $X_i = 1$ and the $(Y_{ij},\, j \ge 1)$ are i.i.d. Bernoulli(x) r.v.'s. Let us define

$$P^g_x := \sum_{i\ge 1} P^g_{i,x}$$

Finally, let us consider multiple species mergers, with possible simultaneous gene mergers. Let $x \in (0,1]$ and $\mu \in \mathcal{M}_1([0,1])$. Let $(X_i,\, i \ge 1)$ be a sequence of i.i.d. Bernoulli(x) r.v.'s and let $(Q_i,\, i \ge 1)$ be an independent sequence of i.i.d. r.v.'s of $[0,1]$ with distribution μ. Then for each $i \ge 1$, conditional on X_i and Q_i, let $(Y_{ij},\, j \ge 1)$ be an independent sequence of i.i.d. Bernoulli(Q_i) r.v.'s. if $X_i = 1$ and the null array otherwise. Let us write $P^s_{x,\mu}$ for the distribution of the array $\mathbf{Z}$ thus defined.

Our main result is that for any simple nested exchangeable coalescent (SNEC) process $\mathcal{R}$, there are

- two non-negative real numbers a_s and a_g;
- a σ-finite measure ν_g on $(0,1]$;
- a σ-finite measure ν_s on $(0,1] \times \mathcal{M}_1([0,1])$,

such that $\mathcal{R}$ can be constructed from a Poisson point process M with intensity $dt \otimes \nu(d\mathbf{Z})$ where

$$\nu := a_s \mathtt{K}_s + a_g \mathtt{K}_g + \int_{(0,1]} \nu_g(dx)\, P^g_x + \int_{(0,1]\times\mathcal{M}_1([0,1])} \nu_s(dx, d\mu)\, P^s_{x,\mu}.$$

Similarly as explained previously, at each atom $(t, \mathbf{Z})$ of M, the double array $\mathbf{Z}$ prescribes which blocks have to merge at time t. For the finite restrictions of $\mathcal{R}$ to have positive holding times, the measures ν_s and ν_g are required to satisfy the forthcoming conditions (3.7) and (3.8) respectively, which are the analogs to (3.1).

Note that coagulations of the Kingman type cannot occur simultaneously in the species partition and in the gene partition.

We now give a couple of examples of SNEC processes.

If $\nu_s(dp, d\mu) = \nu'_s(dp)\, \delta_{\delta_0}(d\mu)$, species and genes never coalesce simultaneously and the nested coalescent is a multispecies coalescent (see Introduction), where the species tree is given by the Λ-coalescent with coagulation measure ν'_s and Kingman coefficient a_s, while the genes in the same species block undergo independent Λ-coalescents with coagulation measure ν_g and Kingman coefficient a_g. In particular, when ν'_s and ν_g are zero, the SNEC process is a nested Kingman coalescent (Kingman-in-Kingman).

Whenever ν_s is not under the form $\nu_s(dp, d\mu) = \nu'_s(dp)\, \delta_{\delta_0}(d\mu)$, species blocks and gene blocks can coalesce simultaneously. For example if $\nu_s(dp, d\mu) = \nu'_s(dp)\, \delta_{\delta_x}(d\mu)$ for $x \in (0,1]$, at each species coalescence event, a proportion x of gene blocks contained in the

species blocks participating in the coalescence event, are simultaneously merged together. In particular, if $x = 1$, the gene tree coincides with the species tree on lineages situated after a species coalescence event. Recall that there are conditions (see (3.7)) for ν_s to be a correct SNEC measure, which in this case translate to

$$\int_{(0,1]} \nu_s'(dp)\, p^2 < \infty \quad \text{and} \quad \int_{(0,1]} \nu_s'(dp)\, p\, x^2 < \infty,$$

which is simply equivalent to

$$\int_{(0,1]} \nu_s'(dp)\, p < \infty.$$

Otherwise the simplest sort of measure ν_s can be obtained by parameterizing its second component μ, for example if μ is a Beta distribution $\mu_{a,b}(dq) = c_{a,b}\, q^{a-1}(1-q)^{b-1}\, dq$, where $a, b > 0$ and $c_{a,b} = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}$, we can consider ν_s under the form

$$\nu_s(dp, d\mu) = \nu_s'(dp, da, db)\, \delta_{\mu_{a,b}}(d\mu).$$

In this case, the condition (3.7a) reads

$$\int_{(0,1]\times(0,\infty)\times(0,\infty)} \nu_s'(dp, da, db)\, p^2 < \infty,$$

and (3.7b) becomes

$$\int_{(0,1]\times(0,\infty)\times(0,\infty)} \nu_s'(dp, da, db)\, p \int_{[0,1]} c_{a,b}\, q^{a+1}(1-q)^{b-1}\, dq < \infty,$$

which can be rewritten

$$\int_{(0,1]\times(0,\infty)\times(0,\infty)} \nu_s'(dp, da, db)\, \frac{pa(a+1)}{(a+b)(a+b+1)} < \infty.$$

Note that the idea to use a Beta distribution here is inspired by the Λ-coalescent setting [76], where Beta distributions appear as natural candidates for the parametrization of the measure Λ, as the coalescence rate of each k-tuple of blocks among a total number of b blocks is expressed in the form

$$\int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx).$$

3.2.2 Notation

For any $n \in \bar{\mathbb{N}} := \mathbb{N} \cup \{+\infty\}$, let $\mathcal{P}_n$ be the set of partitions of $[n]$. A partition π is called *simple* if at most one of its non-empty blocks is not a singleton. We denote the set of simple partitions of $[n]$ by $\mathcal{P}_n'$, that is,

$$\mathcal{P}_n' = \{\pi \in \mathcal{P}_n,\ \text{Card}\{i,\ |\pi_i| > 1\} \leq 1\}$$

where $\pi_1, \pi_2, \ldots$ denote the blocks of π ordered by their least element and $|\pi_i|$ stands for the number of elements in the block π_i. Recall that a partition π can be viewed as an equivalence relation, in the sense that $i \overset{\pi}{\sim} j$ if and only if i and j belong to the same block

of the partition π. If π^g and π^s belong to $\mathcal{P}_n$, we will say that the bivariate partition $\pi = (\pi^s, \pi^g)$ is *nested* (or equivalently that π^g is finer than π^s) when

$$i \overset{\pi^g}{\sim} j \implies i \overset{\pi^s}{\sim} j.$$

Note that this is defines a natural partial order on $\mathcal{P}_n$, and we can write $\pi^g \preceq \pi^s$ if (π^g, π^s) is nested. The set of nested partitions of $[n]$ is denoted in the sequel by $\mathcal{N}_n$. We will sometimes use the notation $\mathbf{1}_n := \{[n]\}$ for the coarsest partition of $[n]$, and $\mathbf{0}_n := \{\{1\}, \{2\}, \ldots\}$ for the finest partition of $[n]$.

Example 3.1. An example of nested partition of $\{1, 2, \ldots, 10\}$ is given by

$$\pi^s = \{\{1, 5, 7\},\ \{2, 4, 8, 10\},\ \{3, 6, 9\}\}$$
$$\pi^g = \{\{1\},\ \{2, 4\},\ \{3\},\ \{5, 7\},\ \{6, 9\},\ \{8\},\ \{10\}\}.$$

The notation (π^s, π^g) owes to our modeling inspiration (see Introduction) where gene lineages are enclosed into species lineages.

Notation related to and properties of $\mathcal{P}_n$ can naturally be extended to the framework of bivariate partitions. For the sake of completeness we specify here the ones we will use repeatedly. The number of non-empty blocks of a bivariate partition $\pi = (\pi_1, \pi_2) \in \mathcal{P}_{n_1} \times \mathcal{P}_{n_2}$ is merely $|\pi| := (|\pi_1|, |\pi_2|)$. If $m_1 < n_1$ and $m_2 < n_2$, we write $\pi_{|m_1 \times m_2}$ for the restriction of π to $\mathcal{P}_{m_1} \times \mathcal{P}_{m_2}$, that is, $\pi_{|m_1 \times m_2} = (\pi_{1\,|m_1}, \pi_{2\,|m_2})$. If $m \leq \min(n_1, n_2)$, we will write $\pi_{|m} := \pi_{|m \times m}$ for its restriction to $\mathcal{P}_m^2 := \mathcal{P}_m \times \mathcal{P}_m$. A sequence $\pi^{(1)}, \pi^{(2)}, \ldots$ of elements of $\mathcal{P}_1^2, \mathcal{P}_2^2, \ldots$ is called *consistent* if for all integers $k' \leq k$, $\pi^{(k')}$ coincides with the restriction of $\pi^{(k)}$ to $[k']^2$. Moreover, a sequence of partitions $(\pi^{(n)} : n \in \mathbb{N})$ is consistent if and only if there exists $\pi \in \mathcal{P}_\infty^2$ such that $\pi_{|n} = \pi^{(n)}$ for every $n \in \mathbb{N}$.

Given a nested partition we can use the coagulation operator Coag (more details in Chapter 3 in Bertoin [10]) to write the species partition in terms of the labels of the gene partition. Recall that if $\pi \in \mathcal{P}_n$ and $\tilde{\pi} \in \mathcal{P}_m$ with $m \geq |\pi|$, then define $\pi' = \mathrm{Coag}(\pi, \tilde{\pi})$ as the partition of $\mathcal{P}_n$ such that

$$\pi'_j = \bigcup_{i \in \tilde{\pi}_j} \pi_i.$$

For every $n \in \bar{\mathbb{N}}$, let $\pi = (\pi^s, \pi^g)$ be an element of $\mathcal{N}_n$ and write $m = |\pi^g|$. The unique partition $\bar{\pi} \in \mathcal{P}_m$ such that $\pi^s = \mathrm{Coag}(\pi^g, \bar{\pi})$ is called the *link* partition of π. We sometimes say that π is linked by $\bar{\pi}$. To illustrate the previous definition, observe that the nested partition defined in Example 3.1 has link partition $\bar{\pi} = \{\{1, 4\}, \{2, 6, 7\}, \{3, 5\}\}$.

We can next get a partition of $\mathcal{P}_{n_1} \times \mathcal{P}_{n_2}$ through the coagulation of two pairs of partitions. More precisely, if $(\pi^1, \tilde{\pi}^1) \in \mathcal{P}_{n_1} \times \mathcal{P}_{n'_1}$ and $(\pi^2, \tilde{\pi}^2) \in \mathcal{P}_{n_2} \times \mathcal{P}_{n'_2}$ with $n'_1 \geq |\pi^1|$ and $n'_2 \geq |\pi^2|$, then $(\mathrm{Coag}(\pi^1, \tilde{\pi}^1), \mathrm{Coag}(\pi^2, \tilde{\pi}^2))$ is well defined and it is an element of $\mathcal{P}_{n_1} \times \mathcal{P}_{n_2}$. If we denote $\pi = (\pi^1, \pi^2)$ and $\tilde{\pi} = (\tilde{\pi}^1, \tilde{\pi}^2)$ we will say that the pair $(\pi, \tilde{\pi})$ is *admissible* and denote the latter operation by $\mathrm{Coag}_2(\pi, \tilde{\pi})$. In the following we will sometimes call the partition $\tilde{\pi}$ as the *recipe* partition.

In the sequel, we are interested in the coagulation of a nested partition, say $\pi = (\pi^s, \pi^g)$, with a pair of simple partitions $\tilde{\pi} = (\tilde{\pi}^s, \tilde{\pi}^g)$. Nevertheless, we should observe that the

resulting partition, $\mathrm{Coag}_2(\pi, \tilde{\pi})$ is not necessarily nested. For instance, if we coagulate the partition π of Example 3.1, with $\tilde{\pi}^s = \{\{1,2\},\{3\}\}$, and $\tilde{\pi}^g = \{\{1,3\},\{2\},\{4\},\{5\},\{6\},\{7\}\}$ then $\mathrm{Coag}(\pi^g, \tilde{\pi}^g)$ is not nested in $\mathrm{Coag}(\pi^s, \tilde{\pi}^s)$. In order to maintain the nested property while coagulating a nested partition we need to watch out the way the gene blocks do merge together and if they respect the species structure. To this end, for any $n \in \bar{\mathbb{N}}$ and $\pi \in \mathcal{N}_n$, we can define the set $\widetilde{\mathcal{P}}(\pi) \subset (\mathcal{P}'_n)^2$ of simple recipe partitions permitting a consistent merger of species and genes, i.e.

$$\widetilde{\mathcal{P}}(\pi) = \left\{\tilde{\pi} = (\tilde{\pi}^s, \tilde{\pi}^g) \in (\mathcal{P}'_n)^2,\ i \overset{\tilde{\pi}^g}{\sim} j \implies i \overset{\bar{\pi}}{\sim} j,\ \text{or } k \overset{\tilde{\pi}^s}{\sim} l,\ \text{where } \pi^g_i \subset \pi^s_k \text{ and } \pi^g_l \subset \pi^s_l\right\},$$

where $\bar{\pi}$ denotes as usual the link partition of π. Simply put, $\widetilde{\mathcal{P}}(\pi)$ is the subset of $(\mathcal{P}'_n)^2$ such that

$$\tilde{\pi} \in \widetilde{\mathcal{P}}(\pi) \iff \mathrm{Coag}_2(\pi, \tilde{\pi}) \in \mathcal{N}_n.$$

Finally the natural partial order on partitions can be extended to bivariate partitions by defining $(\pi^{1,s}, \pi^{1,g}) \preceq (\pi^{2,s}, \pi^{2,g}) \iff \pi^{1,s} \preceq \pi^{2,s}$ and $\pi^{1,g} \preceq \pi^{2,g}$. This partial order allows us to see coalescent processes as nondecreasing processes in the space of nested partitions.

3.3 Simple nested exchangeable coalescents

In the aim to describe the joint dynamics of the species and gene partitions, we will now define a nondecreasing process with values in the nested partitions, called *nested coalescent process*. In this work we are only interested in *simple* nested coalescents in the sense that at any jump event, called coalescence event, all blocks undergoing a modification merge into one single block. Simple exchangeable coalescent processes were first introduced independently by Pitman [76] and Sagitov [79], and are usually called in the literature Λ-coalescents (see Introduction). Here we use the term *simple* as in [10], to denote the analog of a Λ-coalescent process in the case of (nested) bivariate partitions.

Note that for any partition $\pi \in \mathcal{P}_\infty$ and any *injection* $\sigma : \mathbb{N} \to \mathbb{N}$, there is a partition $\sigma(\pi)$ defined by

$$i \overset{\sigma(\pi)}{\sim} j \iff \sigma(i) \overset{\pi}{\sim} \sigma(j).$$

For bivariate partitions we define in the same way $\sigma(\pi^s, \pi^g) := (\sigma(\pi^s), \sigma(\pi^g))$. For random partitions, exchangeability is usually defined as invariance under the action of permutations $\sigma : \mathbb{N} \to \mathbb{N}$. Here, to avoid degenerate processes we will define our processes as being invariant under the action of all injections $\sigma : \mathbb{N} \to \mathbb{N}$. Indeed, by making this assumption we avoid dependence on, for instance, the total number of blocks of the partition. An example of what we consider here a degenerate process with values in $\mathcal{P}_\infty$ would be a modified Kingman coalescent where any pair of blocks merge at rate $a = a(n)$, a function of n the total number of blocks. While this process would be invariant under permutations of $\mathbb{N}$, it is in general not invariant under injections, as their action can change the total number of blocks in a partition of $\mathbb{N}$. Furthermore, given $(\Pi(t),\ t \geq 0)$ such a process and n an integer, the restriction $(\Pi(t)_{|n},\ t \geq 0)$ would not be a Markov process, as the jump rates of $\Pi(t)_{|n}$ depend on the whole partition $\Pi(t)$. Invariance under injections ensures us that processes can be consistently defined, i.e. that $(\Pi(t)_{|n},\ t \geq 0)$ will always be a

Markov process. It will also be useful in forthcoming proofs to consider invariance under injections rather than only permutations.

Since we consider processes with values in the space $\mathcal{P}_\infty$, let us endow it with the natural topology generated by the sets of the form $\{\pi' \in \mathcal{P}_\infty,\ \pi'_{|n} = \pi\}$ for $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_n$. It is readily checked that this topology is metrizable and makes $\mathcal{P}_\infty$ compact. Also, note that the product topology on $\mathcal{P}_\infty^2$, and that induced on $\mathcal{N}_\infty$ also makes them compact.

Definition 3.2. Let $\mathcal{R} := ((\mathcal{R}^s(t), \mathcal{R}^g(t)),\ t \geq 0)$ be a càdlàg Markov process with values in $\mathcal{P}_\infty^2$. This process is called a *simple nested exchangeable coalescent*, SNEC for short, if

i) For any $t \geq 0$, $\mathcal{R}(t)$ is nested;

ii) The process $(\mathcal{R}(t),\ t \geq 0)$ evolves with simple coalescence events, that is for any time $t \geq 0$ such that $\mathcal{R}(t-) \neq \mathcal{R}(t)$, there is a random bivariate partition $\widetilde{\mathcal{R}}(t) = (\widetilde{\mathcal{R}}^s(t), \widetilde{\mathcal{R}}^g(t))$ taking values in $\widetilde{\mathcal{P}}(\mathcal{R}(t-))$ such that

$$\mathcal{R}(t) = \mathrm{Coag}_2(\mathcal{R}(t-), \widetilde{\mathcal{R}}(t));$$

iii) The semigroup of the process $(\mathcal{R}(t),\ t \geq 0)$ is exchangeable, in the sense that for any $t, t' \geq 0$ and any *injection* $\sigma : \mathbb{N} \to \mathbb{N}$,

$$\left(\sigma(\mathcal{R}(t+t')) \mid \mathcal{R}(t) = \pi\right) \overset{(d)}{=} \left(\mathcal{R}(t+t') \mid \mathcal{R}(t) = \sigma(\pi)\right). \tag{3.2}$$

To start the analysis of SNEC processes we would like to make some observations related to Definition 3.2. First note that $\mathcal{R}$ is a $\mathcal{N}_\infty$-valued process such that for every $t, t' \geq 0$, the conditional distribution of $\mathcal{R}(t+t')$ given $\mathcal{R}(t) = \pi$ is the law of $\mathrm{Coag}_2(\pi, \tilde{\pi})$, where $\tilde{\pi} \in \widetilde{\mathcal{P}}(\pi)$, hence the law of $\tilde{\pi}$ depends on t' but also on π. Also, it will be clear from our main result (see Theorem 3.5) that $(\mathcal{R}^s(t),\ t \geq 0)$ is an exchangeable coalescent, however $(\mathcal{R}^g(t),\ t \geq 0)$ is not a Markov process in general, because the distribution of $\mathcal{R}^g(t+t')$ may depend on $\mathcal{R}^s(t)$.

We now turn to investigate the transitions of the restrictions of a SNEC to finite partitions, which relies on the following lemma.

Lemma 3.3 (Projective Markov property)**.** *Let $\mathcal{R} = (\mathcal{R}(t),\ t \geq 0)$ be a process with values in $\mathcal{N}_\infty$ and for every integer n, write $\mathcal{R}_{|n} = (\mathcal{R}_{|n}(t),\ t \geq 0)$ for its restriction to $\mathcal{N}_n$. Then $\mathcal{R}$ is a SNEC in $\mathcal{N}_\infty$ if and only if for all $n \in \mathbb{N}$, $\mathcal{R}_{|n}$ is a continuous-time Markov chain on the space $\mathcal{N}_n$ satisfying the analog of statements i) – iii) of Definition 3.2, namely:*

i) For all $t \geq 0$, $\mathcal{R}_{|n}$ is nested;

ii) For $\varrho, \pi \in \mathcal{N}_n$, the rate from ϱ to π is zero if π can not be obtained from a simple coalescence event;

iii) The Markov chain $(\mathcal{R}_{|n}(t),\ t \geq 0)$ is exchangeable, in the sense that for any $t, t' \geq 0$, $\varrho, \pi \in \mathcal{N}_n$ and σ permutation *of n, the rate from ϱ to π is equal to that from $\sigma(\varrho)$ to $\sigma(\pi)$.*

Proof. Let $\mathcal{R}$ be a SNEC in $\mathcal{N}_\infty$ and let $n \in \mathbb{N}$. Let us prove that $\mathcal{R}_{|n}$ satisfies the claimed properties. Let $\varrho \in \mathcal{N}_n$. Pick $\varrho^\star \in \mathcal{N}_\infty$ such that $\varrho^\star_{|n} = \varrho$, and which contains an infinite number of species blocks, each of which containing an infinite number of gene blocks, each of them being an infinite subset of $\mathbb{N}$. Now for any $\varrho' \in \mathcal{N}_\infty$ such that $\varrho'_{|n} = \varrho$, there is an injection $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\sigma(\varrho^\star) = \varrho'$ and such that $\sigma_{|[n]} = \mathrm{id}_{[n]}$, so for any $t, t' \geq 0$,

$$\begin{aligned}(\mathcal{R}_{|n}(t+t') \mid \mathcal{R}(t) = \varrho') &\overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}(t) = \sigma(\varrho^\star)) \\ &\overset{(d)}{=} (\sigma(\mathcal{R})_{|n}(t+t') \mid \mathcal{R}(t) = \varrho^\star) \\ &\overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}(t) = \varrho^\star).\end{aligned}$$

Since this is valid for any ϱ' such that $\varrho'_{|n} = \varrho$, this conditional distribution depends only on $\{\mathcal{R}_{|n}(t) = \varrho\}$, which proves that $\mathcal{R}_{|n}$ is a Markov process. Now the assumption that $\mathcal{R}$ has càdlàg paths ensures us that the process $\mathcal{R}_{|n}$ stays some positive time in each visited state *a.s.* Therefore $\mathcal{R}_{|n}$ is a continuous-time Markov chain. Now statements *i) – iii)* are easily deduced from Definition 3.2.

Conversely, let $\mathcal{R} = (\mathcal{R}(t),\, t \geq 0)$ be a process with values in $\mathcal{N}_\infty$ such that for all $n \in \mathbb{N}$, $\mathcal{R}_{|n}$ is a Markov chain satisfying *i) – iii)* of the lemma. Then *i)* and *ii)* of Definition 3.2 follow immediately, and it remains to check that for any injection $\sigma : \mathbb{N} \to \mathbb{N}$, the equality in distribution (3.2) holds.

Let $\sigma : \mathbb{N} \to \mathbb{N}$ be an injection and fix $n \in \mathbb{N}$. Define $N = \max\{\sigma(1), \sigma(2), \ldots, \sigma(n)\}$, and consider $\widetilde{\sigma} : [N] \to [N]$ a permutation such that for all $1 \leq i \leq n$, $\widetilde{\sigma}(i) = \sigma(i)$. For instance, one can define inductively for $n+1 \leq i \leq N$,

$$\widetilde{\sigma}(i) := \min([N] \setminus \{\sigma(1), \sigma(2), \ldots, \sigma(i-1)\}).$$

Now notice that for any $t \geq 0$ and any $\pi \in \mathcal{N}_\infty$,

$$\sigma(\pi)_{|n} = \widetilde{\sigma}(\pi_{|N})_{|n},$$

which enables us to write, for any $t, t' \geq 0$,

$$\begin{aligned}(\sigma(\mathcal{R})_{|n}(t+t') \mid \mathcal{R}(t) = \pi) &\overset{(d)}{=} (\widetilde{\sigma}(\mathcal{R}_{|N}(t+t'))_{|n} \mid \mathcal{R}(t) = \pi) \\ &\overset{(d)}{=} (\mathcal{R}_{|N}(t+t')_{|n} \mid \mathcal{R}_{|N}(t) = \widetilde{\sigma}(\pi_{|N})) \\ &\overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}_{|n}(t) = \widetilde{\sigma}(\pi_{|N})_{|n}) \\ &\overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}_{|n}(t) = \sigma(\pi)_{|n}) \\ &\overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}(t) = \sigma(\pi)).\end{aligned}$$

The passage to the second line in the last display is a consequence of *iii)* of the lemma, and we used the fact that restrictions are Markov chains, i.e. $(\mathcal{R}_{|n}(t+t') \mid \mathcal{R}_{|n}(t) = \pi_{|n}) \overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}(t) = \pi)$. Since n is arbitrary in $(\sigma(\mathcal{R})_{|n}(t+t') \mid \mathcal{R}(t) = \pi) \overset{(d)}{=} (\mathcal{R}_{|n}(t+t') \mid \mathcal{R}(t) = \sigma(\pi))$, we have shown (3.2), concluding the proof. □

This key lemma enables us to give the following first properties of SNEC processes.

Proposition 3.4. *Let $\mathcal{R}$ be a SNEC.*

- *If the process $\mathcal{R}$ starts from an exchangeable nested partition $\mathcal{R}(0)$, then for any $t \geq 0$, $\mathcal{R}^g(t)$ and $\mathcal{R}^s(t)$ are exchangeable partitions.*
- *The process $\mathcal{R}$ is a Feller process, so in particular it satisfies the strong Markov property.*
- *Conditional on $\mathcal{R}(t)$, if $\bar{\mathcal{R}}(t)$ denotes the link partition of $\mathcal{R}(t)$ then for any $t, t' \geq 0$, the distribution of $\mathcal{R}^g(t+t')$ is the law of $Coag(\mathcal{R}^g(t), \tilde{\pi}^g)$, where $\tilde{\pi}^g$ is a random partition such that $\sigma(\tilde{\pi}^g) \overset{d}{=} \tilde{\pi}^g$ for any permutation σ preserving $\bar{\mathcal{R}}(t)$ i.e., such that*

$$i \overset{\bar{\mathcal{R}}(t)}{\sim} j \Rightarrow \sigma(i) \overset{\bar{\mathcal{R}}(t)}{\sim} \sigma(j). \tag{3.3}$$

Another property is that the process $(\mathcal{R}^s(t),\ t \geq 0)$ is a simple exchangeable coalescent process, but we do not prove it at this point as it will be clear from Theorem 3.5.

Proof. The first point of the proposition is immediate considering *iii)* of Definition 3.2.

As for the second point, recall that $\mathcal{N}_\infty$ is endowed with the topology generated by the sets of the form $\{\pi \in \mathcal{N}_\infty,\ \pi_{|n} = \widehat{\pi}\}$, for $n \in \mathbb{N}$, $\widehat{\pi} \in \mathcal{N}_n$. It is easy to see that this topology is metrized by $d(\pi, \pi') := (\sup\{n \in \mathbb{N},\ \pi_{|n} = \pi'_{|n}\})^{-1}$ (with $(\sup \mathbb{N})^{-1} = 0$) and that $(\mathcal{N}_\infty, d)$ is compact.

We need to show that for any continuous (then bounded) function $f : \mathcal{N}_\infty \to \mathbb{R}$, the function $P_t f : \pi \mapsto \mathbb{E}_\pi f(\mathcal{R}(t))$ (where $\mathbb{E}_\pi(\cdot) = \mathbb{E}(\cdot \,|\, \mathcal{R}(0) = \pi)$) is continuous, and that $P_t f(\pi) \to f(\pi)$ as $t \to 0$. By definition the process is càdlàg so we have almost surely $f(\mathcal{R}(t)) \to f(\mathcal{R}(0))$ so clearly by taking expectations $P_t f(\pi) \to f(\pi)$ as $t \to 0$. Now to show that $P_t f$ is continuous, consider $n \in \mathbb{N}$ and let $\{\widehat{\pi}^1, \dots, \widehat{\pi}^k\}$ be an enumeration of $\mathcal{N}_n$. We pick $\pi^1, \dots, \pi^k \in \mathcal{N}_\infty$ such that $\pi^i_{|n} = \widehat{\pi}^i$, and define $\widehat{f}_n : \mathcal{N}_\infty \to \mathbb{R}$ by

$$\widehat{f}_n(\pi) = f(\pi^i) \quad \text{ if } \pi_{|n} = \widehat{\pi}^i.$$

Now since f is continuous on $(\mathcal{N}_\infty, d)$ which is compact, f is uniformly continuous, which means that

$$\omega_n := \sup_{\pi \in \mathcal{N}_\infty} |f(\pi) - \widehat{f}_n(\pi)| \to 0 \quad \text{as } n \to \infty.$$

For $t > 0$ and $\pi, \pi' \in \mathcal{N}_\infty$, we have

$$|P_t f(\pi) - P_t f(\pi')| \leq \left| \mathbb{E}_\pi \widehat{f}_n(\mathcal{R}(t)) - \mathbb{E}_{\pi'} \widehat{f}_n(\mathcal{R}(t)) \right| + 2\omega_n. \tag{3.4}$$

Now suppose $\pi_{|n} = \pi'_{|n}$. Since $\widehat{f}_n$ depends only on $\pi_{|n}$ and by Lemma 3.3 the process $\mathcal{R}_{|n}$ has the same distribution under $\mathbb{P}_\pi$ or $\mathbb{P}_{\pi'}$, we have the equality $\mathbb{E}_\pi \widehat{f}_n(\mathcal{R}(t)) = \mathbb{E}_{\pi'} \widehat{f}_n(\mathcal{R}(t))$, and plugging that into (3.4), we get

$$\sup\{|P_t f(\pi) - P_t f(\pi')|,\ \pi, \pi' \in \mathcal{N}_\infty,\ \pi_{|n} = \pi'_{|n}\} \leq 2\omega_n \to 0 \quad \text{as } n \to \infty,$$

showing that $P_t f$ is continuous.

For the third point of the proposition, $\mathcal{R}^g(t+t')$ is clearly of the form $\mathrm{Coag}(\mathcal{R}^g(t), \tilde{\pi}^g)$, where $\tilde{\pi} = (\tilde{\pi}^s, \tilde{\pi}^g)$ is a random recipe partition whose distribution depends on $\mathcal{R}(t)$ and

t'. Let us show that the conditional distribution of $\tilde{\pi}$ given $\mathcal{R}(t)$ is invariant under the action of permutations preserving $\bar{\mathcal{R}}(t)$.

Without loss of generality, we can work under the conditioning $\{\mathcal{R}(t) = (\varrho, \mathbf{0}_\infty)\}$, where ϱ is any partition and $\mathbf{0}_\infty$ is the partition into singletons, so that for all π, we have $\mathrm{Coag}(\mathcal{R}^g(t), \pi) = \pi$. In particular, note that in this case we have $\mathcal{R}^g(t + t') = \tilde{\pi}^g$, and $\bar{\mathcal{R}}(t) = \varrho$. Let σ be a permutation such that $\sigma(\varrho) = \varrho$. The problem then reduces to showing that

$$\left(\sigma(\mathcal{R}^g(t + t')) \mid \mathcal{R}(t) = (\varrho, \mathbf{0}_\infty)\right) \overset{(d)}{=} \left(\mathcal{R}^g(t + t') \mid \mathcal{R}(t) = (\varrho, \mathbf{0}_\infty)\right),$$

which is now an immediate consequence of *iii)* in Definition 3.2. □

Let us now investigate the transition rates of the Markov chains $\mathcal{R}_{|n}$ appearing in Lemma 3.3, for every $n \in \mathbb{N}$. In this direction fix $n \in \mathbb{N}$, let $\varrho \in \mathcal{N}_\infty$ and $\pi \in \mathcal{N}_n$ and denote the jump rate of $\mathcal{R}_{|n}$ from $\varrho_{|n}$ to π by

$$q_n(\varrho, \pi) := \lim_{t \to 0+} \frac{1}{t} \mathbb{P}_\varrho(\mathcal{R}_{|n}(t) = \pi) \tag{3.5}$$

where $\mathbb{P}_\varrho(\cdot) = \mathbb{P}(\cdot \,|\, \mathcal{R}(0) = \varrho)$. The index n is not necessary in the notation as it can be read in the partition π. However we keep it as it will ease reading. Remind that $q_n(\varrho, \pi)$ only depends on ϱ through $\varrho_{|n}$. As is remarked in Lemma 3.3, $q_n(\varrho, \pi)$ equals zero if π is not obtained from $\varrho_{|n}$ by coagulating blocks according to a partition in $\widetilde{\mathcal{P}}(\varrho_{|n})$, that is by merging some species blocks of $\varrho^s_{|n}$ into one and some gene blocks of the new species into one. Also observe that the rates do not depend on the sizes of the gene blocks in the starting configuration so there is no loss of generality if we consider that $\varrho^g = \mathbf{0}_\infty$, the trivial partition made of singletons. Of course changing the starting partition ϱ has some effect on the arrival partition π. This is why we will need to write transition rates in another way, giving more emphasis on the dependence of the coagulation mechanism upon the starting partition.

Fix $n \in \bar{\mathbb{N}}$ and suppose that $\mathcal{R}_{|n}$ starts from n singleton gene blocks allocated into b species. Since labels of genes do not affect the transition rates, we will keep the data of the number of genes in each species in a vector $\mathbf{g} = (g_1, \ldots, g_b)$. This vector suffices to describe the starting position. Indeed $|\mathbf{g}| = b$ gives the number of species and $\sum_{i=1}^b g_i = n$ gives the number of genes.

Now the coagulation mechanism will be described by two terms. We will say that a gene block *participates in the coalescence event* if it merges with other gene blocks. We will say that a species block *participates in the coalescence event* if it merges with other species blocks or if it contains gene blocks that participate in the coalescence event.

The behaviour of the species blocks will be encoded in a vector $\mathbf{s} = (s_1, \ldots, s_b)$ with coordinates taking values in $\{0, 1\}$. Namely, $s_i = 1$ if the i-th species participates in the coalescence event and $s_i = 0$ otherwise. The total number of species involved in the event is $k = \sum_{i=1}^b s_i$.

The behaviour of the gene blocks will be encoded by an array $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_b)$ where $\mathbf{c}_i$ is a vector describing which gene blocks in the i-th species participate in the coalescence event.

If $s_i = 1$ (i-th species block participating in the coalescence event), then $\mathbf{c}_i = (c_{i1}, \dots, c_{ig_i})$ is such that $c_{ij} = 1$ if j-th gene block inside i-th species block participates in the coalescence event and $c_{ij} = 0$ otherwise. If $s_i = 0$, the i-th species block is not participating in the event and so neither will the gene blocks within it. In this case we set $\mathbf{c}_i = (0, 0, \dots, 0) = \mathbf{0}$ and nothing happens at the gene level. Note that the number of gene blocks participating in the coalescence event is $\sum_{i,j} c_{ij}$.

Note that all such arrays $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ do not necessarily code for observable coalescence events, so we will define a restricted set of arrays of interest for our study. First, note that one needs to have $\sum_i s_i \geq 2$ in order to observe a species merger. If $\sum_i s_i = 1$, then there is a gene coalescence if and only if $\sum_{i,j} c_{ij} \geq 2$. Also, we will restrict ourselves to the arrays $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ such that $\sum_{i,j} c_{ij} \neq 1$, because *a sole gene coalescing* is not distinguishable from *no gene coalescing*.

Formally, we consider finite arrays $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ satisfying the assumptions

$$\text{If } |\mathbf{g}| = b, \text{ then } \mathbf{s} \in \{0,1\}^b \text{ and } \mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_b),\ \mathbf{c}_i \in \{0,1\}^{g_i} \text{ with } s_i = 0 \implies \forall j, c_{ij} = 0 \tag{\textbf{H1}}$$

$$\sum_{i,j} c_{ij} < 2 \implies \sum_i s_i \geq 2, \tag{\textbf{H2}}$$

$$\text{and } \sum_{i,j} c_{ij} \neq 1. \tag{\textbf{H3}}$$

We denote by $\mathcal{C}$ the set of arrays $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ satisfying (**H1**), (**H2**) and (**H3**).

We then denote the transition rate of $\mathcal{R}_{|n}$ from a partition described by $\mathbf{g}$ (such that $\sum g_i = n$) to a new partition obtained by merging species and genes according to $\mathbf{s}$ and $\mathbf{c}$ by

$$q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c}).$$

Here again indices b and k are not necessary but permit to read easily the coalescence event at the species level ($k = \sum s_i$ species merging among $b = |\mathbf{g}|$). We insist on the fact that we consider only arrays $(\mathbf{g}, \mathbf{s}, \mathbf{c}) \in \mathcal{C}$ when we study the rates $q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c})$, and that these quantities determine uniquely the law of a SNEC $\mathcal{R}$, since they describe completely the rates associated to each finite-space continuous-time Markov chain $\mathcal{R}_{|n}$.

We introduce a notation that we will use in the next result for ease of writing. For μ a probability on $[0,1]$, consider any probability space where $Z_1, Z_2, \dots$ are i.i.d. with distribution μ and denote the expectation $\mathbb{E}_\mu$. Now take a vector $(g_i,\ i \in S)$ of integers, where S is a finite subset of $\mathbb{N}$. We define

$$\begin{aligned} \mathcal{U}(\mu, (g_i, i \in S)) &= \mathbb{E}_\mu\Big[\sum_{i \in S} g_i Z_i (1 - Z_i)^{g_i - 1} \prod_{j \in S\,:\, j \neq i} (1 - Z_j)^{g_j}\Big] \\ &= \sum_{i \in S} g_i \int_{[0,1]} \mu(dq) q (1-q)^{g_i - 1} \prod_{j \in S\,:\, j \neq i} \int_{[0,1]} \mu(dq) (1-q)^{g_j}. \end{aligned} \tag{3.6}$$

This can be thought of as the probability that a random array $(c_{ij},\ i \in S, 1 \leq j \leq g_i)$ does not satisfy (**H3**), where conditional on $(Z_i,\ i \in S)$ the variables (c_{ij}) are independent, and for all i, j, $c_{ij} = 1$ with probability Z_i. We can now state our main result.

Theorem 3.5. *There exist two non-negative real numbers $a_s, a_g \geq 0$ and two measures:*

- *ν_s on $E = (0,1] \times \mathcal{M}_1([0,1])$;*
- *ν_g on $(0,1]$;*

such that

$$\int_E \nu_s(dp, d\mu) p^2 < \infty, \tag{3.7a}$$

$$\int_E \nu_s(dp, d\mu) p \int_{[0,1]} \mu(dq) q^2 < \infty, \tag{3.7b}$$

$$\text{and } \int_{(0,1]} \nu_g(dq) q^2 < \infty, \tag{3.8}$$

and such that for any array $(\mathbf{g}, \mathbf{s}, \mathbf{c}) \in \mathcal{C}$ such that $|\mathbf{g}| = b$, $\sum_i s_i = k$ and $\sum_j c_{ij} = l_i$,

$$\begin{aligned} q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c}) = \int_E \nu_s(dp, d\mu) p^k (1-p)^{b-k} \Bigg(& \prod_{i\,:\,s_i=1} \int_{[0,1]} \mu(dq) q^{l_i} (1-q)^{g_i - l_i} \\ & + \mathbb{1}\{\mathbf{c} = \mathbf{0}\} \mathcal{U}(\mu, (g_i,\ 1 \leq i \leq b \text{ with } s_i = 1)) \Bigg) \\ + a_s \mathbb{1}\{k = 2, \mathbf{c} = \mathbf{0}\} & \\ + \mathbb{1}\{k = 1\} \left(a_g \mathbb{1}\{l_I = 2\} + \int_{(0,1]} \nu_g(dq) q^{l_I} (1-q)^{g_I - l_I} \right), & \end{aligned} \tag{3.9}$$

where the functional $\mathcal{U}$ is defined in (3.6) and $I = I(\mathbf{g}, \mathbf{s}, \mathbf{c})$, in the case $k = 1$, is the unique index in $\{1, 2, \ldots, b\}$ such that $s_I = 1$.

Furthermore, this correspondence between laws of SNEC processes and quadruplets (a_s, a_g, ν_s, ν_g) satisfying (3.7) and (3.8) is bijective.

Remark 3.6. We will show the *surjective* part of the theorem's last statement in Section 3.5, using an explicit Poissonian construction. For now we prove the existence and uniqueness of the characteristics (a_s, a_g, ν_s, ν_g).

3.4 Proof of Theorem 3.5

Consider a SNEC process $\mathcal{R} = ((\mathcal{R}^s(t), \mathcal{R}^g(t)),\ t \geq 0)$ with values in $\mathcal{N}_\infty$ and recall its jump rates $q_n(\varrho, \pi)$ defined in (3.5). Also recall the alternative notation $q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c})$. Here, $\mathbf{g}$ is a vector of size b such that $\sum g_i = n$, $\mathbf{s}$ is a vector having the same size as $\mathbf{g}$ with coordinates in $\{0,1\}$ such that $\sum s_i = k$, and $\mathbf{c}$ is a family of $|\mathbf{g}|$ elements denoted by $\mathbf{c}_1, \mathbf{c}_2 \ldots$ where $\mathbf{c}_i$ is a vector of $\{0,1\}^{g_i}$ if $s_i = 1$ and $\mathbf{c}_i = \mathbf{0}$ if $s_i = 0$.

Lemma 3.7. *For any initial value $\varrho = (\varrho_s, \varrho_g) \in \mathcal{N}_\infty$, there exists a unique measure μ_ϱ on $\mathcal{N}_\infty$ such that*

$$\mu_\varrho(\{\varrho\}) = 0 \quad \text{and} \quad \forall n \geq 1,\ \mu_\varrho(\Pi_{|n} \neq \varrho_{|n}) < \infty \tag{3.10}$$

and such that the transition rate of the Markov chain $\mathcal{R}_{|n}$ from $\varrho_{|n}$ to $\pi \in \mathcal{N}_n$ is given by

$$q_n(\varrho, \pi) = \mu_\varrho(\Pi_{|n} = \pi). \tag{3.11}$$

Furthermore, for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$,

$$\mu_\varrho(\sigma(\Pi) \in \cdot) = \mu_{\sigma(\varrho)}(\Pi \in \cdot). \tag{3.12}$$

Note that we write $\mu_\varrho(\Pi \in A)$ instead of $\mu_\varrho(A)$ because we implicitly work on the canonical space $\mathcal{N}_\infty$ and we denote by Π the generic element of $\mathcal{N}_\infty$.

Proof. Let $n < m$. We first note that since $\mathcal{R}_{|m}$ and $\mathcal{R}_{|n} = (\mathcal{R}_{|m})_{|n}$ are Markov chains, the transition rates can be expressed, for any $\pi \in \mathcal{N}_n \setminus \{\varrho_{|n}\}$,

$$q_n(\varrho, \pi) = \sum_{\pi' \in \mathcal{N}_m \,:\, \pi'_{|n} = \pi} q_m(\varrho, \pi'). \tag{3.13}$$

Let us now check that this consistency property along with Carathéodory's extension theorem ensures us that there exists a measure μ_ϱ on $\mathcal{N}_\infty \setminus \{\varrho\}$ satisfying (3.11).

Here the family $\mathcal{A} := \{\{\Pi_{|n} = \pi\},\ n \in \mathbb{N}, \pi \in \mathcal{N}_n \setminus \{\varrho_{|n}\}\} \cup \{\varnothing\}$ clearly forms a semi-ring of subsets of $\mathcal{N}_\infty$, and it remains to check that the functional $\widetilde{\mu} : \mathcal{A} \to [0, +\infty]$, defined by

$$\widetilde{\mu}(\varnothing) := 0 \quad \text{and} \quad \widetilde{\mu}(\{\Pi_{|n} = \pi\}) := q_n(\varrho, \pi),$$

is a pre-measure. Equation (3.13) shows that $\widetilde{\mu}$ is finitely additive, and the only difficulty lies in understanding that $\widetilde{\mu}$ is countably additive. Now observe that the topology of $\mathcal{N}_\infty \setminus \{\varrho\}$ is generated by $\mathcal{A}$, and that each of the non-empty sets in $\mathcal{A}$ is both open and closed (thus compact), because

$$\mathcal{N}_\infty \setminus \{\Pi_{|n} = \pi\} = \bigcup_{\varrho \in \mathcal{N}_n \setminus \{\pi\}} \{\Pi_{|n} = \varrho\}.$$

This implies that if $(A_n)_{n \geq 1}$ is a family of pairwise disjoint elements of $\mathcal{A}$ such that $\bigcup_n A_n \in \mathcal{A}$, then at most a finite number of the A_n are non-empty (because since $\bigcup_n A_n$ is compact, there is a finite subcover), so countable additivity reduces to finite additivity. Therefore Carathéodory's extension theorem applies, hence the existence of a measure μ_ϱ on $\mathcal{N}_\infty \setminus \{\varrho\}$ satisfying (3.11).

Considering μ_ϱ as a measure on $\mathcal{N}_\infty$ such that $\mu_\varrho(\{\varrho\}) = 0$, we check easily (3.10) by noticing that

$$\mu_\varrho(\Pi_{|n} \neq \varrho_{|n}) = \sum_{\pi \in \mathcal{N}_n \setminus \{\varrho_{|n}\}} q_n(\varrho, \pi) < \infty.$$

Furthermore, for any n, $\pi \in \mathcal{N}_n \setminus \{\varrho_{|n}\}$ and $\sigma : \mathbb{N} \to \mathbb{N}$ permutation, we have by the exchangeability property (3.2) of a SNEC, that

$$\mu_\varrho(\sigma(\Pi)_{|n} = \pi) = \lim_{t \to 0} \frac{1}{t} \mathbb{P}_\varrho(\sigma(\mathcal{R}(t))_{|n} = \pi) = \lim_{t \to 0} \frac{1}{t} \mathbb{P}_{\sigma(\varrho)}(\mathcal{R}_{|n}(t) = \pi) = \mu_{\sigma(\varrho)}(\Pi_{|n} = \pi),$$

which proves that (3.12) holds on $\mathcal{A}$. Since the topology of $\mathcal{N}_\infty$ is generated by $\mathcal{A}$, the proof is complete. □

The latter lemma implies that there exists a family of exchangeable measures on $\mathcal{N}_\infty$ characterizing (i.e. acting as an analog of a Markov kernel for continuous-space pure-jump Markov chains) the SNEC process $\mathcal{R}$. Furthermore, since we are dealing with a simple coalescent, it is clear from the characterization (3.11) that μ_ϱ is simple in the sense that it is supported by all the possible bivariate partitions obtained from a simple coalescence from ϱ. To put it simply,

$$\mu_\varrho\left(\mathcal{N}_\infty \setminus \{\mathrm{Coag}_2(\varrho, \tilde{\pi}),\ \tilde{\pi} \in \widetilde{\mathcal{P}}(\varrho)\}\right) = 0.$$

The measure μ_ϱ can be translated as a measure on arrays of random variables in $\{0,1\}$. Informally, we can associate to each species in ϱ a 1 entry if it participates in the coalescence and a 0 entry otherwise. Inside the species participating to the coalescence event, we can also associate a 1 entry to the genes participating in the coalescence event and a 0 entry otherwise. To tally with the definition of the SNEC we will need a certain partial exchangeability structure for this array. This picture can be formalized as follows. Let $((X_i, (Y_{ij},\ j \in \mathbb{N})),\ i \in \mathbb{N})$ be an array of Bernoulli random variables and denote by Z_i the i-th line vector $(X_i, (Y_{ij},\ j \in \mathbb{N}))$. We say that this array is *hierarchically exchangeable* if

(A1) the family $(Z_i,\ i \in \mathbb{N})$ is exchangeable;

(A2) for any $i \in \mathbb{N}$, the family $(X_i, (Y_{ij},\ j \in \mathbb{N}))$ is invariant under any permutation over the j's.

We also naturally extend this definition to measures on the space $(\{0,1\} \times \{0,1\}^{\mathbb{N}})^{\mathbb{N}}$. We say that such a measure μ is *hierarchically exchangeable* if it is invariant both under the permutations of the rows, and the permutations within a row.

For an initial state $\varrho = (\varrho^s, \varrho^g) \in \mathcal{N}_\infty$ and an arrival state $\pi = \mathrm{Coag}_2(\varrho, \tilde{\pi}) \in \mathcal{N}_\infty$, with $\tilde{\pi}$ a simple bivariate partition $\tilde{\pi} = (\tilde{\pi}^s, \tilde{\pi}^g) \in (\mathcal{P}'_\infty)^2$, define the array $\mathbf{Z}(\varrho, \pi) = (\mathbf{X}, \mathbf{Y}_1, \mathbf{Y}_2, \ldots)$ by

$$\begin{aligned} &X_i = 1 \text{ if the } i\text{-th block in } \varrho^s \text{ has coalesced in } \pi, \\ &Y_{ij} = 1 \text{ if the } I(i,j)\text{-th block in } \varrho^g \text{ has coalesced in } \pi, \end{aligned} \tag{3.14}$$

where $I(i,j) := k$ if the k-th block of ϱ^g is the j-th gene block of the i-th species block.

Now choose a state ϱ with an infinite number of species blocks, each containing an infinite number of gene blocks. Let ν be the push-forward of μ_ϱ by the application

$$\pi \longmapsto \mathbf{Z}(\varrho, \pi).$$

Then the exchangeability property of μ_ϱ (3.12) implies that ν is a hierarchically exchangeable measure on $(\{0,1\} \times \{0,1\}^{\mathbb{N}})^{\mathbb{N}}$, and (3.10) implies that

$$\nu(\mathbf{Z} = \mathbf{0}) = 0, \quad \text{and} \quad \nu\left(\sum_{i=1}^{n} X_i \geq 2 \text{ or } \exists i \leq n, \sum_{j=1}^{n} Y_{ij} \geq 2\right) < \infty, \tag{3.15}$$

where $\mathbf{0}$ denotes the null array on $(\{0,1\} \times \{0,1\}^{\mathbb{N}})^{\mathbb{N}}$. Also, note that the map $(\mu_\varrho,\ \varrho \in \mathcal{N}_\infty) \mapsto \nu$ is one-to-one. Indeed, we can conversely define for any $\mathbf{Z}$ and any nested partition $\varrho \in \mathcal{N}_\infty$, the nested partition $C(\varrho, \mathbf{Z}) \in \mathcal{N}_\infty$ obtained from ϱ by merging exactly the blocks that *participate in the coalescence* where

- The i-th block of ϱ^s participates iff $X_i = 1$;
- The j-th block in ϱ^g of the i-th block of ϱ^s participates iff $X_i = 1$ and $Y_{ij} = 1$.

With this definition, μ_ϱ is obtained as the push-forward of ν by the application $\mathbf{Z} \mapsto C(\varrho, \mathbf{Z})$.

Now recall the alternative notation $q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c})$ for the transition rate of $\mathcal{R}_{|n}$ (where $n = \sum_i g_i$) from a nested partition with b species blocks and $g_1, \dots, g_b$ gene blocks inside them, to a nested partition obtained by merging k species blocks according to the vector $\mathbf{s}$ and gene blocks inside those species according to the array $\mathbf{c}$. For any array $(\mathbf{g}, \mathbf{s}, \mathbf{c}) \in \mathcal{C}$, note that (3.11) translates in terms of our push-forward ν in the following way:

$$\begin{aligned} q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c}) &= \nu(\forall 1 \le i \le b,\, X_i = s_i, \text{ and } \forall 1 \le j \le g_i, Y_{ij} = c_{ij}) \\ &\quad + \mathbb{1}_{\{\mathbf{c}=\mathbf{0}\}} \nu\left(\forall 1 \le i \le b,\, X_i = s_i, \text{ and } \sum_{i=1}^{b} \sum_{j=1}^{g_i} Y_{ij} = 1\right). \end{aligned} \tag{3.16}$$

Indeed, the first line is quite straightforward and comes from our representation of coalescence events by those arrays $(\mathbf{g}, \mathbf{s}, \mathbf{c}) \in \mathcal{C}$ (see Section 3.3) which basically means that blocks participating in a coalescence event are those associated with a 1. However in the case when $\mathbf{c} = \mathbf{0}$, there is an additional probability to observe the coalescence of species blocks associated to $\mathbf{s}$ with no coalescence of gene blocks (the case when all the Y_{ij}'s are 0 is included in the first term), which is when exactly one of the Y_{ij}'s is equal to 1. This gives rise to the second line of (3.16).

We now have to establish a de Finetti representation of hierarchically exchangeable arrays to express the measure of an event of the form $\{\forall 1 \le i \le b,\, X_i = s_i, \text{ and } \forall 1 \le j \le g_i, Y_{ij} = c_{ij}\}$. Note that we consider random measures in the following, but only on Borel spaces $(S, \mathcal{S})$ (i.e. spaces isomorphic to a Borel subset of $\mathbb{R}$ endowed with the Borel σ-algebra), which will enable us to use de Finetti's theorem [55]. For this we write $\mathcal{M}_1(S)$ for the space of probability measures on S, which is endowed with the σ-algebra generated by the maps $\mu \mapsto \mu(B)$ for all $B \in \mathcal{S}$. The spaces $(S, \mathcal{S})$ that we consider will be for instance $[0,1]$ with its Borel sets or $\{0,1\}^{\mathbb{N}}$ equipped with the product σ-algebra, which are clearly Borel spaces.

Proposition 3.8. *Let $\mathbf{Z} = (Z_i,\ i \in \mathbb{N}) = ((X_i, (Y_{ij},\ j \in \mathbb{N})),\ i \in \mathbb{N})$ be a hierarchically exchangeable array (with variables in $\{0,1\}$). Then there exists a unique probability measure Λ on $E' = [0,1] \times \mathcal{M}_1([0,1]) \times \mathcal{M}_1([0,1])$ (and we will write any element μ of E' as (p, μ_0, μ_1)) such that for all $n \ge 1$*

$$\begin{aligned} &\mathbb{P}(X_i = x_i, Y_{ij} = y_{ij},\ i,j \in [n]) \\ &= \int_{E'} \Lambda(d\mu) \prod_{i=1}^{n} \left[(p\mathbb{1}_{\{x_i=1\}} + (1-p)\mathbb{1}_{\{x_i=0\}}) \int_{[0,1]} \mu_{x_i}(dq_i) \prod_{j=1}^{n} (q_i \mathbb{1}_{\{y_{ij}=1\}} + (1-q_i)\mathbb{1}_{\{y_{ij}=0\}}) \right]. \end{aligned} \tag{3.17}$$

Proof. Let us first observe that if a sequence $(X, (Y_j,\ j \in \mathbb{N}))$ satisfies Hypothesis **(A2)**, then, conditional on $X = x \in \{0,1\}$, the sequence $(Y_j,\ j \in \mathbb{N})$ is exchangeable. We

can thus apply de Finetti's theorem: conditional on $X = x$ there is a unique probability measure μ_x giving the distribution of the asymptotic frequency q of the variables $(Y_j,\ j \in \mathbb{N})$, and conditional on q they are i.i.d. Bernoulli with parameter q. This implies that, for any $\{0,1\}$-valued finite sequence $(x, y_1, y_2, \ldots, y_k)$,

$$\mathbb{P}(X = x, Y_1 = y_1, \ldots, Y_k = y_k) = \mathbb{P}(X = x) \int_{[0,1]} \mu_x(dq) \prod_{j=1}^{k} (q\mathbb{1}_{\{y_j=1\}} + (1-q)\mathbb{1}_{\{y_j=0\}}). \tag{3.18}$$

Also observe that since X is binary, there exists $p \in [0,1]$ such that $\mathbb{P}(X = x) = p\mathbb{1}_{\{x=1\}} + (1-p)\mathbb{1}_{\{x=0\}}$.

As a consequence of Hypothesis **(A1)**, we can apply once again de Finetti's theorem: there exists a unique law $\tilde{\Lambda}$ on $\mathcal{M}_1(\{0,1\}^{\mathbb{N}})$ such that the law of $(Z_i,\ i \in \mathbb{N})$ equals $\int_{\mathcal{M}_1(\{0,1\}^{\mathbb{N}})} \tilde{\Lambda}(d\tilde{\mu}) \bigotimes_{i\geq 1} \tilde{\mu}$. Furthermore it has been seen that $\tilde{\mu}$ can be expressed as in (3.18).

Now let F stand for the measurable mapping such that $F(\tilde{\mu}) = (p, \mu_0, \mu_1) \in E'$ and let Λ be the push-forward of $\tilde{\Lambda}$ by the mapping F. We obtain that if A and $(B_i,\ i \in A)$ are finite subsets of $\mathbb{N}$, then

$$\begin{aligned}
&\mathbb{P}(X_i = x_i, Y_{ij_i} = y_{ij_i},\ i \in A, j_i \in B_i) \\
&= \int_{E'} \Lambda(d\mu) \prod_{i\in A} \left[(p\mathbb{1}_{\{x_i=1\}} + (1-p)\mathbb{1}_{\{x_i=0\}}) \int_{[0,1]} \mu_{x_i}(dq_i) \prod_{j_i \in B_i} (q_i \mathbb{1}_{\{y_{ij_i}=1\}} + (1-q_i)\mathbb{1}_{\{y_{ij_i}=0\}}) \right].
\end{aligned}$$

This ends the proof. □

This result is almost enough to express (3.16) but one has to be careful because the measure ν might not be finite. However, it is σ-finite because by (3.15),

$$\nu = \lim_{n\to\infty} \uparrow \nu \left(\left\{ \sum_{i=1}^{n} X_i \geq 2 \text{ or } \exists i \leq n, \sum_{j=1}^{n} Y_{ij} \geq 2 \right\} \cap \cdot \right),$$

and those events have finite measure. The idea behind the following lemma is to make use of those events and hierarchical exchangeability to express ν as a limit of finite measures which, thanks to an application of Proposition 3.8, have a representation under the form (3.17).

Let us introduce some notation that will enable us to make this argument formal. For a fixed vector $(\mathbf{g}, \mathbf{s}, \mathbf{c}) \in \mathcal{C}$, such that $|\mathbf{g}| = b$, let us examine the event

$$A = A(\mathbf{g}, \mathbf{s}, \mathbf{c}) := \{\forall 1 \leq i \leq b,\ X_i = s_i,\ \text{and } \forall 1 \leq j \leq g_i, Y_{ij} = c_{ij}\}$$

and its measure $\nu(A)$. Let us define, for all $n \geq 1$ the shifted random array

$$\mathbf{Z}_n := (X_{i+n}, Y_{(i+n)j},\ i, j \in \mathbb{N}). \tag{3.19}$$

We decompose naturally $A = (A \cap \{\mathbf{Z}_b \neq \mathbf{0}\}) \cup (A \cap \{\mathbf{Z}_b = \mathbf{0}\})$, where $b = |\mathbf{g}|$.

Recall that the array $\mathbf{Z}$ encodes which species blocks and which gene blocks are participating in a coalescence. Therefore the event $A \cap \{\mathbf{Z}_b \neq 0\}$ indicates that there are merging

species blocks outside of the first b blocks. In fact we will see that this implies that such merging blocks are infinitely many (a random proportion p of them), and within each of these blocks, a random proportion q of gene blocks are also participating in the coalescence event. The following technical lemma makes this statement rigorous.

Lemma 3.9. *For an array $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ satisfying assumptions (**H1**) and (**H2**), there exists a unique measure ν_s on $E = (0,1] \times \mathcal{M}([0,1])$ such that*

$$\nu(A \cap \{\mathbf{Z}_b \neq \mathbf{0}\}) = \int_E \nu_s(dp, d\mu) p^k (1-p)^{b-k} \prod_{i\,:\,s_i=1} \int_{[0,1]} \mu(dq) q^{l_i} (1-q)^{g_i - l_i}, \tag{3.20}$$

where $b := |\mathbf{g}|$, $k := \sum_i s_i$ and $l_i := \sum_j c_{ij}$. Moreover, ν_s satisfies (3.7).

Proof. We define some events that will be used to express $\nu(A \cap \{\mathbf{Z}_b \neq \mathbf{0}\})$.

$$A_n = A_n(\mathbf{g}, \mathbf{s}, \mathbf{c}) := \{\forall 1 \leq i \leq b,\, X_{i+n} = s_i, \text{ and } \forall 1 \leq j \leq g_i, Y_{(i+n)j} = c_{ij}\}$$

$$B_n := \left\{ \sum_{i=1}^{n} X_i \geq 2 \right\}$$

$$B'_n = B'_n(\mathbf{g}, \mathbf{s}, \mathbf{c}) := \left\{ \sum_{i=b+1}^{b+n} X_i \geq 2 \right\}.$$

Note that $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ satisfies (**H1**) and (**H2**), so we have $A \subset \{\sum_{i=1}^m X_i \geq 2 \text{ or } \sum_{i,j=1}^m Y_{ij} \geq 2\}$ for $m = \max(b, g_1, \dots g_b)$. Now because ν satisfies (3.15), necessarily $\nu(A) < \infty$, which implies that

$$\nu(A \cap \{\mathbf{Z}_b \in \cdot\})$$

is a finite hierarchically exchangeable measure on $(\{0,1\} \times \{0,1\}^{\mathbb{N}})^{\mathbb{N}}$. The de Finetti representation (Proposition 3.8) implies that on the event A, $\mathbf{Z}_b$ is either $\mathbf{0}$, or has an infinite number of entries with value 1. In particular, $A \cap \{\mathbf{Z}_b \neq \mathbf{0}\} = A \cap \{\mathbf{Z}_b \text{ has at least two entries at } 1\}$ therefore, there is the equality

$$A \cap \{\mathbf{Z}_b \neq \mathbf{0}\} = \bigcup_{n \geq 1} A \cap B'_n,$$

where the union is increasing. Therefore,

$$\begin{aligned} \nu(A \cap \{\mathbf{Z}_b \neq \mathbf{0}\}) &= \lim_{n \to \infty} \nu(A \cap B'_n). \\ &= \lim_{n \to \infty} \nu(B_n \cap A_n), \end{aligned}$$

where we used the hierarchical exchangeability of ν to get the second equality. Now we know from (3.15) and because ν is exchangeable that the measure

$$\nu(B_n \cap \{\mathbf{Z}_n \in \cdot\})$$

is a finite hierarchically exchangeable measure on $(\{0,1\} \times \{0,1\}^{\mathbb{N}})^{\mathbb{N}}$. Because it is finite we can apply Proposition 3.8 to deduce that there exists a finite measure Λ_n on $E' = (0,1] \times \mathcal{M}([0,1])^2$ such that

$$\begin{aligned} &\nu(B_n \cap A_n) \\ &= \int_{E'} \Lambda_n(dp, d\mu_0, d\mu_1) \prod_{i=1}^{b} \left[(p \mathbb{1}_{\{s_i=1\}} + (1-p)\mathbb{1}_{\{s_i=0\}}) \int_{[0,1]} \mu_{s_i}(dq_i) \prod_{j=1}^{g_i} (q_i \mathbb{1}_{\{c_{ij}=1\}} + (1-q_i)\mathbb{1}_{\{c_{ij}=0\}}) \right]. \end{aligned}$$

We can simplify this expression since ν is supported by the set $\{\forall i \in \mathbb{N}, X_i = 0 \Rightarrow \forall j \in \mathbb{N}, Y_{ij} = 0\}$. This implies that Λ_n-a.e. the measure μ_0 is δ_0 the Dirac measure at 0. Therefore we write $\widetilde{\Lambda}_n$ for the push forward measure on $E := (0,1] \times \mathcal{M}([0,1])$ of Λ_n by the application $(p, \mu_0, \mu_1) \mapsto (p, \mu_1)$. We now have

$$\nu(B_n \cap A_n) = \int_E \widetilde{\Lambda}_n(dp, d\mu) p^k (1-p)^{b-k} \prod_{i\,:\,s_i=1} \int_{[0,1]} \mu(dq) q^{l_i} (1-q)^{g_i - l_i}. \tag{3.21}$$

To be able to pass to the limit, let us check that the sequence of measures $(\widetilde{\Lambda}_n)$ is increasing. Indeed, recall that Λ_n is obtained from two applications of de Finetti's theorem to the exchangeable array $\mathbf{Z}_n$, so the asymptotic parameters p and μ appearing in (3.21) are a deterministic, measurable functional of $\mathbf{Z}_n$. Let us write this functional $F(\mathbf{Z}_n) = (p, \mu)$, so now $\widetilde{\Lambda}_n$ is simply the measure

$$\nu(B_n \cap \{F(\mathbf{Z}_n) \in \cdot\}).$$

But p and μ are asymptotic quantities of the array $\mathbf{Z}_n$, which do not depend on the first row of $\mathbf{Z}_n$, so $F(\mathbf{Z}_{n+1}) = F(\mathbf{Z}_n)$ and we have

$$\begin{aligned}
\widetilde{\Lambda}_n &= \nu(B_n \cap \{F(\mathbf{Z}_n) \in \cdot\}) \\
&= \nu(B_n \cap \{F(\mathbf{Z}_{n+1}) \in \cdot\}) \\
&\leq \nu(B_{n+1} \cap \{F(\mathbf{Z}_{n+1}) \in \cdot\}) \\
&= \widetilde{\Lambda}_{n+1},
\end{aligned}$$

where the passage from the second to the third line is simply because $B_n \subset B_{n+1}$. Therefore there is a limiting measure ν_s on E such that

$$\nu(A \cap \{\mathbf{Z}_b \neq \mathbf{0}\}) = \lim_{n\to\infty} \nu(B_n \cap A_n) = \int_E \nu_s(dp, d\mu) p^k (1-p)^{b-k} \prod_{i\,:\,s_i=1} \int_{[0,1]} \mu(dq) q^{l_i} (1-q)^{g_i - l_i},$$

so we recover (3.20). To prove the uniqueness of this measure, consider any measure ν_s' on E such that (3.20) holds. Then we have simply

$$\tilde{\Lambda}_n(dp, d\mu) = \nu(B_n \cap \{F(\mathbf{Z}_n) \in (dp, d\mu)\}) = (1 - (1-p)^n - np(1-p)^{n-1}) \nu_s'(dp, d\mu),$$

where the first equality is by definition and the second because we assumed that (3.20) holds for ν_s'. Taking limits on both sides yields

$$\nu_s(dp, d\mu) = \nu_s'(dp, d\mu).$$

It remains to prove (3.7). Note that the condition (3.15) implies that

$$\nu(X_1 = X_2 = 1) < \infty \quad \text{and} \quad \nu(X_1 = 1, Y_{1,1} = Y_{1,2} = 1) < \infty.$$

Translating these conditions with the formula (3.20), we recover exactly (3.7). □

Let us now examine $\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\})$. Recall that the event $A \cap \{\mathbf{Z}_b = 0\}$ indicates that there are no other merging species blocks than those within the first b blocks. The next lemma shows that this implies that we are either in a Kingman-type coalescence (a pair

of species blocks are merging, occurring at rate a_s, or a pair of gene blocks within one species are merging, occurring at rate a_g), or in a multiple gene coalescence within a single species block (in which case a random proportion q of gene blocks are merging).

The key idea is to use exchangeability and the σ-finiteness property (3.15) of the measure ν to show by contradiction that $\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\})$ is zero in certain cases.

Lemma 3.10. *For an array $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ satisfying assumptions (H1) and (H2), there exist unique real numbers $a_s, a_g \geq 0$ and a unique measure ν_g on $(0,1]$ satisfying (3.8) such that*

$$\begin{aligned}\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\}) &= a_s \mathbb{1}\{k = 2, \mathbf{c} = \mathbf{0}\} \\ &\quad + \mathbb{1}\{k = 1\} \left(a_g \mathbb{1}\{l_I = 2\} + \int_{(0,1]} \nu_g(dq) q^{l_I} (1-q)^{g_I - l_I} \right),\end{aligned} \tag{3.22}$$

where $b := |\mathbf{g}|$, $k := \sum_i s_i$, $l_i := \sum_j c_{ij}$ and in the case $k = 1$, I is the unique index in $\{1, 2, \ldots, b\}$ such that $s_I = 1$.

Proof. The measure $\nu(\mathbf{X} \in \cdot)$ is an exchangeable measure on $\{0,1\}^{\mathbb{N}}$ such that, because of (3.15), $\nu(X_1 = X_2 = X_3 = 1) < \infty$. Note that exchangeability implies that for any $n, i \geq 3$,

$$\nu(\{X_1 = X_2 = X_3 = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) = \nu(\{X_1 = X_2 = X_i = 1\} \cap \{\mathbf{Z}_i = \mathbf{0}\}), \tag{3.23}$$

But the events $(\{X_1 = X_2 = X_i = 1\} \cap \{\mathbf{Z}_i = \mathbf{0}\}, i \geq 3)$ are disjoint and all included in $\{X_1 = X_2 = 1\}$, so that

$$\sum_{i \geq 3} \nu(\{X_1 = X_2 = X_i = 1\} \cap \{\mathbf{Z}_i = \mathbf{0}\}) \leq \nu(\{X_1 = X_2 = 1\}) < \infty.$$

From (3.23) we deduce $\nu(\{X_1 = X_2 = X_3 = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) = 0$. This implies immediately that for a finite array $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ such that $k = \sum_i s_i > 2$, we have $\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\}) = 0$.

- In the case $k = 2$ (suppose $s_1 = s_2 = 1$), one must examine several cases.
  - Suppose first $c_{1,1} = c_{1,2} = 1$. This means that the first two gene blocks of the first species block coalesce while the first two species blocks coalesce. Then we note that for any $n, i \geq 2$,

    $$\begin{aligned}&\nu(\{X_1 = X_2 = 1,\, Y_{1,1} = Y_{1,2} = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) \\ &= \nu(\{X_1 = X_i = 1,\, Y_{1,1} = Y_{1,2} = 1\} \cap \{\mathbf{Z}_i = \mathbf{0}\}).\end{aligned}$$

    However,

    $$\sum_{i \geq 2} \nu(\{X_1 = X_i = 1,\, Y_{1,1} = Y_{1,2} = 1\} \cap \{\mathbf{Z}_i = \mathbf{0}\}) \leq \nu(\{Y_{1,1} = Y_{1,2} = 1\}) < \infty,$$

    so that necessarily $\nu(\{X_1 = X_2 = 1,\, Y_{1,1} = Y_{1,2} = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) = 0$. So in the case $c_{1,1} = c_{1,2} = 1$, we have $\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\}) = 0$.

– Now suppose $c_{1,1} = c_{2,1} = 1$, and all the other c_{ij} are zero. From our previous point, note that

$$\nu(\{X_1 = X_2 = 1,\, Y_{1,1} = Y_{2,1} = 1, \text{ and } \exists j \geq 2, Y_{1,j} = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) = 0,$$

which implies that the events $(\{X_1 = X_2 = 1,\, Y_{1,j} = Y_{2,1} = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\},\, j \geq 1)$ are ν-a.e. disjoint. Therefore for any $n \geq 2$,

$$\sum_{j\geq 1} \nu(\{X_1 = X_2 = 1,\, Y_{1,j} = Y_{2,1} = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) \leq \nu(\{X_1 = X_2 = 1\}) < \infty,$$

So necessarily $\nu(\{X_1 = X_2 = 1,\, Y_{1,j} = Y_{2,1} = 1\} \cap \{\mathbf{Z}_n = \mathbf{0}\}) = 0$. This implies that in the case $c_{1,1} = c_{2,1} = 1$, we have $\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\}) = 0$.

– The previous two points show that in the case $k = 2$, the only way to have $\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\}) > 0$ is if $\mathbf{c} = \mathbf{0}$. In that case, define

$$\begin{aligned} a_s &:= \nu(\{X_1 = X_2 = 1\} \cap \{\mathbf{Z}_2 = \mathbf{0}\}) \\ &= \nu(\{X_1 = X_2 = 1\} \cap \{\mathbf{Y} = \mathbf{0} \text{ and } \forall k \notin \{1,2\},\, X_k = 0\}). \end{aligned}$$

Then by exchangeability, for all $i, j \in \mathbb{N}$ with $i \neq j$, we have

$$a_s = \nu(\{X_i = X_j = 1\} \cap \{\mathbf{Y} = \mathbf{0} \text{ and } \forall k \notin \{i,j\},\, X_k = 0\}),$$

and in conclusion, for any array $(\mathbf{g}, \mathbf{s}, \mathbf{c})$ such that $k = 2$, we have

$$\nu(A \cap \{\mathbf{Z}_b = \mathbf{0}\}) = \mathbb{1}_{\{\mathbf{c}=\mathbf{0}\}} a_s.$$

- In the case $k = 1$, suppose that $s_1 = 1$. On the event

$$\{X_1 = 1,\, X_2 = X_3 = \cdots = X_b = 0\} \cap \{\mathbf{Z}_b = \mathbf{0}\},$$

we have simply $\mathbf{Z}_1 = \mathbf{0}$, and then the measure

$$\nu' := \nu(\{(Y_{1,j})_{j\in\mathbb{N}} \in \cdot\} \cap \{X_1 = 1,\, \mathbf{Z}_1 = \mathbf{0}\})$$

is an exchangeable measure on $\{0,1\}^{\mathbb{N}}$ such that for all $n \in \mathbb{N}$, $\nu'\left(\sum_{j=1}^n Y_j \geq 2\right) < \infty$. Therefore (see for instance Bertoin [10]) there exist a unique constant $a_g \geq 0$ and ν_g a unique measure on $(0,1]$ satisfying (3.8) such that ν' can be written

$$\nu'(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n) = a_g \mathbb{1}_{l=2} + \int_{(0,1]} \nu_g(dq) q^l (1-q)^{n-l},$$

for any vector $(y_1, y_2, \ldots, y_n) \in \{0,1\}^n \setminus \{\mathbf{0}\}$ such that $l := \sum_i y_i \geq 2$.

Putting all the previous considerations together yields (3.22). □

Now it remains to put together Lemma 3.9 and Lemma 3.10. Recall that we restricted the rate function q to arrays in $\mathcal{C}$, i.e. satisfying (**H1**) to (**H3**). The reason for assuming (**H3**) is that then we can always write $q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c})$ as in (3.16), that is

$$\begin{aligned} q_{b,k}(\mathbf{g}, \mathbf{s}, \mathbf{c}) &= \nu(\forall 1 \leq i \leq b,\, X_i = s_i, \text{ and } \forall 1 \leq j \leq g_i, Y_{ij} = c_{ij}) \\ &\quad + \mathbb{1}_{\{\mathbf{c}=\mathbf{0}\}} \nu\left(\forall 1 \leq i \leq b,\, X_i = s_i, \text{ and } \sum_{i=1}^b \sum_{j=1}^{g_i} Y_{ij} = 1\right). \end{aligned}$$

Using the previous two lemmas to decompose the two lines on the events $\{\mathbf{Z}_b \neq \mathbf{0}\}$ and $\{\mathbf{Z}_b = \mathbf{0}\}$, we obtain the formula (3.9), concluding the proof of Theorem 3.5.

3.5 Poissonian construction

The goal of the present section is to show how any simple nested exchangeable coalescent can be constructed from a Poisson point process. Consider two real coefficients $a_s, a_g \geq 0$ and two measures: ν_s on $E = (0,1] \times \mathcal{M}_1([0,1])$ satisfying (3.7), and ν_g on $(0,1]$, satisfying (3.8). Recall the measures $\mathtt{K}_s, \mathtt{K}_g, P_x^g$ and $P_{x,\mu}^s$ introduced in Section 3.2, and the measure $\nu(d\mathbf{Z})$ defined on the space $\widehat{E}$ of doubly indexed arrays of 0's and 1's $\mathbf{Z} = (\mathbf{X}, (\mathbf{Y}_i,\, i \geq 1)) = (X_i, Y_{ij},\, i,j \geq 1)$

$$\nu := a_s \mathtt{K}_s + a_g \mathtt{K}_g + \int_{(0,1]} \nu_g(dx)\, P_x^g + \int_{(0,1]\times\mathcal{M}_1([0,1])} \nu_s(dx, d\mu)\, P_{x,\mu}^s.$$

Note that ν characterizes the distribution of the SNEC through the relation (3.16). The key idea of the construction is that ν necessarily satisfies (3.15), which is easily shown using exchangeability and conditions (3.7) and (3.8). First, note that $\nu(\mathbf{Z} = \mathbf{0}) = 0$ is trivial from our definitions, and that a straightforward union bound yields

$$\begin{aligned}
&\nu\Big(\textstyle\sum_{i=1}^n X_i \geq 2 \text{ or } \exists i \leq n, \sum_{j=1}^n Y_{ij} \geq 2\Big) \\
&\qquad \leq \sum_{1\leq i<i'\leq n} \nu(X_i = X_{i'} = 1) + \sum_{i=1}^n \sum_{1\leq j<j'\leq n} \nu(X_i = Y_{ij} = Y_{ij'} = 1) \\
&\qquad = \frac{n(n-1)}{2}\nu(X_1 = X_2 = 1) + \frac{n^2(n-1)}{2}\nu(X_1 = Y_{1,1} = Y_{1,2} = 1),
\end{aligned}$$

therefore we need only check that these two quantities are finite. Now by definition, we have

$$\begin{aligned}
\mathtt{K}_s(X_1 = X_2 = 1) = 1, &\quad \mathtt{K}_s(X_1 = Y_{1,1} = Y_{1,2} = 1) = 0, \\
\mathtt{K}_g(X_1 = X_2 = 1) = 0, &\quad \mathtt{K}_g(X_1 = Y_{1,1} = Y_{1,2} = 1) = 1, \\
P_x^g(X_1 = X_2 = 1) = 0, &\quad P_x^g(X_1 = Y_{1,1} = Y_{1,2} = 1) = x^2, \\
P_{x,\mu}^s(X_1 = X_2 = 1) = x^2, &\quad P_{x,\mu}^s(X_1 = Y_{1,1} = Y_{1,2} = 1) = x\int_{[0,1]} \mu(dq) q^2.
\end{aligned}$$

Integrating the last two lines with respect to ν_g and ν_s and summing, we see that (3.7) and (3.8) imply that both $\nu(X_1 = X_2 = 1)$ and $\nu(X_1 = Y_{1,1} = Y_{1,2} = 1)$ are finite, proving (3.15).

Now to start the construction of our process, consider an initial partition $\pi_0 \in \mathcal{N}_\infty$. Let M be a Poisson point process on $(0,\infty) \times \widehat{E}$ with intensity $dt \otimes \nu(d\mathbf{Z})$. We will construct on the same probability space the processes $\mathcal{R}^n = (\mathcal{R}^n(t),\, t \geq 0)$, for $n \in \mathbb{N}$ thanks to M.

Recall that for any $\mathbf{Z} = (\mathbf{X}, (\mathbf{Y}_i,\, i \geq 1)) = (X_i, Y_{ij},\, i,j \geq 1)$ and any nested partition $\pi \in \mathcal{N}_n$, we denote by $C(\pi, \mathbf{Z})$ the nested partition of $\mathcal{N}_n$ obtained from π by merging exactly the blocks that *participate in the coalescence* where

- The i-th block of π^s participates iff $X_i = 1$;
- The j-th block in π^g of the i-th block of π^s participates iff $X_i = 1$ and $Y_{ij} = 1$.

Fix $n \in \mathbb{N}$, and let M_n denote the subset of M consisting of points $(t, \mathbf{Z})$ such that $\sum_{i=1}^n X_i \geq 2$ or $\exists i \leq n$, $\sum_{j=1}^n Y_{ij} \geq 2$. Because of (3.15), there are only a finite number of

such points with t in a compact set of $[0, +\infty)$. Therefore one can label the atoms of the set $M_n := \{(t_k, \mathbf{Z}^{(k)}), k \in \mathbb{N}\}$ in increasing order, i.e. such that $0 \leq t_1 \leq t_2 \dots$

We set $\mathcal{R}^n(t) = (\pi_0)_{|n}$ for $t \in [0, t_1)$. Then define recursively

$$\mathcal{R}^n(t) = C(\mathcal{R}^n(t_i-), \mathbf{Z}^{(i)}), \quad \text{for every } t \in [t_i, t_{i+1}).$$

These processes are consistent in n as we show in the following result.

Proposition 3.11. *For every $t \geq 0$, the sequence of random bivariate partitions $(\mathcal{R}^n(t), n \in \mathbb{N})$ is consistent. If we denote by $\mathcal{R}(t)$ the unique partition of $\mathcal{N}_\infty$ such that $\mathcal{R}_{|n}(t) = \mathcal{R}^n(t)$ for every $n \in \mathbb{N}$, then the process $\mathcal{R} = (\mathcal{R}(t), t \geq 0)$ is a SNEC started from π_0, with rates given as in Theorem 3.5.*

The proof uses similar arguments as in the proof of consistency of exchangeable coalescents given in Proposition 4.5 of [10].

Proof. The key idea (basically (4.4) in [10]) is that by definition, the coagulation operator satisfies

$$\mathrm{Coag}_2(\pi, \tilde{\pi})_{|n} = \mathrm{Coag}_2(\pi_{|n}, \tilde{\pi}) = \mathrm{Coag}_2(\pi_{|n}, \tilde{\pi}_{|n}) \tag{3.24}$$

for any $\pi, \tilde{\pi}$ and n for which this is well defined.

Recall that we defined M_n as the subset of M consisting of points $(t, \mathbf{Z})$ such that $\sum_{i=1}^n X_i \geq 2$ or $\exists i \leq n$, $X_i \sum_{j=1}^n Y_{ij} \geq 2$. Fix $n \geq 2$ and write $(t_1, \mathbf{Z}^{(1)})$ for the first atom of M_n on $(0, \infty) \times \widehat{E}$. Plainly, $\mathcal{R}^{n-1}(t) = \mathcal{R}^n_{|n-1}(t) = (\pi_0)_{|n-1}$ for every $t \in [0, t_1)$.

Consider first the case when $\sum_{i=1}^{n-1} X_i^{(1)} \geq 2$ or $\exists i \leq n-1$, $X_i^{(1)} \sum_{j=1}^{n-1} Y_{ij}^{(1)} \geq 2$. Then $(t_1, \mathbf{Z}^{(1)})$ is also the first atom of M_{n-1} and by definition and using (3.24), $\mathcal{R}^{n-1}(t_1) = \mathcal{R}^n_{|n-1}(t)$.

Now suppose $\sum_{i=1}^{n-1} X_i^{(1)} \leq 1$ and $\forall i \leq n-1$, $X_i^{(1)} \sum_{j=1}^{n-1} Y_{ij}^{(1)} \leq 1$. This implies that at time t_1, there is no species (resp. genes) coalescence between the $n-1$ first species (resp. genes) of $\mathcal{R}^n(t_1-)$. Therefore the coalescence event in $\mathcal{R}^n$ at time t_1 leaves the first $n-1$ blocks of $\mathcal{R}^n(t_1-)^s$ or $\mathcal{R}^n(t_1-)^g$ unchanged, though there may be a coalescence involving the n-th block (in that case, necessarily a singleton $\{n\}$) and one of the $n-1$ first blocks. So finally $\mathcal{R}^n(t_1)_{|n-1} = \mathcal{R}^n(t_1-)_{|n-1} = \mathcal{R}^{n-1}(t_1)$.

In both cases we have $\mathcal{R}^n(t_1)_{|n-1} = \mathcal{R}^{n-1}(t_1)$, and by an obvious induction this is true for any further jump of the process $\mathcal{R}^n$, so that for all $t \geq 0$,

$$\mathcal{R}^n(t)_{|n-1} = \mathcal{R}^{n-1}(t).$$

This shows the existence of $\mathcal{R}$ such that for all n, $\mathcal{R}_{|n} = \mathcal{R}^n$.

From this Poissonian construction $\mathcal{R}^n$ is a Markov process, and by definition the arrays $\mathbf{Z}^{(i)}_{|[n]^2}$ are hierarchically exchangeable, which implies that $\mathcal{R}^n$ is an exchangeable process. Clearly by construction $\mathcal{R}^n(t)$ is nested for all t, and the only jumps of the process $\mathcal{R}^n$ are coalescence events. According to Lemma 3.3, the process $\mathcal{R}$ is a SNEC process. Because the arrays $\mathbf{Z}$, where $(t, \mathbf{Z}) \in M$, are the same arrays that appear in the proof of Theorem 3.5, it is clear that the jump rates of $\mathcal{R}^n$ are those given in Theorem 3.5. □

3.6 Marginal coalescents – Coming down from infinity

Consider a SNEC process $\mathcal{R} = (\mathcal{R}^s, \mathcal{R}^g)$, with rates given as in Theorem 3.5 by two coefficients $a_s, a_g \geq 0$ and two measures, ν_s on $E = (0,1] \times \mathcal{M}_1([0,1])$ and ν_g on $(0,1]$ satisfying (3.7) and (3.8).

It is obvious from Proposition 3.11 that $(\mathcal{R}^s(t),\ t \geq 0)$ is a simple coalescent process, with Kingman coefficient a_s and coagulation measure $\widehat{\nu}_s$ satisfying (3.1) which is the push-forward of $\nu_s(dp, d\mu)$ by the application $(p, \mu) \mapsto p$. Let us call this univariate coalescent the *(marginal) species coalescent* of the SNEC process $\mathcal{R}$.

Now, notice that under an initial condition with a unique species block (i.e., $\mathcal{R}^s$ is constant to the coarsest partition $\mathbf{1}_\infty$), the process $(\mathcal{R}^g(t),\ t \geq 0)$ also behaves as a simple coalescent process, with Kingman coefficient a_g and coagulation measure $\widehat{\nu}_g$ defined by

$$\forall B \text{ Borel set of } (0,1], \quad \widehat{\nu}_g(B) := \nu_g(B) + \int_{(0,1]\times\mathcal{M}_1([0,1])} \nu_s(dp, d\mu) p\, \mu(B).$$

We call the simple coalescent thus defined the *(marginal) gene coalescent* of the SNEC process $\mathcal{R}$.

Equivalently, in terms of Λ-coalescents, the marginal species coalescent is a Λ_s-coalescent with Λ_s defined by

$$\forall B \text{ Borel set of } [0,1], \quad \Lambda_s(B) = a_s \delta_0(B) + \int_{B\times\mathcal{M}_1([0,1])} \nu_s(dp, d\mu) p^2, \tag{3.25}$$

and the marginal gene coalescent is a Λ_g-coalescent with Λ_g defined for all B Borel set of $[0,1]$ by

$$\Lambda_g(B) = a_g \delta_0(B) + \int_B \nu_g(dq) q^2 + \int_{(0,1]\times\mathcal{M}_1([0,1])} \nu_s(dp, d\mu) p \int_B \mu(dq) q^2. \tag{3.26}$$

These two marginal processes allow us to express properties of the initial bivariate SNEC process. Consider an initial state $\varrho_0 \in \mathcal{N}_\infty$ with infinitely many species blocks, each containing infinitely many gene blocks. In a way analogous to the one-dimensional case, recalling that $|\mathcal{R}^g(t)| \geq |\mathcal{R}^s(t)|$ for all $t \geq 0$, we will say that a SNEC *comes down from infinity* (CDI) if for all $t > 0$

$$|\mathcal{R}^g(t)| < \infty \qquad \mathbb{P}_{\varrho_0}\text{-a.s.}$$

In the univariate case, characterizing which coalescent processes come down from infinity has been solved [81] for Λ-coalescents, with the following necessary and sufficient condition for coming down from infinity:

$$\sum_{n\geq 2} \left(\sum_{k=2}^{n} (k-1) \binom{n}{k} \int_{[0,1]} \Lambda(dp) p^{k-2} (1-p)^{n-k} \right)^{-1} < \infty.$$

Note that the previous condition is true as soon as Λ has an atom at 0 ($\Lambda(\{0\})$ is the Kingman coefficient of the process). An equivalent criterion (see [16], and [7] for a probabilistic proof) is the integrability of $1/\psi$ near $+\infty$, where

$$\psi(q) := \int_{[0,1]} \left(e^{-qx} - 1 + qx\right) x^{-2}\, \Lambda(dx). \tag{3.27}$$

We will now see that in the case of simple nested coalescents, we can give a general characterization of the different CDI properties of a SNEC process, depending only on the properties of the marginal species and marginal gene coalescents.

First notice that if the marginal gene coalescent does not CDI, then any species block with infinitely many gene blocks at some time t clearly keeps infinitely many gene blocks for any $t' \geq t$. Also in any case the process $\mathcal{R}^s$ has the distribution of the marginal species coalescent, so determining whether the number of species comes down from infinity is trivial.

Proposition 3.12. *We assume here that $\widehat{\nu}_s(\{1\}) = \widehat{\nu}_g(\{1\}) = 0$ and that the marginal gene coalescent comes down from infinity (CDI). Then we have the following three cases.*

i) If the marginal species coalescent CDI as well, then $\mathcal{R}$ CDI.

ii) If the marginal species coalescent does not CDI but $\int_{[0,1]} \widehat{\nu}_s(dx)\, x = \infty$, then for any initial condition with infinitely many species blocks and for each time $t > 0$, the number of gene blocks in each species block of $\mathcal{R}(t)$ is infinite a.s.

iii) If the marginal species coalescent does not CDI and $\int_{[0,1]} \widehat{\nu}_s(dx)\, x < \infty$, then for any initial condition and for each time $t > 0$, the number of gene blocks in each species block of $\mathcal{R}(t)$ is finite a.s.

As a consequence of this proposition, it is clear that $\mathcal{R}$ comes down from infinity if and only if both the marginal species coalescent and the marginal gene coalescent come down from infinity.

A simple example of a SNEC process coming down from infinity is the nested Kingman coalescent ('Kingman in Kingman'), given by its marginal rates $a_s, a_g > 0$, defined so that each pair of species coalesces at rate a_s independently of the others, and each pair of genes within the same species coalesces at rate a_g independently of the rest. Since the marginal coalescents are precisely two Kingman coalescents, they both come down from infinity.

Note that the Bolthausen-Sznitman coalescent [20] (denoted U-coalescent in [76] because the measure Λ is uniform on $[0,1]$) satisfies the conditions of the peculiar case *ii)*. So for a SNEC $\mathcal{R}$ defined by a Kingman gene coalescent evolving within a species U-coalescent, at each positive time the number of gene blocks within a species block is infinite (if the initial state ϱ_0 has an infinite number of species blocks).

Case *iii)* can easily be obtained by considering a "slow" species coalescent, such as a δ_x-coalescent for $x \in (0,1)$, or any $\beta(a,b)$-coalescent with $a > 1, b > 0$ (that is a Λ-coalescent with $\Lambda(dx) = c_{a,b} x^{a-1}(1-x)^{b-1} dx$).

Proof. ***i)*** Suppose both marginal coalescents come down from infinity, and consider an initial state $\varrho \in \mathcal{N}_\infty$ with infinitely many species blocks, each containing infinitely many gene blocks.

Choose $t > 0$. Since $\mathcal{R}^s$ comes down from infinity, we have $\mathbb{P}_\varrho(|\mathcal{R}^s(t/2)| < \infty) = 1$, and necessarily, $\mathcal{R}^s$ stays constant on an interval $[t/2, T[$, where T is its next jump time. Now

on the interval $[t/2, \min(T,t)[$, within each of the $|\mathcal{R}^s(t/2)|$ species block, the gene blocks undergo independent coalescent processes which CDI, therefore there are finitely many gene blocks in each species at time $\min(T,t)$, which implies

$$\mathbb{P}_\varrho(|\mathcal{R}^g(t)| < \infty) = 1.$$

Let us say a few words before proving *ii)* and *iii)*. Pick $t > 0$ and focus on the species containing 1 (the first species). To this aim, write $M(t)$ for the number of genes within the first species, at time t. By exchangeability, to show *ii)* it is sufficient to show $\mathbb{P}_\varrho(M(t) = \infty) = 1$, for any initial condition ϱ with infinitely many species blocks, and to show *iii)* it is sufficient to show $\mathbb{P}_\varrho(M(t) < \infty) = 1$, for any initial condition ϱ.

ii) Suppose $\int_{[0,1]} \widehat{\nu}_s(dx)\, x = \infty$. First, note that since the species coalescent does not CDI and $\widehat{\nu}_s(\{1\}) = 0$, there are at all times $t \geq 0$ infinitely many species blocks (see for instance [76, Proposition 23]). Now let us fix $\delta \in (0,t]$ and $\varepsilon \in (0,1]$, and investigate the random number of coalescence events in the time interval $[t-\delta, t]$ involving the first species and at least a proportion ε of all other species. More precisely, we consider the number of atoms $(s, \mathbf{Z})$ in the Poissonian construction such that $s \in [t-\delta, t]$, $X_1 = 1$ and $\lim_{n\to\infty} \sum_{i=1}^n X_i/n \geq \varepsilon$. From the Poissonian construction, it is easy to see that this number is a Poisson random variable with mean

$$\delta \int_{[\varepsilon,1]} \widehat{\nu}_s(dx)\, x.$$

Pick any $A \in \mathbb{N}$ and $\eta > 0$. We will show $\mathbb{P}_\varrho(M(t) \leq A) < 2\eta$, which is sufficient to conclude that $M(t) = \infty$ a.s. Note that we assumed that the marginal gene coalescent CDI, so for $\Pi = (\Pi(t),\, t \geq 0)$ a version of this univariate coalescent started from $\mathbf{0}_\infty$, we have $\mathbb{P}(|\Pi(\delta)| < \infty) = 1$ for all $\delta > 0$. In addition, Π is right-continuous, so $|\Pi(\delta)| \uparrow \infty$ as $\delta \to 0$. Therefore, one can choose $\delta > 0$ small enough, and then $\varepsilon > 0$ such that

$$\mathbb{P}(|\Pi(\delta)| \leq A) < \eta \quad \text{and} \quad e(\varepsilon) := \int_{[\varepsilon,1]} \widehat{\nu}_s(dx)\, x \geq \frac{-\log(\eta)}{\delta}. \tag{3.28}$$

Now consider the stopping time defined by

$$T := \inf\{s \geq t-\delta,\ \text{the first species participates at time } s \text{ in a coalescence event involving at least a proportion } \varepsilon \text{ of other species}\}.$$

By the Poisson construction, $T - (t-\delta)$ is an exponential random variable with parameter $e(\varepsilon)$, so from (3.28) we deduce

$$\mathbb{P}_\varrho(T \geq t) \leq \eta.$$

Now since T is a coalescence time for the first species, we have $M(T) = \infty$ almost surely. Indeed, the assumption $\widehat{\nu}^g(\{1\}) = 0$ implies that not every gene participates in the coalescence. But since an infinite number of species participate in the coalescence, the law of large numbers implies that in the newly formed species, there is an infinite number of genes which do not coalesce at time T. Since $M(T) = \infty$, we can define a random injection $\sigma : \mathbb{N} \to \mathbb{N}$ mapping k to the first element of the k-th gene of the first species at time T.

We then define $\widetilde{\Pi}(u) := \sigma(\mathcal{R}^g(T+u))$, which has by the strong Markov property the distribution of a marginal gene coalescent started from $\mathbf{0}_\infty$, independent of T. Furthermore, by construction we have $M(T+u) \geq |\widetilde{\Pi}(u)|$ a.s., so that finally

$$\begin{aligned}\mathbb{P}_\varrho(M(t) \leq A) &\leq \mathbb{P}_\varrho(T > t) + \mathbb{P}_\varrho(t-\delta \leq T \leq t)\, \mathbb{P}_\varrho(|\widetilde{\Pi}(t-T)| \leq A \mid t-\delta \leq T \leq t) \\ &\leq \mathbb{P}_\varrho(T > t) + \mathbb{P}(|\Pi(\delta)| \leq A) \\ &\leq 2\eta.\end{aligned}$$

iii) Now supposing $\int_{[0,1]} \widehat{\nu}_s(dx)\, x < \infty$, with the same argument as previously, the first species participates in coalescence events at some random times $0 < T_1 < T_2 < \ldots$, distributed as a Poisson process with parameter $\int_{[0,1]} \widehat{\nu}_s(dx)\, x$, and all these events involve infinitely many species blocks (recall the marginal species coalescent does not CDI and so in particular has $a_s = 0$). Let $T_0 := 0$ by convention and for each i, we can define a random injection $\sigma_i : \mathbb{N} \to \mathbb{N}$ mapping k to the first element of the k-th gene of the first species at time T_i. Now because the first species does not change during the intervals $[T_i, T_{i+1})$, the process $\widetilde{\Pi}_i$ defined by

$$\widetilde{\Pi}_i(u) := \sigma_i(\mathcal{R}^g(T_i + u))$$

is a marginal gene coalescent (and so CDI by assumption), which is independent of T_i, and there is the following equality between processes, for $u < T_{i+1} - T_i$,

$$M(T_i + u) = \widetilde{\Pi}_i(u).$$

Finally, we have for any $t > 0$, and any initial $\varrho \in \mathcal{N}_\infty$,

$$\begin{aligned}\mathbb{P}_\varrho(M(t) < \infty) &= \sum_{i \geq 0} \mathbb{P}_\varrho(T_i < t < T_{i+1})\, \mathbb{P}_\varrho(M(t) < \infty \mid T_i < t < T_{i+1}) \\ &= \sum_{i \geq 0} \mathbb{P}_\varrho(T_i < t < T_{i+1})\, \mathbb{P}_\varrho(\widetilde{\Pi}_i(t - T_i) < \infty \mid T_i < t < T_{i+1}) \\ &= \sum_{i \geq 0} \mathbb{P}_\varrho(T_i < t < T_{i+1}) = 1,\end{aligned}$$

which concludes the proof. □

References for Chapter 3

[7] J. Berestycki, N. Berestycki, and V. Limic. A Small-Time Coupling between Λ-Coalescents and Branching Processes. *Ann. Appl. Probab.*, 24.2 (Apr. 2014), pp. 449–475. DOI: 10.1214/12-AAP911 (see pp. 15, 82).

[8] J. Berestycki, N. Berestycki, and J. Schweinsberg. The Genealogy of Branching Brownian Motion with Absorption. *Ann. Probab.*, 41.2 (Mar. 2013), pp. 527–618. DOI: 10.1214/11-AOP728 (see p. 60).

[10] J. Bertoin. *Random Fragmentation and Coagulation Processes.* Cambridge University Press, 2006. DOI: 10.1017/CBO9780511617768 (see pp. 10, 11, 60, 61, 64, 65, 79, 81, 90, 91, 94, 95, 99, 117, 126–128, 135, 153).

[14] J. Bertoin. The Structure of the Allelic Partition of the Total Population for Galton–Watson Processes with Neutral Mutations. *Ann. Probab.*, 37.4 (July 2009), pp. 1502–1523. doi: 10.1214/08-AOP441 (see pp. 16, 59, 90).

[15] J. Bertoin and J.-F. Le Gall. Stochastic Flows Associated to Coalescent Processes. *Probab. Theory Related Fields*, 126.2 (2003), pp. 261–288. doi: 10.1007/s00440-003-0264-4 (see p. 60).

[16] J. Bertoin and J.-F. Le Gall. Stochastic Flows Associated to Coalescent Processes. III. Limit Theorems. *Illinois J. Math.*, 50.1-4 (2006), pp. 147–181. doi: 10.1215/ijm/1258059473 (see pp. 60, 82).

[18] A. Blancas, J.-J. Duchamps, A. Lambert, and A. Siri-Jégousse. Trees within Trees: Simple Nested Coalescents. *Electron. J. Probab.*, 23.0 (2018). doi: 10.1214/18-EJP219 (see pp. 11, 58, 133).

[19] A. Blancas, T. Rogers, J. Schweinsberg, and A. Siri-Jégousse. The Nested Kingman Coalescent: Speed of Coming down from Infinity. *Ann. Appl. Probab.*, 29.3 (June 2019), pp. 1808–1836. doi: 10.1214/18-AAP1440 (see pp. 59, 60, 90).

[20] E. Bolthausen and A.-S. Sznitman. On Ruelle's Probability Cascades and an Abstract Cavity Method. *Comm. Math. Phys.*, 197.2 (1998), pp. 247–276. doi: 10.1007/s002200050450 (see p. 83).

[21] É. Brunet and B. Derrida. Genealogies in Simple Models of Evolution. *J. Stat. Mech. Theory Exp.*, 2013.01 (Jan. 16, 2013), P01006. doi: 10.1088/1742-5468/2013/01/P01006 (see p. 60).

[31] D. A. Dawson. Multilevel Mutation-Selection Systems and Set-Valued Duals. *J. Math. Biol.*, 76.1-2 (Jan. 2018), pp. 295–378. doi: 10.1007/s00285-017-1145-2 (see pp. 59, 60).

[32] J. H. Degnan and N. A. Rosenberg. Gene Tree Discordance, Phylogenetic Inference and the Multispecies Coalescent. *Trends Ecol. Evol.*, 24.6 (2009), pp. 332–340. doi: 10.1016/j.tree.2009.01.009 (see p. 59).

[34] M. M. Desai, A. M. Walczak, and D. S. Fisher. Genetic Diversity and the Structure of Genealogies in Rapidly Adapting Populations. *Genetics*, 193.2 (2013), pp. 565–585. doi: 10.1534/genetics.112.147157 (see p. 60).

[35] J. J. Doyle. Trees within Trees: Genes and Species, Molecules and Morphology. *Syst. Biol.*, 46.3 (Sept. 1, 1997), pp. 537–553. doi: 10.1093/sysbio/46.3.537 (see pp. 59, 90).

[38] R. Durrett and J. Schweinsberg. A Coalescent Model for the Effect of Advantageous Mutations on the Genealogy of a Population. *Stochastic Process. Appl.*, 115.10 (2005), pp. 1628–1657. doi: 10.1016/j.spa.2005.04.009 (see p. 60).

[39] B. Eldon and J. Wakeley. Coalescent Processes When the Distribution of Offspring Number among Individuals Is Highly Skewed. *Genetics*, 172.4 (Apr. 1, 2006), pp. 2621–2633. doi: 10.1534/genetics.105.052175 (see p. 60).

[40] A. Etheridge. *Some Mathematical Models from Population Genetics: École d'ete de Probabilités de Saint-Flour XXXIX-2009*. Lecture Notes in Mathematics 2012. Heidelberg ; New York: Springer, 2011 (see pp. 59, 90).

[43] J. Felsenstein. *Inferring Phylogenies*. Vol. 2. Sinauer associates Sunderland, MA, 2004 (see p. 58).

[46] F. Foutel-Rodier, A. Lambert, and E. Schertzer. Exchangeable Coalescents, Ultrametric Spaces, Nested Interval-Partitions: A Unifying Approach (July 13, 2018). arXiv: 1807.05165 [math] (see p. 60).

[50] B. T. Grenfell et al. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*, 303.5656 (2004), pp. 327–332. doi: 10.1126/science.1090727 (see p. 59).

[54] J. Heled and A. J. Drummond. Bayesian Inference of Species Trees from Multilocus Data. *Mol. Biol. Evol.*, 27.3 (2009), pp. 570–580. doi: 10.1093/molbev/msp274 (see p. 59).

[55] O. Kallenberg. *Probabilistic Symmetries and Invariance Principles*. Probability and Its Applications. New York: Springer-Verlag, 2005 (see p. 74).

[58] J. Kingman. The Coalescent. *Stochastic Process. Appl.*, 13.3 (1982), pp. 235–248. doi: 10.1016/0304-4149(82)90011-4 (see pp. 8, 15, 22, 59, 90, 95).

[61] A. Lambert. Population Dynamics and Random Genealogies. *Stoch. Models*, 24 (sup1 2008), pp. 45–163. doi: 10.1080/15326340802437728 (see pp. 8, 58, 90).

[63] A. Lambert. Random Ultrametric Trees and Applications. *ESAIM Proc. Surveys*, 60 (2017), pp. 70–89. doi: 10.1051/proc/201760070 (see pp. 58, 60).

[65] A. Lambert and E. Schertzer. Coagulation-Transport Equations and the Nested Coalescents. *Probab. Theory Related Fields*, (Apr. 15, 2019). doi: 10.1007/s00440-019-00914-4 (see pp. 59, 60, 90).

[68] W. P. Maddison. Gene Trees in Species Trees. *Syst. Biol.*, 46.3 (Sept. 1, 1997), pp. 523–536. doi: 10.1093/sysbio/46.3.523 (see pp. 59, 90).

[70] S. Matuszewski, M. E. Hildebrandt, G. Achaz, and J. D. Jensen. Coalescent Processes with Skewed Offspring Distributions and Nonequilibrium Demography. *Genetics*, 208.1 (2017), pp. 323–338. doi: 10.1534/genetics.117.300499 (see p. 60).

[71] R. A. Neher and O. Hallatschek. Genealogies of Rapidly Adapting Populations. *Proc. Natl. Acad. Sci. USA*, 110.2 (Jan. 8, 2013), pp. 437–442. doi: 10.1073/pnas.1213113110 (see p. 60).

[72] M. Nei and S. Kumar. *Molecular Evolution and Phylogenetics*. Oxford university press, 2000 (see p. 58).

[73] R. D. Page and M. A. Charleston. From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem. *Mol. Phylogenet. Evol.*, 7.2 (Apr. 1997), pp. 231–240. doi: 10.1006/mpev.1996.0390 (see pp. 59, 90).

[74] R. D. Page and M. A. Charleston. Trees within Trees: Phylogeny and Historical Associations. *Trends Ecol. Evol.*, 13.9 (Sept. 1998), pp. 356–359. doi: 10.1016/S0169-5347(98)01438-4 (see pp. 59, 90).

[76] J. Pitman. Coalescents with Multiple Collisions. *Ann. Probab.*, 27.4 (Oct. 1999), pp. 1870–1902. doi: 10.1214/aop/1022874819 (see pp. 60, 63, 65, 83, 84, 90).

[78] N. A. Rosenberg. The Probability of Topological Concordance of Gene Trees and Species Trees. *Theor. Popul. Biol.*, 61.2 (2002), pp. 225–247. doi: 10.1006/tpbi.2001.1568 (see p. 59).

[79] S. Sagitov. The General Coalescent with Asynchronous Mergers of Ancestral Lines. *J. Appl. Probab.*, 36.4 (Dec. 1999), pp. 1116–1125. doi: 10.1017/S0021900200017903 (see pp. 60, 65, 90).

[81] J. Schweinsberg. A Necessary and Sufficient Condition for the Λ-Coalescent to Come Down from Infinity. *Electron. Commun. Probab.*, 5 (2000), pp. 1–11. doi: 10.1214/ECP.v5-1013 (see p. 82).

[82] J. Schweinsberg. Coalescent Processes Obtained from Supercritical Galton–Watson Processes. *Stochastic Process. Appl.*, 106.1 (July 2003), pp. 107–139. doi: 10.1016/S0304-4149(03)00028-0 (see p. 60).

[83] J. Schweinsberg. Coalescents with Simultaneous Multiple Collisions. *Electron. J. Probab.*, 5 (2000). doi: 10.1214/EJP.v5-68 (see p. 60).

[84] J. Schweinsberg. Rigorous Results for a Population Model with Selection II: Genealogy of the Population. *Electron. J. Probab.*, 22 (2017). doi: 10.1214/17-EJP58 (see p. 60).

[85] C. Semple and M. Steel. *Phylogenetics*. Oxford Lecture Series in Mathematics and Its Applications 24. Oxford ; New York: Oxford University Press, 2003 (see pp. 58, 90).

[88] G. J. Szöllősi, E. Tannier, V. Daubin, and B. Boussau. The Inference of Gene Trees with Species Trees. *Syst. Biol.*, 64.1 (2014), e42–e62. doi: 10.1093/sysbio/syu048 (see p. 59).

[90] A. Tellier and C. Lemaire. Coalescence 2.0: A Multiple Branching of Recent Theoretical Developments and Their Applications. *Mol. Ecol.*, 23.11 (2014), pp. 2637–2652. doi: 10.1111/mec.12755 (see p. 60).

[91] E. M. Volz, K. Koelle, and T. Bedford. Viral Phylodynamics. *PLoS Comput. Biol.*, 9.3 (2013), e1002947. doi: 10.1371/journal.pcbi.1002947 (see p. 59).

Chapter 4

Trees within trees II: Nested fragmentations

This chapter is accepted for publication in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* [36].

Contents

4.1 Introduction

Evolutionary biology aims at tracing back the history of species, by identifying and dating the relationships of ancestry between past lineages of extant individuals. This information

is usually represented by a tree or phylogeny [62, 85], species corresponding to leaves of the tree and speciation events (point in time where several species descend from a single one) corresponding to internal nodes.

In modern methods, one analyzes genetic data from samples of individuals to statistically infer their phylogenetic tree. Probabilistic tree models have been well-developed in the last decades – either from individual-based population models like the classical Wright-Fisher model [14, 40, 61, 85], or from forward-in-time branching processes, where the branching particles are species (see for instance Aldous's Markov branching models [3] and the surrounding literature [26, 27, 44, 53]) – allowing for inference from genetic data. A challenge is that trees inferred from different parts of the genome generally fail to coincide, each of them being understood as an alteration of a "true" underlying phylogeny (which we call the *species tree*).

To understand the relation between *gene trees* and the species tree, our goal is to identify a class of Markovian models coupling the evolution of both trees, making the assumption that in general, several gene lineages coexist within the same species, and at speciation events one or several gene lineages diverge from their neighbors to form a new species, i.e. we model the problem as a *tree within a tree* [35, 68, 73, 74], or *nested tree*. See Figure 4.1 for an instance of a simple nested genealogy where discrepancies arise between the resulting gene tree and species tree.

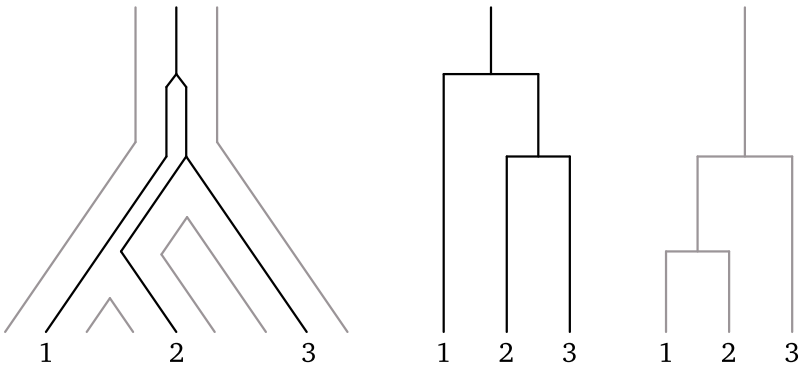


Figure 4.1 – Example of a nested tree where the gene tree (in black) does not coincide with the species tree (in gray).

Recent research aims at defining mathematical processes giving rise to such nested trees, generalizing several well-studied univariate (we will sometime use this term as opposed to *nested*) processes. Some work in progress involves a nested version [19, 65] of the Kingman coalescent [58] (considered the neutral model for evolution, appearing as a scaling limit of many individual-based population models). In Chapter 3 we have studied a nested generalization of Λ-coalescent processes [10, 76, 79] and characterize their distribution. Our present goal is to generalize the forward-in-time branching models originated from Aldous [3]. His assumptions (which will be formally defined for our context in Section 4.3) are basically that the random process of evolution is homogeneous in time and that the law of the process is invariant under both relabeling and resampling of individuals (we then say the process is *exchangeable* and *sampling consistent*). We are interested in the partition-valued processes satisfying these assumptions, i.e. the so-called fragmentation

processes [10, 53], and in this article we generalize their definition to *nested partition-valued* processes to model jointly a gene tree within a species tree.

Crane [27] also generalizes Aldous's Markov branching models to study the gene tree/species tree problem but uses a different approach to the one we use here. Indeed, his model is such that first the entire species tree $\mathbf{t}$ is drawn according to some probability, and then the gene tree $\mathbf{t}'$ is constructed thanks to a generalized Markov branching model that depends on $\mathbf{t}$. In the meantime, our goal is to characterize the class of models in which there is a joint Markov branching construction of both the gene tree and the species tree, under the assumptions of exchangeability and sampling consistency.

In particular our main result Theorem 4.14, which will be formally stated in Section 4.5, shows that nested fragmentation processes satisfying natural branching properties are uniquely characterized by

- three *erosion parameters* $c_{\text{out}}, c_{\text{in},1}$ and $c_{\text{in},2}$ (rates at which a unique lineage can fragment out of its mother block, in three different situations);
- two *dislocation measures* ν_{out} and ν_{in} that are Poissonian intensities of how blocks instantaneously fragment into several new blocks with macroscopic frequencies.

The article is organized as follows. Section 4.2 introduces some notation used throughout the paper, and the definition of nested fragmentations. We also recall some results in the univariate case which we seek to generalize to the nested case. In Section 4.3 we study our so-called *strong exchangeability* assumption, and show its relation to a *projective Markov property*, in order to define characteristic kernels of nested fragmentation processes. In Section 4.4 we use the so-called *outer branching property*, simplifying the representation of characteristic kernels of fragmentations, and giving a natural Poissonian construction of such processes. Focusing on the *inner branching property*, Section 4.5 is dedicated to the full characterization of the semi-group of nested fragmentations, in terms of *erosion* and *dislocation measures*. It is shown that dislocations, similarly as in the univariate case, can be understood as (bivariate) paintbox processes. Finally Section 4.6 briefly shows how our main result, Theorem 4.14, translates in simpler terms when we make the classical biological assumption that all splits are binary.

4.2 Definitions and examples

4.2.1 Definitions, notation

For a set S, write $\mathcal{P}_S$ for the set of partitions of S:

$$\mathcal{P}_S := \{\pi \subset \mathfrak{P}(S) \setminus \{\varnothing\},\ \forall A \neq B \in \pi, A \cap B = \varnothing \text{ and } \textstyle\bigcup_{A\in\pi} A = S\},$$

where $\mathfrak{P}(S)$ denotes the power set of S. Throughout the paper, whenever a subset $\pi' \subset \mathfrak{P}(S)$ is defined in a way such that $\pi' = \pi \cup \{\varnothing\}$ for a certain $\pi \in \mathcal{P}_S$, we will implicitly identify π' and π to avoid the formal and cumbersome notation $\pi' \setminus \{\varnothing\}$.

For S, S' two sets, $\pi \in \mathcal{P}_S$ and $\sigma : S' \to S$ an **injection**, we write

$$\pi^\sigma := \{\sigma^{-1}(A),\ A \in \pi\},$$

and if μ is a measure on $\mathcal{P}_S$ then we write μ^σ for the push-forward of μ by the map $\pi \mapsto \pi^\sigma$.

Note that if $S'' \xrightarrow{\tau} S' \xrightarrow{\sigma} S$ are injections, then we have $\pi^{\sigma\tau} = (\pi^\sigma)^\tau$, and $\mu^{\sigma\tau} = (\mu^\sigma)^\tau$.

For $S' \subset S$, there is a natural surjective map $r_{S,S'} : \mathcal{P}_S \to \mathcal{P}_{S'}$ called the restriction, defined by

$$r_{S,S'}(\pi) = \pi_{|S'} := \{A \cap S',\ A \in \pi\}.$$

Note that $\pi_{|S'} = \pi^\sigma$ for $\sigma : S' \to S, x \mapsto x$ the canonical injection.

There is always a partial order on $\mathcal{P}_S$, denoted $\preceq$ and defined as:

$$\pi \preceq \pi' \quad \text{if} \quad \forall (A,B) \in \pi \times \pi',\ A \cap B \neq \varnothing \Rightarrow A \subset B,$$

that is $\pi \preceq \pi'$ if π is finer than π'. From now on, we prefer to write ζ or ξ for partitions and π for pairs of partitions. Also, throughout the paper we will say if $\zeta \preceq \xi$ that the pair (ζ, ξ) is nested. Let us introduce the space of pairs of nested partitions,

$$\mathcal{P}_S^{2,\preceq} := \{(\zeta,\xi) \in \mathcal{P}_S^2,\ \zeta \preceq \xi\},$$

which we equip with a partial order $\preceq$ defined naturally as

$$(\zeta,\xi) \preceq (\zeta',\xi') \text{ if } \zeta \preceq \zeta' \text{ and } \xi \preceq \xi'.$$

We will use $\mathbf{0}_S$ or sometimes, with some abuse of notation, $\mathbf{0}$ when the context is clear, to denote the partition of S into singletons. Similarly, we will denote $\mathbf{1}_S$ or $\mathbf{1}$ the partition in one block $\{S\}$. For $S' \subset S$ and $\pi = (\zeta,\xi) \in \mathcal{P}_S^{2,\preceq}$, we define naturally the restriction

$$\pi_{|S'} := (\zeta_{|S'}, \xi_{|S'}) \in \mathcal{P}_{S'}^{2,\preceq}.$$

Let us now define, for $n \in \mathbb{N}$, $[n] := \{1,\dots,n\}$ and $[\infty] := \mathbb{N}$, and for $n \in \mathbb{N} \cup \{\infty\}$:

$$\mathcal{P}_n := \mathcal{P}_{[n]} \quad \text{and} \quad \mathcal{P}_n^{2,\preceq} := \mathcal{P}_{[n]}^{2,\preceq}$$

We will generally label the blocks of a partition $\xi = \{\xi_1, \xi_2, \dots\}$, in the unique way such that

$$\min \xi_1 < \min \xi_2 < \dots$$

The space $\mathcal{P}_\infty^{2,\preceq}$ is endowed with a distance d which makes it compact, defined as follows:

$$d(\pi,\pi') = \Big(\sup\{n \in \mathbb{N},\ \pi_{|[n]} = \pi_{|[n]}\}\Big)^{-1},$$

with the convention $(\sup \mathbb{N})^{-1} = 0$. Note that the same expression can be used to define a distance on $\mathcal{P}_\infty$, making it a compact space as well.

For $k \leq n \leq \infty$, $\sigma : [k] \to [n]$ an injection and $\pi = (\zeta,\xi) \in \mathcal{P}_n^{2,\preceq}$, we write

$$\pi^\sigma := (\zeta^\sigma, \xi^\sigma) \in \mathcal{P}_k^{2,\preceq}.$$

A key property of the space $\mathcal{P}_\infty^{2,\preceq}$ is that for any $n \in \mathbb{N}$, and any $\pi \in \mathcal{P}_n^{2,\preceq}$, there is a $\pi^\star \in \mathcal{P}_\infty^{2,\preceq}$ satisfying:

- $\pi^\star_{|[n]} = \pi$;

- for any $\pi' \in \mathcal{P}_\infty^{2,\preceq}$ such that $\pi'_{|[n]} = \pi$, there is an injection $\sigma : \mathbb{N} \to \mathbb{N}$ which satisfies $\sigma_{|[n]} = \mathrm{id}_{[n]}$ and $(\pi^\star)^\sigma = \pi'$.

Indeed, it is clear that one can choose a $\pi^\star = (\zeta^\star, \xi^\star)$ such that $\pi^\star_{|[n]} = \pi$, such that $\zeta^\star$ has infinitely many infinite blocks and no finite blocks, $\xi^\star$ has infinitely many blocks, and each of them contains infinitely many distinct blocks of $\zeta^\star$. This partition immediately satisfies the required property. We will call any such $\pi^\star$ a **universal element of $\mathcal{P}_\infty^{2,\preceq}$ with initial part** π whenever we need to use one.

A measure μ on $\mathcal{P}_n$ or on $\mathcal{P}_n^{2,\preceq}$ is said to be **exchangeable** if for any permutation $\sigma : [n] \to [n]$, we have

$$\mu^\sigma = \mu.$$

A random variable Π taking values in $\mathcal{P}_n$ or in $\mathcal{P}_n^{2,\preceq}$ is said to be exchangeable if for any permutation $\sigma : [n] \to [n]$, we have

$$\Pi^\sigma \overset{(d)}{=} \Pi,$$

that is if its distribution is exchangeable. Similarly, a random process $(\Pi(t), t \geq 0)$ taking values in $\mathcal{P}_n$ or in $\mathcal{P}_n^{2,\preceq}$ is said to be exchangeable if for any initial state π_0 and any permutation $\sigma : [n] \to [n]$, we have

$$(\Pi(t)^\sigma, t \geq 0) \text{ under } \mathbb{P}_{\pi_0} \quad \overset{(d)}{=} \quad (\Pi(t), t \geq 0) \text{ under } \mathbb{P}_{\pi_0^\sigma}, \tag{4.1}$$

where $\mathbb{P}_\pi$ is the distribution of the process started from π.

Finally, a measure or a random process with values in $\mathcal{P}_\infty$ or $\mathcal{P}_\infty^{2,\preceq}$ will be called **strongly exchangeable** if its distribution is invariant under the action of *injections* $\mathbb{N} \to \mathbb{N}$. Note that while it is easily checked that for measures the two properties are equivalent, for processes this is a strictly stronger assumption than being exchangeable. Indeed, since the number of blocks of a partition is invariant under the action of permutations but not under the action of injections, one can define exchangeable Markov jump processes $(\Pi(t), t \geq 0)$ with jump rates depending on the total number of blocks of $\Pi(t)$, preventing strong exchangeability. The reason we prefer to assume strong exchangeability is the following. Consider a strong exchangeable process Π (say with values in $\mathcal{P}_\infty^{2,\preceq}$) and a universal initial state π. Then for any $\pi' \in \mathcal{P}_\infty^{2,\preceq}$, there is an injection $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\pi' = \pi^\sigma$, so strong exchangeability (4.1) ensures us that if $\Pi \sim \mathbb{P}_\pi$, then $\Pi^\sigma \sim \mathbb{P}_{\pi'}$. In other words, the process Π under $\mathbb{P}_\pi$ – i.e. started from π – is a coupling of all possible distributions $\mathbb{P}_{\pi'}$, for $\pi' \in \mathcal{P}_\infty^{2,\preceq}$, which will often be convenient.

In the following we only consider time-homogeneous Markov processes. We can now define nested fragmentation processes in a way that extends naturally the definition of fragmentation processes in the univariate case.

Definition 4.1. Let $\Pi = (\Pi(t), t \geq 0) = ((\zeta(t), \xi(t)), t \geq 0)$ be a Markov process with values in $\mathcal{P}_\infty^{2,\preceq}$. We say Π is a **nested fragmentation process** if:

(i) Π is strongly exchangeable, with nonincreasing càdlàg sample paths.

(ii) **Outer branching property.** For any initial state $\pi = (\zeta, \xi)$ with $\xi = \{\xi_1, \xi_2, \ldots\}$ and given bijections $\sigma_i : [\#\xi_i] \to \xi_i$, where $\#\xi_i$ denotes the cardinality of block ξ_i,

the processes

$$\Big((\Pi^{\sigma_i}(t),\, t \geq 0),\ i \geq 1\Big)$$

are mutually independent under $\mathbb{P}_\pi$.

(iii) **Inner branching property.** The process $(\zeta(t), t \geq 0)$, with values in $\mathcal{P}_\infty$, is a homogeneous univariate fragmentation process, as in [10, Definition 3.2].

In words, the branching properties (ii) and (iii) imply that different blocks at a given time undergo independent fragmentations in the future. Throughout the rest of the paper, unless stated otherwise, we consider an alternative, more convenient definition, which we will prove to be equivalent to Definition 4.1, and whose idea is the following: distinct blocks fragment at distinct times.

Definition 4.1'. Let $\Pi = (\Pi(t), t \geq 0) = ((\zeta(t), \xi(t)), t \geq 0)$ be a Markov process with values in $\mathcal{P}_\infty^{2,\preceq}$. We say Π is a **nested fragmentation process** if:

(i) Π is strongly exchangeable, with nonincreasing càdlàg sample paths.

(ii') Π satisfies the **outer branching property**:

Almost surely for all t such that $\Pi(t-) \neq \Pi(t)$, there is a unique block $B \in \xi(t-)$ such that $\Pi(t-)_{|B} \neq \Pi(t)_{|B}$.

(iii') Π satisfies the **inner branching property**:

Almost surely for all t such that $\zeta(t-) \neq \zeta(t)$, there is a unique block $B \in \zeta(t-)$ such that $\zeta(t-)_{|B} \neq \zeta(t)_{|B}$.

Note that we will show in Section 4.3 that a nested fragmentation process according to Definition 4.1 satisfies also Definition 4.1', and then in Corollary 4.15 it will appear that the converse is true.

Before describing our results in the setting of nested fragmentations, let us recall the concepts of mass partitions and paintbox processes in the univariate setting. These ideas, which will ultimately be extended to the nested case, are paramount in understanding the possible transitions of fragmentation processes.

4.2.2 Univariate results, mass partitions

Random exchangeable partitions $\pi \in \mathcal{P}_\infty$ and their relation to random mass partitions is well known [see 10, Chapter 2]. We denote the space of mass partitions by

$$\mathscr{S}^\downarrow := \Big\{\mathbf{s} = (s_1, s_2, \ldots) \in [0,1]^{\mathbb{N}},\ s_1 \geq s_2 \geq \ldots,\ \textstyle\sum_k s_k \leq 1\Big\}. \tag{4.2}$$

For $\mathbf{s} \in \mathscr{S}^\downarrow$, one defines an exchangeable distribution on $\mathcal{P}_\infty$, by the following so-called *paintbox construction*:

- for $k \geq 0$, define $t_k = \sum_{k'=1}^{k} s_{k'}$, with $t_0 = 0$ by convention.
- let $(U_i, i \geq 1)$ be an i.i.d. sequence of uniform random variables in $[0,1]$.
- define the random partition $\pi \in \mathcal{P}_\infty$ by setting

$$i \sim^\pi j \iff i = j \text{ or } \exists k \geq 1,\ U_i, U_j \in [t_{k-1}, t_k).$$

Then the distribution of π is exchangeable and is denoted $\varrho_{\mathbf{s}}$. Notice that the set $\pi_0 := \{[t_{k-1}, t_k),\, k \geq 1\} \cup \{\{t\},\ \sum_{k\geq 1} s_k \leq t \leq 1\}$ is a partition of $[0,1]$, and that we have $\pi = \pi_0^\sigma$, where $\sigma : \mathbb{N} \to [0,1]$ is the random injection defined by $\sigma : i \mapsto U_i$. Also, note that by definition some blocks are singletons (blocks $\{i\}$ such that $U_i \in [\sum_{k\geq 1} s_k, 1]$), and by construction we have

$$\frac{\#\{i \in [n],\, \{i\} \in \pi\}}{n} \underset{n\to\infty}{\longrightarrow} s_0 := 1 - \textstyle\sum_{k\geq 1} s_k.$$

These integers that are singleton blocks are called the *dust* of the random partition π and the last display tells us there is a frequency s_0 of dust.

Conversely, any random exchangeable partition π has a distribution that can be expressed with these paintbox constructions $\varrho_{\mathbf{s}}$. Indeed, π has **asymptotic frequencies**, i.e.

$$|B| := \lim_{n\to\infty} \frac{\#(B \cap [n])}{n} \quad \text{exists a.s. for all } B \in \pi.$$

Let us write $|\pi|^\downarrow \in \mathscr{S}^\downarrow$ for the nonincreasing reordering of $(|B|, B \in \pi)$, ignoring the zero terms coming from the dust. It is known [58, Theorem 2] that the conditional distribution of π given $|\pi|^\downarrow = \mathbf{s}$ is $\varrho_{\mathbf{s}}$, so we have

$$\mathbb{P}(\pi \in \cdot\,) = \int \mathbb{P}(|\pi|^\downarrow \in \mathrm{d}\mathbf{s}) \varrho_{\mathbf{s}}(\,\cdot\,).$$

This means that any exchangeable probability measure on $\mathcal{P}_\infty$ is of the form ϱ_ν where ν is a probability measure on $\mathscr{S}^\downarrow$, and

$$\varrho_\nu(\,\cdot\,) := \int_{\mathscr{S}^\downarrow} \varrho_{\mathbf{s}}(\,\cdot\,)\nu(\mathrm{d}\mathbf{s}).$$

Furthermore, Bertoin [10, Theorem 3.1] shows that any exchangeable measure μ on $\mathcal{P}_\infty$ such that

$$\mu(\{\mathbf{1}\}) = 0 \quad \text{and} \quad \forall n \geq 1, \quad \mu(\pi_{|[n]} \neq \mathbf{1}_{[n]}) < \infty \tag{4.3}$$

can be written $\mu = c\mathfrak{e} + \varrho_\nu$, where $c \geq 0$, ν is a measure on $\mathscr{S}^\downarrow$ satisfying

$$\nu(\{(1,0,0,\ldots)\}) = 0 \quad \text{and} \quad \int_{\mathscr{S}^\downarrow} (1 - s_1)\nu(\mathrm{d}\mathbf{s}) < \infty, \tag{4.4}$$

and $\mathfrak{e}$ is the so-called *erosion measure*, defined by

$$\mathfrak{e} := \textstyle\sum_{i\in\mathbb{N}} \delta_{\{\{i\}, \mathbb{N}\setminus\{i\}\}}.$$

As a result, each fragmentation process with values in $\mathcal{P}_\infty$ is characterized by its erosion coefficient c and characteristic measure ν, in such a way that its rates can be described as follows:

> A block of size n fragments, independently of the other blocks, into a partition with k different blocks of sizes $n_1, n_2, \ldots, n_k$ at rate
>
> $$c\mathbb{1}\{k = 2, \text{ and } n_1 = 1 \text{ or } n_2 = 1\} + \int_{\mathscr{S}^\downarrow} \nu(\mathrm{d}\mathbf{s}) \sum_{\mathbf{i}} s_{i_1}^{n_1} \cdot s_{i_2}^{n_2} \cdots s_{i_k}^{n_k},$$
>
> where s_0 is defined to be $1 - \sum_{i\geq 1} s_i$, and the sum is over the vectors $\mathbf{i} = (i_1, \ldots, i_k) \in \{0, 1, \ldots\}^k$ such that i_j may be 0 only if $n_j = 1$, and if $j \neq j'$ and $i_j \neq 0$, then $i_{j'} \neq i_j$.

A similar result will be shown in the setting of nested fragmentations.

4.2.3 Transitions of nested fragmentation processes

In this article we show that nested fragmentations are processes for which five different fragmentation events – jumps for the Markov process Π – need to be distinguished. All nested fragmentation processes are entirely characterized by the rates at which those fragmentation events occur. While the main result, Theorem 4.14, cannot be stated at this time because much notation needs to be introduced first, let us briefly explain what the five typical events of a nested fragmentation are with an example. Assume that the nested fragmentation $\Pi = (\zeta, \xi)$ jumps at time t, with (restricting each partition to $\{1, \dots, 12\}$)

$$\begin{array}{lll} \xi(t-) = \{1,4,6\}, & \{2,5,7,8,9,10,11,12\}, & \{3\} \\ \zeta(t-) = \{1,4\},\{6\}, & \{2,9,10,12\},\{5\},\{7,8\},\{11\}, & \{3\}. \end{array}$$

Then the five following events may occur:

- **Outer erosion:** Each inner block erodes out of its outer block at a constant rate. For example, if the block $\{7,8\}$ erodes out of its outer block at time t, then we have

$$\begin{array}{llll} \xi(t) = \{1,4,6\}, & \{2,5,9,10,11,12\}, & \{7,8\}, & \{3\} \\ \zeta(t) = \{1,4\},\{6\}, & \{2,9,10,12\},\{5\},\{11\}, & \{7,8\}, & \{3\}. \end{array}$$

  Note that a macroscopic – i.e. non-singleton – inner block can erode out of its outer block. This may seem counterintuitive as erosion is usually seen as a continuous loss of mass, but here the idea is simply that a single inner block – not a macroscopic *proportion* of blocks – separates from its outer block.

- **Inner erosion:** Each integer erodes out of its inner block at a constant rate. For example, if the integer 2 erodes out of its inner block at time t, then we have

$$\begin{array}{lll} \xi(t) = \{1,4,6\}, & \{2,5,7,8,9,10,11,12\}, & \{3\} \\ \zeta(t) = \{1,4\},\{6\}, & \{2\},\{9,10,12\},\{5\},\{7,8\},\{11\}, & \{3\}. \end{array}$$

- **Inner erosion with creation of new species:** Each integer erodes out of its inner and outer blocks at a constant rate. If the integer 2 erodes out of its inner and outer blocks at time t, then we have

$$\begin{array}{llll} \xi(t) = \{1,4,6\}, & \{2\}, & \{5,7,8,9,10,11,12\}, & \{3\} \\ \zeta(t) = \{1,4\},\{6\}, & \{2\}, & \{9,10,12\},\{5\},\{7,8\},\{11\}, & \{3\}. \end{array}$$

- **Outer dislocation:** An outer block can split into two or more outer blocks. Each of the inner blocks then decides, according to a Kingman paintbox procedure [57], which outer block to join. For example, if the outer block containing 2 splits into three outer blocks, then the partitions at time t can be

$$\begin{array}{lllll} \xi(t) = \{1,4,6\}, & \{2,9,10,11,12\}, & \{5\}, & \{7,8\}, & \{3\} \\ \zeta(t) = \{1,4\},\{6\}, & \{2,9,10,12\},\{11\}, & \{5\}, & \{7,8\}, & \{3\}. \end{array}$$

  Recall that a paintbox process is a way to draw random exchangeable partitions of a (countable) set I: given a partition of $[0,1]$ into intervals, throw a sequence $(U_i)_{i\in I}$ of i.i.d. uniform random variables on $[0,1]$; the blocks of the random partition are composed of the i that lie in the same interval. A paintbox procedure corresponding to the example would be Figure 4.2.

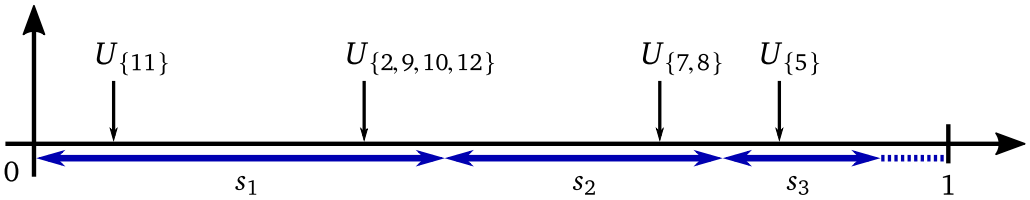


Figure 4.2 – Usual paintbox process, where the interval partition is composed of three intervals of lengths $s_1 \geq s_2 \geq s_3$.

- **Inner dislocation:** An inner block could split into two or more inner blocks, with each of the new inner blocks choosing either to stay in the outer block in which it resided before – its *mother block* –, or move to one of two or more new outer blocks that are created. For example, if the block $\{2, 9, 10, 12\}$ splits into four singletons, with $\{2\}$ choosing to stay in the mother block while the other three integers move to one of two newly created outer blocks. Then the partitions at time t can be

$$\begin{array}{lllll} \xi(t) = \{1,4,6\}, & \{2,5,7,8,11\}, & \{9,12\}, & \{10\}, & \{3\} \\ \zeta(t) = \{1,4\},\{6\}, & \{2\},\{5\},\{7,8\},\{11\}, & \{9\},\{12\}, & \{10\}, & \{3\}. \end{array}$$

  Note that a bivariate paintbox process is needed to construct inner dislocation events: see Figure 4.3 for a paintbox corresponding to this example.

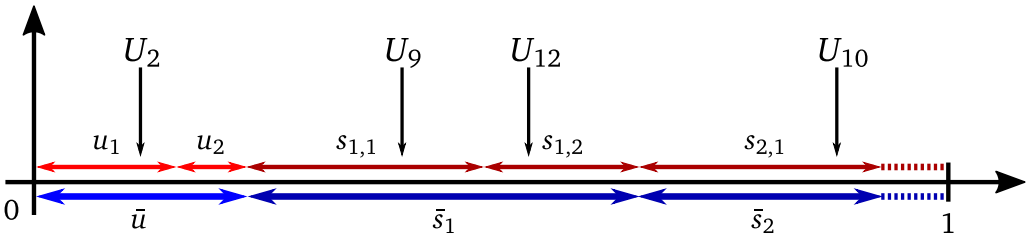


Figure 4.3 – Bivariate paintbox process, built from two nested interval partitions, the coarser (drawn in blue) with interval lengths $\bar{u}, \bar{s}_1, \bar{s}_2, \ldots$, and the finer (drawn in red) with interval lengths u_1, u_2, $s_{1,1}$, $s_{1,2}$, $s_{2,1}$, etc. In this example, the variable U_2 falls into the distinguished interval with length $\bar{u}$, meaning that the integer 2 remains in its mother outer block. The variables U_9 and U_{12} fall in the same outer interval but in distinct inner intervals so an outer block $\{9, 12\}$ is formed, containing two inner blocks $\{9\}$ and $\{12\}$. Similarly, $\{10\}$ forms a new outer and inner block.

Note that not all decreasing transitions are valid. For instance, consider the transition from the initial state above to

$$\begin{array}{llll} \xi(t) = \{1,4\}, & \{6\}, & \{2,5,7,8,9,10,11,12\}, & \{3\} \\ \zeta(t) = \{1,4\}, & \{6\}, & \{2\},\{9,10,12\},\{5\},\{7,8\},\{11\}, & \{3\}, \end{array}$$

where both the inner block $A = \{2, 9, 10, 12\}$ and the outer block $B = \{1, 4, 6\}$ simultaneously undergo fragmentation. In fact since they are not nested ($A \not\subset B$) we will see that this transition is impossible. Also, consider the transition

$$\begin{array}{llll} \xi(t) = \{1,4,6\}, & \{2,5,9,10,12\}, & \{7,8,11\} & \{3\} \\ \zeta(t) = \{1,4\},\{6\}, & \{2\},\{9,10,12\},\{5\}, & \{7,8\},\{11\}, & \{3\}. \end{array}$$

Now inner block A undergoes fragmentation at the same time as its mother block $B = \{2, 5, 7, 8, 9, 10, 11, 12\}$. However, the transition is invalid because the fragmentation of block B separates, in particular, sites 5 and 7, while neither of them is in A. It will be clear along the proof of Theorem 4.14 that such events are impossible for nested fragmentation processes (essentially because if such transitions had positive rates, exchangeability would imply that those rates are infinite).

Let us now start the analysis of nested fragmentation processes by exploiting their strong exchangeability property.

4.3 Projective Markov property – characteristic kernel

The goal of this section is to show that nested fragmentations are processes Π for which the following **projective Markov property** holds:

> For all $n \geq 1$, the process $\Pi^n := (\Pi(t)_{|[n]}, t \geq 0)$ is a continuous-time Markov chain in the finite state space $\mathcal{P}_n^{2,\preceq}$, whose distribution under $\mathbb{P}_\pi$ depends only on $\pi_{|[n]}$.

We already made use of this property in Lemma 3.3 in the context of nested coalescent processes. Here it is exposed in a slightly more general way since we show that for a large class of Markov processes with values in $\mathcal{P}_\infty^{2,\preceq}$ or $\mathcal{P}_\infty$ (not only coalescent or fragmentation processes, but any càdlàg exchangeable process), the projective Markov property is in fact equivalent to strong exchangeability.

Proposition 4.2. *Let $\Pi = (\Pi(t), t \geq 0)$ be an exchangeable Markov process taking values in $\mathcal{P}_\infty^{2,\preceq}$ or $\mathcal{P}_\infty$ with càdlàg sample paths. The following propositions are equivalent:*

(i) Π *is strongly exchangeable.*

(ii) Π *has the projective Markov property, i.e.* $\Pi^n := (\Pi(t)_{|[n]}, t \geq 0)$ *is a Markov chain for all* $n \in \mathbb{N}$.

Remark 4.3. Crane and Towsner [28, Theorem 4.26] show that the projective Markov property is equivalent to the Feller property for exchangeable Markov process taking values in a Fraïssé space (i.e. a space satisfying general "stability and universality" assumptions [see 28, Definitions 4.4 to 4.11]). In particular the space of partitions and the space of nested partitions are Fraïssé spaces (the argument essentially being the existence of so-called universal elements $\pi^\star$ defined in Section 4.2), so for the processes we consider, strong exchangeability is equivalent to the Feller property.

Proof. $(i) \Rightarrow (ii)$: Let $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_n^{2,\preceq}$. Fix a universal $\pi^\star \in \mathcal{P}_\infty^{2,\preceq}$ with initial part π. Now take any $\pi_0 \in \mathcal{P}_\infty^{2,\preceq}$ such that $(\pi_0)_{|[n]} = \pi$, and an injection $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\sigma_{|[n]} = \mathrm{id}_{|[n]}$ and $(\pi^\star)^\sigma = \pi_0$. Now we have

$$\begin{aligned}\mathbb{P}_{\pi_0}(\Pi^n \in \cdot) &= \mathbb{P}_{\pi^\star}((\Pi^\sigma)^n \in \cdot) \\ &= \mathbb{P}_{\pi^\star}(\Pi^n \in \cdot),\end{aligned}$$

so this distribution depends only on π, which proves that Π^n is a Markov process. Now the assumption that Π has càdlàg sample paths ensures that the process Π^n stays some positive time in each visited state *a.s.* Therefore Π^n is a continuous-time Markov chain.

$(ii) \Rightarrow (i)$: Let $\sigma : \mathbb{N} \to \mathbb{N}$ be an injection. For $n \in \mathbb{N}$, let τ be a permutation of $\mathbb{N}$ such that $\tau_{|[n]} = \sigma_{|[n]}$. This property implies $(\pi^\tau)_{|[n]} = (\pi^\sigma)_{|[n]}$ for any $\pi \in \mathcal{P}_\infty^{2,\preceq}$. We deduce

$$\begin{aligned}\mathbb{P}_\pi((\Pi^\sigma)^n \in \cdot) &= \mathbb{P}_\pi((\Pi^\tau)^n \in \cdot)\\ &= \mathbb{P}_{\pi^\tau}(\Pi^n \in \cdot)\\ &= \mathbb{P}_{\pi^\sigma}(\Pi^n \in \cdot),\end{aligned}$$

where the last equality is a consequence of the projective Markov property (the distribution of Π^n under $\mathbb{P}_\pi$ depends only on the initial segment $\pi_{|[n]}$). Since it is true for all n, we have $\mathbb{P}_\pi(\Pi^\sigma \in \cdot) = \mathbb{P}_{\pi^\sigma}(\Pi \in \cdot)$, which proves the property of strong exchangeability. □

Corollary 4.4. *A nested fragmentation as defined by Definition 4.1 satisfies the assumptions of Definition 4.1'.*

Proof. Consider a nested fragmentation process $\Pi = (\zeta, \xi)$ satisfying Definition 4.1. Note that (i) of Definition 4.1 implies that Π satisfies the projective Markov property. Fix any initial state $\pi = (\zeta, \xi) \in \mathcal{P}_\infty^{2,\preceq}$ and an integer $n \in \mathbb{N}$, and write $\xi_{|[n]} = \{\xi_1, \xi_2, \ldots, \xi_k\}$, for some $1 \le k \le n$. Now define bijections $\sigma_i : [\#\xi_i] \to \xi_i$ for each integer $1 \le i \le k$. Assumption (ii) and the projective Markov property imply that the processes

$$\Big((\Pi^{\sigma_i}(t),\, t \ge 0),\ 1 \le i \le k\Big)$$

are mutually independent under $\mathbb{P}_\pi$, and such that Π^{σ_i} has distribution $\Pi^{\#\xi_i}$ started from $\pi^{\sigma_i} = (\zeta^{\sigma_i}, \mathbf{1})$. Independent continuous-time Markov chains have distinct jump times almost surely, so in particular the first jump time T_1^n of Π^n started from $\pi_{|[n]}$ is the first jump time of some Π^{σ_i}, for a unique i. So there is a unique block $B \in \xi(0)_{|[n]} = \xi(T_1^n-)_{|[n]}$ such that $\Pi(T_1^n-)_{|B} \neq \Pi(T_1^n)_{|B}$. By induction and the Markov property applied to successive jumps times $T_1^n, T_2^n, \ldots$ of the Markov chain Π^n, it is clear that almost surely, for all $t \ge 0$ such that $\Pi^n(t-) \neq \Pi^n(t)$, there is a unique block $B \in \xi^n(t-)$ such that $\Pi(t-)_{|B} \neq \Pi(t)_{|B}$. Since this is true for all $n \in \mathbb{N}$, the outer branching property as described in (ii') holds.

It is a result of the univariate theory of fragmentations [10], that (iii) implies (iii'). □

The next proposition is the direct consequence of the projective Markov property in the space $\mathcal{P}_\infty^{2,\preceq}$. It is essentially Lemma 3.7, from which the proof is easily adapted, the argument being entirely independent from any monotonicity (coalescence or fragmentation) assumption.

Proposition 4.5. *Let $\Pi = (\Pi(t), t \ge 0)$ be a stochastic process with values in $\mathcal{P}_\infty^{2,\preceq}$ which satisfies the **projective Markov property**. Then Π is a Markov process, whose distribution is characterized by a transition kernel K from $\mathcal{P}_\infty^{2,\preceq}$ to $\mathcal{P}_\infty^{2,\preceq}$ (i.e. $K_\pi(\,\cdot\,)$ is a nonnegative measure on $\mathcal{P}_\infty^{2,\preceq}$ for all $\pi \in \mathcal{P}_\infty^{2,\preceq}$ and $\pi \mapsto K_\pi(B)$ is measurable for any B Borel set of $\mathcal{P}_\infty^{2,\preceq}$) such that*

- *for all* $\pi \in \mathcal{P}_\infty^{2,\preceq}$*, we have* $K_\pi(\{\pi\}) = \infty$*,*
- *for all* $\pi \in \mathcal{P}_\infty^{2,\preceq}$*,* $n \in \mathbb{N}$ *and* $\pi' \in \mathcal{P}_n^{2,\preceq} \setminus \{\pi_{|[n]}\}$*, the Markov chain* Π^n *has a transition rate from* $\pi_{|[n]}$ *to* π' *equal to*

$$q^n_{\pi,\pi'} = K_\pi\left(r_n^{-1}(\{\pi'\})\right) < \infty,$$

where $r_n(\,\cdot\,) = \cdot_{|[n]}$ *denotes the restriction operation.*

This kernel K *will be called the* ***characteristic kernel*** *of the process* Π*. Furthermore, if* Π *is exchangeable, then* K *is strongly exchangeable, in the sense that for any* $\pi \in \mathcal{P}_\infty^{2,\preceq}$ *and any injection* $\sigma : \mathbb{N} \to \mathbb{N}$*, we have*

$$K_{\pi^\sigma} = K_\pi^\sigma.$$

Proof. See Lemma 3.7. □

Remark 4.6. Note that the transition rates of the Markov chains Π^n are given by the collection of σ-finite measures $K_\pi(\,\cdot \cap \mathcal{P}_\infty^{2,\preceq} \setminus \{\pi\})$, for $\pi \in \mathcal{P}_\infty^{2,\preceq}$. The value $K_\pi(\{\pi\})$ is irrelevant for the distribution of the process Π, and for uniqueness, we set $K_\pi(\{\pi\}) = \infty$, whereas it is conventional for a transition kernel that this value is taken to be 0. However, setting this value to be infinite is necessary so that strong exchangeability $K_{\pi^\sigma} = K_\pi^\sigma$ holds in general for all *injections* $\sigma : \mathbb{N} \to \mathbb{N}$. Indeed, note that if σ is a bijection, then $K_\pi^\sigma(\{\pi^\sigma\}) = K_\pi(\{\pi\})$, but in general, when σ is an injection, one can have $K_\pi^\sigma(\{\pi^\sigma\}) = K_\pi(\{\pi\}) + a$, where $a > 0$. For instance assume – we will see that it is the case for characteristic kernels of nested fragmentation – that K is such that if $\pi_0 = (\zeta, \xi)$ has at least two outer blocks $B \neq B' \in \xi$, then

$$K_{\pi_0}(\{\pi_{|B} \neq (\pi_0)_{|B}\} \cap \{\pi_{|B'} \neq (\pi_0)_{|B'}\}) = 0.$$

Then if $\sigma : \mathbb{N} \to \mathbb{N}$ is an injection with image $\sigma(\mathbb{N}) \subset B$, then one has

$$K_{\pi_0}(\{\pi^\sigma = \pi_0^\sigma\}) \geq K_{\pi_0}(\{\pi_{|B} = (\pi_0)_{|B}\}) \geq K_{\pi_0}(\{\pi = \pi_0\}) + K_{\pi_0}(\{\pi_{|B'} \neq (\pi_0)_{|B'}\}),$$

where $K_{\pi_0}(\{\pi_{|B'} \neq (\pi_0)_{|B'}\})$ may be greater than 0 if K is not trivial.

Let us emphasize that the kernel K essentially gives us the infinitesimal generator of the Markov process Π. Indeed, note that the generator G_n of the continuous-time finite-space Markov chain Π^n is then given by

$$\begin{aligned} G_n f(\pi_{|[n]}) &= \sum_{\pi' \in \mathcal{P}_n^{2,\preceq} \setminus \{\pi_{|[n]}\}} q^n_{\pi,\pi'}(f(\pi') - f(\pi_{|[n]})) \\ &= \int_{\mathcal{P}_\infty^{2,\preceq}} K_\pi(\mathrm{d}\pi')\,(f(\pi'_{|[n]}) - f(\pi_{|[n]})), \end{aligned}$$

for any function $f : \mathcal{P}_n^{2,\preceq} \to \mathbb{R}$ and $\pi \in \mathcal{P}_\infty^{2,\preceq}$. As an obvious consequence, the generator G of the process Π can be applied to any function $g : \mathcal{P}_\infty^{2,\preceq} \to \mathbb{R}$ of the form $g = f \circ r_n$ for some $n \in \mathbb{N}$ and some function $f : \mathcal{P}_n^{2,\preceq} \to \mathbb{R}$, and is written

$$Gg(\pi) = \int_{\mathcal{P}_\infty^{2,\preceq}} K_\pi(\mathrm{d}\pi')\,(g(\pi') - g(\pi)).$$

4.4 Outer branching property

In this section we study the outer branching property – as stated in Definition 4.1' – to analyze the characteristic kernel of nested fragmentations. The aim is to show that it is entirely characterized by a measure on partitions of $\mathbb{N}^2$ satisfying some invariance property.

4.4.1 Simpler kernel

First, the following proposition expresses that the jump rates from initial states with *a single outer block* are sufficient to characterize the whole process.

Proposition 4.7. *Let $\Pi = (\Pi(t), t \geq 0) = ((\zeta(t), \xi(t)), t \geq 0)$ be a strongly exchangeable Markov process with values in $\mathcal{P}_\infty^{2,\preceq}$ and nonincreasing càdlàg sample paths. Write K for its exchangeable characteristic kernel.*

*If Π satisfies the **outer branching property**, then K is characterized by a simpler kernel κ from $\mathcal{P}_\infty$ to $\mathcal{P}_\infty^{2,\preceq}$ which is defined as*

$$\kappa_\zeta(\,\cdot\,) := K_{(\zeta,\mathbf{1})}(\,\cdot\,),$$

where $\mathbf{1}$ denotes the partition of $\mathbb{N}$ with only one block. The simpler kernel is also strongly exchangeable.

The kernel K is determined by κ in the following way: fix $\pi_0 = (\zeta, \xi) \in \mathcal{P}_\infty^{2,\preceq}$ and for simplicity suppose that all the blocks of ξ are infinite. For all $B \in \xi$, define the injection $\sigma_B : \mathbb{N} \to \mathbb{N}$ as the unique increasing map whose image is B, and $\tau_B : B \to \mathbb{N}$ such that $\sigma_B \circ \tau_B = \mathrm{id}_B$. By definition, $(\pi_0)^{\sigma_B}$ is of the form $(\zeta_B, \mathbf{1})$, with $\zeta_B = \zeta^{\sigma_B}$. Now define f_B as the function which maps $\pi \in \mathcal{P}_\infty^{2,\preceq}$ to the unique $\omega \in \mathcal{P}_\infty^{2,\preceq}$ such that

- $\omega \preceq (\{B, \mathbb{N} \setminus B\}, \{B, \mathbb{N} \setminus B\})$,
- $\omega_{|B} = \pi^{\tau_B}$ *and* $\omega_{|\mathbb{N}\setminus B} = (\pi_0)_{|\mathbb{N}\setminus B}$.

Then for any Borel set $A \subset \mathcal{P}_\infty^{2,\preceq}$, we have

$$K_{\pi_0}(A) = \sum_{B \in \xi} \kappa_{\zeta_B}(\{f_B(\pi) \in A\}). \tag{4.5}$$

Remark 4.8.

- This proposition shows how K_{π_0} is expressed in terms of the kernel κ only for $\pi_0 = (\zeta, \xi)$ such that all the blocks of ξ are infinite. In fact this is enough to characterize K entirely since if π_0 does not satisfy this property, there exists a nested partition $\pi_0' = (\zeta', \xi')$, where ξ' has infinite blocks, and an injection $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\pi_0 = (\pi_0')^\sigma$. Then we have $K_{\pi_0} = K^\sigma_{\pi_0'}$, which is determined by κ.
- This result implies that different outer blocks undergo independent fragmentations, in other words a nested fragmentation (recall that we only assume Definition 4.1') satisfies (ii) of Definition 4.1. Indeed, one interprets the sum (4.5) as: independently for each block $B \in \xi$, $(\pi_0)_{|B}$ is replaced by π^{τ_B} at rate $\kappa_{\zeta_B}(\pi \in \cdot)$, which is a measure which depends only on $(\pi_0)_{|B}$.

Proof. First note that the fact that Π has decreasing sample paths implies that for any $\pi_0 \in \mathcal{P}_\infty^{2,\preceq}$, the support of the measure K_{π_0} is included in $\{\pi \preceq \pi_0\}$. Indeed, since $\{\pi \preceq \pi_0\} = \cap_{n\geq 1}\{\pi_{|[n]} \preceq (\pi_0)_{|[n]}\}$, we have

$$K_{\pi_0}(\{\pi \not\preceq \pi_0\}) = \lim_{n\to\infty} K_{\pi_0}(\pi_{|[n]} \not\preceq (\pi_0)_{|[n]}),$$

where for any $n \geq 1$, the right-hand side is equal to the (finite) transition rate of the Markov chain Π^n from $(\pi_0)_{|[n]}$ to any π for which $\pi \not\preceq (\pi_0)_{|[n]}$. But Π^n is a decreasing process by assumption, so this rate is zero, so we conclude

$$K_{\pi_0}(\pi \not\preceq \pi_0) = 0 \tag{4.6}$$

Using the same argument, it is clear that the outer branching property implies that for any $\pi_0 = (\zeta, \xi) \in \mathcal{P}_\infty^{2,\preceq}$, we have

$$K_{\pi_0}\Big(\bigcup_{B_1\neq B_2\in\xi} \{\pi_{|B_1} \neq (\pi_0)_{|B_1} \text{ and } \pi_{|B_2} \neq (\pi_0)_{|B_2}\}\Big) = 0. \tag{4.7}$$

Now without loss of generality (see Remark 4.8), suppose that all the blocks of ξ are infinite, and let us define for all $B \in \xi$, the maps σ_B, τ_B and f_B as in the proposition. Equations (4.6) and (4.7) imply that for any $B \in \xi$, on the event $\{\pi_{|B} \neq (\pi_0)_{|B}\}$, we have

$$\pi = f_B(\pi^{\sigma_B}) \quad K_{\pi_0}\text{-a.e.},$$

where f_B is the map defined in the proposition. Then to show (4.5) for any Borel set $A \subset \mathcal{P}_\infty^{2,\preceq} \setminus \{\pi_0\}$, we have

$$\begin{aligned}
K_{\pi_0}(A) &= K_{\pi_0}(\cup_{B\in\xi}(A \cap \{\pi_{|B} \neq (\pi_0)_{|B}\})) \\
&= \sum_{B\in\xi} K_{\pi_0}(A \cap \{\pi_{|B} \neq (\pi_0)_{|B}\}) \\
&= \sum_{B\in\xi} K_{\pi_0}(\{f_B(\pi^{\sigma_B}) \in A\} \cap \{\pi^{\sigma_B} \neq (\pi_0)^{\sigma_B}\}) \\
&= \sum_{B\in\xi} K_{(\pi_0)^{\sigma_B}}(\{f_B(\pi) \in A\} \cap \{\pi \neq (\pi_0)^{\sigma_B}\}),
\end{aligned}$$

where we use the strong exchangeability of the kernel K in the last line. Now, note that for every $B \in \xi$, by definition of f_B we have $\{f_B(\pi) \neq \pi_0\} = \{\pi \neq (\pi_0)^{\sigma_B}\}$, therefore $\{f_B(\pi) \in A\} \subset \{\pi \neq (\pi_0)^{\sigma_B}\}$, so one can simply rewrite

$$K_{\pi_0}(A) = \sum_{B\in\xi} K_{(\pi_0)^{\sigma_B}}(\{f_B(\pi) \in A\}),$$

In general, if A is a Borel subset of $\mathcal{P}_\infty^{2,\preceq}$ with $\pi_0 \in A$, we have $K_{\pi_0}(A) = \infty$, and for each $B \in \xi$, $K_{(\pi_0)^{\sigma_B}}(\{f_B(\pi) \in A\}) \geq K_{(\pi_0)^{\sigma_B}}(\{f_B(\pi) = \pi_0\}) = K_{(\pi_0)^{\sigma_B}}(\{\pi = (\pi_0)^{\sigma_B}\}) = \infty$, so the equality still holds. Now by definition of σ_B, $(\pi_0)^{\sigma_B}$ is of the form $(\zeta_B, \mathbf{1})$, which concludes the proof that K_{π_0} can be expressed with the simpler kernel κ. Finally, by definition, it is clear that κ inherits strong exchangeability from K. □

Now, to further analyze the *simplified* characteristic kernel κ of an nested fragmentation, we need to introduce some tools, reducing the problem to study exchangeable (with respect to a particular set of injections M) partitions of $\mathbb{N}^2$.

4.4.2 M-invariant measures

Let M be the monoid of functions $\mathbb{N}^2 \to \mathbb{N}^2$ consisting of injective maps of the form

$$(i,j) \longmapsto (\sigma(i), \sigma_i(j)),$$

where σ and $\sigma_1, \sigma_2, \dots$ are injections $\mathbb{N} \to \mathbb{N}$. Let us write π_{R} for the *rows partition* $\{\{(i,j),\, j \geq 1\},\, i \geq 1\} \in \mathcal{P}_{\mathbb{N}^2}$, which has the property that an injection $\tau : \mathbb{N}^2 \to \mathbb{N}$ is in M if and only if $\pi_{\mathrm{R}}^{\tau} = \pi_{\mathrm{R}}$.

Note that in $\mathcal{P}_\infty$ any universal element π has the property that κ_π characterize κ entirely, but there is no natural choice for π. The reason for studying partitions of $\mathbb{N}^2$ is that the rows partition π_{R} is a natural universal element of $\mathcal{P}_{\mathbb{N}^2}$. The following proposition shows that one can make sense of a measure essentially defined as "$\kappa_{\pi_{\mathrm{R}}}$", which then characterize κ and therefore the distribution of a nested fragmentation.

Proposition 4.9. *Let κ be a strongly exchangeable kernel from $\mathcal{P}_\infty$ to $\mathcal{P}_\infty^{2,\preceq}$, and let π_0 denote a universal element of $\mathcal{P}_\infty$, i.e. a partition of $\mathbb{N}$ with infinitely many infinite blocks (and no finite block). Choose a bijection $\sigma : \mathbb{N}^2 \to \mathbb{N}$ such that $\pi_0^\sigma = \pi_{\mathrm{R}}$.*

Then $\mu := \kappa_{\pi_0}^\sigma$ is a measure on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$ which is **M-*invariant***, *in the sense that for all $\tau \in M$, $\mu = \mu^\tau$. Moreover, μ does not depend on π_0 or σ and the mapping $\kappa \mapsto \mu$ is bijective from the set of strongly exchangeable kernels to the set of M-invariant measures on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$.*

Thinking of κ as the jump kernel of a nested fragmentation process, one can see this measure μ as the measure giving the infinitesimal jump rates from the nested partition $(\pi_{\mathrm{R}}, \mathbf{1})$, where each row of $\mathbb{N}^2$ is an inner block.

Proof. Fix $\tau \in M$ and a Borel set $A \subset \mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$. We need to prove $\mu(\pi^\tau \in A) = \mu(A)$. Consider $\varphi = \sigma \circ \tau \circ \sigma^{-1}$. This map satisfies $\varphi \circ \sigma = \sigma \circ \tau$ and $\pi_0^\varphi = \pi_0$, so we have

$$\begin{aligned}
\mu(\pi^\tau \in A) &= \kappa_{\pi_0}(\pi^{\sigma\circ\tau} \in A) \\
&= \kappa_{\pi_0}(\pi^{\varphi\circ\sigma} \in A) \\
&= \kappa_{\pi_0^\varphi}(\pi^\sigma \in A) \\
&= \mu(A).
\end{aligned}$$

This proves that μ is M-invariant. Let us now prove that μ does not depend on π_0 or σ: fix $\pi_1, \pi_2 \in \mathcal{P}_\infty$ (both with infinitely many infinite blocks and no finite block) and σ_1, σ_2 bijections from $\mathbb{N}^2$ to $\mathbb{N}$ such that $\pi_i^{\sigma_i} = \pi_{\mathrm{R}}$. We need to show

$$\kappa_{\pi_1}(\pi^{\sigma_1} \in \cdot\,) = \kappa_{\pi_2}(\pi^{\sigma_2} \in \cdot\,).$$

Let φ be a bijection such that $\pi_1^\varphi = \pi_2$. Note that $\pi_{\mathrm{R}}^{\sigma_2^{-1}\circ\varphi^{-1}\circ\sigma_1} = \pi_2^{\varphi^{-1}\circ\sigma_1} = \pi_1^{\sigma_1} = \pi_{\mathrm{R}}$, i.e.

$\sigma_2^{-1} \circ \varphi^{-1} \circ \sigma_1 \in M$. Now we have

$$\begin{aligned}\kappa_{\pi_1}(\pi^{\sigma_1} \in \cdot) &= \kappa_{\pi_1}\left((\pi^{\varphi})^{\varphi^{-1}\circ\sigma_1} \in \cdot\right)\\ &= \kappa_{\pi_2}\left(\pi^{\varphi^{-1}\circ\sigma_1} \in \cdot\right)\\ &= \kappa_{\pi_2}\left((\pi^{\sigma_2})^{\sigma_2^{-1}\circ\varphi^{-1}\circ\sigma_1} \in \cdot\right)\\ &= \kappa_{\pi_2}\left(\pi^{\sigma_2} \in \cdot\right),\end{aligned}$$

where the last equality follows from the M-invariance of $\kappa_{\pi_2}(\pi^{\sigma_2} \in \cdot)$. So μ is well defined and depends only on κ.

We now prove that $\kappa \mapsto \mu$ is bijective. For any injection $\sigma : \mathbb{N} \to \mathbb{N}^2$, we write 2σ for the map

$$2\sigma : \begin{cases}\mathbb{N} & \longrightarrow \mathbb{N}^2\\ n & \longmapsto 2\sigma(n) = (2i, 2j) \quad \text{where } \sigma(n) = (i,j).\end{cases}$$

Note that for any injection $\sigma : \mathbb{N} \to \mathbb{N}^2$, we have $\pi_{\mathrm{R}}^{\sigma} = \pi_{\mathrm{R}}^{2\sigma}$. Now let σ_1, σ_2 be any two injections such that $\pi_{\mathrm{R}}^{\sigma_1} = \pi_{\mathrm{R}}^{\sigma_2}$. Then there exists a $\tau \in M$ such that

$$\tau \circ \sigma_1 = 2\sigma_2.$$

Indeed one such τ can be defined in the following way. First let us define an injection $\varphi : \mathbb{N} \to \mathbb{N}$, which will serve as a mapping for rows. For any $i \in \mathbb{N}$, there are two possibilities:

- either there is a $j \in \mathbb{N}$ such that $(i,j) \in \mathrm{im}(\sigma_1)$, and then there is an even integer $i' \in \mathbb{N}$ such that $2\sigma_2(\sigma_1^{-1}(i,j)) = (i',k)$ for some $k \in \mathbb{N}$. This number i' does not depend on j because of the fact that $\pi_{\mathrm{R}}^{\sigma_1} = \pi_{\mathrm{R}}^{\sigma_2}$. Indeed if $j_1, j_2 \in \mathbb{N}$ are such that $(i,j_1), (i,j_2) \in \mathrm{im}(\sigma_1)$, then by definition $\sigma^{-1}(i,j_1)$ and $\sigma^{-1}(i,j_2)$ belong to the same block of $\pi_{\mathrm{R}}^{\sigma_1} = \pi_{\mathrm{R}}^{\sigma_2}$, and so $\sigma_2(\sigma^{-1}(i,j_1))$ and $\sigma_2(\sigma^{-1}(i,j_2))$ belong to the same block of π_{R}. So in that case we can define $\varphi(i) := i'$.
- or $\mathrm{im}(\sigma_1) \cap \{(i,j),\ j \geq 1\} = \varnothing$, and then we define $\varphi(i) = 2i - 1$.

The map φ is a well-defined injection, and we may now define

$$\tau : \begin{cases}(i,j) \in \mathrm{im}(\sigma_1) & \longmapsto 2\sigma_2(\sigma_1^{-1}(i,j))\\ (i,j) \notin \mathrm{im}(\sigma_1) & \longmapsto (\varphi(i), 2j-1)\end{cases}$$

It is easy to check that $\tau \in M$ and that $\tau \circ \sigma_1 = 2\sigma_2$. We can now fix μ an M-invariant measure on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$. Consider a partition $\pi_0 \in \mathcal{P}_\infty$ and an injection $\sigma_0 : \mathbb{N} \to \mathbb{N}^2$ such that $\pi_{\mathrm{R}}^{\sigma_0} = \pi_0$. Now for any other σ_1 such that $\pi_{\mathrm{R}}^{\sigma_1} = \pi_0$, let $\tau \in M$ be such that $\tau \circ \sigma_1 = 2\sigma_0$. By M-invariance of μ, we have

$$\begin{aligned}\mu(\pi^{\sigma_1} \in \cdot) &= \mu(\pi^{\tau\circ\sigma_1} \in \cdot)\\ &= \mu(\pi^{2\sigma_0} \in \cdot).\end{aligned}$$

Therefore this measure does not depend on σ_1 but only on π_0, so we may define

$$\kappa_{\pi_0} := \mu(\pi^{\sigma_0} \in \cdot),$$

which is a measure on $\mathcal{P}_\infty^{2,\preceq}$, for all π_0. Now it remains to check that for any injection $\sigma : \mathbb{N} \to \mathbb{N}$, we have $\kappa_{\pi_0}^\sigma = \kappa_{\pi_0^\sigma}$. But if $\pi_{\mathrm{R}}^{\sigma_0} = \pi_0$, then $\pi_{\mathrm{R}}^{\sigma_0 \circ \sigma} = \pi_0^\sigma$, so

$$\begin{aligned}\kappa_{\pi_0}^\sigma &= \mu((\pi^{\sigma_0})^\sigma \in \cdot\,) \\ &= \mu(\pi^{\sigma_0 \circ \sigma} \in \cdot\,) \\ &= \kappa_{\pi_0^\sigma},\end{aligned}$$

so κ is a strongly exchangeable kernel from $\mathcal{P}_\infty$ to $\mathcal{P}_\infty^{2,\preceq}$, and it is easy to check that the M-invariant measure associated with κ is μ. □

Note that for K a characteristic kernel of a nested fragmentation, we have set (see Remark 4.6) $K_\pi(\{\pi\}) = \infty$ for any $\pi \in \mathcal{P}_\infty^{2,\preceq}$, which implies that $\mu(\{(\pi_{\mathrm{R}}, \mathbf{1})\}) = \infty$ for the corresponding M-invariant measure. This is only technical and for our processes this value $\mu(\{(\pi_{\mathrm{R}}, \mathbf{1})\})$ has no relevance. Therefore we will from now on abuse notation and identify M-invariant measures on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$ with their restriction to $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq} \setminus \{(\pi_{\mathrm{R}}, \mathbf{1})\}$. More precisely, in the rest of the article, we *extend the definition of M-invariance* to all measures μ on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$ such that for all $\tau \in M$, μ and μ^τ coincide on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq} \setminus \{(\pi_{\mathrm{R}}, \mathbf{1})\}$. As such, we will now *only* consider M-invariant measures μ satisfying $\mu\left(\{(\pi_{\mathrm{R}}, \mathbf{1})\}\right) = 0$.

Putting together Proposition 4.7 and Proposition 4.9 gives us:

Theorem 4.10. *Let $\Pi = (\Pi(t), t \geq 0)$ be a nested fragmentation process. Then its distribution is characterized by a unique M-invariant measure μ on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$ satisfying*

$$\begin{aligned} &\mu\left(\pi \not\preceq (\pi_{\mathrm{R}}, \mathbf{1})\right) = 0 \\ \text{and } \forall n \in \mathbb{N}, \qquad &\mu\left(\pi_{|[n]^2} \neq (\pi_{\mathrm{R}}, \mathbf{1})_{|[n]^2}\right) < \infty. \end{aligned} \tag{4.8}$$

The characterization is in the sense that for any $\pi_0, \pi_1 \in \mathcal{P}_\infty$ with infinitely many infinite blocks, for any Borel sets $A \subset \mathcal{P}_{\mathbb{N}^2}^{2,\preceq} \setminus \{(\pi_{\mathrm{R}}, \mathbf{1})\}$ and $B \subset \mathcal{P}_\infty^{2,\preceq} \setminus \{(\pi_1, \mathbf{1})\}$,

$$\mu(A) = \kappa_{\pi_0}^{\sigma_0}(A) \quad \text{ and } \quad \kappa_{\pi_1}(B) = \mu^{\sigma_1}(B),$$

where κ is the simplified characteristic kernel of Π, $\sigma_0 : \mathbb{N}^2 \to \mathbb{N}$ is any injection such that $\pi_0^{\sigma_0} = \pi_{\mathrm{R}}$ and $\sigma_1 : \mathbb{N} \to \mathbb{N}^2$ is any injection such that $\pi_{\mathrm{R}}^{\sigma_1} = \pi_1$.

Conversely, for any such measure μ, there is a strongly exchangeable Markov process with values in $\mathcal{P}_\infty^{2,\preceq}$, nonincreasing càdlàg sample paths and the outer branching property with characteristic measure μ.

Remark 4.11. An explicit construction for the converse part of the theorem is described in the next section (Lemma 4.12).

4.4.3 Poissonian construction

Consider μ an M-invariant measure on $\mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$ satisfying (4.8), and let Λ be a Poisson point process on $\mathbb{N} \times [0, \infty) \times \mathcal{P}_{\mathbb{N}^2}^{2,\preceq}$ with intensity $\# \otimes \mathrm{d}t \otimes \mu$, where $\#$ denotes the counting measure and $\mathrm{d}t$ the Lebesgue measure.

Fix $n \in \mathbb{N}$. Because of (4.8), the points $(k,t,\pi) \in \Lambda$ such that $k \leq n$ and $\pi_{|[n]^2} \neq (\pi_{\mathrm{R}}, \mathbf{1})_{|[n]^2}$ can be numbered

$$(k_i^n, t_i^n, \pi_i^n, i \geq 1) \quad \text{with } t_1^n < t_2^n < \dots \quad \text{and } t_i^n \xrightarrow[i\to\infty]{} \infty.$$

Fix any initial value $\pi_0 \in \mathcal{P}_\infty^{2,\preceq}$. Let us define a process $(\Pi_i^n, i \geq 0)$ with values in $\mathcal{P}_{[n]}^{2,\preceq}$, by $\Pi_0^n = (\pi_0)_{|[n]}$ and by induction, conditional on $\Pi_i^n = (\zeta, \xi)$:

- if ξ has less than k_{i+1}^n blocks, then set $\Pi_{i+1}^n := \Pi_i^n$
- if ξ has a k_{i+1}^n-th block, say B, then let $\tau : B \to [n]^2$ be the injection such that $\tau(k) = (i', j')$ iff $k \in B$ is the j'-th element of the i'-th block of $\zeta_{|B}$.

  Then define Π_{i+1}^n as the only element $\pi \in \mathcal{P}_n^{2,\preceq}$ such that $\pi \preceq \Pi_i^n$, $\pi_{|B} = (\pi_i^n)^\tau$ and $\pi_{|[n]\setminus B} = (\Pi_i^n)_{|[n]\setminus B}$.

Now we define the continuous-time processes $(\Pi^n(t), t \geq 0)$ by

$$\Pi^n(t) := \Pi_i^n \quad \text{ iff } t \in [t_{i-1}^n, t_i^n).$$

Lemma 4.12. *The processes Π^n built from this Poissonian construction are consistent in the sense that we have for all $m \geq n \geq 1$ and $t \geq 0$,*

$$\Pi^m(t)_{|[n]} = \Pi^n(t).$$

Therefore, for all $t \geq 0$, there is a unique random variable $\Pi(t)$ with values in $\mathcal{P}_\infty^{2,\preceq}$ such that $\Pi(t)_{|[n]} = \Pi^n(t)$ for all n, and the process $(\Pi(t), t \geq 0)$ is a strongly exchangeable Markov process with càdlàg, nonincreasing sample paths, satisfying the outer branching property, and whose characteristic M-invariant measure is μ.

Proof. Choose an integer $n \in \mathbb{N}$ and consider the variable $(k_1^{n+1}, t_1^{n+1}, \pi_1^{n+1})$. It is clear from the definition that $(\Pi_0^{n+1})_{|[n]} = \Pi_0^n$. Now let us show that $(\Pi_1^{n+1})_{|[n]} = \Pi^n(t_1^{n+1})$.

We distinguish two cases:
1) If $t_1^{n+1} = t_1^n$, then we have necessarily $k_1^{n+1} = k_1^n \leq n$ and $(\pi_1^{n+1})_{|[n]^2} = (\pi_1^n)_{|[n]^2} \neq (\pi_{\mathrm{R}}, \mathbf{1})_{|[n]^2}$. Let us write $\Pi_0^{n+1} = (\zeta^{n+1}, \xi^{n+1})$ and $\Pi_0^n = (\zeta^n, \xi^n)$. Since $(\Pi_0^{n+1})_{|[n]} = \Pi_0^n$, it is clear that the k_1^n-th block of ξ^{n+1} includes the k_1^n-th block of ξ^n, and may at most contain one other element, the integer $n+1$. In other words we have

$$B^{n+1} \cap [n] = B^n,$$

where B^{n+1} and B^n denote those two blocks. Now let us write τ^{n+1}, τ^n for the respective injections in $\mathbb{N}^2$ defined in the construction. Because we defined the injections according to the ordering of the blocks of ζ and with the natural order on $\mathbb{N}$, it is clear that

$$\tau^{n+1}_{|B^n} = \tau^n.$$

Therefore we deduce $((\pi_1^n)^{\tau^{n+1}})_{|B^n} = (\pi_1^n)^{\tau^n}$, which allows us to conclude $(\Pi_1^{n+1})_{|[n]} = \Pi_1^n = \Pi^n(t_1^{n+1})$.

2) If $t_1^{n+1} < t_1^n$, then we have to further distinguish two possibilities:

a) $k_1^{n+1} = n+1$. In that case the $(n+1)$-th block of ξ^{n+1} can either be empty or the singleton $\{n+1\}$. Then by definition, we necessarily have $\Pi_1^{n+1} = \Pi_0^{n+1}$, so we can conclude $(\Pi_1^{n+1})_{|[n]} = \Pi_0^n = \Pi^n(t_1^{n+1})$.
b) $k_1^{n+1} \leq n$, and then necessarily $(\pi_1^{n+1})_{|[n]^2} = (\pi_{\mathrm{R}}, \mathbf{1})_{|[n]^2}$. In that case, let B be the k_1^{n+1}-th block of ξ and $\tau : B \to [n+1]^2$ the injective map defined in the construction. By definition, we have $(\pi_{\mathrm{R}}, \mathbf{1})^\tau = (\zeta, \xi)_{|B}$. Also by definition of τ, for any $k \leq n$, we have $\tau(k) \in [n]^2$. Therefore, we can conclude that

$$((\pi_1^{n+1})^\tau)_{|B\cap[n]} = ((\pi_1^{n+1})_{|[n]^2})^{\tau_{|B\cap[n]}} = (\pi_{\mathrm{R}}, \mathbf{1})^{\tau_{|B\cap[n]}} = (\zeta, \xi)_{|B\cap[n]}.$$

This shows that $(\Pi_1^{n+1})_{|[n]} = (\Pi_0^{n+1})_{|[n]}$, which allows us to conclude $(\Pi_1^{n+1})_{|[n]} = \Pi_0^n = \Pi^n(t_1^{n+1})$.

By induction and the strong Markov property of the Poisson point process Λ, this proves that $(\Pi_i^{n+1})_{|[n]} = \Pi^n(t_i^{n+1})$ for all $i \geq 1$, so $\Pi^{n+1}(t)_{|[n]} = \Pi^n(t)$ for all $t \geq 0$, which concludes the first part of the proof.

It remains to show that the process $(\Pi(t), t \geq 0)$ is a strongly exchangeable Markov process with the outer branching property, and whose characteristic M-invariant measure is μ.

First, notice that from the construction, we deduce immediately that for any n, Π^n is a Markov chain, and at any jump time t_i^n, the partitions Π_{i-1}^n and Π_i^n differ at most on one block of ξ, where $\Pi_{i-1}^n = (\zeta, \xi)$. Therefore the distribution of the Markov chain Π^n is given by the transition rates of the form

$$q^n_{\pi_0,\pi_1},$$

with $\pi_0 = (\zeta, \xi) \in \mathcal{P}_\infty^{2,\preceq}$, and with $\pi_1 \preceq (\pi_0)_{|[n]}$ such that, for some $B \in \xi_{|[n]}$, $(\pi_1)_{|[n]\setminus B} = (\pi_0)_{|[n]\setminus B}$ and $(\pi_1)_{|B} \prec (\pi_0)_{|B}$. Now for such π_0, π_1, write $\tau : B \to \mathbb{N}^2$ for the injection such that $\tau(k) = (i,j)$ iff k is the j-th element of the i-th block of $\zeta_{|B}$. By elementary properties of Poisson point processes we have

$$q^n_{\pi_0,\pi_1} = \mu\left(\pi^\tau = (\pi_1)_{|B}\right), \tag{4.9}$$

Now recall from Proposition 4.2 that since Π satisfies the projective Markov property and is exchangeable (this is immediate from the M-invariance of μ), Π is strongly exchangeable, with a characteristic kernel K such that with the same notation as in (4.9),

$$K_{\pi_0}(\pi_{|[n]} = \pi_1) = q^n_{\pi_0,\pi_1}. \tag{4.10}$$

Now the outer branching property is immediately deduced from the construction of the process, where it is clear that at any jump time, at most one block of the coarser partition is involved. Therefore by Proposition 4.7, the law of Π is characterized by the simpler kernel κ defined by $\kappa_\zeta = K_{(\zeta,\mathbf{1})}$, for $\zeta \in \mathcal{P}_\infty$. Now putting this together with (4.10) and (4.9), since the coarsest partition $\mathbf{1}_{[n]}$ only contains one block $B = [n]$, we have simply

$$\kappa_\zeta(\pi_{|[n]} = \pi_1) = \mu\left((\pi^\tau)_{|[n]} = \pi_1\right),$$

where τ is an injection such that $\pi_{\mathrm{R}}^\tau = \zeta$. In other words with these definitions, the measures κ_ζ and μ^τ coincide on $\mathcal{P}_\infty^{2,\preceq} \setminus \{(\zeta, \mathbf{1})\}$, which shows that μ is the characteristic M-invariant measure of the process Π. □

4.5 Inner branching property

In the previous section we only exploited the outer branching property of Definition 4.1'. This section will instead focus on the inner branching property, which will allow us to further the analysis of the M-invariant measure μ appearing in Theorem 4.10. To introduce the next theorem and main result of this article, let us first give examples of M-invariant measures that give rise to the types of transitions already discussed in Section 4.2.3.

4.5.1 Some examples

Pure erosion For $i \geq 1$, let $\xi_{\text{out}}^{(i)}$ be the partition of $\mathbb{N}^2$ with two blocks such that one of them is the i-th line $\{i\} \times \mathbb{N}$, i.e.

$$\xi_{\text{out}}^{(i)} := \{\{i\} \times \mathbb{N},\ \mathbb{N}^2 \setminus (\{i\} \times \mathbb{N})\}$$

and define the outer erosion measure $\mathfrak{e}^{\text{out}} := \sum_{i \geq 1} \delta(\pi_{\mathrm{R}}, \xi_{\text{out}}^{(i)})$, where for readability we denote without subscripts $\delta(\zeta, \xi)$ the Dirac measure on (ζ, ξ).

Similarly, for $i, j \geq 1$, we define

$$\zeta_{\text{in}}^{(i,j)} := \{\{(i,j)\}\} \cup \{(\{i\} \times \mathbb{N}) \setminus \{(i,j)\}\} \cup \{\{k\} \times \mathbb{N},\ k \geq 1, k \neq i\},$$
$$\xi_{\text{in}}^{(i,j)} := \{\{(i,j)\},\ \mathbb{N}^2 \setminus \{(i,j)\}\},$$

and the inner erosion measures

$$\mathfrak{e}^{\text{in},1} := \sum_{i,j \geq 1} \delta(\zeta_{\text{in}}^{(i,j)}, \mathbf{1}) \quad \text{and} \quad \mathfrak{e}^{\text{in},2} := \sum_{i,j \geq 1} \delta(\zeta_{\text{in}}^{(i,j)}, \xi_{\text{in}}^{(i,j)}).$$

Now, given three real numbers $c_{\text{out}}, c_{\text{in},1}, c_{\text{in},2} \geq 0$, the M-invariant measure $\mu = c_{\text{out}}\mathfrak{e}^{\text{out}} + c_{\text{in},1}\mathfrak{e}^{\text{in},1} + c_{\text{in},2}\mathfrak{e}^{\text{in},2}$ clearly satisfies (4.8), so by Theorem 4.10 there exists a fragmentation process having μ as M-invariant measure.

From the construction, we see that the rates of such a process can be described informally as follows:

- any inner block erodes out of its outer block at rate c_{out}, i.e. it does not fragment but forms, on its own, a new outer block.
- any integer erodes out of its inner block at rate $c_{\text{in},1}$, forming a singleton inner block, within the same outer block as its parent.
- any integer erodes out of its inner and outer block at rate $c_{\text{in},2}$, forming singleton inner and outer blocks.

Outer dislocation Recall the definition of the space of mass partitions $\mathbf{s} = (s_1, s_2, \ldots) \in \mathscr{S}^{\downarrow}$ and of the measures $\varrho_{\mathbf{s}}$ from Section 4.2.2. We define in a similar way, a collection of probability measure $\widehat{\varrho}_{\mathbf{s}}$ on $\mathcal{P}_{\mathbb{N}^2}^{2,\asymp}$, by constructing $\pi = (\zeta, \xi) \sim \widehat{\varrho}_{\mathbf{s}}$ with the following so-called paintbox procedure:

- for $k \geq 0$, let $t_k := \sum_{k'=1}^{k} s_{k'}$, with $t_0 = 0$ by convention.

- let $U_1, U_2, \dots$ be a sequence of i.i.d. uniform r.v. on $[0,1]$ and define the random partition $\xi \succeq \pi_{\mathrm{R}}$ on $\mathbb{N}^2$ by

$$(i,j) \sim^{\xi} (i',j') \iff i = i' \text{ or } U_i, U_{i'} \in [t_k, t_{k+1}) \text{ for a unique } k \geq 0.$$

- $\widehat{\varrho}_{\mathbf{s}}$ is now defined to be the distribution of the random nested partition $\pi = (\pi_{\mathrm{R}}, \xi)$.

Now for ν_{out} a measure on $\mathscr{S}^{\downarrow}$ satisfying (4.4), we define

$$\widehat{\varrho}_{\nu_{\mathrm{out}}}(\,\cdot\,) := \int_{\mathscr{S}^{\downarrow}} \nu_{\mathrm{out}}(\mathrm{d}\mathbf{s})\, \widehat{\varrho}_{\mathbf{s}}(\,\cdot\,).$$

It is straight-forward to check that $\widehat{\varrho}_{\nu_{\mathrm{out}}}$ is an M-invariant measure measure on $\mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ satisfying (4.8), so there exists a fragmentation process having $\widehat{\varrho}_{\nu_{\mathrm{out}}}$ as M-invariant measure.

In intuitive terms, such a process can be described by saying that the outer blocks independently dislocate *around their inner blocks* with **outer dislocation rate** ν_{out}. In a dislocation event, inner blocks are unchanged, and they are indistinguishable. By construction, each newly created outer block selects a given frequency of inner blocks among those forming the original outer block.

Inner dislocation The upcoming example is the most complex on our list, exhibiting simultaneous inner and outer fragmentations. However, in construction it is very similar to the previous example, and it should pose no difficulties to get a good intuition of the dislocation mechanics.

Let us first formally define a space which will serve as an analog of the space of mass partitions $\mathscr{S}^{\downarrow}$.

Definition 4.13. We define a particular space of *bivariate mass partitions*

$$\mathscr{S}^{\downarrow}_{\preceq} \subset [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}^2} \times [0,1] \times [0,1]^{\mathbb{N}}$$

as the subset consisting of elements $\mathbf{p} = ((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1})$ satisfying the following conditions.

$$\begin{gathered} u_1 \geq u_2 \geq \dots \text{ and } \textstyle\sum_l u_l \leq \bar{u}, \\ \forall k \geq 1,\ s_{k,1} \geq s_{k,2} \geq \dots \text{ and } \textstyle\sum_l s_{k,l} \leq \bar{s}_k, \\ \bar{s}_1 \geq \bar{s}_2 \geq \dots, \\ \bar{u} + \textstyle\sum_k \bar{s}_k \leq 1, \\ \text{if } \bar{s}_k = \bar{s}_{k+1}, \text{ then } (l_0 = \inf\{l \geq 1, s_{k,l} \neq s_{k+1,l}\} < \infty) \Rightarrow (s_{k,l_0} > s_{k+1,l_0}). \end{gathered} \tag{4.11}$$

We claim that $\mathscr{S}^{\downarrow}_{\preceq}$ is Polish with respect to the product topology. Indeed, recall [see e.g. 87, Theorem 2.2.1] that any G_δ subset – i.e. a countable intersection of open sets – of a Polish space is Polish. Now, it is readily checked that every condition in (4.11) is closed in the compact space $X := [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}^2} \times [0,1] \times [0,1]^{\mathbb{N}}$ except the last one, but the subset of X satisfying this condition can be written

$$\bigcap_{k\geq 1} \Big[\{\bar{s}_k \neq \bar{s}_{k+1}\} \cup \Big(\bigcap_{l\geq 1} \{\exists i < l,\ s_{k,i} \neq s_{k+1,i}\} \cup \{s_{k,l} \geq s_{k+1,l}\} \Big) \Big],$$

so finally $\mathscr{S}^{\downarrow}_{\preceq}$ can be written as a countable intersection of open and closed sets in X, which are all G_δ (recall that closed subsets of any *metrizable* space are G_δ). Therefore considering this topology, $\mathscr{S}^{\downarrow}_{\preceq}$ is Polish and we will have no trouble considering measures on $\mathscr{S}^{\downarrow}_{\preceq}$.

Now, given a fixed $i \geq 1$ and $\mathbf{p} = ((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1}) \in \mathscr{S}^{\downarrow}_{\preceq}$, one can define a random element $\pi^{(i)} = (\zeta^{(i)}, \xi^{(i)}) \in \mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ with the following paintbox procedure:

- for $k \geq 0$, define $\bar{t}_k = \bar{u} + \sum_{k'=1}^{k} \bar{s}_{k'}$.
- for $l \geq 0$, define $t_{\star,l} = \sum_{l'=1}^{l} u_{l'}$.
- for $k \geq 1$ and $l \geq 0$, define $t_{k,l} = \bar{t}_{k-1} + \sum_{l'=1}^{l} s_{k,l'}$.
- write $\pi_0 = (\zeta_0, \xi_0)$ for the unique element of $\mathcal{P}^{2,\preceq}_{[0,1]}$ such that the non-dust blocks of ξ_0 are
$$[0, \bar{u}) \text{ and } [\bar{t}_{k-1}, \bar{t}_k),\ k \geq 1,$$
and such that the non-singleton blocks of ζ_0 are
$$[t_{\star,l-1}, t_{\star,l}),\ l \geq 1 \text{ and } [t_{k,l-1}, t_{k,l}),\ k, l \geq 1.$$
- let $(U_j, j \geq 1)$ be an i.i.d. sequence of uniform random variables on $[0, 1]$.
- define the random element $\pi^{(i)} \in \mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ as the unique element $\pi^{(i)} = (\zeta^{(i)}, \xi^{(i)}) \preceq (\pi_{\mathrm{R}}, \mathbf{1})$ such that
  - $(\zeta^{(i)}, \xi^{(i)})_{|(\mathbb{N}\setminus\{i\})\times\mathbb{N}} = (\pi_{\mathrm{R}}, \mathbf{1})_{|(\mathbb{N}\setminus\{i\})\times\mathbb{N}}$, i.e. only the i-th row may dislocate.
  - On the i-th row, we have
$$(i, j) \sim^{\zeta^{(i)}} (i, j') \iff U_j \sim^{\zeta_0} U_{j'},$$
$$(i, j) \sim^{\xi^{(i)}} (i, j') \iff U_j \sim^{\xi_0} U_{j'},$$
and also
$$(i, j) \sim^{\xi^{(i)}} (i + 1, 1) \iff U_j \in [0, \bar{u}),$$
where it should be noted that $(i+1, 1)$ could be replaced by any element (i', j') with $i' \neq i$.

See Figure 4.4 for a representation of the bivariate paintbox process. In words, $\pi^{(i)}$ is a random nested partition such that the outer partition $\xi^{(i)}$ has a *distinguished block* containing $(\mathbb{N} \setminus \{i\}) \times \mathbb{N}$, which also contains a proportion $\bar{u}$ of elements of the i-th row. Other non-singleton blocks of $\xi^{(i)}$ can be indexed by $k \geq 1$, each containing a proportion $\bar{s}_k$ of elements of the i-th row. The blocks of the inner partition $\zeta^{(i)}$ are the entire rows, except for the i-th row where non-singleton blocks can be indexed by $(\star, l)$ and (k, l) for $k, l \geq 1$, each respectively containing a proportion u_l or $s_{k,l}$ of elements of the i-th row. As the notation suggests, inner blocks with frequency $s_{k,l}$ (resp. u_l) are included in the outer block with frequency $\bar{s}_k$ (resp. $\bar{u}$) on the i-th row.

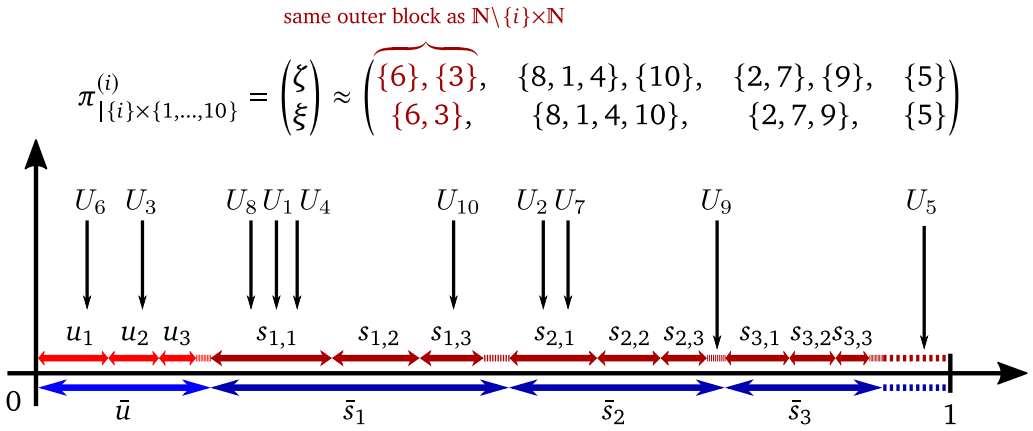


Figure 4.4 – Paintbox construction of $\pi^{(i)}$

The distribution of $\pi^{(i)}$ obtained with this construction is a probability on $\mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ that we denote $\widetilde{\varrho}^{(i)}_{\mathbf{p}}$. We finally define

$$\widetilde{\varrho}_{\mathbf{p}} = \sum_{i\geq 1} \widetilde{\varrho}^{(i)}_{\mathbf{p}}.$$

It is clear from the exchangeability of the sequence $(U_j, j \geq 1)$ that $\widetilde{\varrho}_{\mathbf{p}}$ is M-invariant.

Now consider a measure ν_{in} on $\mathscr{S}^{\downarrow}_{\preceq}$ satisfying

$$\nu_{\text{in}}(\{u_1 = 1 \text{ or } s_{1,1} = 1\}) = 0, \text{ and } \int_{\mathscr{S}^{\downarrow}_{\preceq}} (1-u_1)\,\nu_{\text{in}}(\mathrm{d}\mathbf{p}) < \infty. \tag{4.12}$$

Similarly as in the previous example, we define

$$\widetilde{\varrho}_{\nu_{\text{in}}}(\,\cdot\,) = \int_{\mathscr{S}^{\downarrow}_{\preceq}} \widetilde{\varrho}_{\mathbf{p}}(\,\cdot\,)\,\nu_{\text{in}}(\mathrm{d}\mathbf{p}).$$

It is again straight-forward to check that $\widetilde{\varrho}_{\nu_{\text{in}}}$ is an M-invariant measure on $\mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ satisfying (4.8), so there exists a fragmentation process having $\widetilde{\varrho}_{\nu_{\text{in}}}$ as M-invariant measure.

In intuitive terms, such a process can be described by saying that the inner blocks independently dislocate with **inner dislocation rate** ν_{in}. In a dislocation event, new inner blocks are formed, each with a given proportion of the original block, and regroup, either in the original outer block (with a total proportion $\bar{u}$ with respect to the original inner block) or in newly created outer blocks.

A combination of the above The mechanisms we discussed in the three proposed examples can be added in a parallel way, each event arising at its own independent rate and events from distinct mechanisms occurring at distinct times. More precisely, for a set of erosion coefficients $c_{\text{out}}, c_{\text{in},1}, c_{\text{in},2} \geq 0$, an outer dislocation measure ν_{out} on $\mathscr{S}^{\downarrow}$ satisfying (4.4) and an inner dislocation measure ν_{in} on $\mathscr{S}^{\downarrow}_{\preceq}$ satisfying (4.12), the measure

$$\mu := c_{\text{out}}\mathfrak{e}^{\text{out}} + c_{\text{in},1}\mathfrak{e}^{\text{in},1} + c_{\text{in},2}\mathfrak{e}^{\text{in},2} + \widehat{\varrho}_{\nu_{\text{out}}} + \widetilde{\varrho}_{\nu_{\text{in}}}$$

is a valid M-invariant measure on $\mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ satisfying (4.8), and thus corresponds to a fragmentation process exhibiting simultaneously all the discussed mechanisms at the rates described above. The main result of this article is to prove that any nested fragmentation process admits such a representation.

4.5.2 Characterization of nested fragmentations

Theorem 4.14. *Let $\Pi = (\Pi(t), t \geq 0) = ((\zeta(t), \xi(t)), t \geq 0)$ be a nested fragmentation process. Then there are*

- *an outer erosion coefficient $c_{\mathrm{out}} \geq 0$ and two inner erosion coefficients $c_{\mathrm{in},1}, c_{\mathrm{in},2} \geq 0$;*
- *an outer dislocation measure ν_{out} on $\mathscr{S}^{\downarrow}$ satisfying (4.4);*
- *an inner dislocation measure ν_{in} on $\mathscr{S}^{\downarrow}_{\preceq}$ satisfying (4.12);*

such that the M-invariant measure μ of the process can be written

$$\mu = c_{\mathrm{out}} \mathfrak{e}^{\mathrm{out}} + c_{\mathrm{in},1} \mathfrak{e}^{\mathrm{in},1} + c_{\mathrm{in},2} \mathfrak{e}^{\mathrm{in},2} + \widehat{\varrho}_{\nu_{\mathrm{out}}} + \widetilde{\varrho}_{\nu_{\mathrm{in}}}.$$

Corollary 4.15. *Definition 4.1 is equivalent to Definition 4.1'.*

Proof. We have shown most of the equivalence in Corollary 4.4 and Remark 4.8. What remains is to show that if $\Pi = (\zeta, \xi)$ is a nested fragmentation process according to Definition 4.1', then ζ is a homogeneous fragmentation process in $\mathcal{P}_\infty$. Now if μ is given by the expression of the preceding theorem, using the Poissonian construction of Section 4.4.3 one easily checks that ζ has the same transition rates as a homogeneous fragmentation with erosion coefficient $c = c_{\mathrm{in},1} + c_{\mathrm{in},2}$ and dislocation measure $\nu = \nu_{\mathrm{in}} \circ S^{-1}$, where $S : \mathscr{S}^{\downarrow}_{\preceq} \to \mathscr{S}^{\downarrow}$ is the map given by

$$S(\mathbf{p}) := \text{nonincreasing reordering of } \{u_l,\, l \geq 1\} \cup \{s_{k,l},\, k, l \geq 1\}. \qquad \square$$

The rest of Section 4.5 is dedicated to proving Theorem 4.14. Let μ be the M-invariant characteristic measure on $\mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ associated with Π. Recall that π_{R} denotes the *rows partition*, defined by

$$\pi_{\mathrm{R}} = \{\{(i,j),\, j \geq 1\},\, i \geq 1\}.$$

First, notice that the inner branching property implies that μ-a.e. we have

$$\exists i \in \mathbb{N}, \quad \zeta_{|(\mathbb{N}\setminus\{i\})\times\mathbb{N}} = (\pi_{\mathrm{R}})_{|(\mathbb{N}\setminus\{i\})\times\mathbb{N}},$$

where ζ is the first coordinate in the standard variable $\pi = (\zeta, \xi) \in \mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$. This will enable us to decompose μ further. Let us write

$$\begin{gathered} \mu_{\mathrm{out}} := \mu(\,\cdot \cap \{\zeta = \pi_{\mathrm{R}}\}), \\ \text{for } i \in \mathbb{N}, \quad \mu_{\mathrm{in},i} := \mu\left(\{\zeta_{|\{i\}\times\mathbb{N}} \neq \mathbf{1}_{\{i\}\times\mathbb{N}}\} \cap \cdot\right), \\ \text{such that } \mu_{\mathrm{in}} := \mu(\,\cdot \cap \{\zeta \neq \pi_{\mathrm{R}}\}) = \textstyle\sum_{i\geq 1} \mu_{\mathrm{in},i} \\ \text{and } \mu = \mu_{\mathrm{out}} + \mu_{\mathrm{in}}. \end{gathered} \tag{4.13}$$

On the event $\{\zeta = \pi_{\mathrm{R}}\}$, we have

$$\xi = f(\xi^\sigma),$$

where $\sigma : \mathbb{N} \to \mathbb{N}^2$ is the injection $i \mapsto (i, 1)$, and $f : \mathcal{P}_\infty \to \mathcal{P}_{\mathbb{N}^2}$ is the map such that $(i,j) \sim^{f(\pi_0)} (i',j') \iff i \sim^{\pi_0} i'$. By M-invariance of μ, the measure

$$\widetilde{\mu}_{\mathrm{out}} := \mu\left(\{\zeta = \pi_{\mathrm{R}}\} \cap \{\xi^\sigma \in \cdot\,\}\right)$$

is an exchangeable measure on $\mathcal{P}_\infty$, of which μ_{out} is the push-forward by the map $(\pi_{\text{R}}, f(\,\cdot\,))$.

Also, note that μ satisfies the σ-finiteness assumption (4.8), which implies that $\widetilde{\mu}_{\text{out}}$ satisfies (4.3), showing (see Section 4.2.2) that it can be decomposed

$$\widetilde{\mu}_{\text{out}} = c_{\text{out}}\mathfrak{e} + \varrho_{\nu_{\text{out}}},$$

where $c_{\text{out}} \geq 0$ and ν_{out} is a measure on $\mathscr{S}^{\downarrow}$ satisfying (4.4). Thanks to our definitions, this immediately translates into

$$\mu_{\text{out}} = c_{\text{out}}\mathfrak{e}^{\text{out}} + \widehat{\varrho}_{\nu_{\text{out}}},$$

and to prove Theorem 4.14, it only remains to show that we can write

$$\mu_{\text{in}} = \textstyle\sum_{i\geq 1}\mu_{\text{in},i} = c_{\text{in},1}\mathfrak{e}^{\text{in},1} + c_{\text{in},2}\mathfrak{e}^{\text{in},2} + \widetilde{\varrho}_{\nu_{\text{in}}}.$$

To that aim, note that by exchangeability we have $\mu_{\text{in},i} = \mu_{\text{in},1}^{\tau_{1,i}}$ where $\tau_{1,i} : \mathbb{N}^2 \to \mathbb{N}^2$ denotes the bijection swapping the first and i-th rows, so the measure $\mu_{\text{in},1}$ is sufficient to recover μ_{in} entirely. Let us examine the distribution of ξ under $\mu_{\text{in},1}$. We claim that μ-a.e. on the event $\{\zeta_{|\{1\}\times\mathbb{N}} \neq \mathbf{1}_{\{1\}\times\mathbb{N}}\}$, the equality $\xi_{|(\mathbb{N}\setminus\{1\})\times\mathbb{N}} = \mathbf{1}_{(\mathbb{N}\setminus\{1\})\times\mathbb{N}}$ holds. Indeed, if this was not the case, by M-invariance we would have

$$a := \mu(\zeta_{|\{1\}\times\mathbb{N}} \neq \mathbf{1}_{\{1\}\times\mathbb{N}}, \text{ and } (2,1) \nsim^{\xi} (3,1)) > 0.$$

Let us then show that in fact $a = 0$. By M-invariance of μ, we have for any $i \geq 4$,

$$a = \mu(\zeta_{|\{i\}\times\mathbb{N}} \neq \mathbf{1}_{\{i\}\times\mathbb{N}}, \text{ and } (2,1) \nsim^{\xi} (3,1)),$$

but because of the inner branching property, we have seen that the events $\{\zeta_{|\{i\}\times\mathbb{N}} \neq \mathbf{1}_{\{i\}\times\mathbb{N}}\}$ have μ-negligible intersections. Now we have

$$\begin{aligned}\infty > \mu(\pi_{|[3]^2} \neq (\pi_{\text{R}}, \mathbf{1})_{|[3]^2}) &\geq \mu\left((2,1) \nsim^{\xi} (3,1)\right)\\ &\geq \mu\left(\cup_{i\geq 4}\{\zeta_{|\{i\}\times\mathbb{N}} \neq \mathbf{1}_{\{i\}\times\mathbb{N}}, \text{ and } (2,1) \nsim^{\xi} (3,1)\}\right)\\ &= \textstyle\sum_{i\geq 4} a.\end{aligned}$$

This shows that necessarily $a = 0$.

Now in order to further study $\mu_{\text{in},1}$ we need to introduce exchangeable partitions on a space with a distinguished element. Results in that direction have been established by Foucart [45], where distinguished exchangeable partitions are introduced and used to construct a generalization of Λ-coalescents modeling the genealogy of a population with immigration. Here we need to define in a similar way distinguished partitions in our bivariate setting. Informally, we will see that in a gene fragmentation, certain resulting gene blocks remain in a distinguished species block, that one can interpret as the mother species.

Definition 4.16. For $n \in \mathbb{N}\cup\{\infty\}$, we define $[n]_\star := [n]\cup\{\star\}$, where $\star$ is not an element of $\mathbb{N}$. We define $\mathcal{P}_{n,\star}^{2,\preceq}$ as the set of nested partitions $\pi = (\zeta,\xi) \in \mathcal{P}_{[n]_\star}^{2,\preceq}$ such that $\star$ is isolated in the finer partition ζ:

$$\mathcal{P}_{n,\star}^{2,\preceq} := \left\{\pi = (\zeta,\xi) \in \mathcal{P}_{[n]_\star}^{2,\preceq},\ \{\star\} \in \zeta\right\}.$$

We define the action of an injection $\sigma : [n] \to [n]$ on an element $\pi \in \mathcal{P}^{2,\preceq}_{n,\star}$ as the action of the unique extension $\widetilde{\sigma} : [n]_\star \to [n]_\star$ such that $\widetilde{\sigma}(\star) = \star$, and define **exchangeability** for measures on $\mathcal{P}^{2,\preceq}_{n,\star}$ as invariance under the actions of such injections $\sigma : [n] \to [n]$.

Let us come back to the decomposition of $\mu_{\mathrm{in},1}$. We define an injection

$$\tau : \begin{cases} [\infty]_\star & \longrightarrow \mathbb{N}^2 \\ j \in \mathbb{N} & \longmapsto (1, j) \\ \star & \longmapsto (2, 1). \end{cases}$$

Note that here we could have chosen any value $\tau(\star) = (i, j)$ with $i \geq 2$, since μ-a.e. on the event $\{\zeta_{|\{1\}\times\mathbb{N}} \neq \mathbf{1}_{\{1\}\times\mathbb{N}}\}$ those elements are all in the same block of ξ. The argument above shows that on the event $\{\zeta_{|\{1\}\times\mathbb{N}} \neq \mathbf{1}_{\{1\}\times\mathbb{N}}\}$, we have μ-a.e. the equality

$$\pi = (\zeta, \xi) = g(\pi^\tau),$$

where $g : \mathcal{P}^{2,\preceq}_{\infty,\star} \to \mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ is a deterministic function which we can define by: $g(\pi_0)$ is the only $\pi \in \mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ such that

$$\pi^\tau = \pi_0, \quad \pi \preceq (\pi_{\mathrm{R}}, \mathbf{1}_{\mathbb{N}^2})$$
$$\text{and } \pi_{|(\mathbb{N}\setminus\{1\})\times\mathbb{N}} = (\pi_{\mathrm{R}}, \mathbf{1}_{\mathbb{N}^2})_{|(\mathbb{N}\setminus\{1\})\times\mathbb{N}}.$$

Let us now write

$$\widetilde{\mu}_{\mathrm{in}} := \mu_{\mathrm{in},1}(\pi^\tau \in \cdot\,). \tag{4.14}$$

Note that the push-forward of this exchangeable measure on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ by the map g is $\mu_{\mathrm{in},1}$. Also, note that the σ-finiteness assumption (4.8) and the fact that $\mu_{\mathrm{in},1}$-a.e. we have $\zeta_{|\{1\}\times\mathbb{N}} \neq \mathbf{1}_{\{1\}\times\mathbb{N}}$ imply that $\widetilde{\mu}_{\mathrm{in}}$ satisfies

$$\widetilde{\mu}_{\mathrm{in}}(\{\zeta_{|[\infty]} = \mathbf{1}\}) = 0, \text{ and } \forall n \geq 1, \quad \widetilde{\mu}_{\mathrm{in}}(\pi_{|[n]_\star} \neq \pi_n) < \infty. \tag{4.15}$$

where $\pi_n := (\{\{\star\}, [n]\}, \mathbf{1}_{[n]_\star})$ denotes the coarsest partition on $\mathcal{P}^{2,\preceq}_{n,\star}$.

We can summarize the previous discussion in the following lemma.

Lemma 4.17. *The characteristic M-invariant measure μ of a nested fragmentation process in $\mathcal{P}^{2,\preceq}_{\infty}$ can be decomposed*

$$\mu = c_{\mathrm{out}}\mathfrak{e}^{\mathrm{out}} + \widehat{\varrho}_{\nu_{\mathrm{out}}} + \mu_{\mathrm{in}},$$

where $c_{\mathrm{out}} \geq 0$, ν_{out} is a measure on $\mathscr{S}^\downarrow$ satisfying (4.4), and $\mu_{\mathrm{in}} := \mu(\cdot \cap \{\zeta \neq \pi_{\mathrm{R}}\})$. Also, there exists an exchangeable measure $\widetilde{\mu}_{\mathrm{in}}$ on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ which satisfies (4.15) and such that $\mu_{\mathrm{in}} = \sum_i \mu^{\tau_{1,i}}_{\mathrm{in},1}$, where

- *$\mu_{\mathrm{in},1}$ is a measure on $\mathcal{P}^{2,\preceq}_{\mathbb{N}^2}$ which is the push-forward of $\widetilde{\mu}_{\mathrm{in}}$ by the map g defined in the previous paragraph.*
- *$\tau_{1,i} : \mathbb{N}^2 \to \mathbb{N}^2$ is the bijection swapping the first row with the i-th row.*

In the next section, we will develop tools to analyze and further decompose the measure $\widetilde{\mu}_{\mathrm{in}}$ into terms of erosion and dislocation.

4.5.3 Bivariate mass partitions

Recall our space of bivariate mass partitions defined in Definition 4.13,

$$\mathscr{S}^{\downarrow}_{\preceq} \subset [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}^2} \times [0,1] \times [0,1]^{\mathbb{N}},$$

as the subset consisting of elements $\mathbf{p} = ((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1})$ satisfying conditions (4.11). We wish to match exchangeable measures on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ and measures on $\mathscr{S}^{\downarrow}_{\preceq}$, and to that aim we need some further definitions. We say that an element $\pi = (\zeta, \xi) \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ has **asymptotic frequencies** if ζ and ξ have asymptotic frequencies, and we write

$$|\pi|^{\downarrow} = ((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1}) \in \mathscr{S}^{\downarrow}_{\preceq}$$

for the unique – because of the ordering conditions in (4.11) – element satisfying:

- the block $B \in \xi$ containing $\star$ has asymptotic frequency $|B| = \bar{u}$ and the nonincreasing reordering of the asymptotic frequencies of the blocks of $\zeta \cap B$ is the sequence $(u_l, l \geq 1)$.
- for any other block $B \in \xi$ with a positive asymptotic frequency, there is a $k \in \mathbb{N}$ such that $|B| = \bar{s}_k$ and the nonincreasing reordering of the asymptotic frequencies of the blocks of $\zeta \cap B$ is the sequence $(s_{k,l}, l \geq 1)$.
- the mapping $B \mapsto k$ is injective, and for any k such that $\bar{s}_k > 0$, there is a block $B \in \xi$ such that $|B| = \bar{s}_k$.

4.5.4 A paintbox construction for nested partitions

We first adapt the construction used in our third example of Section 4.5.1 to our new partition space $\mathcal{P}^{2,\preceq}_{\infty,\star}$. Note that if $\mathbf{p} = ((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1}) \in \mathscr{S}^{\downarrow}_{\preceq}$, then one can define a random element $\pi = (\zeta, \xi) \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ with a paintbox procedure very similar to the one described in the *inner dislocation* example in Section 4.5.1. For the sake of readability, let us recall the notation and construction:

- for $k \geq 0$, define $\bar{t}_k = \bar{u} + \sum_{k'=1}^{k} \bar{s}_{k'}$.
- for $l \geq 0$, define $t_{\star,l} = \sum_{l'=1}^{l} u_{l'}$.
- for $k \geq 1$ and $l \geq 0$, define $t_{k,l} = \bar{t}_{k-1} + \sum_{l'=1}^{l} s_{k,l'}$.
- write $\pi_0 = (\zeta_0, \xi_0)$ for the unique element of $\mathcal{P}^{2,\preceq}_{[0,1]}$ such that the non-dust blocks of ξ_0 are

  $$[0, \bar{u}) \text{ and } [\bar{t}_k, \bar{t}_{k+1}),\ k \geq 1,$$

  and such that the non-singleton blocks of ζ_0 are

  $$[t_{\star,l-1}, t_{\star,l}),\ l \geq 1 \text{ and } [t_{k,l-1}, t_{k,l}),\ k, l \geq 1.$$

- let $(U_i, i \geq 1)$ be an i.i.d. sequence of uniform random variables on $[0,1]$ and define the random injection $\sigma : i \in \mathbb{N} \mapsto U_i \in [0,1]$.

- finally define the random element $\pi \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ as the unique $\pi = (\zeta, \xi)$ such that $\pi_{|\mathbb{N}} = \pi_0^\sigma$, and the block of ξ containing $\star$ is equal to:

$$\{\star\} \cup \{i \geq 1,\ U_i < \bar{u}\}.$$

The distribution of π obtained with this construction is a probability on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ that we denote $\bar{\varrho}_{\mathbf{p}}$. It is clear from the exchangeability of the sequence $(U_i, i \geq 1)$ that $\bar{\varrho}_{\mathbf{p}}$ is exchangeable, and from the strong law of large numbers, that $\bar{\varrho}_{\mathbf{p}}$-a.s., π possesses asymptotic frequencies equal to $|\pi|^\downarrow = \mathbf{p}$. For a measure ν on $\mathscr{S}^\downarrow_{\preceq}$, we will define a corresponding exchangeable measure $\bar{\varrho}_\nu$ on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ by

$$\bar{\varrho}_\nu(\,\cdot\,) = \int_{\mathscr{S}^\downarrow_{\preceq}} \bar{\varrho}_{\mathbf{p}}(\,\cdot\,)\, \nu(\mathrm{d}\mathbf{p}).$$

The following lemma shows that every probability measure on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ is of this form.

Lemma 4.18. *Let $\pi = (\zeta, \xi)$ be a random exchangeable element of $\mathcal{P}^{2,\preceq}_{\infty,\star}$. Then π has asymptotic frequencies $|\pi|^\downarrow \in \mathscr{S}^\downarrow_{\preceq}$ a.s. and its distribution conditional on $|\pi|^\downarrow = \mathbf{p}$ is $\bar{\varrho}_{\mathbf{p}}$. In other words, we have*

$$\mathbb{P}(\pi \in \,\cdot\,) = \int_{\mathscr{S}^\downarrow_{\preceq}} \mathbb{P}(|\pi|^\downarrow \in \mathrm{d}\mathbf{p})\, \bar{\varrho}_{\mathbf{p}}(\,\cdot\,).$$

Proof. Independently from π, let $(X_i, i \geq 1)$ and $(Y_i, i \geq 1)$ be i.i.d. uniform random variables on $[0,1]$. Conditional on π, we define a random variable $Z_n \in [0,1] \times ([0,1] \cup \{\star\})$ for each $n \in \mathbb{N}$ by

$$Z_n := \begin{cases} (X_{A_n}, Y_{B_n}) & \text{if } \star \not\sim^\xi n, \\ (X_{A_n}, \star) & \text{if } \star \sim^\xi n, \end{cases} \quad \text{where} \quad \begin{cases} A_n := \min\{m \in \mathbb{N},\ m \sim^\zeta n\} \\ B_n := \min\{m \in \mathbb{N},\ m \sim^\xi n\}. \end{cases}$$

It is straight-forward that we recover entirely π from the sequence $(Z_n, n \geq 1)$ because we have

$$\begin{aligned} n \sim^\zeta m &\iff x(Z_n) = x(Z_m), \\ n \sim^\xi m &\iff y(Z_n) = y(Z_m), \\ n \sim^\xi \star &\iff y(Z_n) = \star, \end{aligned} \tag{4.16}$$

where x and y denote respectively the projection maps from $[0,1] \times ([0,1] \cup \{\star\})$ to the first and second coordinates. Now, notice that the exchangeability of π implies that the sequence $(Z_n, n \geq 1)$ is an exchangeable sequence of random variables. Then, by an application of de Finetti's theorem, we see that there is a random probability measure P on $[0,1] \times ([0,1] \cup \{\star\})$ such that conditional on P, the sequence $(Z_n, n \geq 1)$ is i.i.d. with distribution P.

Now notice that if P is a probability measure on $[0,1] \times ([0,1] \cup \{\star\})$, we can define

$$|P|^\downarrow = ((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1}) \in \mathscr{S}^\downarrow_{\preceq}$$

by setting the following, where everything is numbered in an order compatible with our conditions (4.11).

- $\bar{u} := P(y = \star)$.
- $\bar{s}_k := P(y = y_k)$, where $(y_k, k \geq 1)$ is the injective sequence of points of $[0,1]$ such that $P(y = y_k) > 0$.
- $u_l := P(x = x_{\star,l}, y = \star)$ where $(x_{\star,l}, l \geq 1)$ is the injective sequence of points of $[0,1]$ such that $P(x = x_{\star,l}, y = \star) > 0$.
- $s_{k,l} := P(x = x_{k,l}, y = y_k)$ where $(x_{k,l}, l \geq 1)$ is the injective sequence of points of $[0,1]$ such that $P(x = x_{k,l}, y = y_k) > 0$.

It should now be clear that defining with (4.16) a random $\pi \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ from a sequence $(Z_n, n \geq 1)$ of P-i.i.d. random variables is in fact the same as defining π from a paintbox construction $\bar{\varrho}_{\mathbf{p}}$ with $\mathbf{p} = |P|^{\downarrow}$. Therefore, the distribution of π is given by

$$\mathbb{P}(\pi \in \cdot\,) = \int_{\mathscr{S}^{\downarrow}_{\preceq}} \mathbb{P}(|P|^{\downarrow} \in \mathrm{d}\mathbf{p})\, \bar{\varrho}_{\mathbf{p}}(\,\cdot\,),$$

which concludes the proof since for any $\mathbf{p}$ we have $\bar{\varrho}_{\mathbf{p}}$-a.s. that $|\pi|^{\downarrow}$ exists and is equal to $\mathbf{p}$. □

4.5.5 Erosion and dislocation for nested partitions

As in the standard $\mathcal{P}_\infty$ case, we can decompose any exchangeable measure μ on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ satisfying some finiteness condition similar to (4.3) in a canonical way. To ease the notation, recall that we define for $n \in \mathbb{N} \cup \{\infty\}$, π_n the maximal element in $\mathcal{P}^{2,\preceq}_{n,\star}$

$$\pi_n := (\{\{\star\}, [n]\},\, \mathbf{1}_{[n]_\star}).$$

We also define two erosion measures $\mathfrak{e}^1$ and $\mathfrak{e}^2$ by

$$\mathfrak{e}^1 = \sum_{i \geq 1} \delta_{(\{\{\star\},\{i\},[\infty]\setminus\{i\}\},\mathbf{1}_{[\infty]_\star})},$$
$$\mathfrak{e}^2 = \sum_{i \geq 1} \delta_{(\{\{\star\},\{i\},[\infty]\setminus\{i\}\},\, \{\{i\},[\infty]_\star\setminus\{i\}\})}.$$

Proposition 4.19. *Let μ be an exchangeable measure on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ satisfying (4.15), namely*

$$\mu(\{\zeta_{|[\infty]} = \mathbf{1}\}) = 0, \text{ and } \forall n \geq 1, \quad \mu(\pi_{|[n]_\star} \neq \pi_n) < \infty.$$

Then there are two real numbers $c_1, c_2 \geq 0$ and a measure ν on $\mathscr{S}^{\downarrow}_{\preceq}$ satisfying (4.12), namely

$$\nu(\{u_1 = 1 \text{ or } s_{1,1} = 1\}) = 0, \text{ and } \int_{\mathscr{S}^{\downarrow}_{\preceq}} (1 - u_1)\, \nu(\mathrm{d}\mathbf{p}) < \infty$$

such that $\mu = c_1\mathfrak{e}^1 + c_2\mathfrak{e}^2 + \bar{\varrho}_\nu$. Conversely, any μ of this form is exchangeable and satisfies (4.15).

Proof. The proof follows closely that of Theorem 3.1 in [10], as our result is a straightforward extension of it. We first define $\mu_n := \mu(\,\cdot \cap \{\pi_{[n]_\star} \neq \pi_n\})$ which is a finite measure, and

$$\overleftarrow{\mu}_n := \mu_n^{\theta_n},$$

where $\theta_n : \mathbb{N} \to \mathbb{N}$ is the n-shift defined by $\theta_n(i) = i + n$. We can check that $\overleftarrow{\mu}_n$ is an exchangeable measure on $\mathcal{P}^{2,\preceq}_{\infty,\star}$. Indeed let us take $\sigma : \mathbb{N} \to \mathbb{N}$ a permutation, and consider $\tau : \mathbb{N} \to \mathbb{N}$ the permutation defined by

$$\tau : \begin{cases} i \leq n & \longmapsto i \\ i > n & \longmapsto n + \sigma^{-1}(i-n). \end{cases}$$

We have clearly $\tau \circ \theta_n \circ \sigma = \theta_n$ and $\tau_{|[n]} = \mathrm{id}_{[n]}$, so we can use the τ-invariance of μ to conclude

$$\begin{aligned} \overleftarrow{\mu}_n(\pi^\sigma \in \cdot\,) &= \mu_n(\pi^{\theta_n \circ \sigma} \in \cdot\,) \\ &= \mu(\{\pi^{\theta_n \circ \sigma} \in \cdot\,\} \cap \{\pi_{|[n]_\star} \neq \pi_n\}) \\ &= \mu(\{\pi^{\tau \circ \theta_n \circ \sigma} \in \cdot\,\} \cap \{(\pi^\tau)_{|[n]_\star} \neq \pi_n\}) \\ &= \mu(\{\pi^{\theta_n} \in \cdot\,\} \cap \{\pi_{|[n]_\star} \neq \pi_n\}) \\ &= \overleftarrow{\mu}_n(\,\cdot\,), \end{aligned}$$

which proves that $\overleftarrow{\mu}_n$ is exchangeable. Since it is also finite, Lemma 4.18 implies that $|(\pi^{\theta_n})|^\downarrow = |\pi|^\downarrow$ exists μ-a.e. on the event $\{\mu_{|[n]_\star} \neq \pi_n\}$, and that we have

$$\overleftarrow{\mu}_n(\,\cdot\,) = \int_{\mathscr{S}^\downarrow_\preceq} \mu_n(|\pi|^\downarrow \in \mathrm{d}\mathbf{p})\ \bar{\varrho}_{\mathbf{p}}(\,\cdot\,). \tag{4.17}$$

Now since $\cup_n \{\pi_{|[n]_\star} \neq \pi_n\} = \{\pi \neq \pi_\infty\}$ and $\mu(\{\pi = \pi_\infty\}) \leq \mu(\{\zeta_{|[\infty]} = \mathbf{1}\}) = 0$, necessarily the existence of $|\pi|^\downarrow \in \mathscr{S}^\downarrow_\preceq$ holds μ-a.e.

For simplicity, denote $\mathbf{1} \in \mathscr{S}^\downarrow_\preceq$ as the element $((u_l)_{l\geq 1}, (s_{k,l})_{k,l\geq 1}, \bar{u}, (\bar{s}_k)_{k\geq 1}) \in \mathscr{S}^\downarrow_\preceq$ with $\bar{u} = u_1 = 1$ (note that $\bar{\varrho}_{\mathbf{1}} = \delta_{\pi_\infty}$), and define $\varphi(\,\cdot\,) := \mu(\,\cdot \cap \{|\pi|^\downarrow \neq \mathbf{1}\})$. Fix $k \in \mathbb{N}$, and consider the measure $\varphi(\pi_{|[k]_\star} \in \cdot\,)$ on $\mathcal{P}^{2,\preceq}_{k,\star}$. Note that

$$\{|\pi|^\downarrow \neq \mathbf{1}\} = \bigcup_{n\geq 1} \{|\pi|^\downarrow \neq \mathbf{1},\ (\pi^{\theta_k})_{|[n]_\star} \neq \pi_n\},$$

where the union is increasing, so one can write

$$\begin{aligned} \varphi(\pi_{|[k]_\star} \in \cdot\,) &= \mu(\{\pi_{|[k]_\star} \in \cdot\,\} \cap \{|\pi|^\downarrow \neq \mathbf{1}\}) \\ &= \lim_{n\to\infty} \mu\left(\{\pi_{|[k]_\star} \in \cdot\,\} \cap \{|\pi|^\downarrow \neq \mathbf{1},\ (\pi^{\theta_k})_{|[n]_\star} \neq \pi_n\}\right). \end{aligned} \tag{4.18}$$

Now let us use invariance of μ under the permutation $\sigma : \mathbb{N} \to \mathbb{N}$ defined by

$$\sigma : \begin{cases} i \in \{1, \dots k\} & \mapsto i + n, \\ i \in \{k+1, \dots, k+n\} & \mapsto i - k, \\ i \geq k + n + 1 & \mapsto i, \end{cases}$$

to obtain

$$\begin{aligned} &\mu\left(\{\pi_{|[k]_\star} \in \cdot\,\} \cap \{|\pi|^\downarrow \neq \mathbf{1},\ (\pi^{\theta_k})_{|[n]_\star} \neq \pi_n\}\right) \\ &= \mu\left(\{(\pi^{\theta_n})_{|[k]_\star} \in \cdot\,\} \cap \{|\pi|^\downarrow \neq \mathbf{1},\ \pi_{|[n]_\star} \neq \pi_n\}\right). \end{aligned}$$

Now by definition of μ_n and $\overleftarrow{\mu_n}$, this expression is exactly

$$\mu_n\left(\{(\pi^{\theta_n})_{|[k]_\star} \in \cdot\} \cap \{|\pi|^\downarrow \neq \mathbf{1}\}\right) = \overleftarrow{\mu}_n\left(\{\pi_{|[k]_\star} \in \cdot\} \cap \{|\pi|^\downarrow \neq \mathbf{1}\}\right).$$

Plugging this into (4.18) and then using (4.17), we obtain

$$\begin{aligned}\varphi(\pi_{|[k]_\star} \in \cdot) &= \lim_{n\to\infty} \overleftarrow{\mu}_n\left(\{\pi_{|[k]_\star} \in \cdot\} \cap \{|\pi|^\downarrow \neq \mathbf{1}\}\right)\\ &= \lim_{n\to\infty} \int_{\mathscr{S}^\downarrow_\preceq \setminus \{\mathbf{1}\}} \mu_n(|\pi|^\downarrow \in \mathrm{d}\mathbf{p})\, \bar{\varrho}_{\mathbf{p}}(\pi_{|[k]_\star} \in \cdot).\end{aligned}$$

Finally, note that the sequence of measures μ_n is increasing and converges to μ, in the sense that $\mu_n(B) \uparrow \mu(B)$ when $n \to \infty$ for any Borel set $B \subset \mathcal{P}^{2,\preceq}_{\infty,\star}$. This allows us to take the limit in the last display:

$$\varphi(\pi_{|[k]_\star} \in \cdot) = \int_{\mathscr{S}^\downarrow_\preceq \setminus \{\mathbf{1}\}} \mu(|\pi|^\downarrow \in \mathrm{d}\mathbf{p})\, \bar{\varrho}_{\mathbf{p}}(\pi_{|[k]_\star} \in \cdot).$$

Since this is true for all $k \in \mathbb{N}$, we have

$$\varphi(\,\cdot\,) = \int_{\mathscr{S}^\downarrow_\preceq \setminus \{\mathbf{1}\}} \mu(|\pi|^\downarrow \in \mathrm{d}\mathbf{p})\, \bar{\varrho}_{\mathbf{p}}(\,\cdot\,) = \bar{\varrho}_\nu,$$

with $\nu(\,\cdot\,) = \mu(\{|\pi|^\downarrow \in \cdot\} \cap \{|\pi|^\downarrow \neq \mathbf{1}\})$. Now notice that the paintbox construction of the probability measures $\bar{\varrho}_{\mathbf{p}}$ implies that

$$\bar{\varrho}_\nu(\pi_{|[n]_\star} \neq \pi_n) = \int_{\mathscr{S}^\downarrow_\preceq} \nu(\mathrm{d}\mathbf{p})\Big(1 - \sum_{l\geq 1} u_l^n\Big),$$

and that since $u_1 \geq u_2 \geq \ldots$ and $\sum_l u_l \leq 1$, we have for $n \geq 2$,

$$1 - u_1 \leq 1 - u_1 \textstyle\sum_l u_l^{n-1} \leq 1 - \sum_l u_l^n \leq 1 - u_1^n \leq n(1 - u_1).$$

Integrating this with respect to ν, we find that clearly $\bar{\varrho}_\nu$ satisfies the right-hand side of (4.15) iff ν satisfies the right-hand side of (4.12). For the left-hand side, notice that by construction $\nu(\{u_1 = 1 \text{ or } s_{1,1} = 1\}) = \bar{\varrho}_\nu(\{\zeta_{|[\infty]} = \mathbf{1}\}) = 0$.

We now write $\psi(\,\cdot\,) := \mu(\,\cdot\, \cap \{|\pi|^\downarrow = \mathbf{1}\})$ so that $\mu = \varphi + \psi = \bar{\varrho}_\nu + \psi$. Take an integer $n \in \mathbb{N}$. We know that $\overleftarrow{\psi}_n(\,\cdot\,) := \psi(\{\pi^{\theta_n} \in \cdot\} \cap \{\pi_{|[n]_\star} \neq \pi_n\})$ is a finite exchangeable measure on $\mathcal{P}^{2,\preceq}_{\infty,\star}$ such that $|\pi|^\downarrow = \mathbf{1}$ $\overleftarrow{\psi}_n$-a.e. Now recall that $\bar{\varrho}_{\mathbf{1}} = \delta_{\pi_\infty}$. A consequence of Lemma 4.18 is that $\pi = \pi_\infty$ $\overleftarrow{\psi}_n$-a.e., which in turn implies that ψ-a.e. on the event $\{\pi_{|[n]_\star} \neq \pi_n\}$, we have $\pi^{\theta_n} = \pi_\infty$. Since there is only a finite number of elements $\pi \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ such that $\pi^{\theta_n} = \pi_\infty$, we have

$$\psi(\,\cdot\, \cap \{\pi_{|[n]_\star} \neq \pi_n\}) = \sum_i a_i \delta_{\widehat{\pi}_i},$$

where the sum is finite, and for each i, we have $\widehat{\pi}_i^{\theta_n} = \pi_\infty$. Now suppose we have $\psi(\{\widehat{\pi}\}) > 0$, for a $\widehat{\pi} \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ such that $\widehat{\pi}^{\theta_n} = \pi_\infty$. Let $I(\widehat{\pi}) := \{\widehat{\pi}^\sigma, \sigma \text{ permutation}\}$. By the exchangeability of ψ, we have necessarily $\psi(\{\pi\}) = \psi(\{\widehat{\pi}\}) > 0$ for any $\pi \in I(\widehat{\pi})$. Since for any $m \in \mathbb{N}$ we have $\psi(\pi_{|[m]_\star} \neq \pi_m) < \infty$, we deduce

$$\#\{\pi \in I(\widehat{\pi}),\ \pi_{|[m]_\star} \neq \pi_m\} \leq \psi(\pi_{|[m]_\star} \neq \pi_m)/\psi(\{\widehat{\pi}\}) < \infty. \tag{4.19}$$

We claim that the elements $\widehat{\pi} = (\widehat{\zeta}, \widehat{\xi}) \in \mathcal{P}^{2,\preceq}_{\infty,\star}$ satisfying $\widehat{\pi}^{\theta_n} = \pi_\infty$ and (4.19) for any m are such that $\widehat{\zeta}$ and $\widehat{\xi}$ have no more than two blocks, and in that case one of the blocks is a singleton. Indeed if $1 \sim 2 \nsim 3 \sim 4$ for $\widehat{\xi}$ or $\widehat{\zeta}$, then the permutations $\sigma_i = (2, i+2)(4, i+4)$, written as a composition of two transpositions, are such that for $i \neq j \geq n$ and $m \geq 3$, $\widehat{\pi}^{\sigma_i} \neq \widehat{\pi}^{\sigma_j}$ and $\widehat{\pi}^{\sigma_i}_{|[m]_\star} \neq \pi_m$. So having two blocks with two or more integers contradicts (4.19). One can check in the same way that the situation $1 \nsim 2 \nsim 3$ is also contradictory.

Putting everything together, we necessarily have

- either $\widehat{\pi} = (\{\{\star\}, \{i\}, \mathbb{N} \setminus \{i\}\}, \mathbf{1}_{[\infty]_\star})$ for an $i \in \mathbb{N}$,
- or $\widehat{\pi} = (\{\{\star\}, \{i\}, \mathbb{N} \setminus \{i\}\}, \{\{i\}, [\infty]_\star \setminus \{i\}\})$ for an $i \in \mathbb{N}$.

We conclude using the exchangeability of ψ that there exists two real numbers $c_1, c_2 \geq 0$ such that $\psi = c_1 \mathfrak{e}^1 + c_2 \mathfrak{e}^2$, enabling us to write

$$\mu = \varphi + \psi = \bar{\varrho}_\nu + c_1 \mathfrak{e}^1 + c_2 \mathfrak{e}^2,$$

which concludes the proof. □

Applying this result to $\widetilde{\mu}_{\text{in}}$ implies the existence of $c_{\text{in},1}, c_{\text{in},2} \geq 0$ and ν_{in} a measure on $\mathscr{S}^{\downarrow}_{\preceq}$ satisfying (4.12) such that

$$\widetilde{\mu}_{\text{in}} = c_{\text{in},1} \mathfrak{e}^1 + c_{\text{in},2} \mathfrak{e}^2 + \bar{\varrho}_{\nu_{\text{in}}}.$$

This concludes the proof of Theorem 4.14 because with our definitions in Section 4.5.1, this equality translates into

$$\mu_{\text{in}} = c_{\text{in},1} \mathfrak{e}^{\text{in},1} + c_{\text{in},2} \mathfrak{e}^{\text{in},2} + \widetilde{\varrho}_{\nu_{\text{in}}}.$$

Combining this with Lemma 4.17, we conclude

$$\mu = c_{\text{out}} \mathfrak{e}^{\text{out}} + c_{\text{in},1} \mathfrak{e}^{\text{in},1} + c_{\text{in},2} \mathfrak{e}^{\text{in},2} + \widehat{\varrho}_{\nu_{\text{out}}} + \widetilde{\varrho}_{\nu_{\text{in}}}.$$

4.6 Application to binary branching

Consider a nested fragmentation process $(\Pi(t),\, t \geq 0) = (\zeta(t), \xi(t),\, t \geq 0)$ with only binary branching. The representation given by Theorem 4.14 then becomes quite simpler, because the dislocation measures ν_{out} and ν_{in} necessarily satisfy

$$s_1 = 1 - s_2 \qquad \nu_{\text{out}}\text{-a.e.}$$

and

$$\begin{cases} & u_1 = 1 - u_2 \\ \text{or} & s_{1,1} = 1 - s_{1,2} \qquad \nu_{\text{in}}\text{-a.e.}, \\ \text{or} & u_1 = 1 - s_{1,1} \end{cases}$$

i.e. their support is the set of mass partitions with only two nonzero terms, and no dust. See Figure 4.5 for an example of a nested discrete tree illustrating the three possible dislocation events corresponding to ν_{in}.

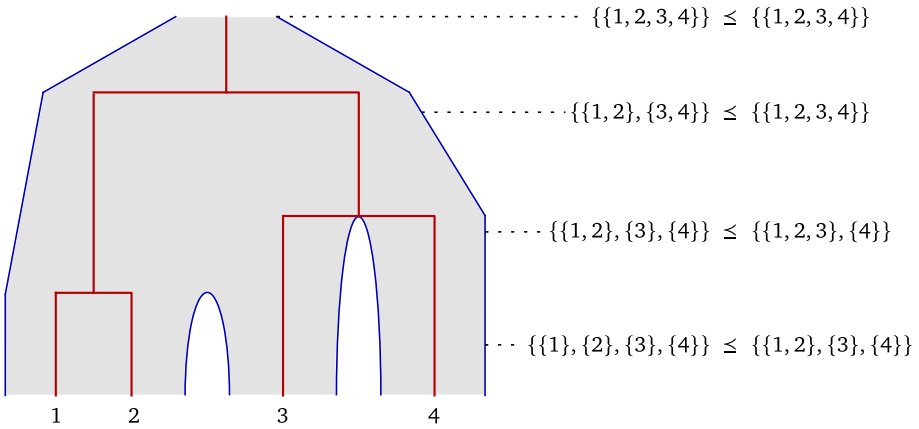


Figure 4.5 – Binary nested tree exhibiting the three different inner dislocation events. Time flows from top to bottom, and the right-hand side of the picture shows the sequence of nested partitions picked at chosen times between events, in the form $\pi = (\zeta \preceq \xi)$. The first event corresponds to the case $u_1 = 1 - u_2$, where the inner block $\{1, 2, 3, 4\}$ splits into two blocks $\{1, 2\}$ and $\{3, 4\}$ and the outer block remains unchanged. The second dislocation is of the type $u_1 = 1 - s_{1,1}$, that is the block $\{3, 4\}$ splits into two distinct blocks, one of which (the singleton $\{3\}$) stays in the *mother* outer block. The other new inner block $\{4\}$ forms a new outer block identical to itself. The last and third dislocation is of the type $s_{1,1} = 1 - s_{1,2}$, meaning that $\{1, 2\}$ splits into $\{1\}$ and $\{2\}$, these two blocks together forming a new outer block, distinct from the mother block – i.e. the one containing $\{3\}$.

Therefore, we can decompose ν_{out} and ν_{in} into four measures on $[0, 1]$ defined by

$$\begin{aligned}
\bar{\nu}_{\text{out}}(\,\cdot\,) &:= \nu_{\text{out}}(s_1 \in \,\cdot\,) + \nu_{\text{out}}(1 - s_1 \in \,\cdot\,) \\
\bar{\nu}_{\text{in},1}(\,\cdot\,) &:= \mathbb{1}\{u_1 = 1 - u_2\}(\nu_{\text{in}}(u_1 \in \,\cdot\,) + \nu_{\text{in}}(1 - u_1 \in \,\cdot\,)) \\
\bar{\nu}_{\text{in},2}(\,\cdot\,) &:= \mathbb{1}\{s_{1,1} = 1 - s_{1,2}\}(\nu_{\text{in}}(s_{1,1} \in \,\cdot\,) + \nu_{\text{in}}(1 - s_{1,1} \in \,\cdot\,)) \\
\bar{\nu}_{\text{in},3}(\,\cdot\,) &:= \mathbb{1}\{u_1 = 1 - s_{1,1}\}\nu_{\text{in}}(u_1 \in \,\cdot\,).
\end{aligned}$$

Thus defined, and because of the σ-finiteness conditions (4.4) and (4.12), those measures satisfy the following

$$\bar{\nu}_{\text{out}}, \bar{\nu}_{\text{in},1} \text{ and } \bar{\nu}_{\text{in},2} \text{ are } (x \mapsto 1 - x)\text{-invariant} \tag{4.20}$$

$$\int_{[0,1]} \nu(\mathrm{d}x)x(1 - x) < \infty, \text{ for } \nu \in \{\bar{\nu}_{\text{out}}, \bar{\nu}_{\text{in},1}\} \tag{4.21}$$

$$\bar{\nu}_{\text{in},2}([0, 1]) < \infty \tag{4.22}$$

$$\int_{[0,1]} \bar{\nu}_{\text{in},3}(\mathrm{d}x)(1 - x) < \infty. \tag{4.23}$$

For the sake of completeness, let us use those measures to express the transition rates $q^n_{\pi,\pi'}$ of the Markov chain $\Pi^n := (\Pi(t)_{|[n]})$ from one nested partition $\pi = (\zeta, \xi) \in \mathcal{P}_n^{2,\preceq}$ to another $\pi' = (\zeta', \xi') \in \mathcal{P}_n^{2,\preceq} \setminus \{\pi\}$ in the following way:

- If π' cannot be obtained from a binary fragmentation of π, then $q^n_{\pi,\pi'} = 0$.
- If π' can be obtained from a binary fragmentation of π, with $B \in \zeta$ and $C \in \xi$ two blocks of π participating in the fragmentation, but such that $B \not\subset C$, then $q^n_{\pi,\pi'} = 0$.

- Otherwise, let us write $B \subset C$, with $B \in \zeta$ and $C \in \xi$ for (the) two blocks of π participating in the fragmentation, and $B_1, B_2 \in \zeta'$, $C_1, C_2 \in \xi'$ the resulting blocks, chosen in a way that $B_1 \subset C_1$. Note that B or C might not fragment, in which case we let B_2 or C_2 be the empty set $\varnothing$. Now define $X_1 := \#B_1$ and $X_2 := \#B_2$ the cardinality of the resulting blocks of ζ'. Also, we define $Y_1 := \#\zeta'_{|C_1}$ the number of inner blocks in C_1 in the resulting partition π', and similarly $Y_2 := \#\zeta'_{|C_2}$.

With those definitions, the transition rates for the Markov chain Π^n can be written

$$\begin{aligned} q^n_{\pi,\pi'} = {} & c_{\text{out}}(\mathbb{1}\{\zeta' = \zeta, Y_1 = 1\} + \mathbb{1}\{\zeta' = \zeta, Y_2 = 1\}) \\ & + c_{\text{in},1}(\mathbb{1}\{\xi' = \xi, X_1 = 1\} + \mathbb{1}\{\xi' = \xi, X_2 = 1\}) \\ & + c_{\text{in},2}(\mathbb{1}\{X_1 = Y_1 = 1\} + \mathbb{1}\{B_2 = C_2 \text{ and } X_2 = Y_2 = 1\}) \\ & + \mathbb{1}\{\zeta' = \zeta\} \int_{[0,1]} \bar{\nu}_{\text{out}}(\mathrm{d}x) x^{Y_1}(1-x)^{Y_2} \\ & + \mathbb{1}\{\xi' = \xi\} \int_{[0,1]} \bar{\nu}_{\text{in},1}(\mathrm{d}x) x^{X_1}(1-x)^{X_2} \\ & + \mathbb{1}\{B_1 \cup B_2 = C_1\} \int_{[0,1]} \bar{\nu}_{\text{in},2}(\mathrm{d}x) x^{X_1}(1-x)^{X_2} \\ & + \mathbb{1}\{\zeta' = \zeta\} \int_{[0,1]} \bar{\nu}_{\text{in},3}(\mathrm{d}x)\big((1-x)^{\#C_1}\mathbb{1}\{Y_1 = 1\} \\ & \qquad + (1-x)^{\#C_2}\mathbb{1}\{Y_2 = 1\}\big) \\ & + \mathbb{1}\{\zeta' \neq \zeta\} \int_{[0,1]} \bar{\nu}_{\text{in},3}(\mathrm{d}x)\big(x^{X_2}(1-x)^{X_1}\mathbb{1}\{Y_1 = 1\} \\ & \qquad + x^{X_1}(1-x)^{X_2}\mathbb{1}\{Y_2 = 1\}\big). \end{aligned} \tag{4.24}$$

Note that several indicator functions in the last display may be equal to 1 for the same pair (π, π'). This explicit formula allows for computer simulations of binary nested fragmentations, although to that aim it might be simpler to adapt the Poissonian construction (Section 4.4.3) and use nested partitions of arrays $[n]^2$. Also, one could exactly compute the probability of a given nested tree under different nested fragmentation models, which would be a first step towards statistical inference.

References for Chapter 4

[3] D. Aldous. Probability Distributions on Cladograms. *Random Discrete Structures*. The IMA Volumes in Mathematics and Its Applications 76. Springer New York, 1996, pp. 1–18. DOI: 10.1007/978-1-4612-0719-1_1 (see pp. 9, 90).

[10] J. Bertoin. *Random Fragmentation and Coagulation Processes*. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511617768 (see pp. 10, 11, 60, 61, 64, 65, 79, 81, 90, 91, 94, 95, 99, 117, 126–128, 135, 153).

[14] J. Bertoin. The Structure of the Allelic Partition of the Total Population for Galton–Watson Processes with Neutral Mutations. *Ann. Probab.*, 37.4 (July 2009), pp. 1502–1523. DOI: 10.1214/08-AOP441 (see pp. 16, 59, 90).

[19] A. Blancas, T. Rogers, J. Schweinsberg, and A. Siri-Jégousse. The Nested Kingman Coalescent: Speed of Coming down from Infinity. *Ann. Appl. Probab.*, 29.3 (June 2019), pp. 1808–1836. doi: 10.1214/18-AAP1440 (see pp. 59, 60, 90).

[26] B. Chen, D. Ford, and M. Winkel. A New Family of Markov Branching Trees: The Alpha-Gamma Model. *Electron. J. Probab.*, 14 (2009), pp. 400–430. doi: 10.1214/EJP.v14-616 (see p. 90).

[27] H. Crane. Generalized Markov Branching Trees. *Adv. in Appl. Probab.*, 49.01 (Mar. 2017), pp. 108–133. doi: 10.1017/apr.2016.81 (see pp. 90, 91).

[28] H. Crane and H. Towsner. The Structure of Combinatorial Markov Processes (Mar. 18, 2016). arXiv: 1603.05954 [math.PR] (see p. 98).

[35] J. J. Doyle. Trees within Trees: Genes and Species, Molecules and Morphology. *Syst. Biol.*, 46.3 (Sept. 1, 1997), pp. 537–553. doi: 10.1093/sysbio/46.3.537 (see pp. 59, 90).

[36] J.-J. Duchamps. Trees within Trees II: Nested Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat. (to appear)*, (2019+). arXiv: 1807.05951 (see pp. 11, 89, 133, 153).

[40] A. Etheridge. *Some Mathematical Models from Population Genetics: École d'ete de Probabilités de Saint-Flour XXXIX-2009*. Lecture Notes in Mathematics 2012. Heidelberg ; New York: Springer, 2011 (see pp. 59, 90).

[44] D. J. Ford. *Probabilities on Cladograms: Introduction to the Alpha Model*. Stanford University, 2006 (see p. 90).

[45] C. Foucart. Distinguished Exchangeable Coalescents and Generalized Fleming-Viot Processes with Immigration. *Adv. in Appl. Probab.*, 43.02 (June 2011), pp. 348–374. doi: 10.1239/aap/1308662483 (see p. 113).

[53] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum Tree Asymptotics of Discrete Fragmentations and Applications to Phylogenetic Models. *Ann. Probab.*, 36.5 (Sept. 2008), pp. 1790–1837. doi: 10.1214/07-AOP377 (see pp. 10, 90, 91, 126).

[57] J. F. C. Kingman. The Representation of Partition Structures. *J. Lond. Math. Soc. (2)*, 18.2 (1978), pp. 374–380. doi: 10.1112/jlms/s2-18.2.374 (see pp. 96, 130).

[58] J. Kingman. The Coalescent. *Stochastic Process. Appl.*, 13.3 (1982), pp. 235–248. doi: 10.1016/0304-4149(82)90011-4 (see pp. 8, 15, 22, 59, 90, 95).

[61] A. Lambert. Population Dynamics and Random Genealogies. *Stoch. Models*, 24 (sup1 2008), pp. 45–163. doi: 10.1080/15326340802437728 (see pp. 8, 58, 90).

[62] A. Lambert. Probabilistic Models for the (Sub)Tree(s) of Life. *Braz. J. Probab. Stat.*, 31.3 (Aug. 2017), pp. 415–475. doi: 10.1214/16-BJPS320 (see p. 90).

[65] A. Lambert and E. Schertzer. Coagulation-Transport Equations and the Nested Coalescents. *Probab. Theory Related Fields*, (Apr. 15, 2019). DOI: 10.1007/s00440-019-00914-4 (see pp. 59, 60, 90).

[68] W. P. Maddison. Gene Trees in Species Trees. *Syst. Biol.*, 46.3 (Sept. 1, 1997), pp. 523–536. DOI: 10.1093/sysbio/46.3.523 (see pp. 59, 90).

[73] R. D. Page and M. A. Charleston. From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem. *Mol. Phylogenet. Evol.*, 7.2 (Apr. 1997), pp. 231–240. DOI: 10.1006/mpev.1996.0390 (see pp. 59, 90).

[74] R. D. Page and M. A. Charleston. Trees within Trees: Phylogeny and Historical Associations. *Trends Ecol. Evol.*, 13.9 (Sept. 1998), pp. 356–359. DOI: 10.1016/S0169-5347(98)01438-4 (see pp. 59, 90).

[76] J. Pitman. Coalescents with Multiple Collisions. *Ann. Probab.*, 27.4 (Oct. 1999), pp. 1870–1902. DOI: 10.1214/aop/1022874819 (see pp. 60, 63, 65, 83, 84, 90).

[79] S. Sagitov. The General Coalescent with Asynchronous Mergers of Ancestral Lines. *J. Appl. Probab.*, 36.4 (Dec. 1999), pp. 1116–1125. DOI: 10.1017/S0021900200017903 (see pp. 60, 65, 90).

[85] C. Semple and M. Steel. *Phylogenetics*. Oxford Lecture Series in Mathematics and Its Applications 24. Oxford ; New York: Oxford University Press, 2003 (see pp. 58, 90).

[87] S. M. Srivastava. *A Course on Borel Sets*. Vol. 180. Graduate Texts in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. DOI: 10.1007/978-3-642-85473-6 (see pp. 109, 129).

Chapter 5

Fragmentations with self-similar branching speeds

This chapter is submitted to *Advances in Applied Probability*.

Contents

5.1 Introduction

A fragmentation process is a system of particles evolving in time in a Markovian way, where each particle is assigned a mass and may dislocate at random times, distributing its mass among newly created particles. It is usually assumed that particles evolve independently of one another, in a way depending only on their mass. Self-similar fragmentations are processes where the speed of fragmentation of a particle is accelerated proportionally to a

function of its mass – which then must be a power function, characterized by an exponent $\alpha \in \mathbb{R}$. These processes are said to be homogeneous when $\alpha = 0$. Homogeneous and self-similar fragmentations have been characterized in the early 2000s (see [6, 11], or [10] for a general introduction), and their connections to random trees have been developed in e.g. [4] or [52, 53].

These studies have been made under a conservative assumption, which prevents the total mass in the system from increasing. This assumption allows for instance the representation of fragmentation processes in terms of exchangeable partition-valued processes, which are convenient objects allowing one to naturally recover *discrete* genealogical structures in fragmentation processes.

The primary goal of this article is to extend the self-similar assumption while staying in a conservative setting. To this aim, we assume that particles are described by a pair mass-mark which evolves jointly in a Markovian way, such that a) the total mass does not increase, and b) it is now the *mark* – which may a priori fluctuate in any way – of a particle which determines the speed at which it fragments. The conservative assumption allows us to model this idea with Markov processes taking values in marked partitions of the integers, with very little restriction concerning marks. Consequently, if one ignores the masses of particles, our processes essentially give constructions for quite general non-conservative fragmentations. Related and inspiring works include self-similar branching Markov chains [59], the recent so-called branching Lévy processes of [17], as well as many recent developments which have been published on self-similar growth-fragmentation processes (see e.g. [29, 49, 86]), introduced by Bertoin [9], which allow masses of particles to fluctuate as a positive Markov process.

The article is organized as follows. In the remainder of the introduction, we recall some definitions and basic results of usual self-similar fragmentations, and define the space of marked partitions in which our processes live. In Section 5.2 we define our extended self-similar fragmentation (ESSF) processes, and point out their basic properties. We characterize ESSF processes with a type of Lévy-Khinchin representation in Section 5.3, and then give sufficient conditions for a process to almost surely a) reach an absorbing state in finite time b) have a genealogy where the sum of lengths of all branches is finite. Because most proofs are somewhat technical, we defer them to Appendix 5.A to ease the exposition.

5.1.1 Self-similar fragmentations

To study processes with values in the space of partitions of $\mathbb{N}$, let us recall some classical notation and definitions. First define $[n] := \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$ and $[\infty] := \mathbb{N} := \{1, 2, \dots\}$. Now for $n \in \mathbb{N} \cup \{\infty\}$, we denote by $\mathcal{P}_n$ the space of partitions of $[n]$. We often see a partition $\pi \in \mathcal{P}_n$ as the equivalence relation $\sim^\pi$ it represents on $[n]$. We will denote by $\mathbf{0}_n$ (resp. $\mathbf{1}_n$) the partition of $[n]$ into singletons (resp. the partition with a single block $\{[n]\}$). We will often omit the subscript n and write only $\mathbf{0}$ or $\mathbf{1}$ when the context is clear.

For $n < m \leq \infty$ and $\pi \in \mathcal{P}_m$, we denote by $\pi_{|[n]}$ its restriction to the set $[n] \subset [m]$. $\mathcal{P}_\infty$ may be understood as the projective limit of the sets $(\mathcal{P}_n, n \in \mathbb{N})$, and as such, a natural

metric which makes this space compact may be defined on it by

$$d(\pi, \pi') = \sup\{n \in \mathbb{N},\ \pi_{|[n]} = \pi'_{|[n]}\}^{-1},$$

where by convention $(\sup \mathbb{N})^{-1} = 0$. We will consider the action of permutations of $\mathbb{N}$ on $\mathcal{P}_\infty$, and more generally we can define, for any $1 \leq n \leq m \leq \infty$, any *injection* $\sigma : [n] \to [m]$ and any $\pi \in \mathcal{P}_m$, the partition $\pi^\sigma \in \mathcal{P}_n$ defined by:

$$i \sim^{\pi^\sigma} j \iff \sigma(i) \sim^{\pi} \sigma(j), \qquad i, j \in [n].$$

Note that in this paper, a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection with *finite* support $\{n \in \mathbb{N},\ \sigma(n) \neq n\}$. We usually label the blocks of a partition $\pi = \{\pi_1, \pi_2, \ldots\}$ in the unique way such that the sequence $(\min \pi_k, k \geq 1)$ is increasing. This way, π_1 is necessarily the block containing 1, π_2 is the block containing the lowest integer not in the same block as 1, etc. By convention, if π has a finite number of blocks, say K, we define $\pi_{K+l} = \varnothing$ for all $l \geq 1$. It will be useful to define a fragmentation operator $\mathrm{Frag} : \mathcal{P}_\infty \times (\mathcal{P}_\infty)^{\mathbb{N}} \to \mathcal{P}_\infty$ by

$$\mathrm{Frag}(\pi, \pi^{(\cdot)}) = \{\pi_k \cap \pi_l^{(k)},\ k, l \geq 1\},$$

where (π_k) are the ordered blocks of π and $(\pi_l^{(k)})$ the ordered blocks of $\pi^{(k)}$. In words, blocks of the new partition are formed from the restriction of the k-th partition of the sequence $\pi^{(\cdot)}$ to π_k, for each $k \geq 1$.

Now let us recall the definition of partition-valued fragmentation processes (see e.g. [10]). For this definition, we restrict ourselves to the space of partitions that have asymptotic frequencies, i.e. $\pi \in \mathcal{P}_\infty$ such that for all $k \geq 1$,

$$|\pi_k| := \lim_{n \to \infty} \frac{\#\pi_k \cap [n]}{n} \quad \text{exists.}$$

In this case, we write $|\pi|^\downarrow$ for the nonincreasing reordering of the sequence $(|\pi_1|, |\pi_2|, \ldots)$. Let us write $\mathcal{P}'_\infty$ for the space of partitions of $\mathbb{N}$ with asymptotic frequencies.

Definition 5.1. A self-similar fragmentation process is a càdlàg Markov process $(\Pi(t), t \geq 0)$ with values in $\mathcal{P}'_\infty$, such that almost surely for all $k \in \mathbb{N}$, the map $t \mapsto |\Pi_k(t)|$ is right-continuous and for which the following properties hold.

(i) *Exchangeability*: for all $\pi \in \mathcal{P}'_\infty$, for all $\sigma : \mathbb{N} \to \mathbb{N}$ permutation,

$$(\Pi(t)^\sigma,\ t \geq 0) \text{ under } \mathbb{P}_\pi \overset{(d)}{=} (\Pi(t),\ t \geq 0) \text{ under } \mathbb{P}_{\pi^\sigma},$$

where $\mathbb{P}_\pi$ denotes the distribution of the process started from π.

(ii) *Self-similar branching*: there exists $\alpha \in \mathbb{R}$ such that if $(\Omega, \mathbb{P})$ is a probability space where $(\Pi^{(\cdot)}(t), t \geq 0)$ is a sequence of independent copies of the process started from $\mathbf{1}$, then for any $\pi \in \mathcal{P}'_\infty$, we have

$$(\Pi(t),\ t \geq 0) \text{ under } \mathbb{P}_\pi \overset{(d)}{=} (\mathrm{Frag}(\pi, \widetilde{\Pi}^{(\cdot)}(t)),\ t \geq 0) \text{ under } \mathbb{P}, \tag{5.1}$$

where $\widetilde{\Pi}^{(\cdot)}$ is the sequence of time-changed processes defined by

$$\widetilde{\Pi}^{(k)}(t) = \Pi^{(k)}(|\pi_k|^\alpha t), \qquad k \geq 1, t \geq 0.$$

Note that a fragmentation with self-similarity index $\alpha = 0$ is called homogeneous. It is well-known (we refer to [10, Section 1 to 3] for a detailed account on the theory of partition-valued fragmentations) that self-similar fragmentations can be characterized in terms of their self-similarity index α, a so-called erosion coefficient $c \geq 0$ and a dislocation measure ν on the (metric and compact when equipped with the uniform distance) space

$$\mathscr{S}^{\downarrow} := \{\mathbf{s} = (s_1, s_2, \ldots) \in [0,1]^{\mathbb{N}} \text{ where } s_1 \geq s_2 \geq \ldots \geq 0 \text{ and } \sum_k s_k \leq 1\},$$

satisfying

$$\int_{\mathscr{S}^{\downarrow}} (1 - s_1)\, \nu(\mathrm{d}\mathbf{s}) < \infty.$$

In words, c is the rate at which each singleton detaches from "macroscopic" blocks and ν is a measure giving the rates of "sudden dislocations", i.e. a block with asymptotic frequency x fragments at rate $\nu(\mathrm{d}\mathbf{s})$ into (possibly infinitely many) blocks with frequencies given by $x\mathbf{s} = (xs_1, xs_2, \ldots)$ – these dislocations of blocks are usually represented by a so-called paintbox process, which we will define in the context of marked partitions in the next section. The self-similarity index α of a fragmentation encodes, through property (5.1), the speed at which blocks fragment, depending on their size. For instance, if α is negative, then there is a random time T which is finite almost surely at which $\Pi(T)$ is the partition into singletons, whereas it is never the case when $\alpha \geq 0$ and $\nu(s_1 = 0) = 0$. Note that $\alpha = 0$ means that there is no time change – in that case the sequence $\widetilde{\Pi}^{(\cdot)}$ in (5.1) is simply $\Pi^{(\cdot)}$ – the process is then said to be homogeneous.

Our goal is to generalize these objects and define processes $(\Pi(t), \mathbf{V}(t), t \geq 0)$, where Π is partition-valued and $\mathbf{V}(t) = (V_n(t), n \geq 1)$ is a random map $\mathbb{N} \to [0, \infty)$ playing the role of $(|\pi_k|^{\alpha}, k \geq 1)$, i.e. dictating the speed of fragmentation of different blocks of Π. To define this we need first to introduce the formalism of marked partitions and processes in this space.

5.1.2 Partitions with marks

Let us consider partitions where each block is decorated with a mark. For convenience, we consider that the space of marks is the space $[0, \infty]$ where 0 is identified with ∞. Topologically it is a circle so we will denote it by S^1, but throughout the paper elements of S^1 will be identified with their unique representative in $[0, \infty)$, and this enables us to consider for instance the maps

$$m_x : v \mapsto xv \qquad \text{and} \qquad p_\alpha : v \mapsto v^\alpha$$

as well-defined and continuous, where x is in S^1 or $[0, \infty)$ and $\alpha \in \mathbb{R} \setminus \{0\}$. Note that for a technical reason, we choose to use throughout the article the convention $0^0 = 0$, so that $0^\alpha = 0$ for any $\alpha \in \mathbb{R}$, and $v^0 = \mathbb{1}_{v \neq 0}$ for any $v \in S^1$. For convenience and with a slight abuse we will often identify an element of $[0, \infty)$ with the corresponding element of S^1.

For $n \in \mathbb{N} \cup \{\infty\}$, we consider the space of marked partitions defined by

$$\mathcal{M}_n := \{x = (\pi, \mathbf{v}) \in \mathcal{P}_n \times (S^1)^{[n]},\ \forall i, j \in [n],\ i \sim^{\pi} j \implies v_i = v_j\}.$$

It is a closed subset of $\mathcal{P}_n \times (S^1)^{[n]}$, which, endowed with the product topology, is compact metrizable, therefore Polish. Note that by definition, if $(\pi, \mathbf{v}) \in \mathcal{M}_n$ where $\pi = \mathbf{1}$ is the partition into a single block, then $\mathbf{v}$ is of the form $(v, v, \ldots)$ for a unique $v \in S^1$. For this reason we will use the abuse of notation $(\mathbf{1}, v)$ to denote this element. We see $x = (\pi, \mathbf{v})$ as the partition π where each block is given a mark. Therefore, we will sometimes say B is *a block of x with mark v* if $B \in \pi$ and $v_i = v$ for some (hence all) $i \in B$. Similarly, we will use the notation $i \sim^x j$ if i and j are in the same block of π.

Note that for $n < m \leq \infty$ and $x = (\pi, \mathbf{v}) \in \mathcal{M}_m$, we can naturally consider the restrictions $x_{|[n]} = (\pi, \mathbf{v})_{|[n]} := (\pi_{|[n]}, (v_1, v_2, \ldots, v_n)) \in \mathcal{M}_n$, which are clearly continuous maps.

Similarly, we can extend the action of injections $\sigma : [n] \to [m]$ to our context and define for $x = (\pi, \mathbf{v}) \in \mathcal{M}_m$,

$$x^\sigma = (\pi, \mathbf{v})^\sigma = (\pi^\sigma, \mathbf{v}^\sigma) := (\pi^\sigma, (v_{\sigma(i)},\, i \in [n])) \in \mathcal{M}_n.$$

We say that a random variable X with values in $\mathcal{M}_\infty$ is exchangeable if for all $\sigma : \mathbb{N} \to \mathbb{N}$ permutations,

$$X^\sigma \overset{(d)}{=} X.$$

Finally we can also extend the fragmentation operator Frag to marked partitions by setting

$$\mathrm{Frag}\big((\pi, \mathbf{v}), (\pi^{(\cdot)}, \mathbf{v}^{(\cdot)})\big) := \big(\mathrm{Frag}(\pi, \pi^{(\cdot)}), \widetilde{\mathbf{v}}\big),$$

where, for $i \geq 1$, $\widetilde{v}_i$ is defined by $v_i v_i^{(k_i)}$, where k_i is the label of the block containing i – so that i is in the k_i-th block of π.

We say that a marked partition $x \in \mathcal{M}_\infty$ is *non-degenerate* if every finite block has mark 0, and we denote the space of non-degenerate marked partitions by

$$\mathcal{M}_\infty^\star := \big\{x = (\pi, \mathbf{v}) \in \mathcal{M}_\infty,\ \forall i \geq 1,\ i \text{ in a finite block of } \pi \implies v_i = 0\big\}.$$

In particular for singleton blocks, $\{i\} \in \pi$ implies $v_i = 0$. Note that this space is still Polish [see e.g. 87, Theorem 2.2.1] as a G_δ-subset – a countable intersection of open sets – of $\mathcal{M}_\infty$. Indeed, letting for all $i \in \mathbb{N}$, $N_i : \mathcal{M}_\infty \to \mathbb{N} \cup \{\infty\}$ be the map associating $(\pi, \mathbf{v})$ with the cardinality of the block of π containing i, then, taking d to be any metric on S^1 compatible with its topology, we can write

$$\begin{aligned}\mathcal{M}_\infty^\star &= \bigcap_{i \geq 1} \{N_i < \infty \implies v_i = 0\} \\ &= \bigcap_{i,j,k \geq 1} \big(\{N_i \geq j\} \cup \{d(v_i, 0) < 1/k\}\big),\end{aligned}$$

which is a countable intersection of open subsets of $\mathcal{M}_\infty$. Note that if n is finite, one cannot define an analogous property of non-degeneracy for marked partitions in $\mathcal{M}_n$.

Now let us define paintbox processes for exchangeable marked partitions. Consider the space $([0,1] \times [0,\infty), \preceq)$ equipped with the lexicographic order, that is if $z = (s, v) \in [0,1] \times [0,\infty)$ and $z' = (s', v') \in [0,1] \times [0,\infty)$, then

$$z \preceq z' \iff s < s' \text{ or } (s = s' \text{ and } v \leq v').$$

Let us define

$$\mathscr{Z}_0^{\downarrow} := \{\mathbf{z} = (z_1, z_2, \ldots) \in \big([0,1] \times [0,\infty)\big)^{\mathbb{N}},\ z_1 \succeq z_2 \succeq \ldots,\ \text{and}\ \sum_k s_k \leq 1\},$$

and note that, endowed with the product topology, it is a Polish space. Indeed, it can be written

$$\mathscr{Z}_0^{\downarrow} = \{\sum_k s_k \leq 1\} \cap \bigcap_{i \geq 1} \Big(\{s_i > s_{i+1}\} \cup \{s_i = s_{i+1} \text{ and } v_i \geq v_{i+1}\}\Big),$$

which is a countable intersection of closed and open subsets of $\big([0,1] \times [0,\infty)\big)^{\mathbb{N}}$. This space being Polish, closed sets are G_δ, and so $\mathscr{Z}_0^{\downarrow}$ is Polish. Because this will be consistent with our previous definition of $\mathcal{M}_\infty^\star$, we want to ignore the possible indices $k \geq 1$ such that $s_k = 0$. Therefore, we will rather use the space

$$\begin{aligned}\mathscr{Z}^{\downarrow} &:= \{\mathbf{z} \in \mathscr{Z}_0^{\downarrow},\ \forall k \geq 1,\ s_k = 0 \implies v_k = 0\} \\ &= \bigcap_{k,l \geq 1} \{\mathbf{z} \in \mathscr{Z}_0^{\downarrow},\ s_k > 0 \text{ or } v_k < 1/l\},\end{aligned}$$

which is still Polish.

Similarly as in the usual case, we say that $x = (\pi, \mathbf{v}) \in \mathcal{M}_\infty$ has asymptotic frequencies if π has asymptotic frequencies. In that case, we define $|x|^{\downarrow} \in \mathscr{Z}^{\downarrow}$ as the nonincreasing reordering (with respect to the lexicographic order $\preceq$ on $[0,1] \times [0,\infty)$) of the sequence of pairs

$$\big((|B_i|, v_i),\ B_i \text{ is the } i\text{-th infinite block of } x \text{ such that } |B_i| > 0 \text{ and with mark } v_i\big).$$

Note that we consider only blocks B satisfying $|B| > 0$ in the previous display since in general the set $\{(|B_i|, v_i),\ B_i \text{ is the } i\text{-th block of } x, \text{ with mark } v_i\}$ may be impossible to enumerate in nonincreasing order.

Now let us introduce a paintbox construction for marked partitions. Consider $\mathbf{z} = (\mathbf{s}, \mathbf{v}) \in \mathscr{Z}^{\downarrow}$, and let $(U_n, n \geq 1)$ be an i.i.d. sequence of $[0,1]$-uniform random variables. Define $X = (\Pi, \mathbf{V})$ as the $\mathcal{M}_\infty^\star$-valued random variable given by the following relation:

$$\begin{aligned} i \sim^{\Pi} j &\iff i = j \text{ or } \exists n \geq 1,\ t_{n-1} \leq U_i, U_j < t_n, \\ V_i &:= \begin{cases} v_n & \text{if } t_{n-1} \leq U_i < t_n, \text{ for } n \geq 1, \\ 0 & \text{if } \sum_k s_k \leq U_i, \end{cases}\end{aligned}$$

where $t_n := \sum_{k=1}^n s_k$, with $t_0 := 0$ by convention. It is easily checked that the random variable X is exchangeable. Also, recall the definition of asymptotic frequencies for a marked partition, and note that the law of large numbers implies $|X|^{\downarrow} = \mathbf{z}$ almost surely. We denote by $\varrho_{\mathbf{z}}$ the distribution of X. We will also make use of the distribution of $X_{|[n]}$ for $n \in \mathbb{N}$, which we denote by $\varrho_{\mathbf{z}}^n$. Note that for any $v \in S^1$, if $\mathbf{z} = (\mathbf{s}, \mathbf{v}) \in \mathscr{Z}^{\downarrow}$ is the unique element such that $s_1 = 1$ and $v_1 = v$, then $\varrho_{\mathbf{z}} = \delta_{(\mathbf{1},v)}$. For this reason, we will again abuse notation and let $(\mathbf{1}, v) \in \mathscr{Z}^{\downarrow}$ denote this element, so that $\varrho_{(\mathbf{1},v)} = \delta_{(\mathbf{1},v)}$.

It is well-known since the work of Kingman [57] that the law of an exchangeable partition can be expressed as a mixture of paintbox processes. Using the same arguments, one obtains the following result for marked partitions.

Proposition 5.2. *Let X be an exchangeable random variable with values in $\mathcal{M}_\infty^\star$. Then there exists a unique probability measure ν on $\mathscr{Z}^\downarrow$ such that*

$$\mathbb{P}(X \in \cdot) = \int_{\mathscr{Z}^\downarrow} \varrho_{\mathbf{z}}(\cdot)\, \nu(\mathrm{d}\mathbf{z}). \tag{5.2}$$

Proof. See Appendix 5.A.1. □

This setting of marked partitions being in place, we can now define our objects of study.

5.2 Extended self-similar fragmentations

5.2.1 Definitions, first properties

Let us now define self-similar fragmentation processes with values in $\mathcal{M}_\infty$. For this, let us introduce a family of *self-similar fragmentation* operators $(\mathrm{ssFrag}_\alpha, \alpha \in \mathbb{R})$, defined as follows. For $n \in \mathbb{N} \cup \{\infty\}$, consider a marked partition $x = (\pi, \mathbf{v}) \in \mathcal{M}_n$ and a sequence $\bar{x}^{(\cdot)}$ of càdlàg *maps* $\bar{x}^{(k)} \colon [0,\infty) \to \mathcal{M}_n$, satisfying $\bar{x}^{(k)}(0) = (\mathbf{1}, 1)$. Writing for all $k \geq 1$ and $t \geq 0$, $\bar{x}^{(k)}(t) = (\bar{\pi}^{(k)}(t), \bar{\mathbf{v}}^{(k)}(t))$, we define

$$\mathrm{ssFrag}_\alpha(x, \bar{x}^{(\cdot)}) := (\hat{\pi}, \hat{\mathbf{v}})$$

as the map $[0,\infty) \to \mathcal{M}_n$ such that

$$\hat{v}_i(t) = v_i \bar{v}^{(k_i)}(v_i^\alpha t)$$
$$i \sim^{\hat{\pi}(t)} j \iff i \sim^{\pi} j \text{ and } i \sim j \text{ in } \bar{\pi}^{(k_i)}(v_i^\alpha t),$$

where k_i is defined as the label of the block of π containing i (i.e. such that i is in the k_i-th block of π). Note that thanks to this definition, if a block B of x has mark 0, then the process $\mathrm{ssFrag}_\alpha(x, \bar{x}^{(\cdot)})$ is frozen at block B, in the sense that for all $t \geq 0$, B is a block of $\hat{\pi}(t)$, and every $j \in B$ will have $\hat{v}_j(t) = 0$. Also, the assumptions on the maps $\bar{x}^{(k)}$ imply that $\mathrm{ssFrag}_\alpha(x, \bar{x}^{(\cdot)})$ is càdlàg and satisfies $\mathrm{ssFrag}_\alpha(x, \bar{x}^{(\cdot)})(0) = x$.

Remark 5.3.

(i) Consider here a convergent sequence $x_n = (\pi_n, \mathbf{v}_n) \to x = (\pi, \mathbf{v}) \in \mathcal{M}_\infty$, and assume that $v_{n,i} = 0$ for all $n \geq 1$ whenever $v_i = 0$ for some i. If additionally we have for some $t \geq 0$, for all $i \geq 1$ such that $v_i > 0$, and for all $k \geq 1$,

$$\bar{x}^{(k)}(v_{n,i}^\alpha t) \underset{n\to\infty}{\longrightarrow} \bar{x}^{(k)}(v_i^\alpha t),$$

then it is a straightforward consequence of the definition that

$$\mathrm{ssFrag}_\alpha(x_n, \bar{x}^{(\cdot)})(t) \underset{n\to\infty}{\longrightarrow} \mathrm{ssFrag}_\alpha(x, \bar{x}^{(\cdot)})(t).$$

(ii) Note that one could define ssFrag_α in terms of Frag because we have the equality

$$\mathrm{ssFrag}_\alpha\big(x, \bar{x}^{(\cdot)}\big)(t) = \mathrm{Frag}\big(x, \bar{x}^{(\cdot)}(w_{(\cdot)}^\alpha t)\big),$$

where $w_{(\cdot)}$ is the vector defined by $w_{(k)} = v_i$, for any i in the k-th block of π.

We can now define the following generalization of self-similar fragmentations.

Definition 5.4. Let $X(t) = (\Pi(t), \mathbf{V}(t), t \geq 0)$ be a stochastic process with values in $\mathcal{M}_\infty$. We say that X is an *extended self-similar fragmentation* (ESSF) process if it is a *stochastically continuous strong Markov* process with *càdlàg sample paths*, for which the following properties hold:

(i) *Exchangeability:* for all permutations $\sigma : \mathbb{N} \to \mathbb{N}$, for all $x \in \mathcal{M}_\infty$,

$$(X(t)^\sigma,\, t \geq 0) \text{ under } \mathbb{P}_x \overset{(d)}{=} (X(t),\, t \geq 0) \text{ under } \mathbb{P}_{x^\sigma},$$

where $\mathbb{P}_x$ denotes the distribution of the Markov process started from x.

(ii) *Self-similar branching:* there exists $\alpha \in \mathbb{R}$ such that for all $x \in \mathcal{M}_\infty$,

$$X \text{ under } \mathbb{P}_x \overset{(d)}{=} \mathrm{ssFrag}_\alpha(x, X^{(\cdot)}),$$

where $X^{(\cdot)}$ is an i.i.d. sequence of copies of the process started from $(\mathbf{1}, 1)$. As usual, we call α the index of self-similarity, and we will say for conciseness that X is an α-ESSF. For the special case $\alpha = 0$, we will sometimes say the process X is *homogeneous*.

An ESSF process X will be called *non-degenerate* if for all $x \in \mathcal{M}_\infty^\star$, the process has sample paths in $\mathcal{M}_\infty^\star$, $\mathbb{P}_x$-almost surely.

Remark 5.5.

(i) Consider $X = (\Pi, \mathbf{V})$ an α-ESSF and $\gamma \in \mathbb{R} \setminus \{0\}$. Let us define $Y := (\Pi, \mathbf{V}^\gamma)$, where $\mathbf{V}^\gamma(t)$ is simply the vector $(V_i(t)^\gamma,\, i \geq 1)$. Then it is easily checked that Y is an ESSF again, with index of self-similarity α/γ. Therefore, if $\alpha \neq 0$ and $\beta \neq 0$, taking $\gamma = \alpha/\beta$, one can transform any α-ESSF into a β-ESSF, but note that one cannot get a homogeneous process with this transformation. As a result, there are really two classes of ESSF processes to consider: the α-ESSF with $\alpha \neq 0$, which are a simple transformation away from being 1-ESSF processes, and the so-called homogeneous 0-ESSF processes.

(ii) Note that this definition extends the classical case of Definition 5.1. Indeed, if Π is a usual α-self-similar fragmentation process started from $\mathbf{1}$, then by definition, almost surely for all $t \geq 0$ and $i \in \mathbb{N}$, $\Pi(t)$ has asymptotic frequencies and one can define $V_i(t) := |B|$ if B is the block containing i in $\Pi(t)$. Now consider an independent sequence $X^{(\cdot)}$ of copies of $(\Pi, \mathbf{V})$, and define for any $x \in \mathcal{M}_\infty$,

$$X_x = \mathrm{ssFrag}_\alpha(x, X^{(\cdot)}).$$

Then X_x is the distribution of an α-ESSF started from x, which extends the usual self-similar fragmentation Π – consider $x = (\mathbf{1}, 1)$ to obtain the original process. Note also that in this case X is non-degenerate, because finite blocks have asymptotic frequency equal to 0.

As a first remark about ESSF processes, let us show a projective Markov property. It is very analogous to [18, Lemma 3.2] and [36, Proposition 2], but we need another statement in the present context.

Lemma 5.6. *Let X be an ESSF process. Then for any $n \in \mathbb{N}$, the process $(X(t)_{|[n]}, t \geq 0)$ is Markovian in $\mathcal{M}_n$. More precisely, there exists a transition kernel $(p_t^n, t \geq 0)$ on $\mathcal{M}_n$ such that for any initial state $x \in \mathcal{M}_\infty$,*

$$\mathbb{P}_x(X(t)_{|[n]} \in \cdot) = p_t^n(x_{|[n]}, \cdot).$$

Proof. We only need to prove that

$$\mathbb{P}_x\big(X(t)_{|[n]} \in \cdot\big) = \mathbb{P}_{x'}\big(X(t)_{|[n]} \in \cdot\big)$$

for any two initial states $x, x' \in \mathcal{M}_\infty$ such that $x'_{|[n]} = x_{|[n]}$.

Consider a probability space $(\Omega, \mathbb{P})$ such that $X^{(\cdot)}$ is a sequence of i.i.d. copies of the ESSF process started from $(\mathbf{1}, 1)$, and let $\alpha \in \mathbb{R}$ be the self-similarity index of X. By the branching property, we have

$$\begin{aligned}\mathbb{P}_x(X(t)_{|[n]} \in \cdot) &= \mathbb{P}\big(\mathrm{ssFrag}_\alpha(x, X^{(\cdot)})(t)_{|[n]} \in \cdot\big)\\ \text{and } \mathbb{P}_{x'}(X(t)_{|[n]} \in \cdot) &= \mathbb{P}\big(\mathrm{ssFrag}_\alpha(x', X^{(\cdot)})(t)_{|[n]} \in \cdot\big).\end{aligned}$$

It remains to notice that by definition, $\mathrm{ssFrag}_\alpha(x, X^{(\cdot)})(t)_{|[n]}$ is in fact a functional which depends only on $x_{|[n]}$ and $X^{(\cdot)}$. Therefore, because $x'_{|[n]} = x_{|[n]}$, we have

$$\mathrm{ssFrag}_\alpha(x, X^{(\cdot)})(t)_{|[n]} = \mathrm{ssFrag}_\alpha(x', X^{(\cdot)})(t)_{|[n]}$$

everywhere on Ω, which implies by the preceding display that

$$\mathbb{P}_x\big(X(t)_{|[n]} \in \cdot\big) = \mathbb{P}_{x'}\big(X(t)_{|[n]} \in \cdot\big),$$

concluding the proof. □

The previous lemma shows that given an ESSF process X, one can define its law started from any $x_0 \in \mathcal{M}_n$, for any $n \in \mathbb{N}$, as the law of the restriction $X_{|[n]}$ of the initial process started from any $x \in \mathcal{M}_\infty$ such that $x_{|[n]} = x_0$.

As a result, the restriction $X_{|[1]}$ of an ESSF process $X = (\Pi, \mathbf{V})$ to $\mathcal{M}_1 = \mathcal{P}_1 \times S^1$ is a Markov process. Since the space $\mathcal{P}_1$ is a singleton, the lemma implies that the real-valued process $V_1 = (V_1(t), t \geq 0)$ is a Markov process in S^1 and note that by exchangeability, the process V_i has the same marginal distribution for all $i \geq 1$. Further, Definition 5.4 implies that it is an a.s. càdlàg strong Markov process satisfying a self-similar property; more precisely, for $v \geq 0$ let P_v denote the distribution of V_1 started at v on the Skorokhod space of càdlàg maps $[0, \infty) \to S^1$, and let V denote the canonical process on that space. Then

$$(V(t),\, t \geq 0) \text{ under } P_v \overset{(d)}{=} (vV(v^\alpha t),\, t \geq 0) \text{ under } P_1,$$

where α is the self-similarity index of X. In other words, V is a positive self-similar Markov process (pssMp). Note that in the literature, the index of self-similarity of a

pssMp refers in general to $-\alpha$ [75] or $-1/\alpha$ when $\alpha \neq 0$, e.g. in [67] where Lamperti calls this the *order* of the process rather than the index. Here we use the convention found in the self-similar fragmentation literature, e.g. [6, 11, 59]. Let us summarize in a proposition some properties of V that can be deduced from the well-developed theory of self-similar Markov processes. First, if $X = (\Pi, \mathbf{V})$ is an ESSF process, for each $i \geq 1$ define $\zeta_i := \inf\{t \geq 0,\, V_i(t) = 0\}$, and for $t \in [0, \zeta_i]$,

$$\varphi_i(t) := \int_0^t V_i(t)^\alpha \, \mathrm{d}s.$$

Note that φ_i is continuous and increasing. We define its right-continuous inverse $\tau_i(t)$, for $t \in [0, \infty)$, by

$$\tau_i(t) := \begin{cases} \varphi_i^{-1}(t) & \text{if } t < \varphi_i(\zeta_i), \\ \infty & \text{if } t \geq \varphi_i(\zeta_i). \end{cases}$$

We need a convention for infinite times, so we let $V_i(\infty) \equiv 0$, so that $V_i(\tau_i(t))$ is always defined. Also, note that the definition of the Frag operator implies that a.s., Π has nonincreasing sample paths for the *finer-than* partial order – π is finer than π' if the blocks of π' can be written as unions of blocks of π. Since a.s. for all $n \geq 1$, $\Pi(t)_{|[n]}$ is nonincreasing in a finite set, it is eventually constant. This implies that $\Pi(t)$ converges a.s. when $t \to \infty$, and we may denote its limit by $\Pi(\infty)$. Let us now state the proposition.

Proposition 5.7. *Let $X = (\Pi, \mathbf{V})$ be an α-ESSF process and $i \geq 1$, and define*

$$\begin{aligned} \zeta_i &:= \inf\{t \geq 0,\, V_i(t) = 0\}, \\ \varphi_i(t) &:= \int_0^t V_i(t)^\alpha \, \mathrm{d}s, \qquad t \in [0, \zeta_i] \\ \tau_i(t) &:= \varphi_i^{-1}(t), \qquad t \geq 0 \end{aligned}$$

Then the following properties hold.

- *Either $\zeta_i < \infty$ $\mathbb{P}_{(\mathbf{1},1)}$-a.s., or $\zeta_i = \infty$ $\mathbb{P}_{(\mathbf{1},1)}$-a.s.*
- *Either $\varphi_i(\zeta_i) < \infty$ $\mathbb{P}_{(\mathbf{1},1)}$-a.s., or $\varphi_i(\zeta_i) = \infty$ $\mathbb{P}_{(\mathbf{1},1)}$-a.s.*
- *In the case $\zeta_i < \infty$, either V_i reaches 0 continuously $\mathbb{P}_{(\mathbf{1},1)}$-a.s., or V_i eventually jumps to 0 $\mathbb{P}_{(\mathbf{1},1)}$-a.s.*
- *$\varphi_i(\zeta_i) = \infty$ iff V_i reaches 0 continuously.*
- *The process $\xi_i := \log(V_i \circ \tau_i)$, – i.e. defined by*

$$\xi_i(t) = \log V_i(\tau_i(t)), \qquad 0 \leq t < \varphi_i(\zeta_i),$$

is a (killed in the case $\varphi_i(\zeta_i) < \infty$) Lévy process called the inverse Lamperti transform *of V_i.*

Proof. These are classical results on pssMp, we refer to [67] for a proof. □

This proposition tells us that it is natural to consider the time-changed processes $V_i \circ \tau_i$ for $n \geq 1$, which behave as exponentials of Lévy processes. However, there is no unique

time-change that could make the whole process X behave nicely. Instead, we have to rely on stopping lines, which are tools generalizing stopping times in the context of branching Markov processes (see e.g. [25] for their use in branching Brownian motion, or [10, 11] in the context of fragmentations).

5.2.2 Stopping lines, changing the index of self-similarity

First let us define some filtrations associated with an ESSF process $X = (\Pi, \mathbf{V})$. To this aim, let us endow the power set $2^{\mathbb{N}} := \{A \subset \mathbb{N}\}$ with the topology generated by the metric $d(A, B) := \big(\sup\{n \in \mathbb{N}, A \cap [n] = B \cap [n]\}\big)^{-1}$, which makes $2^{\mathbb{N}}$ a compact space. Now for $i \in \mathbb{N}$, let us define the block process $(B_i(t), t \geq 0)$ as the $2^{\mathbb{N}}$-valued càdlàg process such that for all $t \geq 0$, $B_i(t)$ is the block of $\Pi(t)$ containing i, that is:

$$B_i(t) = \{j \in \mathbb{N},\ i \sim^{\Pi(t)} j\}.$$

Now we can define a sequence of natural filtrations associated to X by

$$\mathcal{G}_i = (\mathcal{G}_i(t),\ t \geq 0) \quad \text{with } \mathcal{G}_i(t) = \sigma\big(B_i(s), V_i(s),\ s \in [0, t]\big), \qquad i \geq 1, t \geq 0.$$

Definition 5.8. Let $X = (\Pi, \mathbf{V})$ be an ESSF process. A sequence $L = (L_i, i \geq 1)$ of random variables with values in $[0, \infty]$ is called a stopping line if

(i) for all $i \geq 1$, L_i is a $\mathcal{G}_i$-stopping time.

(ii) for $i, j \geq 1$, if $i \sim^{\Pi(L_i)} j$, then $L_i = L_j$.

Since (ii) entails that $i \sim^{\Pi(L_i)} j$ is an equivalence relation, its equivalence classes form a well-defined partition of $\mathbb{N}$ which we denote by $\Pi(L)$ with a slight abuse of notation. Also, denoting $\mathbf{V}(L)$ as the vector $(V_i(L_i), i \geq 1)$, it is clear that $X(L) := (\Pi(L), \mathbf{V}(L))$ is a well-defined (random) element of $\mathcal{M}_\infty$.

Remark 5.9. A fixed time $t \geq 0$ can be seen as a stopping line (an L for which $L_i \equiv t$ for all $i \geq 1$), and it is easily checked that for a stopping line L, one can define $L + t$ and $L \wedge t$ by

$$(L + t)_i = L_i + t \qquad \text{and } (L \wedge t)_i = L_i \wedge t,$$

which are again stopping lines. Thus for a stopping line L we will be able to consider the processes $X(L + \cdot) := (X(L + t),\ t \geq 0)$ and $X(L \wedge \cdot) := (X(L \wedge t),\ t \geq 0)$. Since it will be useful, we define the following σ-algebra:

$$\mathcal{G}_L := \sigma\big(X(L \wedge t),\ t \geq 0\big).$$

We can now state the Markov property for stopping lines, which is analogous to what can be found in [10, Lemma 3.14].

Proposition 5.10 (Stopping line Markov property)**.** *Let X be an α-ESSF, and L be a stopping line. Then conditional on $\mathcal{G}_L$, the following equality in distribution holds:*

$$X(L + \cdot) \overset{(d)}{=} \mathrm{ssFrag}_\alpha\big(X(L), X^{(\cdot)}\big), \tag{5.3}$$

where $X^{(\cdot)}$ is an independent, i.i.d. sequence of copies of the process started from $(\mathbf{1}, 1)$.

Proof. See Appendix 5.A.2. □

The next step in the analysis of ESSF processes is to bring the index of self-similarity to 0. This will be done via the random time changes $(\tau_i(t), i \geq 1, t \geq 0)$ defined above by

$$\tau_i(t) := \varphi_i^{-1}(t), \quad \text{where} \quad \varphi_i(u) := \int_0^u V_i(s)^\alpha \, \mathrm{d}s, \quad u \geq 0.$$

These time changes enable us to turn an α-ESSF into a homogeneous ESSF. The following proposition makes this claim more precise.

Proposition 5.11. *Let $X = (\Pi, \mathbf{V})$ be an α-ESSF, with $\alpha \in \mathbb{R}$. Let $\beta \in \mathbb{R}$ and define the random times*

$$\tau_i^\beta(t) = \left(\int_0^\cdot V_i(s)^\beta \, \mathrm{d}s \right)^{-1}(t), \qquad i \geq 1, \, t \geq 0,$$

Then for each $t \geq 0$, $\tau^\beta(t)$ is a stopping line, and the process $X \circ \tau^\beta := (X(\tau^\beta(t)), \, t \geq 0)$ is an $(\alpha - \beta)$-ESSF. Furthermore, if X is non-degenerate, then $X \circ \tau^\beta$ is also non-degenerate.

Proof. See Appendix 5.A.3. □

By bringing the index of self-similarity to 0 we can transform any ESSF into a homogeneous process. Let us now study further those 0-ESSF.

5.3 Main results

5.3.1 Decomposition of ESSF processes

Let us consider here a homogeneous 0-ESSF process $X = (\Pi, \mathbf{V})$, started from $(\mathbf{1}, 1)$. We know by Lemma 5.6 that it satisfies a projective Markov property, i.e. for all $n \in \mathbb{N}$, $X_{|[n]}$ defines a Markov process with values in $\mathcal{P}_n$. Let $n \in \mathbb{N}$ be fixed, and define the stopping time

$$T_n := \inf\{t \geq 0, \, \Pi(t)_{|[n]} \neq \mathbf{1}_n \text{ or } V_1(t) = 0\},$$

as well as the killed process

$$\tilde{\xi}_n := (\log V_1(t), \, 0 \leq t < T_n).$$

Note that homogeneity implies that the pair $(\tilde{\xi}_n - \log v, T_n)$ has the same distribution under every $\mathbb{P}_{(\mathbf{1},v)}$ for all $v \in S^1 \setminus \{0\}$. Therefore for $t \geq 0$, conditional on $\{T_n > t\}$, the Markov property applied at time t shows that $(\tilde{\xi}_n(t + \cdot) - \tilde{\xi}_n(t), T_n - t)$ has the same distribution as $(\tilde{\xi}_n, T_n)$ under $\mathbb{P}_{(\mathbf{1},1)}$. This shows that the killed process $\tilde{\xi}_n$ is distributed as

$$\tilde{\xi}_n \overset{(d)}{=} (\xi_n(t), \, 0 \leq t < T_n),$$

where ξ_n is a Lévy process and T_n is an independent exponential random variable. Note that this implies that if $T_n < \infty$, then $V_1(T_n-) = \exp(\xi_n(T_n)) > 0$. Now for $n \in \mathbb{N}$ such that $T_n < \infty$ almost surely, consider D_n, the dislocation (or freezing) at time T_n, defined by

$$D_n := (\Pi(T_n), \mathbf{V}(T_n)/V_1(T_n-))_{|[n]} \in \mathcal{M}_n,$$

where the division $\mathbf{V}(T_n)/V_1(T_n-)$ is to be understood coordinate-wise. Equivalently, D_n is the unique random marked partition such that

$$X(T_n)_{|[n]} = \mathrm{Frag}\big(X(T_n-)_{|[n]}, D_n\big),$$

with a slight abuse of notation in this case since $X(T_n-)_{|[n]}$ has only one block (D_n is not a sequence but additional terms are useless to define a fragmentation of a single block).

Note that this implies that D_n has the same distribution under every $\mathbb{P}_{(\mathbf{1},v)}$ for all $v \in S^1 \setminus \{0\}$. Thus for any bounded measurable maps $g : \mathbb{R} \to \mathbb{R}$, $h : \mathcal{M}_n \to \mathbb{R}$ and $t \geq 0$, applying the Markov property at time $t \geq 0$, one gets

$$\mathbb{E}_{(\mathbf{1},1)}[g(\xi_n(t))\mathbb{1}_{T_n>t}h(D_n)] = \mathbb{E}_{(\mathbf{1},1)}[g(\xi_n(t))\mathbb{1}_{T_n>t}]\,\mathbb{E}_{(\mathbf{1},1)}h(D_n),$$

which shows that the killed Lévy process (ξ_n, T_n) and the marked partition D_n are independent. Let us define $\mathcal{D}_n$ as the law of D_n, and notice also that exchangeability of $X_{|[n]}$ implies that $\mathcal{D}_n$ is an exchangeable probability measure on $\mathcal{M}_n$.

Since (ξ_n, T_n) is a killed Lévy process, one can define uniquely $d_n \in \mathbb{R}$, $\beta_n \geq 0$, $J_n \geq 0$ and λ_n a measure on $\mathbb{R} \setminus \{0\}$ satisfying $\int 1 \wedge y^2\, \lambda_n(\mathrm{d}y) < \infty$, such that

- the process ξ_n is a Lévy process with characteristic exponent

$$\psi_n(\theta) := \log \mathbb{E}[\mathrm{e}^{i\theta\xi_n(1)}] = id_n\theta - \frac{\beta_n}{2}\theta^2 + \int_{\mathbb{R}} \left(\mathrm{e}^{i\theta y} - 1 - i\theta y \mathbb{1}_{|y|\leq 1}\right) \lambda_n(\mathrm{d}y),$$

- ξ_n is killed at rate is $J_n = 1/\mathbb{E}T_n$, which may be 0 if $T_n = \infty$ almost surely.

Remark 5.12. Note that knowing $(\psi_n, J_n, \mathcal{D}_n)$ for $n \in \mathbb{N}$ is enough to reconstruct the process X. Indeed, starting from $(\mathbf{1}, v)$, the process $X_{|[n]}$ up to time T_n has distribution equal to that of

$$Y_n := ((\mathbf{1}_n, v\mathrm{e}^{\xi_n(t)}),\, 0 \leq t < T_n),$$

and at time T_n jumps to $(\Pi, v\mathrm{e}^{\xi_n(T_n-)}\mathbf{V})$, where $(\Pi, \mathbf{V})$ is independently drawn according to $\mathcal{D}_n$.

By the branching property, one only needs to iterate this construction at each jump time, independently for each marked block, to get the whole process $X_{|[n]}$. By Kolmogorov's extension theorem – since $(X_{|[m]})_{|[n]} = X_{|[n]}$ for each $n \leq m$ – these distributions characterize the distribution of X.

Let us now state our main result which identifies the form that those characteristics can take.

Theorem 5.13. *Let X be a non-degenerate 0-ESSF and for each n, write $(\psi_n, J_n, \mathcal{D}_n)$ for the characteristics describing the law of $X_{|[n]}$. Then there is a unique quadruple (c, d, β, Λ), where $c, \beta \geq 0$, $d \in \mathbb{R}$, and Λ is a measure on $\mathscr{Z}^{\downarrow} \setminus \{(\mathbf{1}, 1)\}$, which satisfies necessarily*

$$\int_{\mathscr{Z}^{\downarrow}} (1 - s_1\mathbb{1}_{v_1>0} + (\log v_1)^2 \wedge 1)\, \Lambda(\mathrm{d}\mathbf{z}) < \infty, \tag{5.4}$$

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

(i) $\psi_n(\theta) = id\theta - \dfrac{\beta}{2}\theta^2 + \displaystyle\int_{\mathscr{Z}^\downarrow} \sum_{\substack{j\geq 1\\ v_j>0}} s_j^n(\mathrm{e}^{i\theta\log v_j} - 1) - i\theta\log v_1 \mathbb{1}_{|\log v_1|\leq 1}\,\Lambda(\mathrm{d}\mathbf{z})$.

(ii) $J_n = nc + \displaystyle\int_{\mathscr{Z}^\downarrow}\Big(1 - \sum_{\substack{i\geq 1\\ v_i>0}} s_i^n\Big)\,\Lambda(\mathrm{d}\mathbf{z})$.

(iii) *if* $J_n > 0$, $\mathcal{D}_n = \dfrac{1}{J_n}\Big(\displaystyle\sum_{i=1}^n c\delta_{\mathfrak{e}_i^n} + \int_{\mathscr{Z}^\downarrow} \varrho_{\mathbf{z}}^n(\,\cdot\cap\{\pi \neq \mathbf{1}_n \text{ or } (\pi,\mathbf{v}) = (\mathbf{1}_n, 0)\})\,\Lambda(\mathrm{d}\mathbf{z})\Big)$, *where* $\varrho_{\mathbf{z}}^n$ *is the paintbox process defined in Section 5.1.2, and* $\delta_{\mathfrak{e}_i^n}$ *denotes the Dirac point measure on* $\mathfrak{e}_i^n$, *the marked partition defined as*

$$\mathfrak{e}_i^n := \Big(\{[n]\setminus\{i\},\{i\}\},\,(1,\dots,1,\underbrace{0}_{i-th\ index},1,\dots,1)\Big).$$

Conversely if $c,\beta \geq 0$, $d\in\mathbb{R}$ *and* Λ *is a measure on* $\mathscr{Z}^\downarrow\setminus\{(\mathbf{1},1)\}$, *satisfying (5.4), then there exists a* 0*-ESSF with characteristics as above.*

Proof. See Appendix 5.A.4. □

Remark 5.14. It is an immediate consequence of the theorem that the process describing the block of X containing 1 can be constructed in the following Poissonian way. Consider $\mathcal{N}$ a Poisson point process on $[0,\infty)\times\mathcal{M}_\infty^\star$ with intensity

$$\mathrm{d}t\otimes\Big(\sum_{i=1}^\infty c\delta_{\mathfrak{e}_i^n} + \int_{\mathscr{Z}^\downarrow}\varrho_{\mathbf{z}}^n(\,\cdot\,)\,\Lambda(\mathrm{d}\mathbf{z})\Big),$$

and define

$$\mathcal{N}' := \{(t,\log v_1),\ (t,x)\in\mathcal{N} \text{ with } x=(\pi,\mathbf{v}) \text{ and } v_1\notin\{0,1\}\},$$

which has intensity $\mathrm{d}t\otimes\lambda_1$, where λ_1 is defined by

$$\int_{\mathbb{R}} f\,\mathrm{d}\lambda_1 = \int_{\mathscr{Z}^\downarrow}\sum_{\substack{j\geq 1\\ v_j\notin\{0,1\}}} s_j\,f(\log v_j)\,\Lambda(\mathrm{d}\mathbf{z}),$$

and is the Lévy measure of the process ξ_1. It is clear that one can build a Lévy process $(\xi_1(t), t\geq 0)$ having characteristic exponent ψ_1 given by *(i)* in the theorem and whose point process of jumps is exactly $\mathcal{N}'$.

Define $(B(t), t\geq 0)$ as the $2^{\mathbb{N}}$-valued process given by

$$B(t) = \bigcap_{\substack{0\leq s<t\\ (s,x)\in\mathcal{N}}} A(x),$$

where $A(x)\subset\mathbb{N}$ denotes the block of x containing 1. Also, for any $n\in\mathbb{N}$, define

$$\widetilde{T}_n := \inf\{t\geq 0,\ (t,x)\in\mathcal{N} \text{ with } x=(\pi,\mathbf{v}) \text{ such that } \pi_{|[n]}\neq\mathbf{1} \text{ or } v_1 = 0\} \sim \mathrm{Exp}(J_n).$$

Now $(B(t), \mathrm{e}^{\xi_1(t)}, 0 \le t < \widetilde{T}_1)$ is distributed as the marked block containing 1 in X and by construction, we also get the following equality in distribution

$$\Big(X(T_n)_{|[n]},\, n \in \mathbb{N}\Big) \overset{(d)}{=} \Big((\pi_n,\, \mathrm{e}^{\xi_1(\widetilde{T}_n-)}\mathbf{v}_n)_{|[n]},\, n \in \mathbb{N}\Big),$$

where $x_n = (\pi_n, \mathbf{v}_n)$ is the element of $\mathcal{M}_\infty^\star$ such that $(\widetilde{T}_n, x_n) \in \mathcal{N}$.

Combining Theorem 5.13 with Proposition 5.11, we get the following characterization of all ESSF processes.

Corollary 5.15. *Let X be a non-degenerate α-ESSF. Then there exists a unique quadruple (c, d, β, Λ) as in Theorem 5.13 such that if $(\tau^\alpha(t), t \ge 0)$ are the stopping lines as defined in Proposition 5.11, then $X \circ \tau^\alpha$ is a homogeneous ESSF with characteristics (c, d, β, Λ).*

Let us point out that condition (5.4) is surprisingly nonrestrictive. There are no integrability assumptions concerning the marks of the smallest blocks (with labels greater than 1). Consequently, the point measure $\sum_k \delta_{\tilde{V}_k(t)}$, where $\tilde{V}_k(t)$ denotes the mark of the k-th block in $X(t)$, might assign infinite mass to any interval $(a, b) \subset [0, \infty)$ for any $t > 0$. Indeed it suffices for instance that $\Lambda(\mathrm{d}\mathbf{z})$ be of the form

$$\int_{\mathscr{Z}^\downarrow} \prod_{i\ge1} F_i(s_i, v_i)\, \Lambda(\mathrm{d}\mathbf{z}) = \int_{(0,1)\times S^1} F_1(s_1, v_1) \mathbb{E} \prod_{i\ge2} F_i\left(\frac{1-s_1}{2^{i-1}}, Z_i\right) \nu(\mathrm{d}z_1),$$

where ν is a measure on $(0,1) \times S^1$ with infinite mass and satisfying

$$\int_{(0,1)\times S^1} (1 - s_1 \mathbb{1}_{v_1>0} + (\log v_1)^2 \wedge 1)\, \nu(\mathrm{d}z_1) < \infty,$$

and $Z_2, Z_3, \ldots$ are i.i.d. Exp(1) random variables. On the other hand, if one assumes an integrability condition such as

$$\int_{\mathscr{Z}^\downarrow} (\sum_{i\ge1} v_i^\theta) - 1 - \theta \log v_1 \mathbb{1}_{|\log v_1| \le 1}\, \Lambda(\mathrm{d}\mathbf{z}) < \infty$$

for some $\theta \in \mathbb{R}$, then one observes a process of point measures $(\sum_k \delta_{\tilde{V}_k(t)}, t \ge 0)$ that is nice in the sense that for all $t \ge 0$, $\mathbb{E} \sum_k \tilde{V}_k(t)^\theta < \infty$. This is the object of the next section.

5.3.2 Absorption in finite time

Consider here a non-degenerate α-ESSF with characteristics (c, d, β, Λ), started from $(\mathbf{1}, 1)$. We are interested in the case where the pssMp V_1 reaches 0 in finite time and, if T_i denotes the hitting time of 0 by the process V_i, we aim at giving a sufficient condition for which

$$\sup_{i\in\mathbb{N}} T_i < \infty \qquad \text{a.s.}$$

If this holds, then at this random time the process is frozen in a state $X(\infty) = (\Pi(\infty), \mathbf{V}(\infty))$ where $\mathbf{V}(\infty) = \mathbf{0} = (0, 0, \ldots)$, and we say the process X is *absorbed in finite time*. We first put aside a trivial case and assume that

$$c > 0 \quad \text{or} \quad \int_{\mathscr{Z}^\downarrow} (1 - s_1)\, \Lambda(\mathrm{d}\mathbf{z}) > 0,$$

since otherwise Π would be almost surely constant equal to the coarsest partition $\{\mathbb{N}\}$. Recall that a classical self-similar fragmentation which is absorbed in finite time (this is always true when $\alpha < 0$ [13, Proposition 2]) is always totally fragmented in the limit, in the sense that $\Pi(\infty)$ is the partition of $\mathbb{N}$ into singletons. Clearly in our case, because of the possible freezing of blocks at dislocation events, $\Pi(\infty)$ is almost surely totally fragmented if and only if

$$\forall i \geq 1,\ s_i > 0 \implies v_i > 0 \qquad \Lambda\text{-a.e. on } \mathscr{Z}^{\downarrow}.$$

A stronger property than absorption in finite time is the following: we say X has *finite total length* if

$$\int_0^\infty \#X(t)\,\mathrm{d}t < \infty \quad \text{a.s.},$$

where $\#x$ denotes the number of blocks with positive mark in the marked partition x. One can interpret this quantity as the total length of the tree describing the genealogy of blocks in the fragmentation, hence the name. Note that this implies that for a fixed time $t \geq 0$, $\#X(t)$ is almost surely finite, which is well-known [13, Proposition 2] in the classical self-similar fragmentation case for $\alpha < -1$.

In this section our aim is to provide sufficient conditions for ESSF processes to be absorbed in finite time and to have finite total length. The following result extends the classical setting, and makes use of natural martingales appearing in the homogeneous case. In order to be able to state it, we need a couple of additional definitions. For a marked partition $x = (\pi, \mathbf{v}) \in \mathcal{M}_n$ with $n \in \mathbb{N} \cup \{\infty\}$, and $\theta \in \mathbb{R}$, let us write

$$S_\theta(x) := \sum_{k \in \mathbb{N}} \tilde{v}_k^\theta, \tag{5.5}$$

where $\tilde{v}_k$ denotes the mark associated with the k-th block of x. Let us also introduce $\kappa : \mathbb{R} \to (-\infty, \infty]$ defined by

$$\kappa(\theta) := d\theta + \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^{\downarrow}} \Big(\sum_{i \geq 1} v_i^\theta\Big) - 1 - \theta \log v_1 \mathbb{1}_{|\log v_1| \leq 1}\, \Lambda(\mathrm{d}\mathbf{z}). \tag{5.6}$$

Note that the integral in the last display is well-defined with values in $(-\infty, \infty]$, since

$$\begin{aligned} 1 + \theta \log v_1 \mathbb{1}_{|\log v_1| \leq 1} - \sum_{i \geq 1} v_i^\theta &\leq \big(1 + \theta \log v_1 \mathbb{1}_{|\log v_1| \leq 1} - v_1^\theta\big)_+ \\ &\leq C\big((\log v_1^2) \wedge 1\big) \end{aligned}$$

where C is a positive constant which depends on θ, so the negative part of the integrand in the definition of κ is Λ-integrable.

Proposition 5.16. *Let X be a non-degenerate 0-ESSF with characteristics (c, d, β, Λ) started from $(\mathbf{1}, 1)$. For all $\theta \in \mathbb{R}$ and $t \geq 0$,*

$$\mathbb{E}S_\theta(X(t)) = \mathrm{e}^{t\kappa(\theta)},$$

with S_θ and $\kappa(\theta)$ respectively defined as in (5.5) and (5.6), and where these quantities may be infinite. If there is $\theta \in \mathbb{R}$ such that $\kappa(\theta) < \infty$, then the process

$$(\mathrm{e}^{-t\kappa(\theta)} S_\theta(X(t)),\ t \geq 0)$$

is a martingale. If there is $\theta \neq 0$ such that $\kappa(\theta) < 0$, then for any $\alpha \in \mathbb{R}$:

- *if $-\alpha/\theta > 0$, the α-ESSF with characteristics (c, d, β, Λ) is absorbed in finite time.*
- *if $-\alpha/\theta \geq 1$, the α-ESSF with characteristics (c, d, β, Λ) has finite total length.*

Proof. See Appendix 5.A.5. □

Remark 5.17. For a classical self-similar fragmentation with erosion coefficient $c \geq 0$ and dislocation measure ν, we have

$$\kappa(\theta) = -c\theta + \int_{\mathscr{S}^\downarrow} \Big(\sum_{i \geq 1} s_i^\theta - 1 \Big) \nu(\mathrm{d}\mathbf{s}).$$

Since $\sum_i s_i \leq 1$ ν-a.e., for all $\theta > 1$ we have $\kappa(\theta) < 0$, so we recover absorption in finite time for any $\alpha < 0$ and finite total length for any $\alpha < -1$.

Remark 5.18. Let us also mention that one can model branching Brownian motion in our setting. Indeed, consider a homogeneous ESSF where the logarithm of marks follow drifted Brownian motion and blocks dislocate at rate one into two blocks (say both with asymptotic frequency equal to half of the mother block) carrying the same mark. More precisely, take a 0-ESSF with characteristics $c = 0$, $d \in \mathbb{R}$, $\beta = 1$ and with $\Lambda(\mathrm{d}\mathbf{z})$ a Dirac measure on $((\frac{1}{2}, 1), (\frac{1}{2}, 1), 0, \ldots)$.

Then the point process recording the positions of the logarithm of marks

$$\sum_{k \in \mathbb{N}} \delta_{\log \tilde{V}_k(t)}, \qquad t \geq 0$$

is a classical binary branching Brownian motion with drift d. One gets a cumulant function

$$\kappa(\theta) = d\theta + \frac{\theta^2}{2} + 1.$$

This polynomial in θ takes negative values if and only if $d^2 - 2 > 0$, and we essentially recover the well-known fact that if $d > \sqrt{2}$, the lowest particle of a branching Brownian with drift d goes to $+\infty$.

5.A Proofs

5.A.1 Proof of Proposition 5.2

Let us write as usual $X = (\Pi, \mathbf{V})$. First, note that Π is an exchangeable partition with values in $\mathcal{M}_\infty^\star$, therefore it has asymptotic frequencies – so $|X|^\downarrow$ exists almost surely – and the finite blocks of Π (if any) are necessarily singletons. For the uniqueness part of the proposition, notice that any ν satisfying (5.2) must be equal to $\mathbb{P}(|X|^\downarrow \in \cdot)$.

For the existence, let $(U_k, k \geq 1)$ be an i.i.d. sequence of uniform random variables on $[0, 1]$, independent of X. For every $i \in \mathbb{N}$, let $Z_i = (U_k, V_i) \in [0, 1] \times S^1$, where k is the label of the block containing i. Then the sequence $(Z_i, i \geq 1)$ is exchangeable, with values in a Polish space. Therefore by de Finetti's theorem, there is a random probability

measure $\theta \in \mathcal{M}_1([0,1] \times S^1)$ such that conditional on θ, the sequence $(Z_i, i \geq 1)$ is i.i.d. with distribution θ. Let

$$(a_k, k \geq 1) = (u_k, v_k,\ k \geq 1)$$

denote the collection of atoms of θ. Note that $i \sim^X j$ iff $Z_i = Z_j$, and therefore the law of large numbers ensures us that the blocks of X correspond to those atoms, i.e. for each $k \geq 1$, there is a block B of X with an asymptotic frequency $|B| = \theta(a_k)$ and a mark equal to v_k. Conversely any block which is not reduced to a singleton must be formed in this way. Furthermore, note that singleton blocks have mark 0 because of the assumption that $X \in \mathcal{M}_\infty^\star$, so the knowledge of the atoms $(a_k, k \geq 1)$ and their mass is sufficient to reconstruct the sequence $(Z_i, i \geq 1)$, and therefore the marked partition X.

Now define for all $k \in \mathbb{N}$, $z_k := (\theta(a_k), v_k)$. Up to a reordering, we can assume that $\mathbf{z} = (z_k, k \geq 1)$ is in $\mathscr{Z}^\downarrow$ (if the sequence of atoms is finite, we concatenate to $\mathbf{z}$ infinitely many $(0,0)$ terms). The previous discussion means that conditional on θ, the asymptotic frequencies of X are exactly

$$|X|^\downarrow = \mathbf{z} \in \mathscr{Z}^\downarrow,$$

and conditional on $\mathbf{z}$, the marked partition X is drawn according to $\varrho_{\mathbf{z}}$. Note that the map $\theta \mapsto \mathbf{z}$ is measurable. Indeed, by standard point processes arguments [see e.g. 30, Lemma 9.1.XIII], there exists a measurable enumeration $(a_k, k \geq 1)$ of the atoms of θ, and it is elementary that the nonincreasing reordering of this sequence is measurable. Therefore, defining $\nu = \mathbb{P}(|X|^\downarrow \in \cdot)$, which is the push-forward of the distribution of θ by the map $\theta \mapsto \mathbf{z}$, we see that it satisfies (5.2).

5.A.2 Proof of Proposition 5.10

In this section, it will be helpful to consider the restriction of an ESSF process X to a more general (and possibly random) subset $A \subset \mathbb{N}$, considered as a random variable living on the compact space $2^{\mathbb{N}}$. We first consider a fixed – non random – $A \subset \mathbb{N}$ with cardinality $\#A \in \mathbb{N} \cup \{\infty\}$, and define a canonical enumeration of A by

$$\sigma_A : \begin{cases} [\#A] & \to \mathbb{N} \\ i & \mapsto \min\{n \in \mathbb{N},\ \#(A \cap [n]) = i\}, \end{cases}$$

such that $A = \{\sigma_A(1), \sigma_A(2), \ldots\}$, with $\sigma_A(1) < \sigma_A(2) < \ldots$. Now recall the definition of the action of injections on $\mathcal{M}_\infty$. For any $x \in \mathcal{M}_\infty$, one can see $x^{\sigma_A} \in \mathcal{M}_{\#A}$ as the restriction of x to the set A.

As an inverse operation, for any $x' \in \mathcal{M}_{\#A}, x'' \in \mathcal{M}_{\#A^c}$, where $A^c := \mathbb{N} \setminus A$, we can define

$$x' \overset{A}{\oplus} x'' \in \mathcal{M}_\infty$$

as the pair $(\pi, \mathbf{v})$ such that

$$\forall i, j \in \mathbb{N}, \qquad i \sim^\pi j \iff \begin{cases} & i, j \in A \text{ and } \sigma_A^{-1}(i) \sim^{\pi'} \sigma_A^{-1}(j) \\ \text{or} & i, j \in A^c \text{ and } \sigma_{A^c}^{-1}(i) \sim^{\pi''} \sigma_{A^c}^{-1}(j) \end{cases}$$

$$\text{and } v_i = \begin{cases} v'_{\sigma_A^{-1}(i)} & \text{if } i \in A \\ v''_{\sigma_{A^c}^{-1}(i)} & \text{if } i \in A^c. \end{cases}$$

Similarly, for processes $X' = (X'(t),\, t \geq 0)$ and X'', we write for conciseness

$$X' \overset{A}{\oplus} X'' := (X'(t) \overset{A}{\oplus} X''(t),\, t \geq 0)$$

For $x = (\pi, \mathbf{v}) \in \mathcal{M}_\infty$ and $A \subset \mathbb{N}$, we will say that A is x-compatible if it is a union of a family of blocks of π – i.e. if A is such that $i \in A, j \notin A \implies i \nsim^\pi j$. These definitions enable us to reformulate the branching property as follows.

Lemma 5.19. *Let X be an ESSF process, $x = (\pi, \mathbf{v}) \in \mathcal{M}_\infty$, and $A \subset \mathbb{N}$ an x-compatible set. Defining $X' := X^{\sigma_A}$ and $X'' := X^{\sigma_{A^c}}$, then under $\mathbb{P}_x$, X' and X'' are two independent copies of the process X, respectively started at x^{σ_A} and $x^{\sigma_{A^c}}$, and*

$$X = X' \overset{A}{\oplus} X''.$$

Proof. This is an immediate consequence of the branching property *(ii)* of Definition 5.4 and of the definition of the ssFrag operator. □

Let us now tackle the proof of the Markov property for stopping lines (5.3). We write as usual $X = (\Pi, \mathbf{V})$. We first assume that there exist $0 \leq t_1 < t_2 < \ldots < t_k \leq \infty$ such that for all $i \in \mathbb{N}$, L_i takes values in the finite set $\{t_1, \ldots t_k\}$. We prove the Markov property for such stopping lines by induction on k. For $k = 1$ and $t_1 < \infty$, this amounts to the simple Markov property, so (5.3) holds by definition. If $t_1 = \infty$, then for all $t \geq 0$, for all $i \geq 1$, $L_i + t \equiv \infty$. By convention $\mathbf{V}(\infty)$ is the null vector, and by definition $\mathrm{ssFrag}(X(L), X^{(\cdot)})$ is the process which is a.s. constant equal to $X(\infty) = (\Pi(\infty), \mathbf{V}(\infty))$, so (5.3) holds again.

Now assume that $k > 1$, and that the stopping line Markov property has been proven for all stopping lines taking at most $k-1$ distinct values. By Remark 5.9, $L \wedge t_{k-1}$ is a stopping line taking at most $k-1$ distinct values. Therefore, one can apply the induction hypothesis, which says that conditional on $\mathcal{G}_{L \wedge t_{k-1}}$, the process $X(L \wedge t_{k-1} + \cdot)$ has the distribution of a copy of X started from $X(L \wedge t_{k-1})$. Now we define the random set

$$A := \{i \in \mathbb{N},\, L_i = t_k\},$$

which is $\mathcal{G}_{L \wedge t_{k-1}}$-measurable. Indeed let us show that $\{i \in A\} \in \mathcal{G}_{L \wedge t_{k-1}}$. Since $\{L_i = t_k\} = \{L_i > t_{k-1}\} \in \mathcal{G}_i(t_{k-1})$, one can write the indicator of this event as $\mathbb{1}_{\{L_i > t_{k-1}\}} = F(B_i(s), V_i(s),\, s \in [0, t_{k-1}])$ for a measurable functional F – recall that $\mathcal{G}_i(t_{k-1}) = \sigma(B_i(s), V_i(s),\, s \in [0, t_{k-1}])$, where $B_i(s)$ is the block of $\Pi(s)$ containing i. Now on the event $\{L_i \geq t_{k-1}\}$, we have

$$F(B_i(s), V_i(s),\, s \in [0, t_{k-1}]) = F\big(B_i(s \wedge L_i \wedge t_{k-1}), V_i(s \wedge L_i \wedge t_{k-1}),\, s \in [0, t_{k-1}]\big),$$

therefore we can write

$$\{i \in A\} = \{L_i \geq t_{k-1}\} \cap \big\{F\big(B_i(s \wedge L_i \wedge t_{k-1}), V_i(s \wedge L_i \wedge t_{k-1}),\, s \in [0, t_{k-1}]\big) = 1\big\} \in \mathcal{G}_{L \wedge t_{k-1}},$$

so finally A is $\mathcal{G}_{L \wedge t_{k-1}}$-measurable. Now notice that because L is a stopping line, A is compatible with $\Pi(L \wedge t_{k-1})$ in the sense that A is necessarily a union of blocks of $\Pi(L \wedge t_{k-1})$. Therefore, it is immediate by definition that

$$X(L \wedge t_{k-1} + \cdot) = X' \overset{A}{\oplus} X'',$$

with

$$X' = X(L \wedge t_{k-1} + \cdot)^{\sigma_A}$$
$$\text{and } X'' = X(L \wedge t_{k-1} + \cdot)^{\sigma_{A^c}}.$$

Now by Lemma 5.19, conditional on $\mathcal{G}_{L \wedge t_{k-1}}$, X' and X'' are two independent copies of X started respectively from $X(L \wedge t_{k-1})^{\sigma_A}$ and $X(L \wedge t_{k-1})^{\sigma_{A^c}}$.

Also, notice that by definition of the random set A, we have the equality

$$X(L + \cdot) = X'(t_k - t_{k-1} + \cdot) \overset{A}{\oplus} X'', \tag{5.7}$$

and for the same reason, the following equality between σ-algebras holds:

$$\mathcal{G}_L = \mathcal{G}_{L \wedge t_{k-1}} \vee \sigma(X'(s),\, s \in [0, t_k - t_{k-1}]).$$

Clearly X' and X'' are still independent conditional on $\mathcal{G}_L$, and the distribution of $X'(t_k - t_{k-1} + \cdot)$ conditional on $\mathcal{G}_L$ is by the Markov property at time $t_k - t_{k-1}$ the law of X started from $X'(t_k - t_{k-1}) = X(L)^{\sigma_A}$. Finally, using again Lemma 5.19, conditional on $\mathcal{G}_L$, the process

$$X'(t_k - t_{k-1} + \cdot) \overset{A}{\oplus} X''$$

has simply the distribution of a copy of X started at $X(L)$. So by (5.7) the Markov property for stopping lines holds for L, and so by induction it holds for all stopping lines taking at most a finite number of values.

Now fix a general stopping line L, a time $t \geq 0$, and let us assume that our probability space contains an independent sequence $X^{(\cdot)}$ of i.i.d. copies of the process started from $(\mathbf{1}, 1)$. To conclude, it is enough to prove that conditional on $\mathcal{G}_L$,

$$X(L + t) \overset{(d)}{=} \mathrm{ssFrag}_\alpha(X(L), X^{(\cdot)})(t), \tag{5.8}$$

because then for any $0 \leq t_1 < t_2 < \dots t_k$ one can apply successively (5.8) to the stopping lines $L + t_i, 1 \leq i \leq k$, which implies that (5.3) holds for any finite dimensional distributions. Therefore it remains only to prove

$$\mathbb{E}\left[F(X(L + t))Z\right] = \mathbb{E}\left[F\left(\mathrm{ssFrag}_\alpha(X(L), X^{(\cdot)})(t)\right) Z\right]$$

for any fixed continuous bounded map $F : \mathcal{M}_\infty^\star \to \mathbb{R}$ and Z a $\mathcal{G}_L$-measurable bounded random variable. To show this, consider the sequence of stopping lines $(L^n, n \geq 1)$ defined by

$$L_i^n = 2^{-n}\lceil 2^n L_i \rceil \mathbb{1}_{L_i \leq n} + \infty \mathbb{1}_{L_i > n}, \quad i \geq 1,$$

where $\lceil \cdot \rceil$ denotes the usual ceiling function. This is a classical transformation for stopping times, and it is easily checked that L^n is a stopping line for all $n \geq 1$. Furthermore, right-continuity of the process implies that $X(L^n + t)$ converges a.s. to $X(L + t)$ as n tends to ∞. Therefore

$$\mathbb{E}\left[F(X(L^n + t))Z\right] \underset{n \to \infty}{\longrightarrow} \mathbb{E}\left[F(X(L + t))Z\right],$$

Now because L^n only takes values in a finite set for all n, we can apply (5.3), so

$$\mathbb{E}\left[F(X(L^n + t))Z\right] = \mathbb{E}\left[F\left(\mathrm{ssFrag}_\alpha(X(L^n), X^{(\cdot)})(t)\right) Z\right]. \tag{5.9}$$

This holds because Z is $\mathcal{G}_{L^n}$-measurable since $\mathcal{G}_L \subset \mathcal{G}_{L^n}$ for all $n \geq 1$. For the convergence of the right-hand side, recall that $X(L^n + t) \to X(L + t)$ and note also that if $V_i(L_i) = 0$ for $i \geq 1$, then since L_i is a stopping time, by the strong Markov property $V_i(L_i + t)$ is also zero for all $t \geq 0$, and in particular $V_i(L_i^n) = 0$ for all $n \geq 1$. Now by definition an ESSF process is stochastically continuous, so in particular for any $i \geq 1$, on the event $\{V_i(L_i) > 0\}$, we have

$$X^{(j)}(V_i(L_i)^\alpha t) \underset{n\to\infty}{\longrightarrow} X^{(j)}(V_i(L_i)^\alpha t) \qquad \forall j \geq 1 \; a.s.$$

We can now invoke the continuity property of the operator ssFrag_α pointed out in Remark 5.3 and deduce

$$\mathrm{ssFrag}_\alpha(X(L^n), X^{(\cdot)})(t) \underset{n\to\infty}{\longrightarrow} \mathrm{ssFrag}_\alpha(X(L), X^{(\cdot)})(t) \qquad a.s.$$

Taking limits in (5.9) yields the equality needed to end the proof.

5.A.3 Proof of Proposition 5.11

Let $X = (\Pi, \mathbf{V})$ be an α-ESSF process, $t \geq 0$, and recall the definition of $\tau^\beta(t)$ as in the proposition, i.e.

$$\tau_i^\beta(t) = \left(\int_0^\cdot V_i(s)^\beta \, \mathrm{d}s \right)^{-1} (t), \qquad i \geq 1,$$

where the inverse is to be understood as the right-continuous inverse. For conciseness and because β is fixed, let us write simply τ instead of τ^β throughout the proof. First, let us see that $\tau(t)$ is a stopping line. Fix $i \in \mathbb{N}$, and note that for $T \geq 0$,

$$\{\tau_i(t) \leq T\} = \left\{ \int_0^T V_i(s)^\beta \, \mathrm{d}s \geq t \right\} \in \mathcal{G}_i(T),$$

so $\tau_i(t)$ is a $\mathcal{G}_i$-stopping time. Furthermore, conditional on $\tau_i(t) = T$, for any $i, j \in \mathbb{N}$, if $i \sim j$ in $\Pi(T)$, then almost surely for all $0 \leq s \leq T$, we have $i \sim j$ in $\Pi(s)$ so $V_i(s) = V_j(s)$, and necessarily $\tau_j(t) = \tau_i(t) = T$.

Therefore $\tau(t)$ is indeed a stopping line, so the process $X \circ \tau = (X(\tau(t), t \geq 0)$ is well defined. We claim that its sample paths are càdlàg in $\mathcal{M}_\infty$. Indeed, by definition, for each $i \in \mathbb{N}$, τ_i is a non-decreasing right-continuous map. Now almost surely the following holds: X has càdlàg sample paths, so for each $i \in \mathbb{N}$, and $t \in [0, \infty)$,

$$X(\tau_i(s)) \underset{s\downarrow t}{\longrightarrow} X(\tau_i(t)) \quad \text{and} \quad X(\tau_i(s)) \underset{s\uparrow t}{\longrightarrow} X(\tau_i(t)-) \text{ if } t > 0.$$

Now note that for each stopping line L and integer $n \in \mathbb{N}$, by definition $X(L)_{|[n]}$ is a (deterministic) continuous functional of the variables $(X(L_1), \ldots, X(L_n))$. Applying this to $L = \tau(s)$ and letting $s \to t$, it follows that almost surely

$$X(\tau(s))_{|[n]} \underset{s\downarrow t}{\longrightarrow} X(\tau(t))_{|[n]} \quad \text{and} \quad X(\tau(s))_{|[n]} \text{ converges in } \mathcal{M}_n \text{ when } s \uparrow t, \text{ in the case } t > 0.$$

The integer n being generic, this shows that $X \circ \tau$ is an almost surely càdlàg process.

Let $t \geq 0$ be fixed. Since $\tau(t)$ is a stopping line, we can apply Proposition 5.10, and assume that the process $X(\tau(t) + \cdot)$ is given by

$$\mathrm{ssFrag}_\alpha(X(\tau(t)), X^{(\cdot)}),$$

where $X^{(\cdot)}$ is an independent sequence of i.i.d. copies of the process started from $(\mathbf{1}, 1)$. For each $k \in \mathbb{N}$, let $(\tau^{(k)}(s), s \geq 0)$ denote the stopping lines corresponding to $X^{(k)}$, i.e.

$$\tau_i^{(k)}(s) = \left(\int_0^{\cdot} V_i^{(k)}(u)^\beta \, \mathrm{d}u \right)^{-1} (s), \qquad i \geq 1,\, s \geq 0.$$

Our aim is to show that

$$X(\tau(t + \cdot)) = \mathrm{ssFrag}_{\alpha-\beta}(X(\tau(t)), X^{(\cdot)} \circ \tau^{(\cdot)}). \tag{5.10}$$

Now let us fix $i \in \mathbb{N}$, and work conditional on $\mathcal{G}_{\tau(t)}$. On the event $\{V_i(\tau_i(t)) = 0\}$, then by definition of the operators ssFrag_α, the block containing i is constant in time and has mark 0 in both processes in (5.10), so there is equality for index i. Now we condition on $V_i(\tau_i(t)) = v$ with $v > 0$. Note that $\{V_i(\tau_i(t)) > 0\} \subset \{\tau_i(t) < \infty\}$, so in that case we have $\tau_i(t) < \infty$ almost surely, so there is the equality

$$\int_0^{\tau_i(t)} V_i(s)^\beta \, \mathrm{d}s = t.$$

Therefore we can write, for $s \geq 0$,

$$\begin{aligned}
\tau_i(t + s) &= \left(\int_0^{\cdot} V_i(u)^\beta \, \mathrm{d}u \right)^{-1} (t + s) \\
&= \tau_i(t) + \left(\int_0^{\cdot} V_i(\tau_i(t) + u)^\beta \, \mathrm{d}u \right)^{-1} (s).
\end{aligned}$$

Now for all $u \geq 0$, because $X(\tau(t) + \cdot) = \mathrm{ssFrag}_\alpha(X(\tau(t)), X^{(\cdot)})$, we have $V_i(\tau_i(t) + u) = vV_i^{(k)}(v^\alpha u)$, for k such that i is in the k-th block of $\Pi(\tau(t))$. This implies

$$\begin{aligned}
\tau_i(t + s) - \tau_i(t) &= \left(\int_0^{\cdot} V_i(\tau_i(t) + u)^\beta \, \mathrm{d}u \right)^{-1} (s) \\
&= \left(\int_0^{\cdot} v^\beta V_i^{(k)}(v^\alpha u)^\beta \, \mathrm{d}u \right)^{-1} (s) \\
&= \left(v^{\beta-\alpha} \int_0^{v^\alpha \cdot} V_i^{(k)}(u)^\beta \, \mathrm{d}u \right)^{-1} (s) \\
&= v^{-\alpha} \left(\int_0^{\cdot} V_i^{(k)}(u)^\beta \, \mathrm{d}u \right)^{-1} (v^{\alpha-\beta} s) \\
&= V_i(\tau_i(t))^{-\alpha} \tau_i^{(k)}(V_i(\tau_i(t))^{\alpha-\beta} s).
\end{aligned}$$

Now, defining $L_i(s)$ as the quantity given by the preceding display, the definition of the operators ssFrag_α yields for all $s \geq 0$,

$$\begin{aligned}
X(\tau(t + s)) &= \mathrm{ssFrag}_\alpha(X(\tau(t)), X^{(\cdot)})(L_i(s)) \\
&= \mathrm{ssFrag}_{\alpha-\beta}(X(\tau(t)), X^{(\cdot)} \circ \tau^{(\cdot)}))(s),
\end{aligned}$$

showing that $X \circ \tau$ is an $(\alpha - \beta)$-ESSF.

Finally, note that if X is non-degenerate, then for all $x \in \mathcal{M}_\infty^\star$, $\mathbb{P}_x$-almost surely for any time $t \geq 0$, any block of $X(t)$ is either infinite or has mark 0. So $\mathbb{P}_x$-almost surely, for all possible stopping lines L, the blocks of $X(L)$ satisfy the same condition, i.e. $X(L) \in \mathcal{M}_\infty^\star$. Therefore, $\mathbb{P}_x$-almost surely $X \circ \tau$ has sample paths with values in $\mathcal{M}_\infty^\star$.

5.A.4 Proof of Theorem 5.13

Note that the whole process $(X(t), t \geq 0)$ defines a coupling of all ξ_n for $n \geq 1$. By definition, one has

$$\xi_{n+1}(t) = \xi_n(t), \quad \forall t \leq T_{n+1},$$

and at time T_{n+1}, either ξ_n is killed on the event $\{T_n = T_{n+1}\}$, or, conditional on $\{T_n > T_{n+1}\}$, the process ξ_n jumps, independently of the past, according to the probability

$$\eta_{n+1}(\cdot) := \mathcal{D}_{n+1}\big(\log v_1 \in \cdot \mid \pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \neq 0\big),$$

and goes on independently of the past, its remaining part $(\xi_n(T_{n+1} + t) - \xi_n(T_{n+1}), 0 \leq t \leq T_n - T_{n+1})$ being independent from ξ_{n+1} and equal in distribution to ξ_n by the strong Markov property. Let us first compute the probability p_n that $T_{n+1} = T_n$, which is by construction

$$p_n = \mathcal{D}_{n+1}(\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0). \tag{5.11}$$

From the previous description, one can write

$$T_n = T_{n+1} + ZT'_n,$$

where $Z = \mathbb{1}_{\{T_n \neq T_{n+1}\}} \sim \mathrm{Be}(1-p_n)$ is a Bernoulli random variable with parameter $(1-p_n)$, and T'_n is a random variable equal in distribution to T_n, and independent from T_{n+1} and Z. Then, T_{n+1}, Z and T'_n are independent because Z is simply a function of the marked partition $D_{n+1} = (\Pi(T_{n+1}), \mathbf{V}(T_{n+1})/V_1(T_{n+1}-))_{|[n+1]}$ which is independent from ξ_{n+1}, and T'_n is independent of (ξ_{n+1}, D_{n+1}) because of the strong Markov property and the fact that $\alpha = 0$. Taking expectations yields

$$\frac{1}{J_n} = \frac{1}{J_{n+1}} + \frac{1-p_n}{J_n},$$

which implies that $p_n = J_n/J_{n+1}$.

Now let us rebuild the coupling between ξ_n and ξ_{n+1} to show that their respective Lévy measures λ_n and λ_{n+1} satisfy

$$\lambda_n = \lambda_{n+1} + (J_{n+1} - J_n)\widetilde{\eta}_{n+1}, \tag{5.12}$$

where

$$\widetilde{\eta}_{n+1} := \eta_{n+1}\big(\,\cdot \cap \mathbb{R}\backslash\{0\}\big) = \mathcal{D}_{n+1}\big(\{\log v_1 \in \cdot\} \cap \{v_1 \neq 1\} \mid \pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \neq 0\big). \tag{5.13}$$

Consider the process ξ_{n+1} a Lévy process with characteristic exponent ψ_{n+1}, and let $T_n \sim \mathrm{Exp}(J_n)$ independent. Now independently define $T' \sim \mathrm{Exp}(J_{n+1} - J_n)$, and let $T_{n+1} := T_n \wedge T'$. We see that $T_{n+1} \sim \mathrm{Exp}(J_{n+1})$, that it is independent of ξ_{n+1} and of the event $\{T_n = T_{n+1}\}$, which has probability $J_n/J_{n+1} = p_n$. Now conditional on (T_n, T_{n+1}), let us define

$$D_{n+1} = \begin{cases} D' & \text{if } T_n = T_{n+1}, \quad \text{where } D' \sim \mathcal{D}_{n+1}(\cdot \mid \pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0) \\ D'' & \text{if } T_n \neq T_{n+1}, \quad \text{where } D'' \sim \mathcal{D}_{n+1}(\cdot \mid \pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \neq 0), \end{cases}$$

where D' and D'' are mutually independent and independent of everything else. Note that D_{n+1} is independent of T_{n+1} and of ξ_{n+1}, and because of (5.11), D_{n+1} has indeed distribution $\mathcal{D}_{n+1}$. Let us define $J = \log v_1$, where v_1 is the mark associated with the integer 1 in the marked partition D_{n+1}, and ξ_n a Lévy process with characteristic exponent ψ_n. Now putting everything together, define $\tilde{\xi}_{n+1}$ as the killed Lévy process $(\xi_{n+1}(t), 0 \leq t < T_{n+1})$, and define $\tilde{\xi}_n$ as

$$\tilde{\xi}_n(t) = \begin{cases} \xi_{n+1}(t) & \text{if } t < T_{n+1} \\ \xi_{n+1}(T_{n+1}) + J + \xi_n(t - T_{n+1}) & \text{if } T_{n+1} \leq t < T_n. \end{cases}$$

By construction, the joint distribution of $(\tilde{\xi}_n, \tilde{\xi}_{n+1})$ is equal to the one we get from the original process X, and it should now be clear that the point process of jumps of ξ_n is equal in distribution to the point process of jumps of ξ_{n+1} with additional jumps distributed as $J = \log v_1$, arising at rate $(J_{n+1} - J_n)$. Note that by construction, J has distribution η_{n+1}, so finally we have proven (5.12). The fact that $(\lambda_n, n \geq 1)$ is a nonincreasing sequence of σ-finite measures ensures the existence of a limiting measure λ_∞ on $\mathbb{R} \setminus \{0\}$ such that for all $n \in \mathbb{N}$,

$$\lambda_n = \lambda_\infty + \sum_{k>n} (J_k - J_{k-1}) \widetilde{\eta}_k. \tag{5.14}$$

Recall that we wrote the characteristic exponent of ξ_n in the following way:

$$\psi_n(\theta) = i d_n \theta - \frac{\beta_n}{2}\theta^2 + \int_{\mathbb{R}} \left(\mathrm{e}^{i\theta y} - 1 - i\theta y \mathbb{1}_{|y| \leq 1} \right) \lambda_n(\mathrm{d}y).$$

From the previous discussion, one can construct a coupling between the two Lévy processes such that $(\xi_n(t) - \xi_{n+1}(t), t \geq 0)$ is simply a compound Poisson process with jump measure $(J_{n+1} - J_n)\widetilde{\eta}_{n+1}$, so it is clear that necessarily $\beta_n = \beta_{n+1}$, and

$$d_n = d_{n+1} + (J_{n+1} - J_n) \int_{|y| \leq 1} y \, \widetilde{\eta}_{n+1}(\mathrm{d}y) = d_{n+1} + \int_{|y| \leq 1} y \, (\lambda_n - \lambda_{n+1})(\mathrm{d}y).$$

To summarize, letting $\beta := \beta_1$, the following holds for all $n \in \mathbb{N}$

$$\beta_n = \beta \quad \text{and} \quad d_n = d_1 - \int_{|y| \leq 1} y \, (\lambda_1 - \lambda_n)(\mathrm{d}y), \tag{5.15}$$

where $(\lambda_1 - \lambda_n)$ denotes the positive measure given by

$$(\lambda_1 - \lambda_n) = \sum_{k=2}^{n} (J_k - J_{k-1}) \widetilde{\eta}_k.$$

Let us now examine the consistency properties of the measures $\mathcal{D}_n$. From this point on, for the sake of clarity, we decompose the proof in a series of steps.

Step 1. We prove the existence and uniqueness of a measure $\mathcal{D}$ on $\mathcal{M}_\infty$ satisfying

$$\mathcal{D}(\pi = \mathbf{1} \text{ and } v_1 \neq 0) = 0 \tag{5.16}$$

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

$$\mathcal{D}(\{x_{|[n]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}) = J_n \mathcal{D}_n, \tag{5.17}$$

then we show that this measure is exchangeable.

First, note that for the previous construction to be consistent, the random variable $D'_{|[n]}$ must have distribution $\mathcal{D}_n$. Indeed, on the event $\{T_{n+1} < T_n\}$, the strong Markov property at time T_{n+1} implies that the process $X_{|[n]}$ jumps according to $\mathcal{D}_n$, independently of the past, so on the complement this must hold as well, so

$$\mathcal{D}_{n+1}(x_{|[n]} \in \cdot \mid \pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0) = \mathcal{D}_n,$$

which can be rewritten

$$J_{n+1}\mathcal{D}_{n+1}(\{x_{|[n]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}) = J_n\mathcal{D}_n. \tag{5.18}$$

Now for all integers $n \leq m$, let us define a measure on $\mathcal{M}_m$ by

$$\mu_n^m := J_m\mathcal{D}_m(\,\cdot \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}), \quad m \geq n.$$

Let us prove that for all integers $n \leq k \leq m$, we have

$$\mu_n^m(x_{|[k]} \in \cdot) = \mu_n^k. \tag{5.19}$$

Note that there is nothing to prove in the case $k = m$. Now suppose this is proven for fixed $n \leq k \leq m$. Then,

$$\begin{aligned}
\mu_n^{m+1}(x_{|[k]} \in \cdot) &= J_{m+1}\mathcal{D}_{m+1}(\{x_{|[k]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}) \\
&= J_{m+1}\mathcal{D}_{m+1}(\{(x_{|[m]})_{|[k]} \in \cdot\} \cap \{(\pi_{|[m]})_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\} \\
&\qquad\qquad \cap \{\pi_{|[m]} \neq \mathbf{1}_m \text{ or } v_1 = 0\}) \\
&= J_m\mathcal{D}_m(\{x_{|[k]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}) \\
&= \mu_n^m(x_{|[k]} \in \cdot) = \mu_n^k,
\end{aligned}$$

where we have used (5.18) and the fact that $\{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\} \subset \{\pi_{|[m]} \neq \mathbf{1}_m \text{ or } v_1 = 0\}$. By induction on m this proves (5.19) for any integers $n \leq k \leq m$. Note that in particular, taking $k = n$, we see that the total mass of μ_n^m is equal to that of μ_n^n, which is J_n. In summary, for any $n \in \mathbb{N}$, the sequence $(\mathcal{M}_m, \mu_n^m/J_n, m \geq n)$ defines a inverse system of compact probability spaces, and by the Kolmogorov extension theorem, there exists a unique measure μ_n (with total mass J_n) on the inverse limit $\varprojlim_m \mathcal{M}_m = \mathcal{M}_\infty$ such that for each $m \geq n$, $\mu_n(x_{|[m]} \in \cdot) = \mu_n^m$. Now notice that by definition, for any integers $n_1 \leq n_2 \leq m$, we have

$$\mu_{n_2}^m(\cdot \cap \{\pi_{|[n_1]} \neq \mathbf{1}_{n_1} \text{ or } v_1 = 0\}) = \mu_{n_1}^m,$$

which implies by construction

$$\mu_{n_2}(\cdot \cap \{\pi_{|[n_1]} \neq \mathbf{1}_{n_1} \text{ or } v_1 = 0\}) = \mu_{n_1}.$$

This means that the sequence of measures $(\mu_n, n \geq 1)$ on $\mathcal{M}_\infty$ is increasing, and one can define the limit as $\mathcal{D}$. This measure then satisfies by construction

$$\begin{aligned}
\mathcal{D}(\{x_{|[n]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}) &= \mu_n(x_{|[n]} \in \cdot) \\
&= \mu_n^n \\
&= J_n\mathcal{D}_n,
\end{aligned}$$

which is indeed (5.17).

Secondly, note that since for any $n \in \mathbb{N}$, clearly $\mu_n(\pi = \mathbf{1} \text{ and } v_1 \neq 0) = 0$, where μ_n are the measures defined above, so in the limit (5.16) holds. Let us now show uniqueness. If a measure $\mathcal{D}'$ on $\mathcal{M}_\infty$ satisfies (5.17) and (5.16) then

$$\mathcal{D}'(\cdot \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}) = \mu_n,$$

and letting $n \to \infty$, $\mathcal{D}' = \mathcal{D}$, which proves uniqueness.

Finally, $\mathcal{D}$ is exchangeable. Indeed if $\sigma : \mathbb{N} \to \mathbb{N}$ is a permutation, let $m \in \mathbb{N}$ such that $\sigma(k) = k$ for all $k \geq m$. Now for all $n \geq m$, using the exchangeability of the probability measures $(\mathcal{D}_k, k \geq 1)$, we get

$$\begin{aligned}
\mathcal{D}((x^\sigma)_{|[n]} \in \cdot) &= \lim_{k\to\infty} \mathcal{D}(\{(x^\sigma)_{|[n]} \in \cdot\} \cap \{\pi_{|[k]} \neq \mathbf{1}_k \text{ or } v_1 = 0\}) \\
&= \lim_{k\to\infty} J_k \mathcal{D}_k((x^\sigma)_{|[n]} \in \cdot) \\
&= \lim_{k\to\infty} J_k \mathcal{D}_k(x_{|[n]} \in \cdot) \\
&= \lim_{k\to\infty} \mathcal{D}(\{x_{|[n]} \in \cdot\} \cap \{\pi_{|[k]} \neq \mathbf{1}_k \text{ or } v_1 = 0\}) \\
&= \mathcal{D}(x_{|[n]} \in \cdot).
\end{aligned}$$

As this is true for all $n \geq m$, necessarily $\mathcal{D}(x^\sigma \in \cdot) = \mathcal{D}$, i.e. $\mathcal{D}$ is exchangeable.

Step 2. We prove that $\mathcal{D}(\mathcal{M}_\infty \setminus \mathcal{M}_\infty^\star) = 0$ by using that X is non-degenerate. For this, we need to show first that the process $(B_1(t), t \geq 0)$ of the block of X containing 1 is equal in distribution to a process $(B(t), t \geq 0)$ constructed from a Poisson point process of intensity $\mathrm{d}t \otimes \mathcal{D}$.

More precisely, define $\mathcal{N}$ a Poisson point process on $[0, \infty) \times \mathcal{M}_\infty$ with intensity $\mathrm{d}t \otimes (\mathcal{D} + \widetilde{\lambda}_\infty)$, where $\widetilde{\lambda}_\infty$ is the push-forward of λ_∞ by the map $y \in \mathbb{R} \mapsto (\mathbf{1}, \mathrm{e}^y) \in \mathcal{M}_\infty^\star$. Let us define

$$\mathcal{N}' := \{(t, \log v_1),\ (t, x) \in \mathcal{N} \text{ with } x = (\pi, \mathbf{v}) \text{ and } v_1 \notin \{0, 1\}\},$$

which is then a Poisson point process on $[0, \infty) \times \mathbb{R}$ with intensity

$$\mathrm{d}t \otimes \Big(\mathcal{D}(\{\log v_1 \in \cdot\} \cap \{v_1 \notin \{0, 1\}\}) + \lambda_\infty\Big).$$

However, note that using (5.16), (5.17) and finally (5.13), we get

$$\begin{aligned}
&\mathcal{D}(\{\log v_1 \in \cdot\} \cap \{v_1 \notin \{0, 1\}\}) \\
&\quad = \sum_{n\in\mathbb{N}} \mathcal{D}(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \notin \{0, 1\}\} \cap \{\pi_{|[n+1]} \neq \mathbf{1}_{n+1}\}) \\
&\quad = \sum_{n\in\mathbb{N}} J_{n+1} \mathcal{D}_{n+1}(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \notin \{0, 1\}\}) \\
&\quad = \sum_{n\in\mathbb{N}} (J_{n+1} - J_n)\widetilde{\eta}_{n+1},
\end{aligned}$$

and by (5.14), we find $\mathcal{D}(\{\log v_1 \in \cdot\} \cap \{v_1 \notin \{0, 1\}\}) + \lambda_\infty = \lambda_1$. Therefore, it is possible to define a Lévy process ξ with characteristic exponent ψ_1, such that the point process of

its jumps is precisely $\mathcal{N}'$. Let us also define a process $B = (B(t), t \geq 0)$ with càdlàg sample paths with values in $2^{\mathbb{N}}$ the subsets of $\mathbb{N}$, such that B has the distribution of $(B_1(t), t \geq 0)$ the block containing 1 in X. First define T as the first time $t \in [0, \infty)$ such that there is an atom $(t, (\pi, \mathbf{v})) \in \mathcal{N}$ with $v_1 = 0$. If there is none, then let $T = \infty$. Then, for each $n \in \mathbb{N}$, let $(t_1, x_1), (t_2, x_2), \ldots$ be the whole sequence (finite or infinite) of atoms of $\mathcal{N}$ with $t_1 < t_2 < \ldots \leq T$ such that for each i, $x_i = (\pi_i, \mathbf{v}_i)$ with $(\pi_i)_{|[n]} \neq \mathbf{1}_n$ and $(v_i)_1 > 0$ (or such that $(v_i)_1 = 0$ for the possible last atom, at time T). We can define $\widetilde{B}_0^n = [n]$, and inductively for each $i \geq 1$

$$\widetilde{B}_i^n := \widetilde{B}_{i-1}^n \cap (A_i \cap [n]),$$

where A_i is the block of π_i containing 1. Now define, for $t \in [0, \infty)$,

$$B^n(t) = \widetilde{B}_i^n \text{ if } t \in [t_i, t_{i+1}),$$

where we let $t_0 := 0$, and in the case $T < \infty$, i.e. if the sequence of atoms is finite, say with length $k \in \mathbb{N}$, we let $t_{k+1} := \infty$. It is readily checked that this construction is consistent in the sense that for each $t \geq 0$ there is a single $B(t) \in 2^{\mathbb{N}}$ such that $B^n(t) = B(t) \cap [n]$. Let us show that this process $(B(t), \xi(t), t \geq 0)$ has the same distribution as the marked block containing 1 in X, i.e. $(B_1(t), \log V_1(t), t \geq 0)$. For fixed $n \in \mathbb{N}$ and $x \in \mathcal{M}_n$, recall that

$$X_{|[n]} \text{ under } \mathbb{P}_x \overset{(d)}{=} \mathrm{ssFrag}_0(x, X^{(\cdot)}),$$

where $X^{(\cdot)}$ is an independent i.i.d. sequence of copies of X started from $(\mathbf{1}, 1)$. Using the same notation, for any $A \subset [n]$ with $1 \in A$ and $v > 0$, the law of the process $(B_1(t) \cap [n], \log V_1(t), t \geq 0)$ started from $(A, \log v)$ can be deduced from that of $X^{(1)} = (\Pi^{(1)}, \mathbf{V}^{(1)})$. More precisely, $\log V_1(t)$ behaves as a Lévy process with characteristic exponent ψ_n started from $\log v$, until an independent time $T_n \sim \mathrm{Exp}(J_n)$ when $\Pi^{(1)}_{|[n]}$ first jumps. At that time, $D_n^{(1)} = (\Pi^{(1)}(T_n), \mathbf{V}^{(1)}(T_n)/V_1(T_n-))_{|[n]}$ is independently drawn according to $\mathcal{D}_n$ and then writing $D_n^{(1)} = (\pi, \mathbf{v})$,

$$\log V_1(T_n) = \log V_1(T_n-) + \log v_1 \quad \text{ and } \quad B_1(T_n) = B_1(T_n-) \cap A_1 = A \cap A_1,$$

where A_1 is the block of π containing 1. Note that there is a non-zero probability that $B_1(T_n) = B_1(T_n-)$ (even with $v_1 > 0$) when $A \neq [n]$. Now in our construction, if (t_1, x) is the first atom of $\mathcal{N}$ such that $\pi \neq \mathbf{1}_n$ or $v_1 = 0$, where $x = (\pi, \mathbf{v})$, then t_1 is exponentially distributed with parameter $\mathcal{D}(\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0) = J_n$, and x is independent of t_1, distributed as

$$\frac{1}{J_n} \mathcal{D}(\cdot \cap \{\pi \neq \mathbf{1}_n \text{ or } v_1 = 0\}),$$

so that $x_{|[n]}$ has distribution $\mathcal{D}_n$. It remains only to show that $(\xi(s), 0 \leq s < t_1)$ is distributed as a Lévy process with characteristic exponent ψ_n (killed at t_1). The point process of its jumps is

$$\mathcal{N}' \cap [0, t_1) \times \mathbb{R},$$

which conditional on t_1 has intensity

$$\mathrm{d}t \otimes \Big(\mathcal{D}(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \notin \{0, 1\}\}) + \lambda_\infty \Big).$$

Let us show that this intensity is equal to $\mathrm{d}t \otimes \lambda_n$. Note that

$$\begin{aligned}
&\mathcal{D}\big(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \notin \{0,1\}\}\big) \\
&= \sum_{m \geq n} \mathcal{D}\big(\{\log v_1 \in \cdot\} \cap \{\pi_{|[m]} = \mathbf{1}_m \text{ and } v_1 \notin \{0,1\}\} \cap \{\pi_{|[m+1]} \neq \mathbf{1}_{m+1}\}\big) \\
&= \sum_{m \geq n} J_{m+1} \mathcal{D}_{m+1}\big(\{\log v_1 \in \cdot\} \cap \{\pi_{|[m]} = \mathbf{1}_m \text{ and } v_1 \notin \{0,1\}\}\big) \\
&= \sum_{m \geq n} (J_{m+1} - J_m) \widetilde{\eta}_{m+1},
\end{aligned}$$

so (5.14) shows that

$$\lambda_n = \lambda_\infty + \mathcal{D}\big(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \notin \{0,1\}\}\big) \tag{5.20}$$

therefore the Lévy measure of $(\xi(s), 0 \leq s < t_1)$ is indeed λ_n. In the end, we have shown that

$$(B(t), \xi(t), t \geq 0) \overset{(d)}{=} (B_1(t), \log V_1(t), t \geq 0).$$

From this construction, we see that for each atom $(t,x) \in \mathcal{N}$, the process B jumps, with

$$B(t) = B(t-) \cap A,$$

where A is the block of x containing 1. Let us show that this implies $\mathcal{D}(\mathcal{M}_\infty \setminus \mathcal{M}_\infty^\star) = 0$. Assuming the opposite, there is a non-zero probability that there is an atom $(t,x) \in \mathcal{N}$ with $t < T$ such that x contains a finite block with mark not equal to zero. By exchangeability of $\mathcal{D}$, and from the description of the jumps of B, there is a non-zero probability that there is a jump $B(t) = B(t-) \cap A$ where A is finite and $t < T$. This contradicts the assumption of non-degeneracy of X, as then we would have $X(t) \in \mathcal{M}_\infty \setminus \mathcal{M}_\infty^\star$.

From now on, we view $\mathcal{D}$ as an exchangeable measure on $\mathcal{M}_\infty^\star$, satisfying (5.16) and the σ-finiteness assumption

$$\forall n \in \mathbb{N}, \quad \mathcal{D}\big(\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\big) = J_n < \infty.$$

It remains essentially to study $\mathcal{D}$ in order to express it as a mixture of paintbox processes.

Step 3. Let us decompose

$$\mathcal{D} = \mathcal{D}\big(\,\cdot \cap \{\pi = \mathbf{1}\}\big) + \mathcal{D}\big(\,\cdot \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^\downarrow = \mathbf{1}\}\big) + \mathcal{D}\big(\,\cdot \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^\downarrow \neq \mathbf{1}\}\big)$$

and show that there exist $c \geq 0$ a constant and Λ' a σ-finite measure on $\mathscr{Z}^{\downarrow}$ such that

(a) $\mathcal{D}\big(\,\cdot \cap \{\pi = \mathbf{1}\}\big) = \mathcal{D}(\pi = \mathbf{1})\delta_{(\mathbf{1},0)}$,

(b) $\mathcal{D}\big(\,\cdot \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^\downarrow = \mathbf{1}\}\big) = c \sum_{n \in \mathbb{N}} \delta_{\mathfrak{e}_n}$, where

$$\mathfrak{e}_n := \Big(\{\{n\}, \mathbb{N} \setminus \{n\}\}, (1, \ldots, 1, \underbrace{0}_{n-\text{th index}}, 1, \ldots)\Big) \in \mathcal{M}_\infty^\star,$$

(c) $\mathcal{D}\big(\,\cdot \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^\downarrow \neq \mathbf{1}\}\big) = \displaystyle\int_{\mathscr{Z}^{\downarrow}} \varrho_{\mathbf{z}}(\cdot)\, \Lambda'(\mathrm{d}\mathbf{z})$,

We use similar arguments as in [10, Theorem 3.1], as we have already done in the context of nested fragmentations [36, Proposition 19]. First note that by (5.16), $\mathcal{D}$-a.e. on the event $\{\pi = \mathbf{1}\}$ we have $x = (\mathbf{1}, 0)$, in other words (a) holds. Note also that $\mathcal{D}(\pi = \mathbf{1}) \leq \mathcal{D}(v_1 = 0) = J_1 < \infty$.

Let us now study the measure $\mathcal{D}(\cdot \cap \{\pi \neq \mathbf{1}\})$. Note that $\mathcal{D}(\{\pi \in \cdot\} \cap \{\pi \neq \mathbf{1}\})$ is an exchangeable measure on $\mathcal{P}_\infty$ satisfying for all $n \geq 1$,

$$\mathcal{D}(\pi_{|[n]} \neq \mathbf{1}_n) < \infty.$$

A consequence of [10, Theorem 3.1] is that π has asymptotic frequencies $\mathcal{D}$-a.e. – recall that $|\pi|^\downarrow \in \mathscr{S}^\downarrow \subset [0,1]^{\mathbb{N}}$ denotes the nonincreasing reordering of asymptotic frequencies of blocks of π – and one can write

$$\mathcal{D}(\{\pi \in \cdot\} \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^\downarrow = (1,0,0,\ldots)\}) = c \sum_{n \in \mathbb{N}} \delta_{\{\{n\}, \mathbb{N}\setminus\{n\}\}},$$

where $c \geq 0$. For conciseness – and again with some abuse of notation – we will from now on let $\mathbf{1} := (1,0,0\ldots) \in \mathscr{S}^\downarrow$. Now let us examine the distribution of $x = (\pi, \mathbf{v})$ on the event $\{\pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}$. Since $x \in \mathcal{M}_\infty^\star$ $\mathcal{D}$-a.e., the singleton block must have mark 0, while the other block may have a positive mark. Let η be the distribution of this mark on S^1, that is

$$\eta := \mathcal{D}(\{v_1 \in \cdot\} \cap \{\pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}),$$

which is a measure of total mass c, for any fixed $n > 1$ (by exchangeability, η does not depend on the value of n). First, note that $\eta(\{0\}) = 0$. Indeed, since the events

$$\{\pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}, \qquad n > 1$$

are disjoint, the following holds.

$$\begin{aligned}
\sum_{n>1} \eta(\{0\}) &= \sum_{n>1} \mathcal{D}(\{v_1 = 0 \text{ and } \pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}) \\
&= \mathcal{D}\Big(\bigcup_{n>1} \{v_1 = 0 \text{ and } \pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}\Big) \\
&\leq \mathcal{D}(v_1 = 0) \\
&= J_1 < \infty,
\end{aligned}$$

which implies necessarily $\eta(\{0\}) = 0$. Now recall that $\mathcal{D}(\{\log v_1 \in \cdot\} \cap \{v_1 \notin \{0,1\}\})$ is a Lévy measure, so for all $\varepsilon > 0$,

$$\begin{aligned}
\sum_{n>1} \eta(|\log v_1| > \varepsilon) &= \sum_{n>1} \mathcal{D}(\{|\log v_1| > \varepsilon \text{ and } \pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}) \\
&= \mathcal{D}\Big(\bigcup_{n>1} \{|\log v_1| > \varepsilon \text{ and } \pi = \{\{n\}, \mathbb{N} \setminus \{n\}\}\}\Big) \\
&\leq \mathcal{D}(|\log v_1| > \varepsilon \text{ and } v_1 \neq 0) < \infty,
\end{aligned}$$

which implies necessarily $\eta(|\log v_1| > \varepsilon) = 0$. Letting $\varepsilon \to 0$, we have $\eta(|\log v_1| > 0) = 0$, so in the end $\eta = c\delta_1$, and (b) follows.

Let us now decompose the measure $\mathcal{D}(\cdot \cap \{|\pi|^{\downarrow} \neq \mathbf{1}\})$. Recall that by construction,

$$\mathcal{D}(\cdot \cap \{\pi_{|[n]} \neq \mathbf{1} \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}) \leq \mathcal{D}(\cdot \cap \{\pi_{|[n]} \neq \mathbf{1} \text{ or } v_1 = 0\}) = \mu_n,$$

which is a finite measure with total mass J_n. Now let us introduce the injection $\theta_n : \mathbb{N} \to \mathbb{N}, k \mapsto n + k$, and consider

$$\overleftarrow{\mu}_n := \mathcal{D}(\{x^{\theta_n} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1} \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}),$$

which can be seen as the distribution of the marked partition restricted to $\{n+1, n+2, \ldots\}$, on the event $\{\pi_{|[n]} \neq \mathbf{1} \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}$. It is readily checked that this measure is exchangeable on $\mathcal{M}_{\infty}^{\star}$ and since it is finite, by Proposition 5.2,

$$\overleftarrow{\mu}_n = \int_{\mathscr{Z}^{\downarrow}} \varrho_{\mathbf{z}}(\cdot)\, \Lambda_n(\mathrm{d}\mathbf{z}),$$

with $\Lambda_n = \overleftarrow{\mu}_n(|x|^{\downarrow} \in \cdot)$ a finite measure on $\mathscr{Z}^{\downarrow}$. Asymptotic frequencies are such that $|x|^{\downarrow} = |x^{\theta_n}|^{\downarrow}$ for all $x \in \mathcal{M}_{\infty}^{\star}$, therefore Λ_n can also be written

$$\Lambda_n = \mathcal{D}(\{|x|^{\downarrow} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1} \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}),$$

and taking nondecreasing limits one can define

$$\Lambda' := \mathcal{D}(\{|x|^{\downarrow} \in \cdot\} \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}).$$

To show (c), fix $k, n \in \mathbb{N}$ and consider the permutation $\tau : \mathbb{N} \to \mathbb{N}$ given by

$$\tau(i) = \begin{cases} i + k & \text{if } i \leq n \\ i - n & \text{if } n < i \leq n + k \\ i & \text{otherwise.} \end{cases}$$

Now notice that

$$\mathcal{D}(\{x_{|[k]} \in \cdot\} \cap \{\pi \neq \mathbf{1} \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}) = \lim_{n \to \infty} \mathcal{D}(\{x_{|[k]} \in \cdot\} \cap \{(\pi^{\theta_k})_{|[n]} \neq \mathbf{1}_n \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}),$$

which can be written, using the exchangeability of $\mathcal{D}$,

$$\begin{aligned} &\mathcal{D}(\{x_{|[k]} \in \cdot\} \cap \{(\pi^{\theta_k})_{|[n]} \neq \mathbf{1}_n \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}) \\ &= \mathcal{D}(\{(x^{\tau \circ \theta_n})_{|[k]} \in \cdot\} \cap \{(\pi^{\tau})_{|[n]} \neq \mathbf{1}_n \text{ and } |\pi^{\tau}|^{\downarrow} \neq \mathbf{1}\}) \\ &= \mathcal{D}(\{(x^{\theta_n})_{|[k]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ and } |\pi|^{\downarrow} \neq \mathbf{1}\}) \\ &= \int_{\mathscr{Z}^{\downarrow}} \varrho_{\mathbf{z}}^{k}(\cdot)\, \Lambda_n(\mathrm{d}\mathbf{z}), \end{aligned}$$

where $\varrho_{\mathbf{z}}^{k}$ is the paintbox process restricted to k elements defined in Section 5.1.2. Taking limits and because k is generic, we have indeed (c).

Step 4. It remains to define the measure Λ correctly and we will be able to complete the proof of Theorem 5.13. Recall the definition of $\widetilde{\lambda}_\infty$ as the push-forward of λ_∞ by the map $y \in \mathbb{R} \mapsto (\mathbf{1}, \mathrm{e}^y) \in \mathcal{M}_\infty^\star$, and note that

$$\widetilde{\lambda}_\infty = \int_{\mathscr{Z}^\downarrow} \varrho_{\mathbf{z}}(\cdot)\, \widehat{\lambda}_\infty(\mathrm{d}\mathbf{z}),$$

where $\widehat{\lambda}_\infty$ is the push-forward of λ_∞ by the map $y \in \mathbb{R} \mapsto (\mathbf{1}, \mathrm{e}^y) \in \mathscr{Z}^{\downarrow}$. In the end, let us define

$$\Lambda = \Lambda' + \mathcal{D}(\pi = \mathbf{1})\delta_{(\mathbf{1},0)} + \widehat{\lambda}_\infty.$$

Putting everything together, we have

$$\mathcal{D} + \widetilde{\lambda}_\infty = c \sum_{n\in\mathbb{N}} \delta_{\mathfrak{e}_n} + \int_{\mathscr{Z}^\downarrow} \varrho_{\mathbf{z}}(\cdot)\, \Lambda(\mathrm{d}\mathbf{z}).$$

We can almost complete the proof. First, fix $n \in \mathbb{N}$ and recall that by definition $J_n = \mathcal{D}(\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0)$. Since by definition $\widetilde{\lambda}_\infty(\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0) = 0$ for all n, point (ii) of Theorem 5.13 is proven:

$$J_n = nc + \int_{\mathscr{Z}^\downarrow} \Big(1 - \sum_{\substack{i\geq 1\\ v_i>0}} s_i^n\Big)\, \Lambda(\mathrm{d}\mathbf{z}).$$

This implies that $\big(1 - \sum_{\substack{i\geq 1\\ v_i>0}} s_i^n\big)$ is Λ-integrable for any $n \geq 1$. Furthermore, note that for any $\mathbf{z} = (\mathbf{s}, \mathbf{v}) \in \mathscr{Z}^{\downarrow}$,

$$1 - s_1 \leq 1 - s_1\Big(\sum_{\substack{i\geq 1\\ v_i>0}} s_i\Big) \leq 1 - \sum_{\substack{i\geq 1\\ v_i>0}} s_i^2, \quad \text{and} \quad s_1 \mathbb{1}_{v_1=0} \leq 1 - \sum_{\substack{i\geq 1\\ v_i>0}} s_i,$$

therefore summing the two expressions yields

$$\int_{\mathscr{Z}^\downarrow} 1 - s_1 \mathbb{1}_{v_1>0}\, \Lambda(\mathrm{d}\mathbf{z}) < \infty. \tag{5.21}$$

We keep this in mind for later use and go back to the expression of the measures $\mathcal{D}_n$. If $J_n > 0$, then $\mathcal{D}_n = \mathcal{D}\big(\{x_{|[n]} \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ or } v_1 = 0\}\big)/J_n$, so point (iii) of Theorem 5.13 is proven:

$$\mathcal{D}_n = \frac{1}{J_n}\Big(\sum_{i=1}^n c\delta_{\mathfrak{e}_i^n} + \int_{\mathscr{Z}^\downarrow} \varrho_{\mathbf{z}}^n\big(\,\cdot \cap \{\pi \neq \mathbf{1}_n \text{ or } v_1 = 0\}\big)\, \Lambda(\mathrm{d}\mathbf{z})\Big),$$

where $\mathfrak{e}_i^n \in \mathcal{M}_n$ is defined as

$$\mathfrak{e}_i^n := \Big(\{[n] \setminus \{i\}, \{i\}\}, (1, \dots, 1, \underbrace{0}_{i-\text{th index}}, 1, \dots, 1)\Big).$$

It remains for the first part of the theorem to express ψ_n correctly and to show the integrability condition (5.4). Recall from (5.20) that $\lambda_n = \lambda_\infty + \mathcal{D}\big(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} = \mathbf{1}_n \text{ and } v_1 \notin \{0,1\}\}\big)$, which shows that

$$\psi_n(\theta) = id_n\theta - \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^\downarrow} \sum_{\substack{j\geq 1\\ v_j>0}} s_j^n \left(\mathrm{e}^{i\theta \log v_j} - 1 - i\theta \log v_j \mathbb{1}_{|\log v_j|\leq 1}\right) \Lambda(\mathrm{d}\mathbf{z}),$$

but, from (5.15), we have

$$\begin{aligned} d_n &= d_1 - \int_{|y|\le 1} y\,(\lambda_1 - \lambda_n)(\mathrm{d}y) \\ &= d_1 - \int_{\mathscr{Z}^\downarrow} \sum_{\substack{j\ge 1\\ v_j>0}} s_j(1-s_j^{n-1}) \log v_j \mathbb{1}_{|\log v_j|\le 1}\,\Lambda(\mathrm{d}\mathbf{z}) \end{aligned}$$

where we used $(\lambda_1 - \lambda_n) = \mathcal{D}(\{\log v_1 \in \cdot\} \cap \{\pi_{|[n]} \neq \mathbf{1}_n \text{ and } v_1 \neq 0\})$ which is again deduced from (5.20). Putting the last two displays together, we get

$$\psi_n(\theta) = id_1\theta - \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^\downarrow} \sum_{\substack{j\ge 1\\ v_j>0}} \left(s_j^n\left(\mathrm{e}^{i\theta\log v_j} - 1\right) - i\theta s_j \log v_j \mathbb{1}_{|\log v_j|\le 1}\right)\,\Lambda(\mathrm{d}\mathbf{z}).$$

In order to simplify this notation, note that

$$\begin{aligned} \left|\sum_{\substack{j\ge 1\\ v_j>0}} s_j \log v_j \mathbb{1}_{|\log v_j|\le 1} - \log v_1 \mathbb{1}_{|\log v_1|\le 1}\right| &\le (1 - s_1\mathbb{1}_{v_1>0})|\log v_1|\mathbb{1}_{|\log v_1|\le 1} + \sum_{\substack{j\ge 2\\ v_j>0}} s_j \\ &\le 2(1 - s_1\mathbb{1}_{v_1>0}), \end{aligned}$$

which we proved to be Λ-integrable in (5.21). Therefore, we can finally define

$$d := d_1 + \int_{\mathscr{Z}^\downarrow} \log v_1 \mathbb{1}_{|\log v_1|\le 1} - \sum_{\substack{j\ge 1\\ v_j>0}} s_j \log v_j \mathbb{1}_{|\log v_j|\le 1}\,\Lambda(\mathrm{d}\mathbf{z})$$

in order to get point (i) of Theorem 5.13, that is

$$\psi_n(\theta) = id\theta - \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^\downarrow} \sum_{\substack{j\ge 1\\ v_j>0}} s_j^n\left(\mathrm{e}^{i\theta\log v_j} - 1\right) - i\theta \log v_1 \mathbb{1}_{|\log v_1|\le 1}\,\Lambda(\mathrm{d}\mathbf{z}).$$

Now let us show (5.4). From (5.21), it remains only to check that $(\log v_1)^2 \wedge 1$ is Λ-integrable. Since λ_1 must be a Lévy measure, we have

$$\begin{aligned} \int_{\mathbb{R}} (y^2\wedge 1)\,\lambda_1(\mathrm{d}y) &= \int_{\mathscr{Z}^\downarrow}\int_{\mathbb{R}} (y^2\wedge 1)\,\varrho^1_{\mathbf{z}}(\{\log v_1 \in \mathrm{d}y\} \cap \{v_1 \neq 0\})\,\Lambda(\mathrm{d}\mathbf{z}) \\ &= \int_{\mathscr{Z}^\downarrow} \sum_{\substack{i\ge 1\\ v_i>0}} s_i\left((\log v_i)^2 \wedge 1\right)\Lambda(\mathrm{d}\mathbf{z}) < \infty. \end{aligned}$$

Now note that for all $\mathbf{z} \in \mathscr{Z}^\downarrow$,

$$\begin{aligned} (\log v_1)^2 \wedge 1 &\le s_1\mathbb{1}_{v_1>0}\left((\log v_1)^2\wedge 1\right) + (1 - s_1\mathbb{1}_{v_1>0}) \\ &\le \sum_{\substack{i\ge 1\\ v_i>0}} s_i^n\left((\log v_i)^2\wedge 1\right) + (1 - s_1\mathbb{1}_{v_1>0}), \end{aligned}$$

which is Λ-integrable. This proves (5.4), and ends the proof of the main result of Theorem 5.13.

For the converse part, let (c, d, β, Λ) be a given quadruple, where $c, \beta \geq 0$, $d \in \mathbb{R}$, and Λ is a measure on $\mathscr{Z}^{\downarrow} \setminus \{(\mathbf{1}, 1)\}$ satisfying (5.4). Then $(\psi_n, J_n, \mathcal{D}_n)$ for all $n \in \mathbb{N}$ are well-defined as in the theorem if one checks that

$$\begin{aligned} &\int_{\mathscr{Z}^{\downarrow}} \sum_{\substack{i \geq 1 \\ v_i > 0}} s_i^n \left((\log v_i)^2 \wedge 1 \right) \Lambda(\mathrm{d}\mathbf{z}) < \infty \\ \text{and} \quad &\int_{\mathscr{Z}^{\downarrow}} \Big(1 - \sum_{\substack{i \geq 1 \\ v_i > 0}} s_i^n \Big) \Lambda(\mathrm{d}\mathbf{z}) < \infty. \end{aligned} \tag{5.22}$$

To that aim, note that for all $\mathbf{z} \in \mathscr{Z}^{\downarrow}$,

$$\begin{aligned} \sum_{\substack{i \geq 1 \\ v_i > 0}} s_i^n \left((\log v_i)^2 \wedge 1 \right) + \Big(1 - \sum_{\substack{i \geq 1 \\ v_i > 0}} s_i^n \Big) &\leq s_1^n \left((\log v_1)^2 \wedge 1 \right) + \sum_{\substack{i \geq 2 \\ v_i > 0}} s_i^n + \Big(1 - \sum_{\substack{i \geq 1 \\ v_i > 0}} s_i^n \Big) \\ &\leq \left((\log v_1)^2 \wedge 1 \right) + (1 - s_1^n \mathbb{1}_{v_1 > 0}) \\ &\leq \left((\log v_1)^2 \wedge 1 \right) + n(1 - s_1 \mathbb{1}_{v_1 > 0}), \end{aligned}$$

which is Λ-integrable by (5.4), so (5.22) is proven. Now the construction of a 0-ESSF X with characteristics as above is done via Remark 5.12.

5.A.5 Proof of Proposition 5.16

Consider a non-degenerate homogeneous ESSF $X = (\Pi, \mathbf{V})$ with characteristics (c, d, β, Λ). We make use of a natural genealogy appearing in our construction: jointly for all $n \in \mathbb{N}$, we define processes $(F_n(t), t \geq 0)$ taking values in the subsets of $\mathbb{N}$, starting at $F_n(0) = [n]$, whose role is to "follow" integers along a discrete genealogy of blocks. We will make this statement more precise, but first let us explain the idea. We will build the F_n deterministically from a sample path of X, such that for each time $t \geq 0$,

$$[n] \subset F_n(t) \subset F_{n+1}(t),$$

and so by defining

$$S_\theta^{(n)}(t) = \sum_{B \text{ block of } X(t)_{|F_n(t)}} \tilde{V}_B(t)^\theta, \tag{5.23}$$

where $\tilde{V}_B(t)$ denotes the mark of a block B of $X(t)$, it is clear that

$$S_\theta^{(n)}(t) \leq S_\theta^{(n+1)}(t) \underset{n \to \infty}{\longrightarrow} S_\theta(X(t)) = \sum_{B \text{ block of } X(t)} \tilde{V}_B(t)^\theta.$$

The way to define the sets F_n is the following. Recall that for any $n \in \mathbb{N}$, T_n denotes the first time t when $[n]$ is no longer part of a unique block with positive mark in $X(t)$. Therefore let $F_n(t) = [n]$ for any $0 \leq t < T_n$. Since X is homogeneous, T_n is an exponential random variable with parameter J_n, and conditional on T_n, the mark $(V_1(t), 0 \leq t < T_n)$ behaves as the exponential of a Lévy process ξ_n with characteristic exponent

$$\psi_n : \theta \mapsto id\theta - \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^{\downarrow}} \sum_{\substack{j \geq 1 \\ v_j > 0}} s_j^n (\mathrm{e}^{i\theta \log v_j} - 1) - i\theta \log v_1 \mathbb{1}_{|\log v_1| \leq 1} \Lambda(\mathrm{d}\mathbf{z}).$$

Then at time T_n, $(\Pi(T_n), \mathbf{V}(T_n)/V_1(T_n-))_{|[n]}$ is independent of the past and has distribution $\mathcal{D}_n$. Let us recall that $\mathcal{D}_n$ is expressed in terms of Λ by

$$\mathcal{D}_n = \frac{1}{J_n}\Big(\sum_{i=1}^{n} c\delta_{\mathfrak{e}_i^n} + \int_{\mathscr{Z}^\downarrow} \varrho_{\mathbf{z}}^n(\,\cdot\cap\{\pi \neq \mathbf{1}_n \text{ or } (\pi,\mathbf{v}) = (\mathbf{1}_n, 0)\})\,\Lambda(\mathrm{d}\mathbf{z})\Big).$$

At this time, for each block B among the newly created blocks of $X(T_n)$, if $B \cap [n] \neq \varnothing$ and if B has positive mark, let $F_B \subset B$ be the subset consisting of exactly the first n integers that are part of block B (necessarily B contains infinitely many integers so the first n ones exist). Now we define

$$F_n(T_n) := F_n(T_n-) \cup \bigcup_B F_B,$$

where the union is taken over all newly created blocks of $X(T_n)$ with positive mark and nonempty intersection with $F_n(T_n-)$. After this first step, $X(T_n)_{|F_n(T_n)}$ consists of a random but finite (bounded by n) number of blocks, those with positive marks containing exactly n integers. The construction is recursive: if at time t the marked partition $X(t)_{|F_n(t)}$ contains K blocks of size n, then after an exponential time T with parameter KJ_n, one of them dislocates exactly as in the first step and n integers are selected for each newly created block in this dislocation. At the time of dislocation $F_n(t+T)$ is modified accordingly, and between time t and $t+T$, the branching property ensures us that each block has a mark behaving independently as e^{ξ_n}, where ξ_n is a Lévy process with characteristic exponent ψ_n. This recursion defines the process for all $t \geq 0$, and the construction is designed so that if $S_\theta^{(n)}$ is defined by (5.23), then

$$(S_\theta^{(n)}(t),\, t \geq 0) \overset{(d)}{=} (\sum_i \mathrm{e}^{\theta\xi_n^i(t)},\, t \geq 0),$$

where $\big(\xi_n^i(t),\, i \geq 1,\, t \geq 0\big)$ is a system of branching particles started from a unique particle at position 0, which can be described by:

- particles move independently as Lévy processes equal to ξ_n in distribution.
- a particle branches at rate J_n into a random set of K particles at positions $y + (y_1, \ldots, y_K)$, where y is the position of the mother particle at the time of branching and $(y_1, \ldots, y_K)$ is a vector independent of the past and with distribution given by

  $$\mathbb{E}f(y_1, \ldots, y_K) = \int_{\mathcal{M}_n} f(\log v_1, \ldots, \log v_K)\,\mathcal{D}_n(\mathrm{d}x),$$

  where in the right-hand side integrand, the vector $(v_1, \ldots, v_K)$ denotes the non-zero marks of x.

At this point we need the following lemma, which results from standard branching processes arguments. I could not find a reference which proves this result entirely in this form, so a short, straightforward proof is given below.

Lemma 5.20. *We have*

$$\mathbb{E}S_\theta^{(n)}(t) = \mathrm{e}^{t\kappa^{(n)}(\theta)}, \qquad \textit{with} \qquad \kappa^{(n)}(\theta) = A^{(n)}(\theta) + J_n B^{(n)}(\theta), \tag{5.24}$$

where $A^{(n)}$ corresponds to the movement of particles, with

$$A^{(n)}(\theta) = d\theta + \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^\downarrow} \sum_{\substack{i\geq 1\\ v_i>0}} s_i^n(v_i^\theta - 1) - \theta \log v_1 \mathbb{1}_{|\log v_1|\leq 1}\, \Lambda(\mathrm{d}\mathbf{z}) \tag{5.25}$$

and $B^{(n)}$ corresponds to the branching, with

$$\begin{aligned} B^{(n)}(\theta) &= \int_{\mathcal{M}_n} (S_\theta(x) - 1)\, \mathcal{D}_n(\mathrm{d}x) \\ &= \frac{1}{J_n} \int_{\mathscr{Z}^\downarrow} \int_{\{\pi \neq \mathbf{1}_n \text{ or } v_1 = 0\}} (S_\theta(x) - 1)\, \varrho_{\mathbf{z}}^n(\mathrm{d}x)\, \Lambda(\mathrm{d}\mathbf{z}). \end{aligned}$$

Proof. It is standard in the theory of Lévy processes (see e.g. [80, Theorem 25.17]) that $A^{(n)}(\theta) < \infty$ if and only if $\mathbb{E}\mathrm{e}^{\theta\xi_n(t)} < \infty$ for all $t \geq 0$, and in that case $\mathbb{E}\mathrm{e}^{\theta\xi_n(t)} = \mathrm{e}^{tA^{(n)}(\theta)}$. Now fix $0 < s < t$ and consider the event

$$A_s^t := \{\text{the initial particle branches at time } s \text{ and no other branching occurs before time } t\}.$$

Then conditional on A_s^t, the branching construction yields

$$\begin{aligned} \mathbb{E}\Big[\sum_i \mathrm{e}^{\theta\xi_n(t)} \,\Big|\, A_s^t\Big] &= \mathbb{E}\Big[\int_{\mathcal{M}_n} \sum_i \mathrm{e}^{\theta(\xi_n(s) + \log v_i + \widetilde{\xi}_n^{(i)}(t-s))}\, \mathcal{D}_n(\mathrm{d}x) \,\Big|\, A_s^t\Big] \\ &= \Big(\int_{\mathcal{M}_n} \sum_i v_i^\theta\, \mathcal{D}_n(\mathrm{d}x)\Big) \mathbb{E}\Big[\mathrm{e}^{\theta(\xi_n(s) + \widetilde{\xi}_n^{(1)}(t-s))}\Big] \\ &= \Big(\int_{\mathcal{M}_n} S_\theta(x)\, \mathcal{D}_n(\mathrm{d}x)\Big) \mathbb{E}\mathrm{e}^{\theta\xi_n(t)} \\ &= (B^{(n)}(\theta) + 1)\mathrm{e}^{tA^{(n)}(\theta)}, \end{aligned}$$

where ξ_n and the $\widetilde{\xi}_n^{(i)}$, $i \geq 1$ are i.i.d. Lévy processes started from 0. This quantity does not depend on s, so one may write, if A^t is the event of a single branching occurring before time t,

$$\mathbb{E}\Big[\sum_i \mathrm{e}^{\theta\xi_n^i(t)} \,\Big|\, A^t\Big] = (B^{(n)}(\theta) + 1)\mathrm{e}^{tA^{(n)}(\theta)}.$$

and in particular,

$$\mathbb{E}S_\theta^{(n)}(t) = \mathbb{E}\Big[\sum_i \mathrm{e}^{\theta\xi_n^i(t)}\Big] \geq \mathbb{P}(A^t)(B^{(n)}(\theta) + 1)\mathrm{e}^{tA^{(n)}(\theta)},$$

which shows that if $A^{(n)}(\theta) = \infty$ or $B^{(n)}(\theta) = \infty$, then $\mathbb{E}S_\theta^{(n)}(t) = \infty$. Now let us assume that both quantities are finite, and prove (5.24). First note that the argument above readily extends to

$$\mathbb{E}\Big[\sum_i \mathrm{e}^{\theta\xi_n^i(t)} \,\Big|\, A^{t,k}\Big] = (B^{(n)}(\theta) + 1)^k\, \mathbb{E}\mathrm{e}^{\theta\xi_n(t)}.$$

where $A^{t,k}$ is the event of exactly k particles branching before time t. We now bound from above the probability of $A^{t,k}$. Let $t_0 := 0 < t_1 < t_2 < \dots$ denote the branching times in our particle system. Note that as particles produce at most n offspring, the time between consecutive branching times $t_j - t_{j-1}$ is greater than an exponential random variable with

parameter nJ_nj. Therefore we can compare the process counting branching times in our process and a Yule process with birth rate nJ_n, which yields

$$\mathbb{P}(A^{t,k}) \leq \mathbb{P}(t_k < t) \leq (1 - \mathrm{e}^{-tnJ_n})^k.$$

Now if t^* is small enough so that $(B^{(n)}(\theta) + 1)(1 - \mathrm{e}^{-tnJ_n}) < 1$ for all $0 \leq t \leq t^*$, then we have

$$\forall 0 \leq t \leq t^*, \qquad \mathbb{E}S_\theta^{(n)}(t) = \sum_{k \geq 0} \mathbb{P}(A^{t,k}) \big(B^{(n)}(\theta) + 1\big)^k \, \mathbb{E}\mathrm{e}^{\theta \xi_n(t)} < \infty,$$

so the map $f : t \mapsto \mathbb{E}S_\theta^{(n)}(t)$ takes finite values before time t^*. Now note that for any times $t, s \geq 0$ the branching property applied at time t yields $\mathbb{E}[S_\theta(t+s) \mid S_\theta(t)] = S_\theta(t) f(s)$, and so taking expectations, $f(t+s) = f(t)f(s)$. Since f takes finite value for $0 \leq t \leq t^*$, this shows that $f(t)$ is finite for all $t \geq 0$. Now let us compute $f(t)$ by applying the branching property at the first branching time t_1:

$$\begin{aligned} f(t) &= \mathbb{P}(t_1 > t)\mathbb{E}\mathrm{e}^{\theta \xi_n(t)} + \int_0^t \int_{\mathcal{M}_n} \sum_i \mathbb{E}\mathrm{e}^{\theta(\xi_n(s) + \log v_i)} f(t-s) \, \mathcal{D}_n(\mathrm{d}x) \, \mathbb{P}(t_1 \in \mathrm{d}s) \\ &= \mathrm{e}^{-J_n t} \mathrm{e}^{tA^{(n)}(\theta)} + \int_0^t J_n \mathrm{e}^{-J_n s} \mathrm{e}^{sA^{(n)}(\theta)} \big(B^{(n)}(\theta) + 1\big) f(t-s) \, \mathrm{d}s, \end{aligned}$$

and it is easily checked that the only solution of this equation is indeed

$$\mathbb{E}S_\theta^{(n)}(t) = f(t) = \mathrm{e}^{t(A^{(n)}(\theta) + J_n B^{(n)}(\theta))} = \mathrm{e}^{t\kappa^{(n)}(\theta)},$$

so (5.24) is proved. □

Now note that for any $\mathbf{z} \in \mathscr{Z}^\downarrow$, one can write

$$\sum_{\substack{i \geq 1 \\ v_i > 0}} s_i^n (v_i^\theta - 1) = \int_{\{\pi = \mathbf{1}_n \text{ and } v_1 > 0\}} (S_\theta(x) - 1) \, \varrho_{\mathbf{z}}^n(\mathrm{d}x),$$

so plugging this into the expression (5.25) for $A_\theta^{(n)}$ and putting everything together, we have

$$\kappa^{(n)}(\theta) = d\theta + \frac{\beta}{2}\theta^2 + \int_{\mathscr{Z}^\downarrow} \int_{\mathcal{M}_n} (S_\theta(x) - 1) \, \varrho_{\mathbf{z}}^n(\mathrm{d}x) - \theta \log v_1 \mathbb{1}_{|\log v_1| \leq 1} \, \Lambda(\mathrm{d}\mathbf{z}).$$

Now it is a consequence of the law of large numbers that for all $\mathbf{z} \in \mathscr{Z}^\downarrow$, the following convergence holds (and is nondecreasing)

$$\int_{\mathcal{M}_n} (S_\theta(x) - 1) \, \varrho_{\mathbf{z}}^n(\mathrm{d}x) \underset{n \to \infty}{\longrightarrow} \sum_{i \geq 1} v_i^\theta - 1,$$

and in the end, Lemma 5.20 and monotone convergence yield

$$\mathbb{E}S_\theta(X(t)) = \mathrm{e}^{t\kappa(\theta)}.$$

Now if $\kappa(\theta)$ is finite, it is a simple consequence of the Markov property of the process X that $(\mathrm{e}^{-t\kappa(\theta)} S_\theta(X(t)), t \geq 0)$ is a martingale. Since it is nonnegative it converges almost surely as $t \to \infty$ so it is almost surely bounded by a random variable which we denote by

$C = C_\theta > 0$. Now assume furthermore that $\kappa(\theta) < 0$ for some $\theta \neq 0$. Notice that almost surely for all $i \geq 1$ and $t \geq 0$,

$$V_i(t)^\theta \leq S_\theta(X(t)) \leq C\mathrm{e}^{t\kappa(\theta)},$$

so for any $\alpha \in \mathbb{R}$ such that $-\alpha/\theta > 0$,

$$V_i(t)^{-\alpha} \leq \left(C\mathrm{e}^{t\kappa(\theta)}\right)^{-\alpha/\theta},$$

and so almost surely,

$$\sup_{i\geq 1} \int_0^\infty V_i(t)^{-\alpha}\,\mathrm{d}t \leq \frac{\theta C^{-\alpha/\theta}}{\alpha\kappa(\theta)} < \infty. \tag{5.26}$$

Now recall the stopping lines

$$\tau_i^{-\alpha}(t) = \left(\int_0^\cdot V_i(s)^{-\alpha}\,\mathrm{d}s\right)^{-1}(t), \qquad i \geq 1,\, t \geq 0,$$

which we used to change the self-similarity index. Proposition 5.11 tells us that the time-changed process $X \circ \tau^{-\alpha}$ is an α-ESSF process with characteristics (c, d, β, Λ), and note that for each $i \geq 1$, the integral

$$\zeta_i = \int_0^\infty V_i(s)^{-\alpha}\,\mathrm{d}s$$

is the hitting time of 0 by the pssMp $V_i \circ \tau_i^{-\alpha}$. Clearly (5.26) shows that $X \circ \tau^{-\alpha}$ reaches absorption before time $\frac{\theta C^{-\alpha/\theta}}{\alpha\kappa(\theta)}$.

It remains to show the finite total length property in the case $-\alpha/\theta \geq 1$. Recall that

$$S_\theta(X(t)) = \sum_{k\geq 1} \tilde{V}_k(t)^\theta \leq C\mathrm{e}^{t\kappa(\theta)}.$$

It is elementary (because for any summable sequence u, $\|u\|_p \leq \|u\|_1$ for any $p \geq 1$) that for any α such that $-\alpha/\theta \geq 1$ this implies

$$S_{-\alpha}(X(t)) = \sum_{k\geq 1} \tilde{V}_k(t)^{-\alpha} \leq \left(C\mathrm{e}^{t\kappa(\theta)}\right)^{-\alpha/\theta}.$$

We claim the time change is such that

$$\int_0^\infty \#X \circ \tau^{-\alpha}(t)\,\mathrm{d}t = \int_0^\infty S_{-\alpha}(X(t))\,\mathrm{d}t \leq \frac{\theta C^{-\alpha/\theta}}{\alpha\kappa(\theta)} < \infty \quad \text{a.s.}$$

To make this claim entirely justified, let us define for all $x \in \mathcal{M}_\infty^\star$ the (finite of infinite) set $I(x) = \{i_1, i_2, \ldots\}$ where i_k is the first integer contained in the k-th block with positive mark of x. Notice that by definition for any $i \geq 1$, for all $t \leq \zeta_i$, $\mathrm{d}\tau_i(t) = V_i^\alpha(\tau_i^{-\alpha}(t))\mathrm{d}t$, therefore

$$\begin{aligned}
\int_0^\infty \#X \circ \tau^{-\alpha}(t)\,\mathrm{d}t &= \int_0^\infty \sum_{i\in I(X\circ\tau^{-\alpha}(t))} 1\,\mathrm{d}t \\
&= \sum_{i\geq 1} \int_0^\infty \mathbb{1}_{i\in I(X(\tau_i^{-\alpha}(t)))} \frac{V_i^\alpha(\tau_i^{-\alpha}(t))\,\mathrm{d}t}{V_i^\alpha(\tau_i^{-\alpha}(t))} \\
&= \sum_{i\geq 1} \int_0^\infty \mathbb{1}_{i\in I(X(t))} V_i^{-\alpha}(t)\,\mathrm{d}t \\
&= \int_0^\infty S_{-\alpha}(X(t))\,\mathrm{d}t
\end{aligned}$$

and the proof is complete.

References for Chapter 5

[4] D. Aldous and J. Pitman. The Standard Additive Coalescent. *Ann. Probab.*, 26.4 (Oct. 1998), pp. 1703–1726. doi: 10.1214/aop/1022855879 (see pp. 16, 126).

[6] J. Berestycki. Ranked Fragmentations. *ESAIM Probab. Stat.*, 6 (2002), pp. 157–175. doi: 10.1051/ps:2002009 (see pp. 126, 134).

[9] J. Bertoin. Markovian Growth-Fragmentation Processes. *Bernoulli*, 23.2 (May 2017), pp. 1082–1101. doi: 10.3150/15-BEJ770 (see pp. 12, 126).

[10] J. Bertoin. *Random Fragmentation and Coagulation Processes.* Cambridge University Press, 2006. doi: 10.1017/CBO9780511617768 (see pp. 10, 11, 60, 61, 64, 65, 79, 81, 90, 91, 94, 95, 99, 117, 126–128, 135, 153).

[11] J. Bertoin. Self-Similar Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 38.3 (2002), pp. 319–340. doi: 10.1016/S0246-0203(00)01073-6 (see pp. 12, 126, 134, 135).

[13] J. Bertoin. The Asymptotic Behavior of Fragmentation Processes. *J. Eur. Math. Soc. (JEMS)*, 5.4 (Nov. 1, 2003), pp. 395–416. doi: 10.1007/s10097-003-0055-3 (see p. 140).

[17] J. Bertoin and B. Mallein. Infinitely Ramified Point Measures and Branching Lévy Processes. *Ann. Probab. (to appear)*, (2018+). arXiv: 1703.08078 (see p. 126).

[18] A. Blancas, J.-J. Duchamps, A. Lambert, and A. Siri-Jégousse. Trees within Trees: Simple Nested Coalescents. *Electron. J. Probab.*, 23.0 (2018). doi: 10.1214/18-EJP219 (see pp. 11, 58, 133).

[25] B. Chauvin. Product Martingales and Stopping Lines for Branching Brownian Motion. *Ann. Probab.*, 19.3 (July 1991), pp. 1195–1205. doi: 10.1214/aop/1176990340 (see p. 135).

[29] B. Dadoun. Asymptotics of Self-Similar Growth-Fragmentation Processes. *Electron. J. Probab.*, 22 (2017). doi: 10.1214/17-EJP45 (see p. 126).

[30] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes.* Vol. II. Probability and Its Applications. New York, NY: Springer New York, 2008. doi: 10.1007/978-0-387-49835-5 (see p. 142).

[36] J.-J. Duchamps. Trees within Trees II: Nested Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat. (to appear)*, (2019+). arXiv: 1807.05951 (see pp. 11, 89, 133, 153).

[49] F. G. Ged. Profile of a Self-Similar Growth-Fragmentation. *Electron. J. Probab.*, 24 (2019). doi: 10.1214/18-EJP253 (see p. 126).

[52] B. Haas and G. Miermont. The Genealogy of Self-Similar Fragmentations with Negative Index as a Continuum Random Tree. *Electron. J. Probab.*, 9 (2004), pp. 57–97. DOI: 10.1214/EJP.v9-187 (see p. 126).

[53] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum Tree Asymptotics of Discrete Fragmentations and Applications to Phylogenetic Models. *Ann. Probab.*, 36.5 (Sept. 2008), pp. 1790–1837. DOI: 10.1214/07-AOP377 (see pp. 10, 90, 91, 126).

[57] J. F. C. Kingman. The Representation of Partition Structures. *J. Lond. Math. Soc. (2)*, 18.2 (1978), pp. 374–380. DOI: 10.1112/jlms/s2-18.2.374 (see pp. 96, 130).

[59] N. Krell. Self-Similar Branching Markov Chains. *Séminaire de Probabilités XLII*. Vol. 1979. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 261–280. DOI: 10.1007/978-3-642-01763-6_10 (see pp. 126, 134).

[67] J. Lamperti. Semi-Stable Markov Processes. I. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 22.3 (1972), pp. 205–225. DOI: 10.1007/BF00536091 (see p. 134).

[75] J. C. Pardo and V. Rivero. Self-Similar Markov Processes. *Bol. Soc. Mat. Mex. (3)*, 19.2 (2013), pp. 201–235 (see p. 134).

[80] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics 68. Cambridge, U.K. ; New York: Cambridge University Press, 1999 (see p. 159).

[86] Q. Shi. Growth-Fragmentation Processes and Bifurcators. *Electron. J. Probab.*, 22 (2017). DOI: 10.1214/17-EJP26 (see p. 126).

[87] S. M. Srivastava. *A Course on Borel Sets*. Vol. 180. Graduate Texts in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. DOI: 10.1007/978-3-642-85473-6 (see pp. 109, 129).

Complete bibliography

[1] R. Abraham, J.-F. Delmas, and P. Hoscheit. A Note on the Gromov-Hausdorff-Prokhorov Distance between (Locally) Compact Metric Measure Spaces. *Electron. J. Probab.*, 18 (2013). paper no. 14. DOI: 10.1214/EJP.v18-2116 (see p. 22).

[2] R. Abraham and L. Serlet. Poisson Snake and Fragmentation. *Electron. J. Probab.*, 7 (2002). paper no. 17. DOI: 10.1214/EJP.v7-116 (see pp. 16, 31).

[3] D. Aldous. Probability Distributions on Cladograms. *Random Discrete Structures*. The IMA Volumes in Mathematics and Its Applications 76. Springer New York, 1996, pp. 1–18. DOI: 10.1007/978-1-4612-0719-1_1 (see pp. 9, 90).

[4] D. Aldous and J. Pitman. The Standard Additive Coalescent. *Ann. Probab.*, 26.4 (Oct. 1998), pp. 1703–1726. DOI: 10.1214/aop/1022855879 (see pp. 16, 126).

[5] A.-L. Basdevant and C. Goldschmidt. Asymptotics of the Allele Frequency Spectrum Associated with the Bolthausen-Sznitman Coalescent. *Electron. J. Probab.*, 13 (2008), pp. 486–512. DOI: 10.1214/EJP.v13-494 (see p. 15).

[6] J. Berestycki. Ranked Fragmentations. *ESAIM Probab. Stat.*, 6 (2002), pp. 157–175. DOI: 10.1051/ps:2002009 (see pp. 126, 134).

[7] J. Berestycki, N. Berestycki, and V. Limic. A Small-Time Coupling between Λ-Coalescents and Branching Processes. *Ann. Appl. Probab.*, 24.2 (Apr. 2014), pp. 449–475. DOI: 10.1214/12-AAP911 (see pp. 15, 82).

[8] J. Berestycki, N. Berestycki, and J. Schweinsberg. The Genealogy of Branching Brownian Motion with Absorption. *Ann. Probab.*, 41.2 (Mar. 2013), pp. 527–618. DOI: 10.1214/11-AOP728 (see p. 60).

[9] J. Bertoin. Markovian Growth-Fragmentation Processes. *Bernoulli*, 23.2 (May 2017), pp. 1082–1101. DOI: 10.3150/15-BEJ770 (see pp. 12, 126).

[10] J. Bertoin. *Random Fragmentation and Coagulation Processes*. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511617768 (see pp. 10, 11, 60, 61, 64, 65, 79, 81, 90, 91, 94, 95, 99, 117, 126–128, 135, 153).

[11] J. Bertoin. Self-Similar Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 38.3 (2002), pp. 319–340. DOI: 10.1016/S0246-0203(00)01073-6 (see pp. 12, 126, 134, 135).

[12] J. Bertoin. Subordinators: Examples and Applications. *Lectures on Probability Theory and Statistics: École d'Été de Probabilités de Saint-Flour XXVII*. Springer, 1997, pp. 1–91. doi: 10.1007/978-3-540-48115-7_1 (see pp. 31, 54).

[13] J. Bertoin. The Asymptotic Behavior of Fragmentation Processes. *J. Eur. Math. Soc. (JEMS)*, 5.4 (Nov. 1, 2003), pp. 395–416. doi: 10.1007/s10097-003-0055-3 (see p. 140).

[14] J. Bertoin. The Structure of the Allelic Partition of the Total Population for Galton–Watson Processes with Neutral Mutations. *Ann. Probab.*, 37.4 (July 2009), pp. 1502–1523. doi: 10.1214/08-AOP441 (see pp. 16, 59, 90).

[15] J. Bertoin and J.-F. Le Gall. Stochastic Flows Associated to Coalescent Processes. *Probab. Theory Related Fields*, 126.2 (2003), pp. 261–288. doi: 10.1007/s00440-003-0264-4 (see p. 60).

[16] J. Bertoin and J.-F. Le Gall. Stochastic Flows Associated to Coalescent Processes. III. Limit Theorems. *Illinois J. Math.*, 50.1-4 (2006), pp. 147–181. doi: 10.1215/ijm/1258059473 (see pp. 60, 82).

[17] J. Bertoin and B. Mallein. Infinitely Ramified Point Measures and Branching Lévy Processes. *Ann. Probab. (to appear)*, (2018+). arXiv: 1703.08078 (see p. 126).

[18] A. Blancas, J.-J. Duchamps, A. Lambert, and A. Siri-Jégousse. Trees within Trees: Simple Nested Coalescents. *Electron. J. Probab.*, 23.0 (2018). doi: 10.1214/18-EJP219 (see pp. 11, 58, 133).

[19] A. Blancas, T. Rogers, J. Schweinsberg, and A. Siri-Jégousse. The Nested Kingman Coalescent: Speed of Coming down from Infinity. *Ann. Appl. Probab.*, 29.3 (June 2019), pp. 1808–1836. doi: 10.1214/18-AAP1440 (see pp. 59, 60, 90).

[20] E. Bolthausen and A.-S. Sznitman. On Ruelle's Probability Cascades and an Abstract Cavity Method. *Comm. Math. Phys.*, 197.2 (1998), pp. 247–276. doi: 10.1007/s002200050450 (see p. 83).

[21] É. Brunet and B. Derrida. Genealogies in Simple Models of Evolution. *J. Stat. Mech. Theory Exp.*, 2013.01 (Jan. 16, 2013), P01006. doi: 10.1088/1742-5468/2013/01/P01006 (see p. 60).

[22] N. Champagnat and A. Lambert. Splitting Trees with Neutral Poissonian Mutations I: Small Families. *Stochastic Process. Appl.*, 122.3 (2012), pp. 1003–1033. doi: 10.1016/j.spa.2011.11.002 (see pp. 15, 33, 34).

[23] N. Champagnat and A. Lambert. Splitting Trees with Neutral Poissonian Mutations II: Largest and Oldest Families. *Stochastic Process. Appl.*, 123.4 (2013), pp. 1368–1414. doi: 10.1016/j.spa.2012.11.013 (see p. 15).

[24] N. Champagnat, A. Lambert, and M. Richard. *Birth and Death Processes with Neutral Mutations*. 2012. url: https://www.hindawi.com/journals/ijsa/2012/569081/ (visited on 09/05/2017) (see p. 15).

[25] B. Chauvin. Product Martingales and Stopping Lines for Branching Brownian Motion. *Ann. Probab.*, 19.3 (July 1991), pp. 1195–1205. doi: 10.1214/aop/1176990340 (see p. 135).

[26] B. Chen, D. Ford, and M. Winkel. A New Family of Markov Branching Trees: The Alpha-Gamma Model. *Electron. J. Probab.*, 14 (2009), pp. 400–430. doi: 10.1214/EJP.v14-616 (see p. 90).

[27] H. Crane. Generalized Markov Branching Trees. *Adv. in Appl. Probab.*, 49.01 (Mar. 2017), pp. 108–133. doi: 10.1017/apr.2016.81 (see pp. 90, 91).

[28] H. Crane and H. Towsner. The Structure of Combinatorial Markov Processes (Mar. 18, 2016). arXiv: 1603.05954 [math.PR] (see p. 98).

[29] B. Dadoun. Asymptotics of Self-Similar Growth-Fragmentation Processes. *Electron. J. Probab.*, 22 (2017). doi: 10.1214/17-EJP45 (see p. 126).

[30] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Vol. II. Probability and Its Applications. New York, NY: Springer New York, 2008. doi: 10.1007/978-0-387-49835-5 (see p. 142).

[31] D. A. Dawson. Multilevel Mutation-Selection Systems and Set-Valued Duals. *J. Math. Biol.*, 76.1-2 (Jan. 2018), pp. 295–378. doi: 10.1007/s00285-017-1145-2 (see pp. 59, 60).

[32] J. H. Degnan and N. A. Rosenberg. Gene Tree Discordance, Phylogenetic Inference and the Multispecies Coalescent. *Trends Ecol. Evol.*, 24.6 (2009), pp. 332–340. doi: 10.1016/j.tree.2009.01.009 (see p. 59).

[33] C. Delaporte, G. Achaz, and A. Lambert. Mutational Pattern of a Sample from a Critical Branching Population. *J. Math. Biol.*, 73.3 (Sept. 1, 2016), pp. 627–664. doi: 10.1007/s00285-015-0964-2 (see p. 15).

[34] M. M. Desai, A. M. Walczak, and D. S. Fisher. Genetic Diversity and the Structure of Genealogies in Rapidly Adapting Populations. *Genetics*, 193.2 (2013), pp. 565–585. doi: 10.1534/genetics.112.147157 (see p. 60).

[35] J. J. Doyle. Trees within Trees: Genes and Species, Molecules and Morphology. *Syst. Biol.*, 46.3 (Sept. 1, 1997), pp. 537–553. doi: 10.1093/sysbio/46.3.537 (see pp. 59, 90).

[36] J.-J. Duchamps. Trees within Trees II: Nested Fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat. (to appear)*, (2019+). arXiv: 1807.05951 (see pp. 11, 89, 133, 153).

[37] J.-J. Duchamps and A. Lambert. Mutations on a Random Binary Tree with Measured Boundary. *Ann. Appl. Probab.*, 28.4 (Aug. 2018), pp. 2141–2187. doi: 10.1214/17-AAP1353 (see pp. 11, 14).

[38] R. Durrett and J. Schweinsberg. A Coalescent Model for the Effect of Advantageous Mutations on the Genealogy of a Population. *Stochastic Process. Appl.*, 115.10 (2005), pp. 1628–1657. doi: 10.1016/j.spa.2005.04.009 (see p. 60).

[39] B. Eldon and J. Wakeley. Coalescent Processes When the Distribution of Offspring Number among Individuals Is Highly Skewed. *Genetics*, 172.4 (Apr. 1, 2006), pp. 2621–2633. doi: 10.1534/genetics.105.052175 (see p. 60).

[40] A. Etheridge. *Some Mathematical Models from Population Genetics: École d'ete de Probabilités de Saint-Flour XXXIX-2009*. Lecture Notes in Mathematics 2012. Heidelberg ; New York: Springer, 2011 (see pp. 59, 90).

[41] S. N. Evans. *Probability and Real Trees*. Red. by J. .-.-M. Morel, F. Takens, and B. Teissier. Vol. 1920. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. doi: 10.1007/978-3-540-74798-7 (see p. 17).

[42] W. J. Ewens. The Sampling Theory of Selectively Neutral Alleles. *Theor. Popul. Biol.*, 3.1 (Mar. 1, 1972), pp. 87–112. doi: 10.1016/0040-5809(72)90035-4 (see pp. 9, 15).

[43] J. Felsenstein. *Inferring Phylogenies*. Vol. 2. Sinauer associates Sunderland, MA, 2004 (see p. 58).

[44] D. J. Ford. *Probabilities on Cladograms: Introduction to the Alpha Model*. Stanford University, 2006 (see p. 90).

[45] C. Foucart. Distinguished Exchangeable Coalescents and Generalized Fleming-Viot Processes with Immigration. *Adv. in Appl. Probab.*, 43.02 (June 2011), pp. 348–374. doi: 10.1239/aap/1308662483 (see p. 113).

[46] F. Foutel-Rodier, A. Lambert, and E. Schertzer. Exchangeable Coalescents, Ultrametric Spaces, Nested Interval-Partitions: A Unifying Approach (July 13, 2018). arXiv: 1807.05165 [math] (see p. 60).

[47] F. Freund and M. Möhle. On the Number of Allelic Types for Samples Taken from Exchangeable Coalescents with Mutation. *Adv. in Appl. Probab.*, 41.04 (Dec. 2009), pp. 1082–1101. doi: 10.1017/S000186780000375X (see p. 15).

[48] F. Freund. Almost Sure Asymptotics for the Number of Types for Simple Ξ-Coalescents. *Electron. Commun. Probab.*, 17 (2012). doi: 10.1214/ECP.v17-1704 (see p. 15).

[49] F. G. Ged. Profile of a Self-Similar Growth-Fragmentation. *Electron. J. Probab.*, 24 (2019). doi: 10.1214/18-EJP253 (see p. 126).

[50] B. T. Grenfell et al. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*, 303.5656 (2004), pp. 327–332. doi: 10.1126/science.1090727 (see p. 59).

[51] R. C. Griffiths and A. G. Pakes. An Infinite-Alleles Version of the Simple Branching Process. *Adv. in Appl. Probab.*, 20.3 (Sept. 1988), p. 489. DOI: 10.2307/1427033 (see p. 15).

[52] B. Haas and G. Miermont. The Genealogy of Self-Similar Fragmentations with Negative Index as a Continuum Random Tree. *Electron. J. Probab.*, 9 (2004), pp. 57–97. DOI: 10.1214/EJP.v9-187 (see p. 126).

[53] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum Tree Asymptotics of Discrete Fragmentations and Applications to Phylogenetic Models. *Ann. Probab.*, 36.5 (Sept. 2008), pp. 1790–1837. DOI: 10.1214/07-AOP377 (see pp. 10, 90, 91, 126).

[54] J. Heled and A. J. Drummond. Bayesian Inference of Species Trees from Multilocus Data. *Mol. Biol. Evol.*, 27.3 (2009), pp. 570–580. DOI: 10.1093/molbev/msp274 (see p. 59).

[55] O. Kallenberg. *Probabilistic Symmetries and Invariance Principles*. Probability and Its Applications. New York: Springer-Verlag, 2005 (see p. 74).

[56] D. G. Kendall. On the Generalized "Birth-and-Death" Process. *Ann. Math. Statist.*, 19.1 (Mar. 1948), pp. 1–15. DOI: 10.1214/aoms/1177730285 (see pp. 44, 45, 47).

[57] J. F. C. Kingman. The Representation of Partition Structures. *J. Lond. Math. Soc. (2)*, 18.2 (1978), pp. 374–380. DOI: 10.1112/jlms/s2-18.2.374 (see pp. 96, 130).

[58] J. Kingman. The Coalescent. *Stochastic Process. Appl.*, 13.3 (1982), pp. 235–248. DOI: 10.1016/0304-4149(82)90011-4 (see pp. 8, 15, 22, 59, 90, 95).

[59] N. Krell. Self-Similar Branching Markov Chains. *Séminaire de Probabilités XLII*. Vol. 1979. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 261–280. DOI: 10.1007/978-3-642-01763-6_10 (see pp. 126, 134).

[60] A. Lambert and G. Uribe Bravo. The Comb Representation of Compact Ultrametric Spaces. *p-Adic Numbers Ultrametric Anal. Appl.*, 9.1 (Jan. 2017), pp. 22–38. DOI: 10.1134/S2070046617010034 (see pp. 15, 19, 29).

[61] A. Lambert. Population Dynamics and Random Genealogies. *Stoch. Models*, 24 (sup1 2008), pp. 45–163. DOI: 10.1080/15326340802437728 (see pp. 8, 58, 90).

[62] A. Lambert. Probabilistic Models for the (Sub)Tree(s) of Life. *Braz. J. Probab. Stat.*, 31.3 (Aug. 2017), pp. 415–475. DOI: 10.1214/16-BJPS320 (see p. 90).

[63] A. Lambert. Random Ultrametric Trees and Applications. *ESAIM Proc. Surveys*, 60 (2017), pp. 70–89. DOI: 10.1051/proc/201760070 (see pp. 58, 60).

[64] A. Lambert. The Allelic Partition for Coalescent Point Processes. *Markov Process. Related Fields*, 15 (2009), pp. 359–386. arXiv: 0804.2572 (see pp. 15, 26, 33, 35).

[65] A. Lambert and E. Schertzer. Coagulation-Transport Equations and the Nested Coalescents. *Probab. Theory Related Fields*, (Apr. 15, 2019). DOI: 10.1007/s00440-019-00914-4 (see pp. 59, 60, 90).

[66] A. Lambert and E. Schertzer. Recovering the Brownian Coalescent Point Process from the Kingman Coalescent by Conditional Sampling. *Bernoulli*, 25.1 (Feb. 2019), pp. 148–173. DOI: 10.3150/17-BEJ971 (see p. 22).

[67] J. Lamperti. Semi-Stable Markov Processes. I. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 22.3 (1972), pp. 205–225. DOI: 10.1007/BF00536091 (see p. 134).

[68] W. P. Maddison. Gene Trees in Species Trees. *Syst. Biol.*, 46.3 (Sept. 1, 1997), pp. 523–536. DOI: 10.1093/sysbio/46.3.523 (see pp. 59, 90).

[69] P. Marchal. Nested Regenerative Sets and Their Associated Fragmentation Process. *Mathematics and Computer Science III*. Trends in Mathematics. Birkhäuser Basel, 2004, pp. 461–470. DOI: 10.1007/978-3-0348-7915-6_45 (see pp. 23, 29).

[70] S. Matuszewski, M. E. Hildebrandt, G. Achaz, and J. D. Jensen. Coalescent Processes with Skewed Offspring Distributions and Nonequilibrium Demography. *Genetics*, 208.1 (2017), pp. 323–338. DOI: 10.1534/genetics.117.300499 (see p. 60).

[71] R. A. Neher and O. Hallatschek. Genealogies of Rapidly Adapting Populations. *Proc. Natl. Acad. Sci. USA*, 110.2 (Jan. 8, 2013), pp. 437–442. DOI: 10.1073/pnas.1213113110 (see p. 60).

[72] M. Nei and S. Kumar. *Molecular Evolution and Phylogenetics*. Oxford university press, 2000 (see p. 58).

[73] R. D. Page and M. A. Charleston. From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem. *Mol. Phylogenet. Evol.*, 7.2 (Apr. 1997), pp. 231–240. DOI: 10.1006/mpev.1996.0390 (see pp. 59, 90).

[74] R. D. Page and M. A. Charleston. Trees within Trees: Phylogeny and Historical Associations. *Trends Ecol. Evol.*, 13.9 (Sept. 1998), pp. 356–359. DOI: 10.1016/S0169-5347(98)01438-4 (see pp. 59, 90).

[75] J. C. Pardo and V. Rivero. Self-Similar Markov Processes. *Bol. Soc. Mat. Mex. (3)*, 19.2 (2013), pp. 201–235 (see p. 134).

[76] J. Pitman. Coalescents with Multiple Collisions. *Ann. Probab.*, 27.4 (Oct. 1999), pp. 1870–1902. DOI: 10.1214/aop/1022874819 (see pp. 60, 63, 65, 83, 84, 90).

[77] L. Popovic. Asymptotic Genealogy of a Critical Branching Process. *Ann. Appl. Probab.*, 14.4 (Nov. 2004), pp. 2120–2148. DOI: 10.1214/105051604000000486 (see p. 29).

[78] N. A. Rosenberg. The Probability of Topological Concordance of Gene Trees and Species Trees. *Theor. Popul. Biol.*, 61.2 (2002), pp. 225–247. DOI: 10.1006/tpbi.2001.1568 (see p. 59).

[79] S. Sagitov. The General Coalescent with Asynchronous Mergers of Ancestral Lines. *J. Appl. Probab.*, 36.4 (Dec. 1999), pp. 1116–1125. doi: 10.1017/S0021900200017903 (see pp. 60, 65, 90).

[80] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics 68. Cambridge, U.K. ; New York: Cambridge University Press, 1999 (see p. 159).

[81] J. Schweinsberg. A Necessary and Sufficient Condition for the Λ-Coalescent to Come Down from Infinity. *Electron. Commun. Probab.*, 5 (2000), pp. 1–11. doi: 10.1214/ECP.v5-1013 (see p. 82).

[82] J. Schweinsberg. Coalescent Processes Obtained from Supercritical Galton–Watson Processes. *Stochastic Process. Appl.*, 106.1 (July 2003), pp. 107–139. doi: 10.1016/S0304-4149(03)00028-0 (see p. 60).

[83] J. Schweinsberg. Coalescents with Simultaneous Multiple Collisions. *Electron. J. Probab.*, 5 (2000). doi: 10.1214/EJP.v5-68 (see p. 60).

[84] J. Schweinsberg. Rigorous Results for a Population Model with Selection II: Genealogy of the Population. *Electron. J. Probab.*, 22 (2017). doi: 10.1214/17-EJP58 (see p. 60).

[85] C. Semple and M. Steel. *Phylogenetics*. Oxford Lecture Series in Mathematics and Its Applications 24. Oxford ; New York: Oxford University Press, 2003 (see pp. 58, 90).

[86] Q. Shi. Growth-Fragmentation Processes and Bifurcators. *Electron. J. Probab.*, 22 (2017). doi: 10.1214/17-EJP26 (see p. 126).

[87] S. M. Srivastava. *A Course on Borel Sets*. Vol. 180. Graduate Texts in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. doi: 10.1007/978-3-642-85473-6 (see pp. 109, 129).

[88] G. J. Szöllősi, E. Tannier, V. Daubin, and B. Boussau. The Inference of Gene Trees with Species Trees. *Syst. Biol.*, 64.1 (2014), e42–e62. doi: 10.1093/sysbio/syu048 (see p. 59).

[89] Z. Taïb. *Branching Processes and Neutral Evolution*. Red. by S. A. Levin. Vol. 93. Lecture Notes in Biomathematics. Springer Berlin Heidelberg, 1992. doi: 10.1007/978-3-642-51536-1 (see p. 15).

[90] A. Tellier and C. Lemaire. Coalescence 2.0: A Multiple Branching of Recent Theoretical Developments and Their Applications. *Mol. Ecol.*, 23.11 (2014), pp. 2637–2652. doi: 10.1111/mec.12755 (see p. 60).

[91] E. M. Volz, K. Koelle, and T. Bedford. Viral Phylodynamics. *PLoS Comput. Biol.*, 9.3 (2013), e1002947. doi: 10.1371/journal.pcbi.1002947 (see p. 59).