COALESCENTS WITH MULTIPLE COLLISIONS

Jim Pitman*

Technical Report No. 495

Department of Statistics, University of California
367 Evans Hall # 3860, Berkeley, CA 94720-3860

Revised March 23, 1999

Abstract

For each finite measure Λ on $[0,1]$, a coalescent Markov process, with state space the compact set of all partitions of the set $\mathbb{N}$ of positive integers, is constructed so the restriction of the partition to each finite subset of $\mathbb{N}$ is a Markov chain with the following transition rates: when the partition has b blocks, each k-tuple of blocks is merging to form a single block at rate $\int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx)$. Call this process a Λ-*coalescent*. Discrete measure valued processes derived from the Λ-coalescent model a system of masses undergoing coalescent collisions. Kingman's coalescent, which has numerous applications in population genetics, is the δ_0-coalescent for δ_0 a unit mass at 0. The coalescent recently derived by Bolthausen and Sznitman from Ruelle's probability cascades, in the context of the Sherrington-Kirkpatrick spin glass model in mathematical physics, is the U-coalescent for U uniform on $[0,1]$. For $\Lambda = U$, and whenever an infinite number of masses are present, each collision in a Λ-coalescent involves an infinite number of masses almost surely, and the proportion of masses involved exists as a limit almost surely and is distributed proportionally to Λ. The two-parameter Poisson-Dirichlet family of random discrete distributions derived from a stable subordinator, and corresponding exchangeable random partitions of $\mathbb{N}$ governed by a generalization of the Ewens sampling formula, are applied to describe transition mechanisms for processes of coalescence and fragmentation, including the U-coalescent and its time reversal.

*Research supported in part by N.S.F. Grant DMS97-03961

1 Introduction

Markovian coalescent models for the evolution of a system of masses by a random process of binary collisions were introduced by Marcus [29] and Lushnikov [28]. See Aldous [3] for a recent survey of the scientific literature of these models and their relation to Smoluchowski's mean-field theory of coagulation phenomena. Evans and Pitman [15] gave a general framework for the rigorous construction of partition valued and discrete measure valued coalescent Markov processes allowing infinitely many massses, and treated the binary coalescent model where each pair of masses x and y is subject to a coalescent collision at rate $\kappa(x, y)$ for a suitable rate kernel κ. This paper studies a family of partition valued Markov processes, with state space the compact set of all partitions of $\mathbb{N} := \{1, 2, \ldots\}$, such that the restriction of the partition to each finite subset of $\mathbb{N}$ is a Markov chain with transition rates of a simple form determined by the moments of a finite measure Λ on the unit interval. The case $\Lambda = \delta_0$, a unit mass at 0, is Kingman's coalescent in which every pair of blocks coalesces at rate 1. The case $\Lambda = U$, the uniform distribution on $[0, 1]$ yields the coalescent derived by Bolthausen-Sznitman [9] from Ruelle's probability cascades [39]. See also [8] for another derivation of this coalescent from the genealogy of a continuous-state branching process.

The rest of this paper is organized as follows. Section 2 describes the main results, with pointers to following sections for details. Section 2.1 gives some results for the *partition valued* Λ-*coalescent* for general Λ. Section 2.2 describes an associated discrete measure valued process, the *ranked mass* Λ-*coalescent*. Section 2.3 presents a theorem which shows how certain operations of coagulation and fragmentation act on the two-parameter family of distributions of exchangeable random partitions of $\mathbb{N}$ introduced in [31] and studied further in [33, 34]. Section 2.4 applies this theorem to the U-coalescent to recover some of the results of Bolthausen-Sznitman and to obtain various further developments. The conceptual framework of the paper is provided by Kingman's theory of exchangeable random partitions of $\mathbb{N}$, as reviewed in Appendix A.

2 Summary of Results

For $n \in \mathbb{N} := \{1, 2, \ldots\}$ let $\mathcal{P}_n$ be the finite set of all partitions of the set $[n] := \{1, \ldots, n\}$. Let $\mathcal{P}_\infty$ be the set of all partitions of $\mathbb{N}$. Each $\pi \in \mathcal{P}_\infty$ is identified with the sequence $(\pi_1, \pi_2, \ldots) \in \mathcal{P}_1 \times \mathcal{P}_2 \times \cdots$

where π_n is the restriction of π to $[n]$. Give $\mathcal{P}_\infty$ the topology it inherits as a subset of $\mathcal{P}_1 \times \mathcal{P}_2 \times \cdots$ with the product of discrete topologies. So $\mathcal{P}_\infty$ is compact and metrizable. Following [25, 15], call a $\mathcal{P}_\infty$-valued stochastic process $\Pi_\infty := (\Pi_\infty(t), t \geq 0)$ a *coalescent* if Π_∞ has càdlàg paths and $\Pi_\infty(s)$ a refinement of $\Pi_\infty(t)$ for every $s < t$. That is to say, for each n the restriction $\Pi_n := (\Pi_n(t), t \geq 0)$ of Π_∞ to $[n]$ is a process with right-continuous step function paths such that $\Pi_n(s)$ a refinement of $\Pi_n(t)$ for every $s < t$. The following result is established in Section 3.1:

Theorem 1 *Let $(\lambda_{b,k}, 2 \leq k \leq b < \infty)$ be an array of non-negative real numbers. There exists for each $\pi \in \mathcal{P}_\infty$ a $\mathcal{P}_\infty$-valued coalescent Π_∞ with $\Pi_\infty(0) = \pi$, whose restriction Π_n to $[n]$ is for each n a Markov chain such that when $\Pi_n(t)$ has b blocks each k-tuple of blocks of $\Pi_n(t)$ is merging to form a single block at rate $\lambda_{b,k}$, if and only if*

$$\lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx) \tag{1}$$

for some non-negative and finite measure Λ on the Borel subsets of $[0,1]$. For $(\lambda_{b,k})$ so derived from Λ and $\pi \in \mathcal{P}_\infty$ let $\mathbb{P}^{\Lambda,\pi}$ denote the probability distribution governing Π_∞ with $\Pi_\infty(0) = \pi$ on the space of càdlàg $\mathcal{P}_\infty$-valued paths with the Skorohod topology. Then the collection of laws $(\mathbb{P}^{\Lambda,\pi}, \pi \in \mathcal{P}_\infty)$ defines a strong Markov process with state space $\mathcal{P}_\infty$ and Feller semigroup. Moreover the map $(\Lambda, \pi) \mapsto \mathbb{P}^{\Lambda,\pi}$ is continuous when the spaces of measures are given their weak topologies.

Definition 2 Call this $\mathcal{P}_\infty$-valued Markov process induced by a finite measure Λ on $[0,1]$ the *Λ-coalescent*. Let 1^∞ denote the partition of $\mathbb{N}$ into singletons. Call a Λ-coalescent started in state 1^∞ a *standard Λ-coalescent.*

For $\Lambda = \delta_0$, the transition rates are $\lambda_{b,k} = 1(k=2)$. So the δ_0-coalescent is Kingman's coalescent [25, 27] in which each pair of blocks coalesces at rate 1, and no multiple collisions are allowed. For $r, s > 0$ and $\Lambda = \text{beta}(r,s)$, the probability distribution on $(0,1)$ with density $B(r,s)^{-1}x^{r-1}(1-x)^{s-1}$ at $x \in (0,1)$ where $B(r,s) = \Gamma(r)\Gamma(s)/\Gamma(r+s)$, the rates are $\lambda_{b,k} = B(r+k-2, s+b-k)/B(r,s)$. In particular, if $U = \text{beta}(1,1)$ is the uniform distribution on $(0,1)$, then

$$\lambda_{b,k} = \frac{(k-2)!(b-k)!}{(b-1)!} = \left[(b-1)\binom{b-2}{k-2}\right]^{-1}. \tag{2}$$

These rates identify the U-coalescent with the coalescent studied by Bolthausen-Sznitman [9].

Provided Λ has no mass at 0, it is easily checked that a family of chains Π_n with the transition rates (1) can be constructed as follows from the countable collection of points of a Poisson point process N on $(0,\infty)\times\{0,1\}^\infty$ with intensity $dt\, L(d\xi)$ where

$$L := \int_{(0,1]} x^{-2}\Lambda(dx)P_x \tag{3}$$

with P_x governing $\xi := (\xi_1, \xi_2, \ldots)$ as a sequence of independent Bernoulli trials with $P_x(\xi_i = 1) = x$ for all i. Given an arbitrary partition π of $\mathbb{N}$, let $\Pi_n(0)$ be the restriction of π to $[n]$, and let the process Π_n be allowed the possibility of jumping only at the times t of points (t,ξ) of N such that $\sum_{i=1}^n \xi_i \geq 2$. For such t, if $\Pi_n(t-) = \{A_1, \ldots, A_b\}$ say, where the A_i are in the order of their least elements, let $\Pi_n(t)$ be derived from $\Pi_n(t-)$ by merging those A_i with $\xi_i = 1$. This will result in a transition of Π_n at time t if and only if $\sum_{i=1}^b \xi_i \geq 2$. It follows immediately from the definition (3) of L that Π_n is Markovian with the desired transition rates (1). By construction, Π_n is the restriction to $[n]$ of Π_{n+1} for every n. The Poisson point process N therefore determines a unique $\mathcal{P}_\infty$-valued coalescent process Π_∞ whose restriction to $[n]$ is Π_n for every n. To summarize:

Corollary 3 *Provided Λ has no mass at 0, the above construction of consistent coalescent chains Π_n from a Poisson point process on $(0,\infty)\times\{0,1\}^\infty$ with intensity $dt\, L(d\xi)$ for L in (3) yields a Λ-coalescent process Π_∞.*

In particular, for $\Lambda(dx) = dx$ Corollary 3 gives a new construction of the U-coalescent.

2.1 Some results for general Λ.

Throughout the paper, the notation

$$\mu_r := \int_0^1 x^r \Lambda(dx)$$

is used for the rth moment of the finite measure Λ on $[0,1]$ for arbitrary real r. Note that μ_r is a decreasing function of r with $\infty > \mu_0 \geq \mu_r \geq 0$ for $r \geq 0$, while μ_r may be either finite or $+\infty$ for $r < 0$. For $r = 0, 1, \ldots$ observe from (1) that $\mu_r = \lambda_{r+2,r+2}$ is the rate at which Π_n is jumping to its absorbing state $\{[n]\}$ from any state with $r+2$ blocks. To avoid trivialities, assume from now on that $\mu_0 > 0$. Let F denote a generic probability measure on $[0,1]$, and take $\Lambda = \mu_0 F$. By rescaling the time parameter, there is no loss of generality in supposing $\mu_0 = 1$. So when convenient, results may be presented just for an F-coalescent. Let X denote a random variable with distribution F, defined on some background probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with expectation operator $\mathbb{E}$, so $\mathbb{E}(X^r) = \mu_r/\mu_0$. From (1), the transition rates of the Λ-coalescent are

$$\lambda_{b,k} := \mu_0 \, \mathbb{E}(X^{k-2}(1-X)^{b-k}) \text{ for all } 2 \leq k \leq b. \tag{4}$$

Let Π_∞ be a Λ-coalescent started at π. For $i, j \in \mathbb{N}$ with i and j in different blocks of π, let $\tau_{i,j}$ denote the *collision time of i and j*, meaning the unique time t such that i and j belong to the same block of $\Pi_\infty(t)$ but different blocks of $\Pi_\infty(t-)$. By the exchangeability property of the Λ-coalescent described in Section 3.2, the random time $\tau_{i,j}$ has the same exponential distribution with rate $\lambda_{2,2} = \mu_0$ for all such i, j. Write $\#\pi$ for the number of blocks of a partition π.

Theorem 4 *For an F-coalescent Π_∞ started with i and j in distinct blocks of $\Pi_\infty(0)$, and $\tau_{i,j}$ the collision time of i and j, if the event $(\#\Pi_\infty(\tau_{i,j}-) = \infty)$ has strictly positive probability, then given this event a random variable $X_{i,j}$ with distribution F is recovered as the almost sure relative frequency of blocks of $\Pi_\infty(\tau_{i,j}-)$ which merge at time $\tau_{i,j}$ to form the block containing both i and j.*

According to Proposition 23, provided F has no atom at 1, in a standard Λ-coalescent the probability of the event $(\#\Pi_\infty(\tau_{i,j}-) = \infty)$ is either 0 for all i, j or 1 for all i, j. In particular, it will be seen that for a standard U-coalescent this event has probability one for all i, j. So at each collision time $\tau_{i,j}$ in a standard U-coalescent, the relative frequency of blocks involved has the uniform distribution U.

For any initial partition with a finite number of blocks $b \geq 2$, the total rate of transitions of all kinds in a Λ-coalescent can be variously expressed as

$$\lambda_b := \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k} = \sum_{i=0}^{b-2} (-1)^i (i+1) \binom{b}{i+2} \mu_i \tag{5}$$

$$= \mu_0 \, \mathbb{E}\left[\frac{1-(1-X)^b - bX(1-X)^{b-1}}{X^2}\right] \tag{6}$$

where the ratio is interpreted by continuity to equal $\binom{b}{2}$ if $X = 0$. From (6),

$$\lambda_b \uparrow \mu_{-2} := \int_0^1 x^{-2} \Lambda(dx) \text{ as } b \uparrow \infty \, . \tag{7}$$

It follows that the holding time of the initial state 1^∞ of the standard Λ-coalescent has an exponential distribution with rate μ_{-2}, and that the Λ-coalescent is a Markov process of jump-hold type with bounded transition rates and step-function paths if and only if $\mu_{-2} < \infty$. Example 19 describes more explicitly the simple transition mechanism of the Λ-coalescent when $\mu_{-2} < \infty$.

It was observed by Kingman for $\Lambda = \delta_0$, and is true also for general Λ, that in a standard Λ-coalescent the partition $\Pi_\infty(t)$ is for each t an *exchangeable random partition of* $\mathbb{N}$. That is, for each particular partition $\{B_1, \ldots, B_k\}$ of $[n]$ into k blocks, the probability that $\Pi_n(t) = \{B_1, \ldots, B_k\}$ is a symmetric function of the sizes $n_1, \ldots, n_k$ of the blocks $B_1, \ldots, B_k$, say

$$\mathbb{P}^{\Lambda, 1^\infty}(\Pi_n(t) = \{B_1, \ldots, B_k\}) =: p_t^\Lambda(n_1, \ldots, n_k). \tag{8}$$

For each fixed t and Λ, this function p_t^Λ of finite sequences of positive integers $(n_1, \ldots, n_k)$ is the *exchangeable probability function (EPF)* associated with the $\mathbb{P}^{\Lambda, 1^\infty}$ distribution of $\Pi_\infty(t)$ on $\mathcal{P}_\infty$. This probability distribution on $\mathcal{P}_\infty$ may also be denoted p_t^Λ. Appendix A reviews the basic properties the EPF determining the distribution on $\mathcal{P}_\infty$ of an exchangeable random partition of $\mathbb{N}$. For fixed Λ and n, the EPF $p_t^\Lambda(n_1, \ldots, n_k)$ is determined for all $(n_1, \ldots, n_k)$ with $\sum_{i=1}^k n_i = n$ and all $t \geq 0$ by the $n-1$ moments $\mu_0, \mu_1, \ldots, \mu_{n-2}$ of Λ. For these moments determine the transition rates of the finite state chain $(\Pi_n(t), t \geq 0)$, and these rates in turn determine $p_t(n_1, \ldots, n_k)$ for all such $(n_1, \ldots, n_k)$ and all $t \geq 0$ via Kolmogorov's differential equations. Section 3.8 gives some more explicit expressions.

Definition 5 For a partition π of $[n]$, where $n \in \mathbb{N} \cup \{\infty\}$ and $[\infty] := \mathbb{N}$, write $\pi = \{A_1, A_2, \ldots\}$ to indicate that the blocks of π in increasing order of their least elements are $A_1, A_2, \ldots$, with the convention $A_i = \emptyset$ for $i > \#\pi$. For a partition $\pi = \{A_1, A_2, \ldots\}$ of $\mathbb{N}$ and a partition $\Pi := \{B_1, B_2, \ldots\}$ of $[n]$ with $n \geq \#\pi$ let the Π-*coagulation of* π be the partition of $\mathbb{N}$ whose blocks are the non-empty sets of the form $\cup_{j \in B_i} A_j$ for some $i = 1, 2, \ldots$. For each probability distribution p on $\mathcal{P}_\infty$, define a Markov kernel p-COAG on $\mathcal{P}_\infty$, the *p-coagulation kernel*, as follows: for $\pi \in \mathcal{P}_\infty$ let p-COAG$(\pi, \cdot)$ be the distribution of the Π-coagulation of π for Π with distribution p.

Think of Π as describing a coagulation of singleton subsets into the blocks $B_1, B_2, \ldots$. Then the Π-coagulation of π describes a corresponding coagulation of blocks of π.

Let Π^π_∞ be a $\mathcal{P}_\infty$-valued coalescent process with $\Pi^\pi_\infty(0) = \pi$ for some π with $\#\pi = n \in \mathbb{N} \cup \{\infty\}$. Then it is easily seen that

$$\Pi^\pi_\infty(t) = \text{the } \Pi_n(t)\text{-coagulation of } \pi \text{ for } t \geq 0 \tag{9}$$

for some uniquely defined $\mathcal{P}_n$-valued coalescent process Π_n with initial state 1^n, the partition of $[n]$ into singletons.

Theorem 6 *A coalescent process Π^π_∞ starting at π with $\#\pi = n$ for some $1 \leq n \leq \infty$ is a Λ-coalescent if and only of Π_n defined by* (9) *is distributed as the restriction to $[n]$ of a standard Λ-coalescent. The semigroup of the Λ-coalescent on $\mathcal{P}_\infty$ is thus given by*

$$\mathbb{P}^{\Lambda,\pi}(\Pi_\infty(t) \in \cdot\,) = p^\Lambda_t\text{-COAG}(\pi, \cdot\,) \tag{10}$$

where $p^\Lambda_t(\,\cdot\,) := \mathbb{P}^{\Lambda,1^\infty}(\Pi_\infty(t) \in \cdot\,)$ is the distribution of an exchangeable random partition of $\mathbb{N}$ with the EPF $p^\Lambda_t(n_1, \ldots, n_k)$ which is uniquely determined by Kolmogorov equations for the finite state chains Π_n for $n = 2, 3, \ldots$.

For a standard Λ-coalescent Π_∞ let $\Pi_\infty(t) = \{B_1(t), B_2(t), \ldots\}$. By Kingman's theory of exchangeable random partitions, each block $B_j(t)$ has an almost sure limiting frequency

$$\tilde{f}_j(t) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1(i \in B_j(t)) \tag{11}$$

with $0 \leq \tilde{f}_j(t) \leq 1$ and $\sum_j \tilde{f}_j(t) \leq 1$ almost surely for each t. Define $\mathbf{f}(t) := (f_1(t), f_2(t), \ldots)$ to be the ranked rearrangement of $(\tilde{f}_1(t), \tilde{f}_2(t), \ldots)$, and let P^Λ_t denote the probability distribution of $\mathbf{f}(t)$ on the set $\bar{\mathcal{S}}^\downarrow$ of all non-negative sequences $\boldsymbol{x} = (x_1, x_2, \ldots)$ with $\sum_i x_i \leq 1$ which are *ranked*, meaning $x_1 \geq x_2 \geq \ldots \geq 0$. According to *Kingman's correspondence* $p \leftrightarrow P$ between distributions p of exchangeable random partitions of $\mathbb{N}$ and probability measures P on $\bar{\mathcal{S}}^\downarrow$ (Theorem 36), the distribution p^Λ_t of $\Pi_\infty(t)$ and the distribution P^Λ_t of $\mathbf{f}(t)$ determine each other uniquely: $p^\Lambda_t \leftrightarrow P^\Lambda_t$. It appears that for general Λ there is neither a simple formula for the EPF p^Λ_t, nor any simple description of the corresponding distribution P^Λ_t on $\bar{\mathcal{S}}^\downarrow$. So Theorem 6 is a rather soft generalization of results of Kingman for $\Lambda = \delta_0$ and of Bolthausen-Sznitman for $\Lambda = U$ (recalled in Theorem 14 below) which give explicit descriptions of both p^Λ_t and P^Λ_t for these Λ.

2.2 The ranked mass Λ-coalescent

As in [15, 27, 9], the $\mathcal{P}_\infty$-valued Λ-coalescent can be used to build various discrete measure valued coalescent processes. Section 3.7 gives some results for the *ranked mass Λ-coalescent* $(\boldsymbol{X}(t), t \geq 0)$ with state space $\mathcal{S}^\downarrow := \{\boldsymbol{x} \in \bar{\mathcal{S}}^\downarrow : \sum_i x_i = 1\}$, the set of ranked probability distributions on $\mathbb{N}$, with the topology it inherits as a subset of ℓ_1. In this process, masses labeled by $\mathbb{N}$ collide by the mechanism of the standard Λ-coalescent applied to their labels. The state $\boldsymbol{X}(t) := (X_1(t), X_2(t), \ldots)$ of the process at time t is the ranked rearrangement of the masses. The existence of this process is made precise by the following corollary, which follows from Theorems 1 and 6 by the well known criterion of Dynkin [13, Theorem 10.13] for a function of a Markov process to be Markov. See also [15, §5] where the same construction is applied to other $\mathcal{P}_\infty$-valued coalescents. Variations of the corollary yield corresponding *basic* and *shunted* coalescents with similar regularity properties, as treated in [15].

Corollary 7 *Let Π_∞ be a standard Λ-coalescent. For $\boldsymbol{x} \in \mathcal{S}^\downarrow$ and $\pi \in \mathcal{P}_\infty$ let $(\boldsymbol{x}, \pi) \in \mathcal{S}^\downarrow$ be the decreasing rearrangement of the $\boldsymbol{x}$-masses of the blocks of π. For each $\boldsymbol{x} \in \mathcal{S}^\downarrow$ the process $((\boldsymbol{x}, \Pi_\infty(t)), t \geq 0)$ is an $\mathcal{S}^\downarrow$-valued process with càdlàg paths. Let $\mathbb{Q}^{\Lambda, \boldsymbol{x}}$ be the law of this process on the space of càdlàg $\mathcal{S}^\downarrow$-valued paths with the Skorohod topology. Then $(\mathbb{Q}^{\Lambda, \boldsymbol{x}}, \boldsymbol{x} \in \mathcal{S}^\downarrow)$ is for each Λ the collection of laws of a strong Markov process $\boldsymbol{X}$ with state space $\mathcal{S}^\downarrow$.*

Consider now the sequence of ranked frequencies $\mathbf{f}(t)$ derived from the standard Λ-coalescent at time t. Call the frequencies *proper* if $\sum_i f_i(t) = 1$ almost surely. Recall that $\mu_{-1} := \int_0^1 x^{-1} \Lambda(dx)$.

Theorem 8 *If $\mu_{-1} = \infty$ then the standard Λ-coalescent has proper frequencies almost surely for each $t > 0$; the $\mathcal{S}^\downarrow$-valued process $(\mathbf{f}(t), t > 0)$ defined by these ranked frequencies then has the unique distribution of a càdlàg $\mathcal{S}^\downarrow$-valued process $\boldsymbol{Y} := (\boldsymbol{Y}(t), t > 0)$ governed by the semigroup of the ranked mass Λ-coalescent and such that $Y_1(0+) = 0$ almost surely. Whereas, if $\mu_{-1} < \infty$ then the standard Λ-coalescent does not have proper frequencies almost surely for each $t > 0$, and there exists no such process $\boldsymbol{Y}$.*

Put another way, if $\mu_{-1} = \infty$, the family of distributions $(P_t^\Lambda, t > 0)$ on $\mathcal{S}^\downarrow$ derived from the standard Λ-coalescent Π_∞ defines an *entrance law* for the ranked mass Λ-coalescent semigroup

$(Q_t^\Lambda, t \geq 0)$. That is, for $s, t > 0$ there is the identity $P_s^\Lambda Q_t^\Lambda = P_{s+t}^\Lambda$, where P_t^Λ is the distribution of ranked frequencies of $\Pi_\infty(t)$, and $Q_t^\Lambda(\boldsymbol{x}, \cdot)$ is the distribution of $(\boldsymbol{x}, \Pi_\infty(t))$. The ranked frequencies of the standard Λ-coalescent then define a process with this entrance law which comes in from "dust" at time $0+$, meaning that the largest mass vanishes almost surely as $t \downarrow 0$. The construction of this entrance law was indicated by Kingman [27, §8] for $\Lambda = \delta_0$ and Bolthausen-Sznitman [9, Corollary 2.3] for $\Lambda = U$. The uniqueness property of this entrance law for $\Lambda = \delta_0$ was shown by Aldous [3, A.5]. See Section 3.7 for a description of the process of ranked frequencies $(\mathbf{f}(t), t \geq 0)$ in the improper case $\mu_{-1} < \infty$. Theorem 27 characterizes the entrance boundary of the ranked mass Λ-coalescent in both the proper and improper cases. See also [1, 4, 5, 15] regarding similar entrance laws and the entrance boundary for some particular binary coalescents with time parameter set $(-\infty, \infty)$ instead of $(0, \infty)$.

2.3 The two-parameter family

The following two lemmas recall some known results regarding a two-parameter family of probability distributions of exchangeable random partitions of $\mathbb{N}$. These results turn out to be the basis of various descriptions of the U-coalescent.

Lemma 9 [31, 32] *There exists an exchangeable random partition of* $\mathbb{N}$ *whose block frequencies* $\tilde{f}_n$ *in order of least elements are strictly positive with* $\sum_n \tilde{f}_n = 1$ *almost surely and such that*

$$\tilde{f}_1 = \tilde{Y}_1, \quad \tilde{f}_n = (1 - \tilde{Y}_1) \cdots (1 - \tilde{Y}_{n-1}) \tilde{Y}_n \qquad (n \geq 2) \tag{12}$$

for a sequence of independent random variables $(\tilde{Y}_n)$, *if and only if*

$$\tilde{Y}_n \text{ has beta}(1-\alpha, \theta + n\alpha) \text{ distribution for } n = 1, 2, \ldots \tag{13}$$

for some (α, θ) *with*

$$0 \leq \alpha < 1 \text{ and } \theta > -\alpha; \tag{14}$$

the corresponding EPF is

$$p_{\alpha,\theta}(n_1, \ldots, n_k) := \frac{[\theta/\alpha]_k}{[\theta]_n} \prod_{i=1}^k -[-\alpha]_{n_i} \tag{15}$$

where $[x]_n := \prod_{i=1}^n (x + i - 1)$.

Note that formula (15) contains some factors of θ and α which should be cancelled before evaluation if either $\alpha = 0$ or $\theta = 0$. For (α, θ) subject to (14) call an exchangeable random partition of $\mathbb{N}$ characterized by the EPF (15), or by frequencies of the form (12)-(13), an (α, θ) *partition.* It was shown in [31] how to construct an (α, θ) partition by a simple urn scheme. Following [37], define the *Poisson-Dirichlet distribution with parameters* (α, θ), abbreviated $PD(\alpha, \theta)$, to be the distribution of ranked frequencies of an (α, θ) partition. That is, $PD(\alpha, \theta)$ is the distribution on $\mathcal{S}^{\downarrow}$ obtained after ranking $\tilde{\mathbf{f}}$ generated by (12)-(13).

Lemma 10 *Define* $\boldsymbol{V} := (V_n) \in \mathcal{S}^{\downarrow}$ *by* $V_n := X_n/\xi_1$ *where* $\xi_1 := \sum_m X_m$ *and the* X_n *are the ranked points of a Poisson point process with intensity* $\nu(dx)$ *on* $(0, \infty)$*, as obtained by ranking the jumps of a subordinator* $(\xi_t, 0 \le t \le 1)$*, that is an increasing process with stationary independent increments, with*

$$\mathbb{E}\exp(-\lambda\xi_s) = \exp\left(-s\int_0^\infty (1 - e^{-\lambda x})\nu(dx)\right) \quad \textit{for } \lambda \ge 0.$$

(i)[22, 23] *If* $\nu(dx) = \theta x^{-1}e^{-x}dx$ *for* $\theta > 0$*, corresponding to* ξ_1 *with the gamma*(θ) *distribution* $\mathbb{P}(\xi_1 \in dx) = \Gamma(\theta)^{-1}x^{\theta-1}e^{-x}dx$*, then* $\boldsymbol{V}$ *has* $PD(0, \theta)$ *distribution.*
(ii)[30] *If* $\nu(dx) = cx^{-\alpha-1}dx$ *for* $\alpha \in (0,1)$ *and* $c > 0$*, corresponding to* ξ_1 *with a stable distribution of index* α*, then* $\boldsymbol{V}$ *has* $PD(\alpha, 0)$ *distribution.*

The $PD(0, \theta)$ distribution has well known applications in population genetics, number theory, and combinatorics, as reviewed in [17, 6]. Formula (15) in this case is a variation due to Kingman [26] of the *Ewens sampling formula* [16], [20, Ch. 41]. See [30, 37] for interpretations of $PD(\alpha, 0)$ in terms of excursions of a Markov process such as a Brownian motion or a recurrent Bessel process whose zero set is the closed range of a stable subordinator of index α, and [39, 10, 11, 9] for applications of $PD(\alpha, 0)$ in mathematical physics.

Definition 11 For each probability measure p on $\mathcal{P}_\infty$, define a Markov kernel p-FRAG on $\mathcal{P}_\infty$, the *p-fragmentation kernel* as follows. Let $p\text{-FRAG}(\pi, \cdot)$ be the distribution of a random refinement of π whose restriction to the mth block of π is the restriction of $\Pi^{(m)}$ to that block, where the $(\Pi^{(m)}, m = 1, 2, \ldots)$ are independent random partitions of $\mathbb{N}$ with distribution p.

For $p = p_{\alpha,\theta}$, the distribution of an (α, θ)-partition, the notations (α, θ)-COAG and (α, θ)-FRAG will be used instead of $p_{\alpha,\theta}$-COAG and $p_{\alpha,\theta}$-FRAG. Say that Π' is an (α, θ)-*coagulation of* Π if

$\mathbb{P}(\Pi' \in \cdot \mid \Pi = \pi) = (\alpha, \theta)\text{-COAG}(\pi, \cdot)$ and an *(α, θ)-fragmentation of* Π if $P(\Pi' \in \cdot \mid \Pi = \pi) = (\alpha, \theta)\text{-FRAG}(\pi, \cdot)$. The following theorem is proved in Section 4:

Theorem 12 *For all $0 < \alpha < 1, 0 \leq \beta < 1, \theta > -\alpha\beta$, the following two conditions are equivalent:*
(i) *Π is an (α, θ) partition and Π' is a $(\beta, \theta/\alpha)$-coagulation of Π;*
(ii) *Π' is an $(\alpha\beta, \theta)$ partition and Π is an $(\alpha, -\alpha\beta)$-fragmentation of Π'.*

For each allowed choice of α, β and θ, these equivalent conditions describe a particular joint distribution of a pair (Π, Π') of exchangeable random partitions of $\mathbb{N}$ such that Π is a refinement of Π'. Kingman's correspondence yields parallel descriptions of a joint distribution of a pair $(\boldsymbol{V}, \boldsymbol{V}')$ of random elements of $\mathcal{S}^{\downarrow}$. Recall that $(\boldsymbol{x}, \pi)$ is the ranked rearrangement of the partial sums of $\boldsymbol{x}$ over blocks of π.

Corollary 13 *For all $0 < \alpha < 1, 0 \leq \beta < 1, \theta > -\alpha\beta$, the following two conditions are equivalent:*
(i) *$\boldsymbol{V}$ has $PD(\alpha, \theta)$ distribution and $\boldsymbol{V}' = (\boldsymbol{V}, \Pi'')$ for Π'' a $(\beta, \theta/\alpha)$-partition independent of $\boldsymbol{V}$.*
(ii) *$\boldsymbol{V}'$ has $PD(\alpha\beta, \theta)$ distribution and $\boldsymbol{V}$ is the ranked rearrangement of the collection of products $\{V'_m W_{m,n}, m, n \in \mathbb{N}\}$, where for each m the sequence $\boldsymbol{W}_m := (W_{m,n}, n \in \mathbb{N})$ has $PD(\alpha, -\alpha\beta)$ distribution, and the sequences $\boldsymbol{V}'$ and $\boldsymbol{W}_m, m = 1, 2, \ldots$ are independent.*

To specify the conditional law of $\boldsymbol{V}'$ given $\boldsymbol{V}$ as in (i), say *$\boldsymbol{V}'$ is a $(\beta, \theta/\alpha)$-coagulation of $\boldsymbol{V}$*, and to specify the conditional law of $\boldsymbol{V}$ given $\boldsymbol{V}'$ as in (ii), say *$\boldsymbol{V}$ is an $(\alpha, -\alpha\beta)$-fragmentation of $\boldsymbol{V}'$*. As part of the implication (ii) $\Rightarrow$ (i) in the previous corollary,

> *if $\boldsymbol{V}'$ has $PD(\alpha\beta, \theta)$ distribution and $\boldsymbol{V}$ is an $(\alpha, -\alpha\beta)$-fragmentation of $\boldsymbol{V}'$, then $\boldsymbol{V}$ has $PD(\alpha, \theta)$ distribution.*

For $\beta = 0$ this construction of $PD(\alpha, \theta)$ for $0 < \alpha < 1$ and $\theta > 0$ from the more elementary $PD(0, \theta)$ and $PD(\alpha, 0)$ distributions appears in [37, Proposition 22]. As part of the implication (i) $\Rightarrow$ (ii) in Theorem 12,

> *if Π is an (α, θ) partition and Π' is a $(\beta, \theta/\alpha)$-coagulation of Π, then Π' is an $(\alpha\beta, \theta)$ partition.*

The case $\theta = 0$ of this implication amounts to part (i) of the next theorem. Thus Theorem 12 unifies and generalizes these two known relations involving fragmentation and coagulation operations on the two-parameter family.

2.4 The U-coalescent

Starting from a representation of the U-coalescent in terms of a Poisson process associated with Ruelle's probability cascades [39], and using the characterization of $PD(\alpha, 0)$ in Lemma 10, Bolthausen-Sznitman discovered the following much more explicit form of Theorem 6 for $\Lambda = U$:

Theorem 14 [9]
(i) *The family of Markov kernels* $((e^{-t}, 0)\text{-}\mathrm{COAG}, t \geq 0)$ *on* $\mathcal{P}_\infty$ *forms a semigroup.*
(ii) *The Markov process with this semigroup is the* U*-coalescent:*

$$\mathbb{P}^{U,\pi}(\Pi_\infty(t) \in \cdot) = (e^{-t}, 0)\text{-}\mathrm{COAG}(\pi, \cdot\,). \tag{16}$$

(iii) *For a standard* U*-coalescent,* $\Pi_\infty(t)$ *is an* $(e^{-t}, 0)$*-partition, with EPF*

$$p_t^U(n_1, \ldots, n_k) = \frac{(k-1)!}{(n-1)!} e^{-(k-1)t} \prod_{i=1}^{k} [1 - e^{-t}]_{n_i - 1}. \tag{17}$$

As shown by Bolthausen-Sznitman, (ii) follows easily from (i) by a transition rate calculation, and (iii) can be checked by showing that the right side of (17) solves the system of Kolmogorov backward equations for p_t^U. The EPF in (17) is the instance $(\alpha, \theta) = (e^{-t}, 0)$ of the EPF in (15). So either of (ii) and (iii) can be read from the other by application of Theorem 6 and Lemma 9. Apply Kingman's correspondence to deduce from Theorem 14 that the distribution P_t^U of ranked frequencies of blocks at time t in a standard U-coalescent is $P_t^U = PD(e^{-t}, 0)$.

Part (ii) of Theorem 14 combined with the implication (i) $\Rightarrow$ (ii) of Theorem 12 for

$$\alpha = e^{-s}, \quad \beta = e^{-(t-s)}, \quad \theta = 0$$

yields part (i) of the next corollary, part (ii) of which follows using the implication (ii) $\Rightarrow$ (i) of Theorem 12 for

$$\alpha = e^{-s}, \quad \beta = e^{-(t-s)}, \quad \theta = -e^{-T}.$$

Corollary 15 *Let* Π_∞ *be a standard* U*-coalescent. Then*
(i) *the co-transition probabilities of* Π_∞ *are given for* $0 < s < t$ *by*

$$\mathbb{P}(\Pi_\infty(s) \in \cdot \mid \Pi_\infty(t) = \pi) = (e^{-s}, -e^{-t})\text{-}\mathrm{FRAG}(\pi, \cdot\,).$$

(ii) *Fix $T>0$, and let $\Pi_\infty(T)=\{B_1(T),B_2(T),\ldots\}$. For $0\le t\le T$ and $m=1,2,\ldots$ let $\Pi_\infty^{(m)}(t)$ be the restriction of $\Pi_\infty(t)$ to $B_m(T)$, regarded as a $\mathcal{P}_\infty$-valued process after relabeling $B_m(T)$ by $\mathbb{N}$. Then, independently of $(\Pi_\infty(u),u\ge T)$, the processes $(\Pi_\infty^{(m)}(t),0\le t\le T)$ are independent and identically distributed time-inhomogeneous Markovian coalescents, each with final state $\Pi_\infty^{(m)}(T)=\{\mathbb{N}\}$, and each with the same co-transition probabilities as those of Π_∞ described in* (i). *For each m and $0<t<T$ the partition $\Pi_\infty^{(m)}(t)$ is an $(e^{-t},-e^{-T})$-partition, and the forwards transition probabilities are given for $0<s<t<T$ by*

$$\mathbb{P}(\Pi_\infty^{(m)}(t)\in\cdot\mid\Pi_\infty^{(m)}(s)=\pi)=(e^{-(t-s)},-e^{-(T-s)})\text{-COAG}(\pi,\cdot). \tag{18}$$

Less formally, each of the inhomogeneous Markovian coalescents $\Pi_\infty^{(m)}$ is a copy of the standard U-coalescent conditioned to reach state $\{\mathbb{N}\}$ at time T.

For Π_∞ a standard U-coalescent, let $(\tilde f_1(t),\tilde f_2(t),\ldots)$ denote the frequencies of blocks of $\Pi_\infty(t)$, in order of least elements, as defined by (11). Combine Lemma 10 and Theorem 14 to deduce the representation

$$\tilde f_1(t)=\tilde Y_1(t),\quad \tilde f_n(t)=(1-\tilde Y_1(t))\cdots(1-\tilde Y_{n-1}(t))\tilde Y_n(t)\qquad (n\ge 2) \tag{19}$$

where for each fixed t the $\tilde Y_n(t)$ are independent, and $\tilde Y_n(t)$ has beta$(1-e^{-t},ne^{-t})$ distribution for $n=1,2,\ldots$. As shown in Section 3.9, it follows that the two-dimensional distributions of the process $(\tilde f_1(t),t\ge 0)$ are as described in the following result. What is not at all obvious from this approach is that the process $(\tilde f_1(t),t\ge 0)$ has the Markov property. However, this is an immediate consequence of the description of the time-reversed U-coalescent provided by Corollary 15: the process $(\Pi_\infty^{(1)}(t),0\le t\le T)$ is independent of of $(\Pi_\infty(u),u\ge T)$, and $(\tilde f_1(t),0\le t\le T)$ can be recovered measurably from $\tilde f_1(T)$ and $(\Pi_\infty^{(1)}(t),0\le t\le T)$. See [5, Theorem 6] for a strikingly similar description of the corresponding process derived from the standard additive coalescent with time parameter set $(-\infty,\infty)$.

Corollary 16 *Let $\tilde f_1(t)$ be the frequency of the block containing 1 at time t in a standard U-coalescent. Then*

(i) *the process $(\tilde f_1(t),t\ge 0)$ is Markovian, with the same distribution as the process $(\gamma(1-e^{-t})/\gamma(1),t\ge 0)$ where $(\gamma(s),s\ge 0)$ is a gamma process, with stationary independent increments and $\mathbb{P}(\gamma(s)\in dx)=\Gamma(s)^{-1}x^{s-1}e^{-x}dx, x>0$.*

(ii) *The distribution of* $\tilde{f}_1(t)$ *is* $\text{beta}(1-e^{-t}, e^{-t})$*, and the process* $(-\log(1-\tilde{f}_1(t)), t \geq 0)$ *has non-stationary independent increments.*

(iii) *Let* $J_1 \geq J_2 \geq \ldots$ *be the ranked magnitudes of jumps of the process* $(\tilde{f}_1(t), t \geq 0)$*, and let* T_i *be the time when the jump of magnitude* J_i *occurs. Then the distribution of the sequence* $(J_1, J_2, \ldots)$ *on* $\mathcal{S}^{\downarrow}$ *is* $PD(0,1)$*, and this sequence is independent of the* T_i*, which are independent with standard exponential distribution.*

To restate (i), the random measure on $(0,\infty)$ which assigns mass $\tilde{f}_1(t)$ to $[0,t]$ is a Dirichlet random measure governed by the standard exponential distribution, as in [19]. Parts (ii) and (iii) are equivalents of (i) by well known properties of the Dirichlet random measure [12, Theorem 3.1],[19]. To interpret these results, regard the frequencies of blocks of $\Pi_\infty(t)$ as masses engaged in coalescent collisions governed by the ranked mass U-coalescent with conservation of total mass. Then $\tilde{f}_1(t)$ describes the mass at time t that has coalesced around some particle labeled 1 in the dust at time $0+$. Each jump J_i of the process $(\tilde{f}_1(t), t \geq 0)$ describes the increment of this mass due to a collision at some time T_i. According to Theorem 4, with probability one the collision at each of these times T_i involves an infinite number of other masses besides the mass of magnitude $\tilde{f}_1(T_i-)$ containing particle 1. The sum of all these other masses is J_i. As an application of Corollary 16, consider the increment of mass to the cluster containing 1 at the instant $\tau_{1,2}$ when this cluster first collides with the cluster containing a second particle labeled 2. By exchangeability considerations, for any standard Λ-coalescent there is the formula

$$\mathbb{P}^{\Lambda,1^\infty}(\tau_{1,2} \leq s \,|\, \tilde{f}_1(t), t \geq 0) = \tilde{f}_1(s) \text{ for all } s \geq 0. \tag{20}$$

For the U-coalescent, this fact can be combined with Corollary 16 as follows, to yield an explicit description of the trivariate law of $\tau_{1,2}$ and the random variables $\tilde{f}_1(\tau_{1,2}-)$ and $\tilde{f}_1(\tau_{1,2})$, which represent the mass of the cluster containing 1 just before and just after the collision with the cluster containing 2. Let

$$J_{1,2} := \tilde{f}_1(\tau_{1,2}) - \tilde{f}_1(\tau_{1,2}-)$$

which is the mass added to the cluster containing 1 at the time $\tau_{1,2}$ when that cluster first collides with the cluster containing 2. Then from (20) and Corollary 16 there is the equality of trivariate distributions

$$(\tau_{1,2}, \tilde{f}_1(\tau_{1,2}-), J_{1,2}) \stackrel{d}{=} (X, D(X-), D(X) - D(X-)) \tag{21}$$

where D is the cumulative distribution function of a Dirichlet random discrete distribution on $(0,\infty)$ governed by the standard exponential law, and X is a sample from D. It is well known (see e.g. [33, Corollary 9]) that $D(X) - D(X-)$ has uniform distribution on $[0,1]$ and is independent of X, that the pair $(X, D(X) - D(X-))$ is independent of the the random distribution function D' derived from D by deleting the atom of magnitude $D(X) - D(X-)$ at X and renormalizing to obtain a probability distribution, and that D' has the same distribution as D. So (21) yields the following Corollary:

Corollary 17 *In the standard U-coalescent, let $J_{1,2}$ be the mass added to the cluster containing 1 at the time $\tau_{1,2}$ when that cluster first collides with the cluster containing 2. Then $J_{1,2}$ has the uniform distribution U independent of the standard exponential time $\tau_{1,2}$, and the conditional distribution of $\tilde{f}_1(\tau_{1,2}-)/(1-u)$ given $J_{1,2} = u$ and $\tau_{1,2} = t$ is beta$(1-e^{-t}, e^{-t})$.*

3 The Λ-coalescent

3.1 Construction

For $1 \leq n < N \leq \infty$ and $\pi \in \mathcal{P}_N$ let $\mathtt{R}_n(\pi) \in \mathcal{P}_n$ be the restriction of π to $[n]$. For each finite n let $\Pi_n := (\Pi_n(t), t \geq 0)$ be a $\mathcal{P}_n$-valued coalescent Markov chain defined by the following transition rates: when the partition of $[n]$ has b blocks, each k-tuple of blocks is merging to form a single block at rate $\lambda_{b,k}$, for some array of non-negative real numbers $(\lambda_{b,k})$ indexed by $2 \leq k \leq b$. Call such an array of rates $(\lambda_{b,k})$ *consistent* if for all $n < m < \infty$ and each $\pi_m \in \mathcal{P}_m$, the process $\mathtt{R}_n(\Pi_m)$ given $\Pi_m(0) = \pi_m$ has the same distribution as Π_n given $\Pi_n(0) = \mathtt{R}_n(\pi_m)$.

Lemma 18 *An array of rates $(\lambda_{b,k})$ is consistent if and only if*

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1} \text{ for all } 2 \leq k \leq b. \tag{22}$$

Formula (1) sets up a bijection between consistent arrays $(\lambda_{b,k})$ and finite non-negative measures Λ on the Borel subsets of $[0,1]$.

Proof. It is easily seen that to check consistency for an array of rates it suffices to consider $m = n+1$. Condition (22) then appears from the well known condition in terms of transition rates for a function of a finite state Markov chain Y to be Markovian with some specified rates, no matter what the initial state of Y. See e.g. Rosenblatt [38, §IIId]. Let

$$\mu_{i,j} := \lambda_{i+j+2,i+2} \text{ for } i,j = 0,1,\ldots. \tag{23}$$

Then condition (22) becomes

$$\mu_{i,j} = \mu_{i+1,j} + \mu_{i,j+1} \text{ for } i,j = 0,1,\ldots. \tag{24}$$

It follows from de Finetti's representation of infinite exchangeable sequences of 0's and 1's [18, VII.4] that (24) is the necessary and sufficient condition for an array of non-negative numbers $(\mu_{i,j}, i,j = 0,1,\ldots)$ with $\mu_{0,0} = 1$ to be of the form

$$\mu_{i,j} = \mathbb{E}(X^i(1-X)^j) \qquad (i,j = 0,1,\ldots) \tag{25}$$

for some random variable X with values in $[0,1]$. The conclusion now follows easily. □

Proof of Theorem 1. The necessity of condition (1) follows from the previous Lemma. Assuming (1) holds, the desired process Π_∞ is constructed as in Kingman [27], who carried out this construction in the case $\Lambda = \delta_0$. An application of the Kolmogorov consistency theorem shows that for each initial partition π of $\mathbb{N}$ it is possible to construct the Markov chains Π_n all on the same probability space, each with right continuous step function paths, in such a way that $\Pi_n(0) = \mathtt{R}_n\pi$, and $\Pi_n = \mathtt{R}_n(\Pi_m)$ for $n < m < \infty$. The desired process Π_∞ is then obtained by letting $\Pi_\infty(t)$ be the unique partition of $\mathbb{N}$ whose restriction to $[n]$ is $\Pi_n(t)$ for every n. The claimed regularity properties of Π_∞ then follow by straightforward arguments given in [15, 9] for similar constructions. □

3.2 Symmetry properties

For a bijection σ with domain A and range B, and a partition π of A, let $\sigma\pi$ denote the partition of B whose blocks are the σ-images of the blocks of π. The form of the transition rates of Π_n implies that if Π_n is started in state π_n, then for every permutation σ of $[n]$ the process $\sigma\Pi_n$ is a copy of

Π_n started in state $\sigma\pi_n$. This basic *exchangeability property* of the chains Π_n implies the following exchangeability property of the Λ-coalescent. Let $[\infty] := \mathbb{N}$. For each $n = 1, 2, \ldots, \infty$, and each subset H of $\mathbb{N}$ containing n elements, the restriction of Π_∞ to H, when regarded as a $\mathcal{P}_n$-valued process by labeling H by an arbitrary bijection σ from H to $[n]$, has the same distribution as Π_n started in state $\sigma\pi_H$, where π_H is the restriction to H of the initial state π of Π_∞.

Proof of Theorem 6. The exchangeability for each t of the random partition $\Pi_\infty(t)$ derived from a standard Λ-coalescent is evident from the previous paragraph. The construction of the Λ-coalescent started in state π from the standard Λ-coalescent, and vice versa, are easily established by consideration of restrictions to $[n]$ for each finite n. □

3.3 Generalizations

Note from Definition 5 that no matter what the distribution p on $\mathcal{P}_\infty$, each of the kernels $K = p\text{-COAG}$ acts *locally* on $\mathcal{P}_\infty$, meaning that if Π^π denotes a random partition of $\mathbb{N}$ with distribution $K(\pi, \cdot)$, then for each n the distribution of $\mathtt{R}_n\Pi^\pi$ depends on π only through $\mathtt{R}_n\pi$. It follows that any $\mathcal{P}_\infty$-valued Markov process Π_∞, each of whose transition kernels is of the form $p\text{-COAG}$ for some p, is such that the $\mathcal{P}_n$-valued process $\mathtt{R}_n\Pi_\infty$ is a Markov chain. Such a coalescent process Π_∞ with càdlàg paths could therefore be constructed more generally than in Theorem 1 from a consistent family of Markov chains with more complex transition rules, allowing not just multiple collisions in which several blocks merge to form one block, but simultaneous multiple collisions, in which several new blocks might be formed, each from the merger of two or more smaller blocks. There is a *composition rule* for coagulation kernels associated with exchangeable distributions p_i on $\mathcal{P}_\infty$ which induces a semigroup operation on these distributions, or equivalently on $\bar{\mathcal{S}}^\downarrow$: $(p_1\text{-COAG})(p_2\text{-COAG}) = p_3\text{-COAG}$ where p_3 is determined explicitly by Lemma 34. From this perspective, Theorems 1 and 6 must be special cases of some more general characterization of consistently defined $\mathcal{P}_n$-valued Markov chains with appropriate exchangeability properties, or of one-parameter semigroups of exchangeable coagulation kernels $(p_t\text{-COAG}, t \geq 0)$. Similar remarks apply to Markovian fragmentation processes each of whose transition kernels is $p\text{-FRAG}$ for some p. But such generalizations will not be pursued further here.

3.4 Examples

Example 19 Λ *such that* $\mu_{-2} < \infty$.

Fix $\lambda > 0$, let p be the distribution on $\mathcal{P}_\infty$ of an exchangeable random partition of $\mathbb{N}$ with ranked frequencies $\boldsymbol{Y} = (Y_1, Y_2, \ldots) \in \bar{\mathcal{S}}^\downarrow$, and define for each $t \geq 0$ a transition probability kernel P_t on $\mathcal{P}_\infty$ by

$$P_t := \sum_{m=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^m}{m!} (p\text{-COAG})^m. \tag{26}$$

From the previous general remarks about coagulation kernels, it is clear that this formula defines the semigroup of a $\mathcal{P}_\infty$-valued Markov process Π_∞ whose restrictions Π_n are Markovian for every n. Such a process Π_∞ can be constructed with step function paths by the usual scheme of jumping according to p-COAG at the arrival times of a homogeneous Poisson process with rate λ. If $\mathbb{P}(Y_2 > 0) > 0$ then starting in state 1^∞ there is positive probability of more than two new blocks being formed at the first jump. So the only distributions of $\boldsymbol{Y}$ such that Π_∞ could be a Λ-coalescent for some Λ are those with $\mathbb{P}(Y_2 = 0) = 1$. Assuming $\mathbb{P}(Y_2 = 0) = 1$, let F_1 be the distribution of Y_1, so F_1 could be any probability distribution on $[0, 1]$, and p is the distribution of the exchangeable random partition of $\mathbb{N}$ with ranked frequencies $(Y_1, 0, \ldots)$. By consideration of restrictions to $[n]$ it is easily verified that Π_∞ so constructed is a Λ-coalescent for $\Lambda(dx) = \lambda x^2 F_1(dx)$. This construction can therefore be used to make a Λ-coalescent for any Λ with $\mu_{-2} < \infty$ by use of $\lambda = \mu_{-2}$ and $F_1(dx) := x^{-2}\Lambda(dx)/\mu_{-2}$. Indeed, this is just another way of expressing the result of Corollary 3 when $\mu_{-2} < \infty$. Each jump in this construction according to p-COAG can be described by the following variation of Kingman's paintbox scheme for generating an exchangeable random partition with distribution p: given the current partition is π, to make the next partition π' with distribution p-COAG$(\pi, \cdot)$, first pick Y_1 with distribution F_1. Given $Y_1 = y$ toss a coin which lands heads independently with probability y for each block of π, and let π' be derived from π by merging all the blocks whose coins land heads.

Example 20 Λ *with an atom at* 1. Let $\Lambda_1 = \Lambda + \lambda\delta_1$ where Λ has no atom at 1, for $\lambda > 0$ and δ_1 a unit mass at 1. Let Π_∞ be a Λ-coalescent, let T be an independent exponential time with rate λ, and let $\Pi'_\infty(t)$ equal $\Pi_\infty(t)$ if $t < T$ and $\{\mathbb{N}\}$ if $t \geq T$. Then it is easily seen that Π'_∞ is a Λ_1-coalescent. As the Λ_1-coalescent is so easily described in terms of the Λ-coalescent, to avoid trivial exceptions in the formulation of some later results it may be assumed that Λ has no atom at 1.

3.5 Behavior at collision times

Theorem 4 is an immediate consequence of the following lemma and de Finetti's strong law of large numbers for exchangeable indicator variables. In view of the exchangeability property of the Λ-coalescent, for i and j in distinct blocks of the initial partition, all probabilistic features of the collision at time $\tau_{i,j}$ are identical, modulo relabeling, to corresponding features of the collision at time $\tau_{1,2}$, assuming 1 and 2 are in different blocks of π. So to simplify notation, take $(i,j) = (1,2)$. Say that indicator variables $J_1, \ldots, J_n$ are *exchangeable*(F) *indicators* if the J_i have the same distribution as $K_1, \ldots, K_n$ where the K_i are conditionally independent given X with $\mathbb{P}(K_i = 1 \mid X) = X$ for all i, for some random variable X with distribution F. That is, for every subset A of $[n]$

$$\mathbb{P}(\{i \in [n] : J_i = 1\} = A) = \mathbb{E}(X^{|A|}(1-X)^{n-|A|}) \tag{27}$$

where $|A|$ is the number of elements of A. Recall that $\#\pi$ is the number of blocks of a partition π.

Lemma 21 *Let Π_∞ be an F-coalescent started at π with 1 and 2 in distinct blocks of π. Let $B_1, B_2, \ldots$ denote the blocks of $\Pi_\infty(\tau_{1,2}-)$, in order of their least elements, so $1 \in B_1$ and $2 \in B_2$. Let I_j be the indicator of the event that the collision between B_1 and B_2 at time $\tau_{1,2}$ also involves block B_j, meaning that B_j, B_1 and B_2 all belong to the same block of $\Pi_\infty(\tau_{1,2})$. Then for each $n \geq 3$, conditionally given $\tau_{1,2}$ and $\#\Pi_\infty(\tau_{1,2}-) \geq n$, the $I_3, I_4, \ldots, I_n$ are exchangeable(F) indicators.*

Proof. By (27) it is enough to show for arbitrary $n \geq 3$ and $t > 0$ that for every subset A of $\{3, \ldots, n\}$ with $|A| = a$

$$\mathbb{P}(\{j : 3 \leq j \leq n, I_j = 1\} = A \mid \tau_{1,2} = t, \#\Pi_\infty(t-) \geq n) = \mu_{i,n-i-2} \tag{28}$$

for $\mu_{i,j}$ defined by (23)-(25) with X distributed according to F. For $\pi \in \mathcal{P}_\infty$ with $\#\pi \geq n$ let $\nu_n\pi$ be the least m such that the restriction of π to $[m]$ has at least n blocks. Then it suffices to establish (28) with conditioning on $\tau_{1,2} = t$ and $\#\Pi_\infty(t-) \geq n$ replaced by conditioning on $\tau_{1,2} = t$, $\#\Pi_\infty(t-) \geq n$ and $\nu_n\Pi_\infty(t-) = m$ for arbitrary $m = n, n+1, \ldots$. This variant of (28) is implied another variant of (28) with conditioning on $\tau_{1,2} = t$ and $\#\Pi_\infty(t-) \geq n$ replaced by conditioning on $\tau_{1,2} = t$ and $\Pi_m(t-) = \pi_m$ for some arbitrary partition π_m of $[m]$ into n blocks, say $\{B_{m,1}, B_{m,2}, \ldots, B_{m,n}\}$, with $1 \in B_{m,1}$ and $2 \in B_{m,2}$. But this last form of (28) follows immediately from the description of the transition rates of Π_m: given that Π_m is in such a state π_m, the rate of all transitions that cause blocks $B_{m,1}$ and $B_{m,2}$ to be merged (perhaps also with other blocks) is $\lambda_{2,2} = \mu_{0,0} = 1$, while the rate of these transitions in which the set of other blocks involved is $\{B_{m,j}, j \in A\}$, where $|A| = a$, is $\lambda_{n,a+2} = \mu_{a,n-a-2}$. $\square$

Example 22 As a check on Theorem 4, consider the F-coalescent Π_∞ constructed for F such that $\mu_{-2} < \infty$, as in Example 19, for F with no atom at 1. Let the initial state π with infinitely many blocks have 1 and 2 in different blocks. From the description of Example 19, it is clear that the coalescent has infinitely many blocks at all times $t > 0$, and that $\tau_{1,2} = T_N$ for some random index N, where T_i is the time of the ith jump of Π_∞. The variable $X_{1,2}$ in Theorem 4 is then $X_{1,2} = Y_N$ where by construction the $Y_i, i = 1, 2, \ldots$ are independent with

$$\mathbb{P}(Y_i \in dx) = x^{-2} F(dx)/\mu_{-2}$$

and for $i = 1, 2, \ldots$

$$P(N = i \,|\, Y_1, Y_2, \ldots) = (1 - Y_1^2) \cdots (1 - Y_{i-1}^2) Y_i^2.$$

It follows easily that $\mathbb{P}(Y_N \in dx) = F(dx)$ as claimed, and that N is independent of Y_N with geometric distribution with mean μ_{-2}. As a final check, $\tau_{1,2} = T_N$ is the sum of N independent exponential variables with mean $1/\mu_{-2}$, hence exponential with mean $\mu_{-2}/\mu_{-2} = 1$, as required since $\lambda_{2,2} = 1$.

3.6 The total number of blocks

Let $\#(t) := \#\Pi_\infty(t)$, the number of blocks at time t in a Λ-coalescent Π_∞. It was observed by Kingman for $\Lambda = \delta_0$ and Bolthausen-Sznitman for $\Lambda = U$, and is easily seen for general Λ, that the process $(\#(t), t \geq 0)$ is a time-homogeneous Markov process relative to the filtration of Π_∞, with state-space $\{1, 2, \ldots, \infty\}$ and only downwards jumps, such that for $2 \leq k \leq b < \infty$ the rate of jumps down from b to $b - k + 1$ is

$$\text{rate}(b \to b - k + 1) = \lambda^{\#}_{b,k} := \binom{b}{k} \lambda_{b,k}. \tag{29}$$

The total rate of downward jumps is $\lambda_b := \sum_{k=2}^{b} \lambda^{\#}_{b,k}$, as in (5)-(6). By (6), the sequence $(\lambda_b, b = 1, 2, \ldots)$ is strictly increasing, no matter what non-zero Λ is considered. And (5) shows that the μ_i for $0 \leq i \leq n-2$ and hence the $\lambda_{b,k}$ for all $2 \leq k \leq b \leq n$ are universal linear combinations of $\lambda_2, \ldots, \lambda_n$.

For the rest of this section, Λ is fixed and suppressed in the notation. Let $\mathbb{P}$ govern Π_∞ as a standard Λ-coalescent. For each $n = 1, 2, \ldots, \infty$ let

$$T_n := \inf\{t : \#\Pi_n(t) = 1\}, \tag{30}$$

call it the *absorbtion time* of Π_n. By construction, the distribution of T_2 is exponential with rate $\lambda_2 > 0$, and

$$0 = T_1 < T_2 \leq T_3 \leq T_4 \leq \cdots \uparrow T_\infty \leq \infty. \tag{31}$$

For $\theta > 0$ and $n = 1, 2, \ldots, \infty$, let

$$\phi_n(\theta) := \mathbb{E}(e^{-\theta T_n}) = \theta \int_0^\infty \mathbb{P}(T_n \leq t) e^{-\theta t} dt. \tag{32}$$

For each $\theta > 0$ the rates $\lambda_{b,k}$ determine $\phi_n(\theta)$ for all finite n by the recursion

$$\phi_n(\theta) = \left(\frac{1}{\lambda_n + \theta}\right) \sum_{k=2}^{n} \lambda^{\#}_{n,k} \phi_{n-k+1}(\theta) \qquad \text{where } \phi_1(\theta) = 1, \tag{33}$$

which follows from (29) by conditioning on the number of blocks after the first collision. By (31), as $n \to \infty$

$$\phi_n(\theta) \downarrow \phi_\infty(\theta) = \mathbb{E}(e^{-\theta T_\infty}). \tag{34}$$

Say the Λ-coalescent *comes down from infinity* if $\mathbb{P}(\#(t) < \infty) = 1$ for each $t > 0$. Say it *stays infinite* if $\mathbb{P}(\#(t) = \infty) = 1$ for each $t > 0$. Kingman showed that the δ_0-coalescent comes down from infinity. The results of Bolthausen-Sznitman reviewed in Theorem 14 show the U-coalescent stays infinite.

Proposition 23 *Suppose that Λ has no atom at 1. Then*
either: *$\phi_\infty(\theta) > 0$ for all $\theta > 0$, in which case the standard Λ-coalescent comes down from infinity in such a way that $\#(0+) = \infty$ a.s., by an infinite number of collisions, each involving only a finite number of blocks almost surely; the absorbtion time T_∞ is then a.s. finite.*

or: *$\phi_\infty(\theta) = 0$ for all $\theta > 0$, in which case the standard Λ-coalescent stays infinite by an infinite number of collisions, each involving an infinite number of blocks almost surely. The absorbtion time T_∞ is then a.s. infinite.*

Proof. Let $T := \inf\{t : \#(t) < \infty\}$. One possibility to be eliminated is that $\mathbb{P}(0 < T < \infty) > 0$. If $\mathbb{P}(0 < T < \infty) > 0$ the strong Markov property implies that $\#(T) < \infty$ a.s. on the event $\{0 < T < \infty\}$. On the same event $\#(T-) = \infty$ by the definition of T. Hence T must be a collision time, and Theorem 4 gives a contradiction. The remaining possibilities are eliminated similarly. Theorem 4 also implies that if the coalescent stays infinite then every collision involves infinitely many blocks almost surely. □

Continuing to suppose Λ has no atom at 1, Lemma 25 in the next section shows that if $\mu_{-1} < \infty$ then the Λ-coalescent stays infinite. But this condition is not necessary, as shown by the U-coalescent. If Λ has an atom at 0 of magnitude $\lambda_0 > 0$, then the Λ-coalescent comes down from infinity. For it is easy to see that $(\#(t), t \geq 0)$ can then be constructed on same probability space as $(\#_0(t), t \geq 0)$ derived from a $\lambda_0\delta_0$-coalescent in such a way that $\#(t) \leq \#_0(t)$ for all $t \geq 0$. In principle, the condition on ϕ_∞ in Proposition 23 is a condition on Λ equivalent to the Λ-coalescent staying infinite. But the condition is not an easy one to check. It is a natural problem to find a simpler equivalent of this condition.

The following proposition develops Theorem 4 to give a more complete description of the nature of collisions in a Λ-coalescent that stays infinite.

Proposition 24 *Suppose that Λ has no atom at either 0 or 1, and that the standard Λ-coalescent Π_∞ stays infinite. Let $\tau_{i,j}$ be the collision time of i, j in Π_∞, and $X_{i,j}$ the almost sure limiting fraction of blocks of $\Pi_\infty(\tau_{i,j}-)$ involved in the collision at time $\tau_{i,j}$, as in Theorem 4. Then the random set*

$$S := \{(\tau_{i,j}, X_{i,j}), i, j \in \mathbb{N}\} \subset (0, \infty) \times (0, 1)$$

is the set of points of a Poisson point process with intensity $dt\, x^{-2}\Lambda(dx)$.

Proof. Without loss of generality, it can be supposed that the standard Λ-coalescent Π_∞ is created by the Poisson construction of Corollary 3. By the assumption that the coalescent stays infinite, the random set S is identical to the set of all points (t, x) where x is the limiting relative frequency of 1's of ξ as (t, ξ) ranges over the points used in the Poisson construction of Π_∞. The conclusion is now evident. $\square$

3.7 The ranked mass coalescent

The following lemma starts the proof of Theorem 8.

Lemma 25 *The standard Λ-coalescent has proper frequencies almost surely for each $t > 0$ if and only if $\mu_{-1} = \infty$.*

Proof. By Kingman's correspondence, the exchangeable random partition $\Pi_\infty(t)$ has proper frequencies iff the singleton set $\{1\}$ is almost surely not a block of $\Pi_\infty(t)$. In the restriction of the Λ-coalescent to $[n]$, when there are b blocks including $\{1\}$, the total rate at which $\{1\}$ is colliding with one or more of the $b-1$ other blocks is found to be

$$\rho_b := \sum_{k=2}^{b} \binom{b-1}{k-1} \lambda_{b,k} = \mu_0 \mathbb{E}\left[\frac{1-(1-X)^{b-1}}{X}\right] \tag{35}$$

where the ratio is interpreted by continuity to equal $b-1$ if $X = 0$. Thus ρ_b increases to $\mu_0 \mathbb{E}(X^{-1}) = \mu_{-1}$ as b increases to ∞. In the standard Λ-coalescent, the rate at which $\{1\}$ is colliding with some other block is therefore always bounded above by μ_{-1}. If this moment is finite, the probability that 1 is still a singleton at time t is at least $\exp(-\mu_{-1}t) > 0$, so the frequencies of $\Pi_\infty(t)$ are not proper. If $\mu_{-1} = \infty$ there are two possibilities. Either $\#(t) = \infty$, in which case $\{1\}$ has been subject to collisions at an infinite rate for time t, so some such collision has a.s. occurred by time t, or $\#(t) < \infty$, in which case also $\{1\}$ is a.s. not a singleton of $\Pi_\infty(t)$, because an exchangeable random partition of $\mathbb{N}$ with a finite number of blocks contains no singletons a.s.. $\square$

If $\mu_{-1} < \infty$, the process of frequencies $(\mathbf{f}(t), t \geq 0)$ derived from a standard Λ-coalescent is an $\bar{\mathcal{S}}^{\downarrow}$-valued process governed by an analog of the Λ-coalescent ranked mass semigroup on $\bar{\mathcal{S}}^{\downarrow}$ which allows creation of mass, starting from mass 0 at time 0 and terminating with mass 1 at time $\infty-$. The *missing mass* in this process at time t, that is $1 - \sum_i f_i(t)$, is the relative frequency of the union of all singleton blocks of $\Pi_\infty(t)$. Only in the case $\mu_{-1} = \infty$ is mass 1 instantaneously created at time $0+$ so that the state-space can be restricted to $\mathcal{S}^\downarrow$ for $t > 0$. The process of creation of mass when $\mu_{-1} < \infty$ is described by the following proposition.

Proposition 26 *Let $S_t := 1 - \sum_i f_i(t)$, which is the frequency of singletons at time t in a standard Λ-coalescent Π_∞. If $\mu_{-1} < \infty$ then the process $(-\log S_t, t \geq 0)$ is a drift-free subordinator whose Lévy measure is the image of $x^{-2}\Lambda(dx)$ via the map $x \mapsto -\log(1-x)$. Consequently, the distribution of S_t on $[0,1]$ is that determined by Mellin transform*

$$\mathbb{E}(S_t^\eta) = \exp\left[-t\int_0^1 (1-(1-x)^\eta)x^{-2}\Lambda(dx)\right] \quad \textit{for } \eta \geq 0. \tag{36}$$

For $\eta \in \mathbb{N}$ this formula gives the probability that each element of $[\eta]$ is still a singleton at time t. In particular, the mean frequency of singletons in $\Pi_\infty(t)$ is

$$\mathbb{E}(S_t) = \mathbb{P}(\{1\} \in \Pi_\infty(t)) = \exp(-\mu_{-1}t). \tag{37}$$

Proof. As in the proof of Proposition 24 it can be supposed that Π_∞ is created by the Poisson construction of Corollary 3. Formula (36) for $\eta \in \mathbb{N}$ can be read directly from this Poisson construction. It is clear from the Markov property of Π_∞ expressed in Theorem 6 that the process $(-\log S_t, t \geq 0)$ has stationary independent increments. But if $\hat{S}_t := \exp(-Y_t)$ where $(Y_t, t \geq 0)$ is a drift free subordinator whose Lévy measure is the image of $x^{-2}\Lambda(dx)$ via the map $x \mapsto -\log(1-x)$, then the Lévy-Khintchine formula shows that $E(\hat{S}_t^\eta)$ equals the right side of (36) for every $\eta \geq 0$. Since a probability distribution on $[0,1]$ is determined by its positive integer moments, S_t and $\hat{S}_t$ must have the same distribution for each $t > 0$, and the conclusion follows. □

Consider now the problem of characterizing all entrance laws $(q_t, t > 0)$ for the ranked mass Λ-coalescent. By general theory [14], each entrance law is an integral mixture over some set of extreme entrance laws, called the *entrance boundary* of the semigroup. The following theorem identifies this entrance boundary with $\mathcal{S}^\downarrow$ if $\mu_{-1} < \infty$ and with $\bar{\mathcal{S}}^\downarrow$ if $\mu_{-1} = \infty$. Theorem 8 is an immediate consequence of this more general result.

Theorem 27 *Let $(\boldsymbol{X}(t), t > 0)$ be a process with càdlàg $\mathcal{S}^{\downarrow}$-valued paths governed by the Λ-coalescent ranked mass semigroup. Then $\boldsymbol{X}(0+) := \lim_{t \downarrow 0} \boldsymbol{X}(t)$ exists almost surely as limit in $\bar{\mathcal{S}}^{\downarrow}$ in the sense of componentwise convergence.*
(i) *If $\mu_{-1} < \infty$ then $\boldsymbol{X}(0+) \in \mathcal{S}^{\downarrow}$ a.s. and the extreme entrance laws are obtained by starting the ranked mass Λ-coalescent at $\boldsymbol{x}$ as $\boldsymbol{x}$ ranges over $\mathcal{S}^{\downarrow}$.*
(ii) *If $\mu_{-1} = \infty$ then $\boldsymbol{X}(0+)$ may have an arbitrary probability distribution on $\bar{\mathcal{S}}^{\downarrow}$. There is then for each $\boldsymbol{x} \in \bar{\mathcal{S}}^{\downarrow}$ a unique extreme entrance law under which $\boldsymbol{X}(0+) = \boldsymbol{x}$ a.s.. The corresponding process may be constructed by defining $\boldsymbol{X}(t)$ to be the ranked frequencies of $\Pi_\infty(t)$, where Π_∞ is a Λ-coalescent with $\Pi_\infty(0)$ an exchangeable random partition of $\mathbb{N}$ whose sequence of ranked frequencies is $\boldsymbol{x}$.*

Proof. Acording to the definition of the Λ-coalescent ranked mass semigroup by Corollary 7, for each $\varepsilon > 0$ the process $(\boldsymbol{X}(t), t \geq \varepsilon)$ has the same distribution as $((\boldsymbol{X}(\varepsilon), \Pi^*_\infty(t-\varepsilon)), t \geq \varepsilon)$ for a standard Λ-coalescent Π^*_∞ that is independent of $\boldsymbol{X}(\varepsilon)$. Let $(\Pi^\varepsilon_\infty(t), t \geq \varepsilon)$ be a Λ-coalescent with initial state $\Pi^\varepsilon_\infty(\varepsilon)$ which is an exchangeable random partition of $\mathbb{N}$ such that $\mathbf{f}(\Pi^\varepsilon_\infty(\varepsilon))$ has the same distribution as $\boldsymbol{X}(\varepsilon)$, where $\mathbf{f}(\Pi)$ denotes the sequence of ranked frequencies of an exchangeable random partition Π. By application of Theorem 6 and [15, Lemma 29], the process $((\boldsymbol{X}(\varepsilon), \Pi^*_\infty(t-\varepsilon)), t \geq \varepsilon)$ has the same law as $(\mathbf{f}(\Pi^\varepsilon_\infty(t)), t \geq \varepsilon)$. Thus

$$(\boldsymbol{X}(t), t \geq \varepsilon) \stackrel{d}{=} (\mathbf{f}(\Pi^\varepsilon_\infty(t)), t \geq \varepsilon)$$

where $\stackrel{d}{=}$ denotes equality in distribution of processes. An application of Kolmogorov's extension theorem now shows that there exists a $\mathcal{P}_\infty$-valued Markov process $(\Pi_\infty(t), t > 0)$ governed by the Λ-coalescent semigroup, with $\Pi_\infty(t)$ an exchangeable random partition of $\mathbb{N}$ for each $t > 0$, such that

$$(\boldsymbol{X}(t), t > 0) \stackrel{d}{=} (\mathbf{f}(\Pi_\infty(t)), t > 0).$$

Since $\Pi_\infty(t)$ is refining as t decreases, its limit $\Pi_\infty(0+)$ exists in $\mathcal{P}_\infty$, and is exchangeable. Since for each m the sum of the m largest frequencies of $\Pi_\infty(t)$ is a non-decreasing function of t, the limit of $\mathbf{f}(\Pi_\infty(t))$ as $t \downarrow 0$ exists in $\bar{\mathcal{S}}^{\downarrow}$ almost surely, hence the limit $\boldsymbol{X}(0+)$ exists in $\bar{\mathcal{S}}^{\downarrow}$ almost surely. Moreover, the continuity of Kingman's correspondence (Theorem 36) implies that $\boldsymbol{X}(0+)$ has the same distribution as the ranked frequencies of $\Pi_\infty(0+)$. By the Feller property of the Λ-coalescent semi-group on $\mathcal{P}_\infty$, the process $(\Pi_\infty(t), t > 0)$ must be just the restriction to the time interval $(0,\infty)$ of a Λ-coalescent started in the random state $\Pi_\infty(0+)$. If $\boldsymbol{X}(0+)$ is improper with positive probability, then $\Pi_\infty(0+)$ has a strictly positive frequency of singletons with positive probability. If $\mu_{-1} < \infty$, formula (37) implies that $\Pi_\infty(t)$ has a positive frequency of singletons with positive probability, contradicting the assumption that $\boldsymbol{X}(t) \in \mathcal{S}^\downarrow$ for every $t > 0$. If on the other hand $\mu_{-1} = \infty$ there is no contradiction. Rather, for any given vector $\boldsymbol{x} \in \bar{\mathcal{S}}^\downarrow$, a process started with $\boldsymbol{X}(0+) = \boldsymbol{x}$ is obtained as indicated in the theorem. Finally it is easily shown that the law of each such process is extreme, by application of Kingman's result that the extreme laws of exchangeable random partitions of $\mathbb{N}$ are the laws $p_{\boldsymbol{x}}$ corresponding to a given sequence of ranked frequencies $\boldsymbol{x} \in \bar{\mathcal{S}}^\downarrow$. $\square$

3.8 The exchangeable probability function

Formulae for the EPF $p_t^\Lambda(n_1, \ldots, n_k)$ derived from the standard Λ-coalescent for general Λ, as in (8) can be found explicitly at least for some particular $(n_1, \ldots, n_k)$. Fix Λ and let $p_t(\cdots)$ stand for $p_t^\Lambda(\cdots)$. Due to symmetry of the EPF, it suffices to consider $p_t(n_1, \ldots, n_k)$ for decreasing sequences $(n_1, \ldots, n_k)$. Write for instance $3^2 2^3 1$ for the decreasing sequence $(3,3,2,2,2,1)$. The decreasing rearrangement of the sizes of blocks of a partition π of $[n]$ is a partition of the integer n, call it the *type* of π. Note that $p_t(3^2 2^3 1)$ is *not* the probability that $\Pi_{13}(t)$ is of type $3^2 2^3 1$, but rather the probability that $\Pi_{13}(t) = \pi$ for each particular partition π of $\{1, \ldots, 13\}$ of type $3^2 2^3 1$. The holding time of $(\Pi_n(t), t \geq 0)$ in its initial state 1^n is exponential with rate λ_n, as in (5), so

$$p_t(1^n) = e^{-\lambda_n t} \quad \text{for } n = 2, 3, \ldots \text{ where} \tag{38}$$

$$\lambda_2 = \mu_0; \quad \lambda_3 = 3\mu_0 - 2\mu_1; \quad \lambda_4 = 6\mu_0 - 8\mu_1 + 3\mu_2$$

and so on. Note that the λ_j simplify to $j(j-1)/2$ for $\Lambda = \delta_0$ and to $j-1$ for $\Lambda = U$. The values of $p_t(n_1, \ldots, n_k)$ for all $(n_1, \ldots, n_k)$ with $\sum_j n_j \leq 3$ can be deduced from (38) for $n = 2, 3$ and the addition rules (66) for an EPF. Thus

$$p_t(1) = 1; \quad p_t(1^2) = e^{-\lambda_2 t}; \quad p_t(2) = 1 - e^{-\lambda_2 t} \tag{39}$$

$$p_t(1^3) = e^{-\lambda_3 t}; \quad p_t(2^1 1) = \tfrac{1}{2} e^{-\lambda_2 t} - \tfrac{1}{2} e^{-\lambda_3 t}; \quad p_t(3) = 1 - \tfrac{3}{2} e^{-\lambda_2 t} + \tfrac{1}{2} e^{-\lambda_3 t} \tag{40}$$

As a check, for Λ such as U with $\mu_0 = 1, \mu_1 = 1/2$, formula (17) is recovered for all $(n_1, \dots, n_k)$ with $\sum_i n_i \le 3$.

For $n \ge 3$ a state of type 21^{n-2} can only be entered directly from 1^n. So by conditioning on the entry time and using (38)

$$p_t(21^{n-2}) = \int_0^t e^{-\lambda_n s}\lambda_{n,2}\, e^{-\lambda_{n-1}(t-s)}ds = \frac{\lambda_{n,2}}{(\lambda_n - \lambda_{n-1})}(e^{-\lambda_{n-1}t} - e^{-\lambda_n t}).$$

For $n \ge 4$ a state of type 31^{n-3} can only be entered directly from 1^n or via one of 3 different states of type 21^{n-2}. Conditioning on these cases and the entry time gives

$$p_t(31^{n-3}) = \int_0^t p_s(1^n)\lambda_{n,3}e^{-\lambda_{n-2}(t-s)}ds + 3\int_0^t p_s(21^{n-2})\lambda_{n-1,2}e^{-\lambda_{n-2}(t-s)}ds$$

and hence by integration

$$p_t(31^{n-3}) = \left\{\frac{\lambda_{n,3}}{(\lambda_n - \lambda_{n-2})} + \frac{3\lambda_{n,2}\lambda_{n-1,2}}{(\lambda_n - \lambda_{n-1})}\left[\frac{1}{\lambda_{n-1} - \lambda_{n-2}} - \frac{1}{\lambda_n - \lambda_{n-2}}\right]\right\}e^{-\lambda_{n-2}t} +$$

$$+\frac{3\lambda_{n,2}\lambda_{n-1,2}}{(\lambda_n - \lambda_{n-1})}\left[\frac{-1}{\lambda_{n-1} - \lambda_{n-2}}\right]e^{-\lambda_{n-1}t} + \left\{-\frac{\lambda_{n,3}}{(\lambda_n - \lambda_{n-2})} + \frac{3\lambda_{n,2}\lambda_{n-1,2}}{(\lambda_n - \lambda_{n-1})}\left[\frac{1}{\lambda_n - \lambda_{n-2}}\right]\right\}e^{-\lambda_n t}$$

Proceeding in this way, it is clear that for given n the values of $p_t(n_1, \dots, n_k)$ with $\sum_i n_i = n$ can be found one by one by repeated integration for $k = n, n-1, n-2$ and so on. So a reverse induction on k for fixed n yields:

Proposition 28 *The EPF p_t^Λ derived from a standard Λ-coalescent is of the form*

$$p_t^\Lambda(n_1, \dots, n_k) = \sum_{j=k}^{n} \frac{a_j^\Lambda(n_1, \dots, n_k)}{\prod_{k \le i < h \le b}(\lambda_h - \lambda_i)} e^{-\lambda_j t} \tag{41}$$

where $n := \sum_i n_i$ and $a_j^\Lambda(n_1, \dots, n_k)$ is a polynomial in the $\lambda_{b,i}$ for $k \le b \le n$ and $2 \le i \le b$, with integer coefficients.

In principle, the inductive derivation of this result yields a recursive description of the $a_j^\Lambda(n_1, \dots, n_k)$, but this recursion seems very complicated. Neither does there appear to be any substantial simplification for Λ with simple rates, such as $\Lambda = \delta_x$ or $\Lambda =$ beta(r, s), except in the special cases $\Lambda = \delta_0$ (see [25, 27] and [40, (6.1)]), and $\Lambda = U$ (Theorem 14).

An unfortunate feature of formula (41) is that for given $n = \sum_i n_i$ the description of $p_t(n_1, \ldots, n_k)$ is most complicated when $k = 1$, which is one of the most interesting cases. A more explicit description of $p_t(n)$ can be obtained as follows. Observe first that for T_n the absorbtion time of Π_n as in (30), there is the formula $p_t(n) = \mathbb{P}(T_n \le t)$, so (32) yields the Laplace transform $\int_0^\infty p_t(n) e^{-\theta t} dt = \phi_n(\theta)/\theta$ where $\phi_n(\theta) = \mathbb{E}(e^{-\theta T_n})$ is determined recursively by (33). The connection with T_n can be exploited to obtain the following formula, by conditioning the range of the process $(\#\Pi_n(t), 0 \le t \le T_n)$ to be a particular subset $\{m_0, \ldots, m_c\}$ of $[n]$:

$$p_t(n) = 1 - \sum_{c=1}^{n-1} \sum_{(m_0,\ldots,m_c)} \left(\prod_{i=1}^{c} \frac{\lambda^{\#}(m_i, m_i - m_{i-1} + 1)}{\lambda_{m_i}} \right) G_t(\lambda_{m_1}, \ldots, \lambda_{m_c}) \tag{42}$$

where $\lambda^{\#}(m,k) := \lambda^{\#}_{m,k}$ as in (29), for each $c = 1, \ldots, n-1$ the sequence $(m_0, \ldots, m_c)$ ranges over the set of $\binom{n-2}{c-1}$ strictly increasing sequences of positive integers with $m_0 = 1$ and $m_c = n$, and for $\theta_1 < \ldots < \theta_c$ as in (42) for $\theta_i = \lambda_{m_i}$

$$G_t(\theta_1, \ldots, \theta_c) := \mathbb{P}\left(\sum_{i=1}^{c} \epsilon_i/\theta_i > t \right) = \sum_{i=1}^{c} \left\{ \prod_{j \neq i} \left(\frac{\theta_j}{\theta_j - \theta_i} \right) \right\} e^{-\theta_i t} \tag{43}$$

where $\epsilon_i, i = 1, 2, \ldots$ is a sequence of independent standard exponential variables. This calculation yields:

Proposition 29 *For a Λ-coalescent the probability that any particular set of n blocks in the initial partitition is contained in the same block at time t is*

$$p_t(n) = 1 - \sum_{r=2}^{n} a_{n,r} e^{-\lambda_r t} \tag{44}$$

where for each $2 \le r \le n$ the function $a_{n,r} := a_{n,r}(\lambda_2, \ldots, \lambda_n)$ is a rational function of $\lambda_2, \ldots, \lambda_n$, or of the first $n-2$ moments of Λ, which may be expressed as follows in terms of the λ_i and the $\lambda^{\#}_{b,k}$ which are just linear combinations of the λ_i:

$$a_{n,r} = \sum_{c=1}^{n-1} \sum_{(m_0,\ldots,m_c) \ni r} \prod_{i=1}^{c} \lambda^{\#}(m_i, m_i - m_{i-1} + 1) \frac{1(m_i \neq r)}{(\lambda_{m_i} - \lambda_r)} \tag{45}$$

where for each $c = 1, \ldots, n-1$ the inner sum is over the set of $\binom{n-2}{c-2}$ strictly increasing sequences of positive integers $(m_i, 0 \le i \le c)$ with $m_0 = 1, m_c = n$, and $m_i = r$ for some i, and the factor $1(m_i \neq r)/(\lambda_{m_i} - \lambda_r)$ equals $1/(\lambda_{m_i} - \lambda_r)$ if $m_i \neq r$ and 0 otherwise.

For $n = 2, 3$ it is easily checked that these formulae (44)-(45) are consistent with the previous formulae (39)-(40). For $n = 4$, after simplification using $\lambda_{3,2}/(\lambda_3 - \lambda_2) = 1/2$, it is found that the coefficients are

$$a_{4,2} = \frac{3}{2} + \frac{2\lambda_{4,3} + 3\lambda_{4,2}}{2(\lambda_4 - \lambda_2)}; \; a_{4,3} = -\frac{1}{2} - \frac{3}{2}\frac{\lambda_{4,2}}{(\lambda_4 - \lambda_3)}; \; a_{4,4} = -\frac{\lambda_{4,3}}{\lambda_4 - \lambda_2} + \frac{3\lambda_{3,2}\lambda_{4,2}}{(\lambda_4 - \lambda_2)(\lambda_4 - \lambda_3)} \tag{46}$$

For $\Lambda = U$,

$$\lambda_b = b - 1; \;\; \lambda^{\#}_{b,k} = \frac{b}{k(k-1)}, \tag{47}$$

so in particular $\lambda_{4,2} = 1/3$, $\lambda_{4,3} = 1/6$, $\lambda_{3,2} = 1/2$. Expression (44) for $n = 4$ with the substitutions (46)-(47) rather miraculously reduces to the following instance of (17):

$$p^U_t(4) = \frac{1}{3!}\left(6 - 11e^{-t} + 6e^{-2t} - e^{-3t}\right) = \frac{1}{3!}(1 - e^{-t})(2 - e^{-t})(3 - e^{-t}). \tag{48}$$

It does not seem at all evident from (45) why the substitutions (47) allow formula (44) to be factorized for every n, as implied by (17). As a consequence of (17), for $1 \leq j \leq n-1$ the coefficient of e^{-jt} in the generalization of (48) to n instead of 4 is $(-1)^j \left[{n \atop j+1} \right] /(n-1)!$ where $\left[{n \atop k} \right]$ is the Stirling number of the first kind which is the number of permutations of $[n]$ with k cycles. The only j for which this is obviously consistent with (44) is $j = 0$. Equate coefficents of e^{-jt} for $1 \leq j \leq n-1$ to see that the factorization implied by (17) amounts to the identity

$$a_{n,r} = (-1)^r \left[{n \atop r} \right] /(n-1)! \text{ for } 2 \leq r \leq n \text{ for } \lambda_b, \; \lambda^{\#}_{b,k} \text{ as in (47).} \tag{49}$$

3.9 Distribution of the frequencies

In this section, let $(\Pi_\infty(t), t \geq 0)$ be a standard F-coalescent for some probability distribution F on $[0, 1]$, with moments $\mu_r := \int_0^1 x^r F(dx)$. For $j = 1, 2, \ldots$ let $\tilde{f}_{j,t} := \tilde{f}_j(t)$ as in (11) be the frequencies of blocks of $\Pi_\infty(t)$ in order of least elements.

Let $\tilde{F}_t$ denote the probability distribution of $\tilde{f}_{1,t}$ on $[0, 1]$. As a consequence of known formulae for exchangeable random partitions and associated random discrete distributions [32, 31], this probability distribution $\tilde{F}_t$ on $[0, 1]$ carries a good deal of information about the distribution of $\Pi_\infty(t)$

and the corresponding distribution of ranked frequencies on $\bar{\mathcal{S}}^{\downarrow}$. The distribution $\tilde{F}_t$ is determined by the following formula for its moments [31, Corollary 8]:

$$\int_0^1 x^m \tilde{F}_t(dx) = \mathbb{E}(\tilde{f}_{1,t}^m) = p_t(m+1) \quad \text{for } m = 0, 1, \dots \tag{50}$$

where $p_t(n)$ is given by (44). In particular, $\tilde{f}_{1,t}$ has mean

$$\mathbb{E}(\tilde{f}_{1,t}) = p_t(2) = 1 - e^{-t} \tag{51}$$

no matter what F, and (40) gives the variance

$$\mathrm{Var}(\tilde{f}_{1,t}) = p_t(3) - (p_t(2))^2 = \tfrac{1}{2}e^{-t} + \tfrac{1}{2}e^{-\lambda_3 t} - e^{-2t} \tag{52}$$

where $\lambda_3 = 3 - 2\mu_1$ for μ_1 the mean of F. One other simple general feature of $\tilde{F}_t$ can be read from Proposition 26:

$$\tilde{F}_t\{0\} = \mathbb{P}(\tilde{f}_{1,t} = 0) = \exp(-\mu_{-1} t). \tag{53}$$

The following proposition records some basic properties of the process $(\tilde{f}_{1,t}, t \geq 0)$.

Proposition 30 *No matter what F, the process $(\tilde{f}_{1,t}, t \geq 0)$ is an increasing pure jump process with càdlàg paths, with*

$$\tilde{f}_{1,0} = 0 \text{ and } \tilde{f}_{1,\infty-} = 1 \text{ almost surely.} \tag{54}$$

If $\mu_{-1} = \infty$ then almost surely $\tilde{f}_{1,t} > 0$ for all $t > 0$ while $\tilde{f}_{1,0+} = 0$. Whereas if $\mu_{-1} < \infty$ the process $(\tilde{f}_{1,t}, t \geq 0)$ starts by holding at zero until an exponential time with rate μ_{-1}, when it enters $(0, 1]$ by a jump, and proceeds thereafter by a succession of holds and jumps, with holding rates bounded above by μ_{-1}.

Proof. It is obvious that the process $(\tilde{f}_{1,t}, t \geq 0)$ has increasing sample paths, and the facts (54) follow immediately from this and formula (51). If $\mu_{-1} = \infty$ then by Lemma 25 the exchangeable random partition $\Pi_\infty(t)$ has proper frequencies a.s. for each $t > 0$, hence $\tilde{f}_{1,t} > 0$ for all $t > 0$, and $\tilde{f}_{1,0+} = 0$ by the continuity of Kingman's correspondence. According to [15, Proposition 30] the process $(\tilde{f}_{1,t}, t \geq 0)$ has càdlàg paths, and by [15, Proposition 1] the paths are of pure jump type. If $\mu_{-1} < \infty$ the Poisson construction of Corollary 3 shows that the process $(\tilde{f}_{1,t}, t \geq 0)$ has càdlàg step function paths which can jump only when there is a collision involving 1. The rate of such jumps at time t is bounded above by μ_{-1}, and equal to μ_{-1} as long as 1 remains a singleton of $\Pi_\infty(t)$, meaning as long as $\tilde{f}_{1,t} = 0$. Let T be the time of the first collision involving 1, and let S_{T-} be the frequency of singletons in $\Pi_\infty(T-)$. Proposition 26 shows that $S_{T-} > 0$ almost surely, hence there is a relative frequency $X_1 > 0$ of classes involved in the collision, that is the common value of $X_{1,j}$ in Theorem 4 for all j involved in the collision. Since $\tilde{f}_{1,T} \geq S_{T-}X_{1,j} > 0$ the process $(\tilde{f}_{1,t}, t \geq 0)$ enters $(0,1]$ by a jump at time T. □

Suppose now that $\mu_{-1} = \infty$, so $\tilde{f}_{1,t}$ can be interpreted via Theorem 8 as the mass at time t containing a particle labeled 1 in the ranked mass coalescent derived from frequencies of the standard F-coalescent. The following proposition, which follows easily from the dynamics of Π_∞ described in Theorem 6, shows that the joint law of $\tilde{f}_{1,t}$ and $\tilde{f}_{1,t+u}$ is uniquely determined by the distribution of the sequence $(\tilde{f}_{i,t}, i = 1, 2, \ldots)$ and the distribution $\tilde{F}_u$ of $\tilde{f}_{1,u}$. In principle, a similar description could be given of the conditional distribution of $(\tilde{f}_{i,t+u}, i = 1, 2, \ldots)$ given $(\tilde{f}_{i,t}, i = 1, 2, \ldots)$, which would determine the Markovian dynamics of the *shunted* Λ*-coalescent*, defined as in [15], whose state at time t is $(\tilde{f}_{i,t}, i = 1, 2, \ldots)$.

Proposition 31 *Let Π_∞ be a standard Λ-coalescent for Λ with $\mu_{-1} = \infty$. Let $I_i(t, u)$ be the indicator of the event that the ith block of $\Pi_\infty(t)$, whose frequency is $\tilde{f}_{i,t}$, has merged with the block containing 1 by time $t + u$, so that*

$$\tilde{f}_{1,t+u} = \sum_{i=1}^{\infty} \tilde{f}_{i,t} I_i(t, u) \tag{55}$$

where $I_1(t, u) = 1$. Then the $(I_i(t, u), i = 2, 3, \ldots)$ are exchangeable$(\tilde{F}_u)$ indicators which are independent of the sequence $(\tilde{f}_{i,t}, i = 1, 2, \ldots)$.

As a check, take expectations in (55) and use $\mathbb{E}(\sum_{i=2}^{\infty} \tilde{f}_{i,t}) = 1 - \mathbb{E}(\tilde{f}_{1,t}) = e^{-t}$ and $\mathbb{E}(I_i(t, u)) = \mathbb{E}(\tilde{f}_{1,u}) = 1 - e^{-u}$ to recover $1 - e^{-(t+u)} = (1 - e^{-t}) + e^{-t}(1 - e^{-u})$. Less obvious identities can be obtained from the equality of higher moments in (55).

Applied to the U-coalescent, Proposition 31 yields the following result, which amounts to the description of two-dimensional distributions given in parts (i) and (ii) Corollary 16.

Proposition 32 *In the standard U-coalescent, for each $t, u > 0$ the joint law of $\tilde{f}_{1,t}$ and $\tilde{f}_{1,t+u}$ is determined as follows: the distribution of $\tilde{f}_{1,t}$ is* $\mathrm{beta}(1-e^{-t}, e^{-t})$, *and*

$$1 - \tilde{f}_{1,t+u} = (1 - \tilde{f}_{1,t}) Z_{t,u} \tag{56}$$

where $Z_{t,u}$ is independent of $\tilde{f}_{1,t}$ with $\mathrm{beta}(e^{-t} - e^{-(t+u)}, e^{-(t+u)})$ *distribution.*

Proof. The identification of the distribution $\tilde{F}_t$ of $\tilde{f}_{1,t}$ for the standard U-coalescent was indicated already in (19). Lemma 9 and (19) yield also that $\tilde{f}_{i,t} = (1 - \tilde{f}_{1,t}) \tilde{f}^*_{i-1,t}$ for $i = 2, 3, \ldots$ where $(\tilde{f}^*_{j,t}, j = 1, 2, \ldots)$ is the sequence of frequencies of blocks of an (e^{-t}, e^{-t}) partition, and this sequence is independent of $\tilde{f}_{1,t}$. A manipulation of (55) now yields (56) with

$$Z_{t,u} = 1 - \sum_{j=1}^{\infty} \tilde{f}^*_{j,t} I_{j+1}(t, u) \tag{57}$$

which is independent of $\tilde{f}_{1,t}$. Finally, the distribution of $Z_{t,u}$ is identified by a moment computation. □

In the proof of the previous proposition, the distribution of $Z_{t,u}$ displayed in (57) was identified. Take $\alpha = e^{-t}$ and $\beta = e^{-u}$ in (57) to deduce the following corollary:

Corollary 33 *For α and p in $(0,1)$ let $g_{\alpha,p}(x)$ denote the probability density at $x \in (0,1)$ of the random variable $\sum_{n=1}^{\infty} V_{n,\alpha} Y_{n,p}$ where the sequence $(V_{n,\alpha}, n = 1, 2, \ldots)$ has $PD(\alpha, \alpha)$ distribution, $(Y_{n,p}, n = 1, 2, \ldots)$ is a sequence of independent Bernoulli(p) indicators, and the two sequences are independent. Then for each $\alpha, x \in (0,1)$, the the function $p \mapsto g_{\alpha,p}(x)$ is characterized by the following formula: for $\beta \in (0,1)$*

$$\frac{1}{\Gamma(1-\beta)\Gamma(\beta)} \int_0^1 p^{-\beta} (1-p)^{\beta-1} g_{\alpha,p}(x) dp = \frac{\Gamma(\alpha)}{\Gamma(\alpha - \alpha\beta)\Gamma(\alpha\beta)} x^{\alpha - \alpha\beta - 1} (1-x)^{\alpha\beta - 1}. \tag{58}$$

That is to say, the mixture over p of the distribution of $\sum_n V_{n,\alpha} Y_{n,p}$, for p given a $\mathrm{beta}(1-\beta, \beta)$ distribution, is the $\mathrm{beta}(\alpha - \alpha\beta, \alpha\beta)$ distribution. For fixed α and x the left side of (58) is essentially a Mellin transform in β, so this formula determines the the function $p \mapsto g_{\alpha,p}(x)$ by uniqueness of

Mellin transforms. The probability density $x \mapsto g_{\alpha,p}(x)$ was characterized in a different way in [7, (4.b')] by the following formula: for $\lambda > 0$

$$\int_0^1 \frac{g_{\alpha,p}(x)dx}{(1+\lambda x)^\alpha} = \frac{1}{1+((1+\lambda)^\alpha - 1)p}. \tag{59}$$

This probability density $g_{\alpha,p}(x)$ is the density of the time spent positive by a skew-Bessel bridge $(b_t, 0 \leq t \leq 1)$ of dimension $2-2\alpha$, with $p = \mathbb{P}(X_t > 0)$ for each fixed $t \in (0,1)$. See also [36, §4]. For $\alpha = p = 1/2$ this is the density of the time spent positive by a standard Brownian bridge, which, according to a famous result of Lévy, is uniform. That is $g_{1/2,1/2}(x) = 1$ for all x. No explicit formula for $g_{\alpha,p}(x)$ seems to be known for other values of (α, p), but one should be obtainable from (58) by inversion of the Mellin transform. See also Tsilevich [41] for study of related distributions.

As a check on Corollary 33, formula (58) can be derived from (59) as follows. For $k = 0,1,2,\ldots$ a random variable $Z_{r,s}$ with beta(r,s) distribution has kth moment $\mathbb{E}(Z_{r,s}^k) = [r]_k/[r+s]_k$ where $[x]_k = x(x+1)\ldots(x+k-1)$. So for $\beta \in (0,1)$ and $|a| < 1$

$$\mathbb{E}\left(\frac{1}{1+aZ_{1-\beta,\beta}}\right) = \sum_{k=0}^{\infty}(-a)^k \mathbb{E}(Z_{1-\beta,\beta}^k) = \sum_{k=0}^{\infty}(-a)^k \frac{[1-\beta]_k}{k!} = (1+a)^{-(1-\beta)}.$$

Apply this with $a = (1+\lambda)^\alpha - 1$ to see that for Z whose distribution is the mixture of the distributions characterized by formula (59), for p given a beta$(1-\beta,\beta)$ distribution,

$$\mathbb{E}\left[\frac{1}{(1+\lambda Z)^\alpha}\right] = (1+\lambda)^{-(\alpha-\alpha\beta)}. \tag{60}$$

But also

$$\mathbb{E}\left[\frac{1}{(1+\lambda Z_{\alpha-\alpha\beta,\alpha\beta})^\alpha}\right] = \sum_{k=0}^{\infty}(-\lambda)^k \frac{[\alpha]_k}{k!}\frac{[\alpha-\alpha\beta]_k}{[\alpha]_k} = (1+\lambda)^{-(\alpha-\alpha\beta)}. \tag{61}$$

From the identity of these transforms it follows that Z has the same distribution as $Z_{\alpha-\alpha\beta,\alpha\beta}$, as claimed in Corollary 33.

4 Coagulation and fragmentation operations on the two parameter family

Theorem 12 will be proved after two lemmas which follow easily from Definitions 5 and 11.

Lemma 34 *Let Π^i_∞ for $i = 1,2$ be two random partitions of $\mathbb{N}$, with restrictions Π^i_n to $[n]$. Let p_1 and p be two exchangeable probability distributions on $\mathcal{P}_\infty$. Then the following two conditions are equivalent:*

(i) *Π^1_∞ is exchangeable with distribution p_1 and for all $\pi \in \mathcal{P}_\infty$*

$$\mathbb{P}(\Pi^2_\infty \in \cdot \mid \Pi^1_\infty = \pi) = p\text{-}\mathrm{COAG}(\pi, \cdot);$$

(ii) *for each $n = 1, 2, \ldots$ the joint law of (Π^1_n, Π^2_n) on $\mathcal{P}_n \times \mathcal{P}_n$ is given by the following formula: for each pair of partitions $\pi^1 := \{A_1, \ldots, A_K\}$ and $\pi^2 := \{B_1, \ldots, B_k\}$ of $[n]$ such that π^1 is a refinement of π^2, and $\#\{\ell : A_\ell \subseteq B_i\} = j_i$ for each $1 \le i \le k$*

$$\mathbb{P}(\Pi^1_n = \pi^1, \Pi^2_n = \pi^2) = p_1(a_1, \ldots, a_K)p(j_1, \ldots, j_k) \tag{62}$$

where $a_i = |A_i|$, and $\sum_{i=1}^k j_i = K$.

When the conditions of the lemma hold, Π^2_∞ is evidently exchangeable with EPF $p_2(n_1, \ldots, n_k)$ obtained for $(n_1, \ldots, n_k)$ with $\sum_i n_i = n$ by summing formula (62) over all $\pi^1 \in \mathcal{P}_n$ which are refinements of π^2 for any particular π^2 with $|B_i| = n_i$ for all i. In principle then, Lemma 34 describes the action of p-COAG on an arbitrary exchangeable distribution p_1. This induces an operation on the set of probability measures on $\bar{\mathcal{S}}^\downarrow$ via Kingman's correspondence. But this operation seems difficult to describe more explicitly. Similar remarks apply to p-FRAG instead of p-COAG, as a consequence of the following analog of the previous lemma. But the action of p-FRAG on $\bar{\mathcal{S}}^\downarrow$ is much simpler: this is just the operation described in a particular case in Corollary 13, and considered more generally as an action on $\mathcal{S}^\downarrow$ in [35].

Lemma 35 *Let Π^i_∞ for $i = 1,2$ be two random partitions of $\mathbb{N}$, with restrictions Π^i_n to $[n]$. Let p_2 and $\hat{p}$ be two exchangeable probability distributions on $\mathcal{P}_\infty$. Then the following two conditions are equivalent:*

(i) *Π^2_∞ is exchangeable with distribution p_2 and for all $\pi \in \mathcal{P}_\infty$*

$$\mathbb{P}(\Pi^1_\infty \in \cdot \mid \Pi^2_\infty = \pi) = \hat{p}\text{-}\mathrm{FRAG}(\pi, \cdot);$$

(ii) *for each $n = 1, 2, \ldots$ the joint law of (Π^1_n, Π^2_n) on $\mathcal{P}_n \times \mathcal{P}_n$ is given by the following formula: for each pair of partitions $\pi^1 := \{A_1, \ldots, A_K\}$ and $\pi^2 := \{B_1, \ldots, B_k\}$ of $[n]$ such that π^1 is a* *refinement of π^2 obtained by breaking each B_i into j_i blocks of sizes $a_{i,1}, \dots, a_{i,j_i}$ for some $j_i \geq 1$ with $\sum_{i=1}^k j_i = K$,*

$$\mathbb{P}(\Pi_n^1 = \pi^1, \Pi_n^2 = \pi^2) = \prod_{i=1}^k \hat{p}(a_{i,1}, \dots, a_{i,j_i}) p_2(b_1, \dots, b_k) \tag{63}$$

where $b_i = |B_i| = \sum_{\ell=1}^{j_i} a_{i,\ell}$.

Proof of Theorem 12. Let $\text{PAR} := \{(\alpha, \theta) : 0 \leq \alpha < 1, \theta > -\alpha\}$, which is the set of all (α, θ) such that formula (15) defines an EPF corresponding to a random discrete distribution with an infinite number of atoms almost surely. The argument yields the following sharper form of the result: *for $(\alpha, \theta), (\alpha_c, \theta_c), (\alpha_1, \theta_1)$ and (α_f, θ_f) in* PAR *the joint distribution of (Π, Π') defined by*

(i) Π is an (α, θ) partition and Π' is an (α_c, θ_c) coagulation of Π

is identical to the joint distribution of (Π, Π') defined instead by

(ii) Π' is an (α_1, θ_1) partition and Π is an (α_f, θ_f)-fragmentation of Π'

if and only if the parameters are of the form allowed in Theorem 12, that is

$$\alpha_f = \alpha; \quad \theta_f = -\alpha_1 = -\alpha\alpha_c; \quad \theta_c = \theta/\alpha, \quad \theta_1 = \theta. \tag{64}$$

Indeed, by application of Lemmas 34 and 35, and the formula (15) for the EPF of an (α, θ) partition, the two joint distributions in question are identical if and only if for all $1 \leq k \leq K \leq n < \infty$ and all choices of $(a_1, \dots, a_K)$, $(b_1, \dots, b_k)$ and $(j_1, \dots, j_k)$ subject to the obvious constraints,

$$\frac{[\theta/\alpha]_K}{[\theta]_n} \left(\prod_{\ell=1}^K -[-\alpha]_{a_\ell} \right) \frac{[\theta_c/\alpha_c]_k}{[\theta_c]_K} \prod_{i=1}^k -[-\alpha_c]_{j_i}$$
$$= \frac{[\theta_1/\alpha_1]_k}{[\theta_1]_n} \left(\prod_{\ell=1}^K -[-\alpha_f]_{a_\ell} \right) \prod_{i=1}^k \frac{-[-\alpha_1]_{b_i} [\theta_f/\alpha_f]_{j_i}}{[\theta_f]_{b_i}}.$$

If (64) holds this equality is evident by inspection. The necessity of (64) can be deduced using the fact that if $\prod_{i=1}^k [-\alpha]_{n_i} = c_{n,k} \prod_{i=1}^k [-\beta]_{n_i}$ for all $1 \leq k \leq n$ and all $(n_1, \dots, n_k)$ with $\sum_i n_i = n$ for some constants $c_{n,k}$, then $\alpha = \beta$. $\square$

Looking at (62) and (63), there appears to be little hope of matching the two formulae unless all the EPF's involved are of the Gibbs form

$$p(n_1, \dots, n_k) = \frac{b_k}{c_n} \prod_{i=1}^k w(n_i) \tag{65}$$

for some sequences of weights $(b_1, b_2, \ldots)$ and $(w_1, w_2, \ldots)$ and some sequence of normalization constants $(c_1, c_2, \ldots)$ determined by these weights. As shown by Kerov [21], the (α, θ) formula (15) and its limiting cases yield every EPF of this form. So it might be that Theorem 12 describes the only possible choices of non-degenerate laws p_1 and p of exchangeable random partitions of $\mathbb{N}$ such that the action of the p-COAG on p_1 to obtain p_2 can be inverted by $\hat{p}$-FRAG for some $\hat{p}$.

A Exchangeable Random Partitions

This appendix recalls the basic results of Kingman's theory of exchangeable random partitions, which are used throughout the paper.

Let Π be a random partition of $\mathbb{N}$ with restrictions $\mathtt{R}_n\Pi$ to $[n]$ for $n = 1, 2, \ldots$. Call Π *exchangeable* iff for each particular partition $\{B_1, \ldots, B_k\}$ of $[n]$ into k blocks, the probability $\mathbb{P}(\mathtt{R}_n\Pi = \{B_1, \ldots, B_k\})$ is a symmetric function of the sizes $n_1, \ldots, n_k$ of the sets $B_1, \ldots, B_k$, say

$$\mathbb{P}(\mathtt{R}_n\Pi = \{B_1, \ldots, B_k\}) = p(n_1, \ldots, n_k).$$

Then p is a non-negative symmetric function of sequences of positive integers $(n_1, \ldots, n_k)$ of arbitrary finite length, subject to $p(1) = 1$ and a sequence of addition rules with obvious probabilistic interpretations, the first few of which are

$$p(1) = p(2) + p(1,1); \quad p(2) = p(3) + p(2,1); \quad p(1,1) = 2p(2,1) + p(1,1,1). \tag{66}$$

Following [31], call p the *exchangeable probability function (EPF)* of Π. The same symbol p may denote the probability distribution of Π on $\mathcal{P}_\infty$. So $p(n_1, \ldots, n_k)$ is the p measure of the set $\{\pi \in \mathcal{P}_\infty : \mathtt{R}_n\pi = \pi_n\}$ for each particular partition π_n of $[n]$ into k blocks of sizes $(n_1, \ldots, n_k)$. For a sequence of random variables $Y_1, Y_2, \ldots$ let $\Pi(Y_1, Y_2, \ldots)$ denote the random partition of $\mathbb{N}$ whose blocks are the sets $\{i : Y_i = y\}$ as y ranges over all values of the Y_i. Let $\bar{\mathcal{S}}^{\downarrow}$ be the set of all non-negative sequences $\boldsymbol{x} = (x_1, x_2, \ldots)$ with $x_1 \geq x_2 \geq \ldots \geq 0$ and $\sum_i x_i \leq 1$, and give $\bar{\mathcal{S}}^{\downarrow}$ the topology of coordinatewise convergence. For $\boldsymbol{x} = (x_1, x_2, \ldots) \in \bar{\mathcal{S}}^{\downarrow}$ let $p_{\boldsymbol{x}}$ denote the distribution of $\Pi(Y_1, Y_2, \ldots)$ for Y_n that are independent and identically distributed according to some probability distribution on the line whose nth largest atom is x_n, and whose continuous component has probability $1 - \sum_n x_n$.

Theorem 36 (Kingman's correspondence [24, 25]). *A bijective correspondence $p \leftrightarrow P$ between probability distributions p of exchangeable random partitions of $\mathbb{N}$ and probability distributions P on $\bar{\mathcal{S}}^{\downarrow}$ is determined as follows. Each block B_n of an exchangeable random partition $\Pi = \{B_1, B_2, \ldots\}$ of $\mathbb{N}$ with distribution p has an almost sure limiting relative frequency $\tilde{f}_n$. Let P be the distribution on $\bar{\mathcal{S}}^{\downarrow}$ of the ranked rearrangement $\mathbf{f}$ of these frequencies. Then the conditional distribution of Π given $\mathbf{f} = \boldsymbol{x}$ is $p_{\boldsymbol{x}}$, so $p = \int_{\bar{\mathcal{S}}^{\downarrow}} P(d\boldsymbol{x}) p_{\boldsymbol{x}}$. This correspondence is continuous in the sense that a sequence of EPF's p_n has a pointwise limit p if and only if the corresponding sequence of probability distributions P_n on $\bar{\mathcal{S}}^{\downarrow}$ has a weak limit P, and then p corresponds to P.*

The most general distribution p of an exchangeable random partition of $\mathbb{N}$ is thus obtained as the distribution of $\Pi(Y_1, Y_2, \ldots)$ for a sequence of exchangeable random variables (Y_n). For, according to de Finetti's theorem, such $Y_1, Y_2, \ldots$ are conditionally independent with distribution G given some random probability distribution G. The corresponding P is then the distribution of ranked sizes of atoms of G. Aldous [2] gave a quick proof of Kingman's correspondence based on de Finetti's theorem. If the sequence of ranked frequencies $\mathbf{f} = (f_1, f_2, \ldots)$ of the exchangeable random partition Π is *proper*, meaning $\sum_n f_n = 1$ almost surely, the EPF p is determined by the distribution P of $\mathbf{f}$ via the formula

$$p(n_1, \ldots, n_k) = \sum_{(i_1, \ldots, i_k)} \mathbb{E}\left(\prod_{j=1}^{k} f_{i_j}^{n_j} \right) \tag{67}$$

where the sum is over all sequences of distinct positive integers $(i_1, \ldots, i_k)$. See also [31, 33] for simpler characterizations of the EPF in terms of the distribution of frequencies of blocks of Π in order of their least elements.

Acknowledgment

Thanks to Steven Evans and the referee for a number of helpful suggestions, and to Vlada Limic and Warren Ewens for comments and corrections.

References

[1] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25:812–854, 1997.

[2] D.J. Aldous. Exchangeability and related topics. In P.L. Hennequin, editor, *École d'Été de Probabilités de Saint-Flour XII, Springer Lecture Notes in Mathematics, Vol. 1117*. Springer-Verlag, 1985.

[3] D.J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5:3–48, 1999.

[4] D.J. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.*, 3:1–59, 1998.

[5] D.J. Aldous and J. Pitman. The standard additive coalescent. Technical Report 489, Dept. Statistics, U.C. Berkeley, 1997. To appear in *Ann. Probab.*. Available via http://www.stat.berkeley.edu/users/pitman.

[6] R. Arratia, A.D. Barbour, and S. Tavaré. Random combinatorial structures and prime factorizations. *Notices of the AMS*, 44:903–910, 1997.

[7] M. Barlow, J. Pitman, and M. Yor. Une extension multidimensionnelle de la loi de l'arc sinus. In *Séminaire de Probabilités XXIII*, pages 294–314. Springer, 1989. Lecture Notes in Math. 1372.

[8] J. Bertoin and J. F. Le Gall. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. Preprint, 1999.

[9] E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2):247–276, 1998.

[10] B. Derrida. Statistical properties of randomly broken objects and of multivalley structures in disordered systems. *J. Phys. A: Math. Gen.*, 20:5273–5288, 1987.

[11] B. Derrida. Non-self-averaging effects in sums of random variables, spin glasses, random maps and random walks. In M. Fannes, C. Maes, and A. Verbeure, editors, *On Three Levels: Micro-, Meso-, and Macro-Approaches in Physics*, NATO ASI Series, pages 125–137. Plenum Press, New York and London, 1994.

[12] K. Doksum. Tailfree and neutral random probabilities and their posterior distributions. *The Annals of Probability*, 2:183 – 201, 1974.

[13] E. B. Dynkin. *Markov Processes, Vol. I.* Springer, Berlin, 1965.

[14] E.B. Dynkin. Sufficient statistics and extreme points. *Ann. Probability*, 6:705–730, 1978.

[15] S.N. Evans and J. Pitman. Construction of Markovian coalescents. *Ann. Inst. Henri Poincaré*, 34:339–383, 1998.

[16] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.*, 3:87 – 112, 1972.

[17] W.J. Ewens. Population genetics theory – the past and the future. In S. Lessard, editor, *Mathematical and Statistical Problems in Evolution.* University of Montreal Press, Montreal, 1988.

[18] W. Feller. *An Introduction to Probability Theory and its Applications, Vol 2.* Wiley, 1966.

[19] T.S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1:209–230, 1973.

[20] N.L. Johnson, S. Kotz, and N. Balakrishnan. *Discrete Multivariate Distributions.* Wiley, New York, 1997.

[21] S. Kerov. Coherent random allocations and the Ewens-Pitman formula. PDMI Preprint, Steklov Math. Institute, St. Petersburg, 1995.

[22] J. F. C. Kingman. Random discrete distributions. *J. Roy. Statist. Soc. B*, 37:1–22, 1975.

[23] J. F. C. Kingman. Random partitions in population genetics. *Proc. R. Soc. Lond. A.*, 361:1–20, 1978.

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

[25] J. F. C. Kingman. The coalescent. *Stochastic Processes and their Applications*, 13:235–248, 1982.

[26] J.F. C. Kingman. Exchangeability and the evolution of large populations. In G. Koch and F. Spizzichino, editors, *Exchangeability in Probability and Statistics*. North-Holland Publishing Company, 1982.

[27] J.F.C. Kingman. On the genealogy of large populations. In J. Gani and E.J. Hannan, editors, *Essays in Statistical Science*, volume 19A of *J.A.P. Special vol.*, pages 27–43. Applied Probability Trust, 1982.

[28] A.A. Lushnikov. Coagulation in finite systems. *J. Colloid and Interface Science*, 65:276–285, 1978.

[29] A.H. Marcus. Stochastic coalescence. *Technometrics*, 10:133 – 143, 1968.

[30] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Th. Rel. Fields*, 92:21–39, 1992.

[31] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Th. Rel. Fields*, 102:145–158, 1995.

[32] J. Pitman. Random discrete distributions invariant under size-biased permutation. *Adv. Appl. Prob.*, 28:525–539, 1996.

[33] J. Pitman. Some developments of the Blackwell-MacQueen urn scheme. In T.S. Ferguson et al., editor, *Statistics, Probability and Game Theory; Papers in honor of David Blackwell*, volume 30 of *Lecture Notes-Monograph Series*, pages 245–267. Institute of Mathematical Statistics, Hayward, California, 1996.

[34] J. Pitman. Partition structures derived from Brownian motion and stable subordinators. *Bernoulli*, 3:79–96, 1997.

[35] J. Pitman and M. Yor. Random discrete distributions derived from self-similar random sets. *Electronic J. Probability*, 1:Paper 4, 1–28, 1996.

[36] J. Pitman and M. Yor. On the relative lengths of excursions derived from a stable subordinator. In *Séminaire de Probabilités XXXI*, pages 287–305. Springer, 1997. Lecture Notes in Math. 1655.

[37] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25:855–900, 1997.

[38] M. Rosenblatt. *Random Processes.* Springer-Verlag, New York, 1974.

[39] D. Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108:225–239, 1987.

[40] S. Tavaré. Line-of-descent and genealogical processes and their applications in population genetics. *Theoret. Population Biol.*, 26:119–164, 1984.

[41] N. V. Tsilevich. Distribution of the mean value for certain random measures. *Zapiski Nauchn. Sem. POMI*, 240:269–280, 1997.