CONSTRUCTION OF MARKOVIAN COALESCENTS

Steven N. Evans* and Jim Pitman†

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

Revised December 9, 1997

Abstract

Partition-valued and measure-valued coalescent Markov processes are constructed whose state describes the decomposition of a finite total mass m into a finite or countably infinite number of masses with sum m, and whose evolution is determined by the following intuitive prescription: each pair of masses of magnitudes x and y runs the risk of a binary collision to form a single mass of magnitude $x+y$ at rate $\kappa(x,y)$, for some non-negative, symmetric collision rate kernel $\kappa(x,y)$. Such processes with finitely many masses have been used to model polymerization, coagulation, condensation, and the evolution of galactic clusters by gravitational attraction. With a suitable choice of state space, and under appropriate restrictions on κ and the initial distribution of mass, it is shown that such processes can be constructed as Feller or Feller-like processes. A number of further results are obtained for the *additive coalescent* with collision kernel $\kappa(x,y) = x + y$. This process, which arises from the evolution of tree components in a random graph process, has asymptotic properties related to the stable subordinator of index 1/2.

*Research supported in part by Presidential Young Investigator Award and N.S.F. Grant DMS-9703845

†Research supported in part by N.S.F. Grants MCS94-04345 and DMS97-03961

Résumé

Cet article propose une construction des processus markoviens de coalescence dont l'espace d'état - un espace de mesures ou une partition ensembliste - décrit la décomposition d'une masse totale finie m en un ensemble fini ou dénombrable de masses dont la somme reste constante et égale à m, et dont l'évolution est déterminée par la règle suivante: chaque paire de masses de magnitudes x et y court le risque d'une collision binaire pour former une masse unique de magnitude $x + y$ avec un taux $\kappa(x, y)$ où κ est un noyau positif et symétrique décrivant le taux de collisions. De tels processus impliquant un nombre fini de masses ont servi de modèle à des phénomènes de polymérisation, de coagulation, de condensation ou encore pour décrire l'évolution d'amas galactiques sous l'influence du champ gravitationnel. Avec un espace d'état convenablement choisi, et sous réserve des restrictions adéquates sur κ et la distribution initiale de masse, on démontre que ces processus peuvent être construits comme des processus de Feller (ou similaires à ces processus). On obtient plusieurs autres résultats pour le processus de coalescence additive, dont le noyau est $\kappa(x, y) = x + y$. Ce processus, qui émerge de l'évolution des arbres au sein d'un processus de graphe aléatoire, a des propriétés asymptotiques liées au subordinateur stable d'indice 1/2.

Contents

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 [27]. Such models have been applied to chemical processes of polymerization [20], and other physical processes of coagulation and condensation such as the evolution of galactic clusters by gravitational attraction [40]. See Aldous [6] for a recent survey of the literature of these models and their relation to Smoluchowski's mean-field theory of coagulation phenomena.

While our interest in these models is mathematical, we use cosmological terms, and imagine a stochastic mechanism in which smaller galaxies merge through collisions to form larger galaxies. We suppose that at any given time, each pair of galaxies of masses say x and y runs the risk of a binary collision to form a single galaxy of mass $x+y$ at rate $\kappa(x,y)$, where κ is some non-negative, symmetric function. We write this intuitive prescription symbolically as

$$\{x,y\} \to x+y \text{ at rate } \kappa(x,y) \tag{1}$$

Assuming that the universe consists of a finite number of galaxies, each containing a finite number of particles of equal mass, the state of the universe is commonly represented as a *partition of* n, that is an unordered collection of positive integers with sum n, where n is the total number of particles in the universe. Transition rates between various partitions of n implied by (1) then determine the distribution of the state of the universe at time $t>0$ given some initial state at $t=0$ via the Kolmogorov forward equations for the finite-state Markov chain [29, 27, 20].

It is of interest in many settings to study limiting models in which $n\to\infty$. One limiting regime which has been extensively studied [20, 46, 45, 6] is the *thermodynamic limit*, in which the n particles are supposed to occupy some volume V, the collision rate is understood as a rate per unit time per unit volume, and n and V are allowed to tend to infinity in such a way that in the limit there is at each time t a deterministic density per unit volume of galaxies containing i particles, say $c_i(t)$ for $i=1,2,3,\ldots$. These densities then satisfy a system of differential equations known as *Smoluchowski's coagulation equations* [47]. In this limit, the resulting process is essentially deterministic rather than stochastic. Normal approximations to fluctuations of the concentrations in large finite volumes relative to means determined by the Smoluchowski equations have also been obtained [45, 16].

Our concern here is with a different limiting scheme, in which the number of interacting galaxies tends to infinity, but a fixed total mass m is maintained. After passage to the limit, the state at time t is a random decomposition of the total mass m into a countable number of masses with sum m. The problem is to construct a Markovian evolution of masses subject to the intuitive prescription of rates (1), allowing the interaction of a countably infinite number of masses instead of just a finite number. With appropriate assumptions on κ and the initial distribution of masses, we establish the existence of such a process, which we call a *κ-coalescent*, as a limit in distribution of a finite-state chain defined by a finite number of masses evolving with the same collision rate kernel κ. We assume throughout that our system has a finite total mass m. By scaling, we can assume $m = 1$. But see also Aldous [1], who obtains interesting results for the *multiplicative coalescent* with collision rate $\kappa(x, y) = xy$ in a system with infinite total mass.

Informally, we regard a κ-coalescent as an evolving family of agglomerating galaxies with total mass 1. The only distinguishing feature of a galaxy is its mass. However, to rigorize this notion it is convenient to label the galaxies present by elements of the set $\mathbb{N} := \{1, 2, \dots\}$, and to think of the κ-coalescent as taking values in the set $\mathcal{S}$ of probability measures on $\mathbb{N}$. Different labeling conventions then lead to different "codings" of essentially the same object as an $\mathcal{S}$-valued processes. This point of view is introduced in Section 2, where we formulate a general definition of an $\mathcal{S}$-valued coalescent process, and relate this definition to the Marcus-Lushnikov model.

Section 3 presents another formalization of coalescents as partition–valued processes. This perspective encompasses Kingman's coalescent [24]. Each block of the partition at time t represents a collection of initially present galaxies that have succesively merged by some sequence of binary coalescences into a single galaxy. Section 3.2 records some explicit formulae for the semigroup of the additive coalescent (that is, the κ–coalescent with $\kappa(x, y) = x + y$) viewed as a partition–valued Markov chain.

Partition–valued coalescent processes with infinitely many galaxies are constructed in Section 4. Various codings of measure–valued coalescent processes with infinitely many galaxies are then constructed in Section 5 as deterministic transformations of corresponding partition–valued processes.

Section 6.1 presents asymptotics as $n \to \infty$ for the additive coalescent when the initial state is n galaxies of mass $1/n$. These asymptotics provide one motivation for the rigorous construction of such an additive coalescent with an arbitrary initial state consisting a countable number of masses. Both multiplicative and additive coalescents arise combinatorially from the study of random graphs [1, 33], and this limiting regime arises naturally in that work.

Section 6.2 investigates a particularly interesting feature of two of the codings of the additive coalescent as a measure–valued process described in Section 5. For a large class of initial mass distributions we show that the asymptotic ratio between a remote tail of the mass distribution at some later time t and the corresponding tail at time 0 is $e^{-\gamma t}$ for a suitable constant γ. Consequently, the value of the time parameter can be reconstructed from the current and initial states of the process.

Section 7 records some connections between our approach to coalescent processes and Kingman's theory of exchangeable random partitions. We conclude in Section 8 by mentioning some open problems. See [18] for a treatment of infinitely–many–species analogues of the classical Lotka–Volterra equations which appear as hydrodynamic limits of the kinds of coalesent processes studied here, and [19, 36] for other recent developments.

Measure-valued Markov processes have recently been the subject of considerable study, particularly those that arise in population genetics (see, for instance [17, 14]). These processes have arisen as high-density limits of the empirical measure for a particle system in which there is some sort of Markovian motion of the individual particles combined with between-particle interactions involving a small number of particles. But in these models, even when the values of the limiting process are discrete measures, it is usually the case that mass moves between atoms in a continuous manner. By contrast, in the processes we study here, mass transfers occur by a purely discontinuous process.

2 Measure-valued coalescents

The term *coalescent* has been applied to various mathematical models for a system of masses evolving over time in such a way that smaller masses collide to form larger masses, with conservation of mass [29, 1, 6]. Kingman [24] developed a coalescent model in mathematical genetics to describe lines of descent in a large population. This section offers a general framework for coalescent processes which is adequate for the construction of Markovian coalescent processes with collision rate κ for a variety of kernels κ.

2.1 Partial orderings of measures on $\mathbb{N}$

We will describe the *state* of the system of coalescing galaxies at a given time by a sequence of non-negative components

$$\boldsymbol{x} := (x_1, x_2, \dots) \tag{2}$$

We interpret x_i as the mass of the ith galaxy in some inventory of galaxies. If $x_i = 0$ it is understood that there is no galaxy labeled i in this inventory. Set $\mathbb{N} := \{1, 2, \dots\}$. We regard $\boldsymbol{x}$ a measure on $\mathbb{N}$ by defining

$$x_I := \sum_{i \in I} x_i \tag{3}$$

for subsets I of $\mathbb{N}$. So x_I is the total mass of all galaxies with labels in the set I. We assume further that the total mass $x_{\mathbb{N}}$ is finite, and reduce by scaling to the case $x_{\mathbb{N}} = 1$. Thus the state space of our coalescent processes will be identified as the set $\mathcal{S}$ of all probability measures on $\mathbb{N}$, or some suitable subset of $\mathcal{S}$.

Given two states $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$, say $\boldsymbol{x}$ *is finer than* $\boldsymbol{y}$, or $\boldsymbol{y}$ *is coarser than* $\boldsymbol{x}$, and write $\boldsymbol{x} \lesssim \boldsymbol{y}$, if there is a map Ψ from $\mathbb{N}$ to $\mathbb{N}$ such that $\Psi(\boldsymbol{x}) = \boldsymbol{y}$. Here $\Psi(\boldsymbol{x})$ denotes the *push-forward of* $\boldsymbol{x}$ *by* Ψ, defined by

$$(\Psi(\boldsymbol{x}))_j = \sum_i x_i 1(\Psi(i) = j) \tag{4}$$

We call $\lesssim$ the relation of *refinement* on $\mathcal{S}$. Recall that a binary relation $\leq$ defined on a set S is called a *partial ordering* of S if $\leq$ is *reflexive* ($x \leq x$), *antisymmetric* ($x \leq y$ and $y \leq x$ implies $x = y$) and *transitive* ($x \leq y$ and $y \leq z$ implies $x \leq z$). The relation of refinement on $\mathcal{S}$ is reflexive and transitive, but not antisymmetric. Write $\boldsymbol{x} \sim \boldsymbol{y}$ and say $\boldsymbol{x}$ is a *rearrangement* of $\boldsymbol{y}$ if $\boldsymbol{x} \lesssim \boldsymbol{y}$ and $\boldsymbol{y} \lesssim \boldsymbol{x}$. Then $\sim$ is an equivalence relation on $\mathcal{S}$. It is easily seen that $\boldsymbol{x} \sim \boldsymbol{y}$ iff there exists a bijection $\beta : \{i : x_i > 0\} \to \{j : x_j > 0\}$ such that $x_i = y_{\beta(i)}$ for all i with $x_i > 0$. Each $\sim$ equivalence class has a unique representative

$\boldsymbol{y}$ which is *ranked*, meaning that $y_1 \geq y_2 \geq \ldots \geq 0$. Let $\mathcal{S}^{\downarrow}$ denote the subset of $\mathcal{S}$ comprising all ranked states, and define $\mathtt{RANK} : \mathcal{S} \to \mathcal{S}^{\downarrow}$ by $\mathtt{RANK}(\boldsymbol{x}) = \boldsymbol{x}^{\downarrow}$ where $\boldsymbol{x}^{\downarrow}$ is the unique ranked state that is a rearrangement of $\boldsymbol{x}$. The restriction of $\lesssim$ to $\mathcal{S}^{\downarrow}$ defines a partial ordering of $\mathcal{S}^{\downarrow}$. Let $\overset{sto}{\leq}$ denote the *stochastic ordering* on $\mathcal{S}$, that is the partial ordering of $\mathcal{S}$ defined by

$$\boldsymbol{x} \overset{sto}{\leq} \boldsymbol{y} \text{ iff } \textstyle\sum_{i=1}^{n} x_i \geq \sum_{i=1}^{n} y_i \text{ for all } n \in \mathbb{N}. \tag{5}$$

It can be shown that $\boldsymbol{x} \lesssim \boldsymbol{y}$ implies $\boldsymbol{y}^{\downarrow} \overset{sto}{\leq} \boldsymbol{x}^{\downarrow}$ but the converse is false. For example, $(2/3, 1/3, 0, \ldots)$ is stochastically smaller than, but not coarser than $(1/2, 1/2, 0, \ldots)$. Note also that if $\boldsymbol{x} \lesssim \boldsymbol{y}$ where $\boldsymbol{y} = \Psi(\boldsymbol{x})$ with $\Psi(k) \leq k$ for all $k \in \mathbb{N}$ (we call such a map *leftward*), then $\boldsymbol{y} \overset{sto}{\leq} \boldsymbol{x}$.

2.2 Coalescent evolutions and processes

Let the *time parameter set* $\mathbb{I} \subseteq \mathbb{R}$ be a possibly infinite interval which may be open or closed at either end. Consider a map $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ from $\mathbb{I}$ into some subset $\mathcal{S}'$ of $\mathcal{S}$. Write $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots)$. Interpret $x_i(t)$ as the mass of the galaxy labeled i at time t.

Given a topology on $\mathcal{S}'$, say that $(\boldsymbol{x}(t), t \in \mathbb{I})$ is an *$\mathcal{S}'$-coalescent evolution* if it is càdlàg and for some family of *tracking functions* $\Psi_{s,t} : \mathbb{N} \to \mathbb{N}$, $s, t \in \mathbb{I}$, $s < t$, satisfying the *composition rule*

$$\Psi_{s,u} = \Psi_{t,u} \circ \Psi_{s,t} \text{ for } s, t, u \in \mathbb{I} \text{ with } s < t < u \tag{6}$$

there is *conservation of mass*:

$$\boldsymbol{x}(t) = \Psi_{s,t}(\boldsymbol{x}(s)) \text{ for } s, t \in \mathbb{I} \text{ with } s < t\ . \tag{7}$$

In other words, for each pair of times $s, t \in \mathbb{I}$ with $s < t$, the mass $x_i(s)$ of each galaxy in existence at time s is identified as part of the mass $x_j(t)$ of some unique galaxy in existence at the subsequent time t, where $j = \Psi_{s,t}(i)$. The value of $\Psi_{s,t}(i)$ is of no significance if $x_i(s) = 0$. As a consequence of (7),

$$\boldsymbol{x}(s) \text{ is finer than } \boldsymbol{x}(t) \text{ whenever } s < t. \tag{8}$$

Call an $\mathcal{S}'$-coalescent evolution *leftward* if it admits a system of leftward tracking functions. It is not hard to show that if the topology on $\mathcal{S}'$ is at least as strong as the topology of weak convergence, and $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ is a càdlàg $\mathcal{S}'$-valued function such that (8) holds, then there is a leftward $\mathcal{S}'$-coalescent evolution $(\boldsymbol{y}(t),\, t \in \mathbb{I})$ such that $\boldsymbol{x}(t) \sim \boldsymbol{y}(t)$ for all $t \in \mathbb{I}$.

The notion of a family of tracking functions satisfying the composition rule captures mathematically the intuitive idea of a *merger history tree*, developed less formally in [26, 39, 40] in the cosmological setting. Loosely speaking, a leftward $\mathcal{S}'$-coalescent evolution describes the evolution of a universe using a labeling scheme such that when galaxies coalesce the label of the resulting galaxy is no greater than the labels of any of the participants in the coalescence.

For $\boldsymbol{x} \in \mathcal{S}$ and $1 \le i < j < \infty$ define $\boldsymbol{x}^{i \leftarrow j} \in \mathcal{S}$ by

$$\boldsymbol{x}^{i \leftarrow j} = \boldsymbol{y} \text{ where } y_i = x_i + x_j;\ y_j = 0 \text{ and } y_k = x_k \text{ for } k \notin \{i, j\} \tag{9}$$

Thinking of $\boldsymbol{x}$ as a sequence of masses placed on the positive integers, $\boldsymbol{x}^{i \leftarrow j}$ is derived from $\boldsymbol{x}$ by removing mass x_j from place j and adding it to the mass x_i at place i. Call an $\mathcal{S}'$-coalescent evolution $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ *basic* if it is leftward and for all $t \in \mathbb{I}$ with $\boldsymbol{x}(t-) \neq \boldsymbol{x}(t)$ there exist $i < j$ (depending on t) such that $\boldsymbol{x}(t) = \boldsymbol{x}(t-)^{i \leftarrow j}$. Call an $\mathcal{S}'$-coalescent evolution $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ *binary* if there exists a basic $\mathcal{S}'$-coalescent evolution $(\boldsymbol{y}(t),\, t \in \mathbb{I})$ such that $\boldsymbol{x}(t) \sim \boldsymbol{y}(t)$ for all $t \in \mathbb{I}$. Intuitively, a binary $\mathcal{S}'$-coalescent evolution describes the evolution of a universe in which galaxies only coalesce in pairs. Moreover, in the basic case the galaxies are labeled so that each new galaxy formed by a binary collision is given the smaller of the labels of the two colliding galaxies while the labels of all other galaxies remain unchanged.

It is clear that if $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ is a basic $\mathcal{S}'$-coalescent evolution, `RANK` maps $\mathcal{S}'$ into $\mathcal{S}''$ and both $\mathcal{S}'$ and $\mathcal{S}''$ are equipped with topologies that make `RANK` continuous, then $(\mathtt{RANK} \circ \boldsymbol{x}(t),\, t \in \mathbb{I})$ is a binary $\mathcal{S}''$-coalescent evolution. A similar comment holds with the map `RANK` replaced by the map `SHUNT`, where $\mathtt{SHUNT} : \mathcal{S} \to \mathcal{S}$ is the map that "squeezes out" 0 masses; for example, for $a, b, c, d > 0$ with $a + b + c + d = 1$

$$\mathtt{SHUNT}(0, a, 0, b, c, 0, d, 0, \dots) = (a, b, c, d, 0, \dots). \tag{10}$$

In this case the resulting $\mathcal{S}''$-coalescent evolution is leftward as well as binary.

Proposition 1 *Suppose that $(\boldsymbol{x}(t), t \in \mathbb{I})$ is an $\mathcal{S}'$-coalescent evolution for a topology on $\mathcal{S}' \subseteq \mathcal{S}$ which is at least as strong as the topology of weak convergence, and that either $\boldsymbol{x}(t) \in \mathcal{S}^{\downarrow}$ for each $t \in \mathbb{I}$, or that $(\boldsymbol{x}(t), t \in \mathbb{I})$ is leftward. Then for all $k \in \mathbb{N}$, the function $t \mapsto \sum_{\ell=k}^{\infty} x_\ell(t)$ is non-increasing, and the function $t \mapsto x_k(t)$ is of bounded variation, with total variation at most 2. Moreover, these functions have no continuous component (that is, they are pure jump functions).*

Proof. The following argument shows that the function $t \mapsto \sum_{\ell=k}^{\infty} x_\ell(t)$ is non-increasing with no continuous component, assuming that $(\boldsymbol{x}(t), t \in \mathbb{I})$ is a leftward $\mathcal{S}'$-coalescent. The rest is left to the reader. Without loss of generality, it can be supposed that $\mathbb{I} = \mathbb{R}_+$. Consider the sub-probability measure valued function $\boldsymbol{x}^{(n)}(t)$ defined by $\boldsymbol{x}^{(n)}(t) := \Psi_{0,t}(\boldsymbol{x}^{(n)}(0))$ where $\Psi_{0,t}$ is the tracking function such that $\boldsymbol{x}(t) = \Psi_{0,t}(\boldsymbol{x}(0))$ and $x_i^{(n)}(0) = x_i(0)1(i \leq n)$. For $k \in \mathbb{N}$, $n \in \mathbb{N}$ and $t \geq 0$ let

$$a_{k,n}(t) := \sum_{\ell=1}^{k-1} x_\ell^{(n)}(t) \tag{11}$$

Then $t \mapsto a_{k,n}(t)$ is an non-decreasing pure jump function and so is $t \mapsto a_{k,n+1}(t) - a_{k,n}(t)$ for each $n \in \mathbb{N}$. Since

$$\lim_{n\to\infty} a_{k,n}(t) = \sum_{\ell=1}^{k-1} x_\ell(t) = 1 - \sum_{\ell=k}^{\infty} x_\ell(t) \tag{12}$$

this function too is non-decreasing with no continuous component.

□

Given an $\mathcal{S}'$-valued stochastic process $(\boldsymbol{X}(t,\omega), t \in \mathbb{I}, \omega \in \Omega)$ defined on some probability space $(\Omega, \mathcal{F}, P)$, call the process an *$\mathcal{S}'$-coalescent* if for all $\omega \in \Omega$ the sample path $t \mapsto \boldsymbol{X}(t,\omega)$ is an $\mathcal{S}'$-coalescent evolution, and the associated tracking functions $\Psi_{s,t}^{\omega}$ from $\mathbb{N}$ to $\mathbb{N}$ can be chosen such that $\omega \mapsto \Psi_{s,t}^{\omega}(i)$ is $\mathcal{F}$-measurable for all $s, t \in \mathbb{I}$ and $i \in \mathbb{N}$. The measurability assumption means that for each pair of times s and t, and each pair of labels i and j, the event $\{\Psi_{s,t}(i) = j\}$, that the galaxy labeled i at time s is contained in the galaxy labeled j at time t, has a well defined probability $P\{\Psi_{s,t}(i) = j\}$.

2.3 Markovian $\mathcal{S}$-coalescents

For $\boldsymbol{x} \in \mathcal{S}$ let $\#\boldsymbol{x}$ be the number of non-zero components of $\boldsymbol{x}$. We interpret $\#\boldsymbol{x}$ as the number of galaxies present when the universe is in state $\boldsymbol{x} = (x_1, x_2, \dots)$. Let $\mathcal{S}^K := \{\boldsymbol{x} \in \mathcal{S} : \#\boldsymbol{x} < \infty\}$ be the set of finitely supported probability measures. Note that if $(\boldsymbol{x}(t), t \in \mathbb{I})$ is an $\mathcal{S}^K$-coalescent evolution, then the function $t \mapsto \boldsymbol{x}(t)$ has only finitely many jumps on each compact sub-interval of $\mathbb{I}$.

Let $\kappa(x, y)$ be a non-negative measurable symmetric function of $x, y \in]0,1[$. Call such a function κ a *collision kernel.* Call a process $\boldsymbol{X} := (\boldsymbol{X}(t), t \in \mathbb{I})$ an $(\mathcal{S}^K, \kappa)$*-coalescent* if $\boldsymbol{X}$ is a time-homogeneous Markov $\mathcal{S}^K$-coalescent process of jump-hold type, with state space $\mathcal{S}^K \cap \mathcal{S}'$ for some appropriate subset $\mathcal{S}'$ of $\mathcal{S}$, and transition rates of the form

$$\mathrm{rate}_\kappa(\boldsymbol{x} \to \boldsymbol{y}) = \sum_{i<j} 1(x_i > 0, x_j > 0, \boldsymbol{x}^{ij} = \boldsymbol{y})\kappa(x_i, x_j) \tag{13}$$

where $\boldsymbol{x}^{ij}$ is some rearrangement of $\boldsymbol{x}^{i \leftarrow j}$.

Three different $(\mathcal{S}^K, \kappa)$-coalescents, which we describe as *basic, shunted* and *ranked* are defined by the following choices of $\boldsymbol{x}^{ij}$ and $\mathcal{S}'$:

	$\boldsymbol{x}^{ij}$	$\mathcal{S}'$
basic	$\boldsymbol{x}^{i \leftarrow j}$	$\mathcal{S}^1 := \{\boldsymbol{x} \in \mathcal{S} : x_1 > 0\}$
shunted	$\mathtt{SHUNT}\, \boldsymbol{x}^{i \leftarrow j}$	$\mathcal{S}^* := \mathtt{SHUNT}(\mathcal{S})$
ranked	$\mathtt{RANK}\, \boldsymbol{x}^{i \leftarrow j}$	$\mathcal{S}^\downarrow := \mathtt{RANK}(\mathcal{S})$

(14)

It is easily shown that if $\boldsymbol{X}$ is any $(\mathcal{S}^K, \kappa)$-coalescent then $\mathtt{RANK} \circ \boldsymbol{X}$ is a ranked $(\mathcal{S}^K, \kappa)$-coalescent. So the various $(\mathcal{S}^K, \kappa)$-coalescents differ only in the way that galaxies are relabeled after collisions. Note also that if $\boldsymbol{X}$ is a basic $(\mathcal{S}^K, \kappa)$-coalescent, then the sample paths of $\boldsymbol{X}$ are basic $\mathcal{S}^K$-coalescent evolutions, and $\mathtt{SHUNT} \circ \boldsymbol{X}$ is a shunted $(\mathcal{S}^K, \kappa)$-coalescent whose sample paths are binary, leftward $\mathcal{S}^K$-coalescent evolutions. For $n \in \mathbb{N}$ let $\mathcal{S}^\downarrow_n$ denote the set of elements of $\mathcal{S}^\downarrow$ each of whose coordinates is a multiple of $1/n$. A ranked $(\mathcal{S}^K, \kappa)$-coalescent with time parameter set $\mathbb{I} = \mathbb{R}_+$ and initial state in $\mathcal{S}^\downarrow_n$ takes all of its values in $\mathcal{S}^\downarrow_n$, so we may call it an $(\mathcal{S}^\downarrow_n, \kappa)$-coalescent. By multiplication by n, the non-zero terms of an element of $\mathcal{S}^\downarrow_n$ identify a non-increasing sequence of positive integers with sum n, that is a *partition of* n. With $\mathcal{S}^\downarrow_n$ so identified with the set of all partitions of n, what we call here an $(\mathcal{S}^{\downarrow}_n, \kappa)$-coalescent is identical to the stochastic coalescent model of Marcus [29] and Lushnikov [27], with collision rate $\kappa(i/n, j/n)$ between each pair of galaxies of i and j particles.

Our aim now is to construct Markov processes which are appropriately continuous extensions of the basic, shunted and ranked coalescents, under suitable assumptions on the collision kernel κ. Ideally, we would like to extend the state space of the basic coalescent to all of $\mathcal{S}^1$, and that of the shunted and ranked coalescents to $\mathcal{S}^*$ and $\mathcal{S}^{\downarrow}$ respectively. We achieve this in the important special case of the *additive* kernel $\kappa(x, y) = K(x + y)$ for some constant $K > 0$. See also Example 7 regarding the case $\kappa(x, y) = K$, which is much more elementary. We consider also the case of a *Lipschitz kernel*, that is a κ subject to the conditions

$$\begin{gathered}\kappa(x, y) = \kappa(y, x), \;\; \kappa(0, 0) = 0 \text{ and} \\ |\kappa(a, b) - \kappa(c, d)| \leq K(|a - c| + |b - d|)\end{gathered} \tag{15}$$

for some constant K, where it is supposed that $\kappa(x, y)$ is defined if either x or y is equal to 0, though $\kappa(x, y)$ is not then interpreted as a jump rate. For a Lipschitz kernel, we are only able to extend the three coalescents to the state spaces $\mathcal{S}^1_1$, $\mathcal{S}^*_1$ and $\mathcal{S}^{\downarrow}_1$, where for $\beta \geq 0$ and $\mathcal{S}' \subseteq \mathcal{S}$ we set $\mathcal{S}'_\beta := \mathcal{S}' \cap \mathcal{S}_\beta$ with $\mathcal{S}_\beta$ the set of probability measures $\boldsymbol{x}$ on $\mathbb{N}$ such that $\sum_{k \in \mathbb{N}} k^\beta x_k < \infty$. For treatment of shunted and ranked coalescents we equip $\mathcal{S}_\beta$ with the metric

$$\delta_\beta(\boldsymbol{x}, \boldsymbol{y}) := \sum_{k \in \mathbb{N}} k^\beta |x_k - y_k| \tag{16}$$

Note that δ_0 is the restriction to $\mathcal{S}$ of the ℓ^1 metric, and the topology induced on $\mathcal{S}$ by δ_0 is the topology of weak convergence. For the basic coalescent we work with

$$\Delta_\beta(\boldsymbol{x}, \boldsymbol{y}) := \delta_\beta(\boldsymbol{x}, \boldsymbol{y}) + \sup_{k \in \mathbb{N}} 2^{-k} \{1(x_k = 0, y_k \neq 0) + 1(x_k \neq 0, y_k = 0)\}.$$

Note that $(\mathcal{S}_\beta, \delta_\beta)$ is a complete, separable, metric space for each $\beta \geq 0$, and that the topology induced by Δ_β is strictly stronger than that induced by δ_β. The space $\mathcal{S}_\beta$ with the topology induced by Δ_β is homeomorphic to

$$\{(\boldsymbol{x}, \varepsilon) \in \mathcal{S} \times \{0, 1\}^{\mathbb{N}} \;:\; x_i = 0 \text{ if and only if } \varepsilon_i = 0\}$$

when this set is equipped with the relative topology inherited from the product of the δ_β topology on $\mathcal{S}_\beta$ and the product topology on $\{0,1\}^{\mathbb{N}}$. In particular, $(\mathcal{S}_\beta, \Delta_\beta)$ is a Lusin space. Given a metric space (E, ζ), let $D(\mathbb{R}_+, E)$ or $D(\mathbb{R}_+, E, \zeta)$ denote the space of càdlàg functions from $\mathbb{R}_+$ into E equipped with the Skorohod topology.

Define kernels μ_κ, μ_κ^* and $\mu_\kappa^\downarrow$ on $\mathcal{S}^1$, $\mathcal{S}^*$ and $\mathcal{S}^\downarrow$ respectively, by setting $\mu_\kappa(\boldsymbol{x}, \{\boldsymbol{y}\})$, $\mu_\kappa^*(\boldsymbol{x}, \{\boldsymbol{y}\})$, and $\mu_\kappa^\downarrow(\boldsymbol{x}, \{\boldsymbol{y}\})$ equal to $\mathrm{rate}_\kappa(\boldsymbol{x} \to \boldsymbol{y})$ as in (13) for the appropriate choice of $\boldsymbol{x}^{ij}$ as in (14). The following theorem is proved in Section 5:

Theorem 2 *Suppose either that $\beta = 0$ and $\kappa(a,b) = K(a+b)$ for some constant $K \geq 0$, or that $\beta = 1$ and κ is subject to the Lipschitz condition (15). There exist Hunt processes $\boldsymbol{X}$, $\boldsymbol{X}^*$, and $\boldsymbol{X}^\downarrow$, with state-spaces $(\mathcal{S}_\beta^1, \Delta_\beta)$, $(\mathcal{S}_\beta^*, \delta_\beta)$, and $(\mathcal{S}_\beta^\downarrow, \delta_\beta)$, laws $(\mathbb{Q}^{\boldsymbol{x}}, \boldsymbol{x} \in \mathcal{S}_\beta^1)$, $(\mathbb{Q}_*^{\boldsymbol{x}}, \boldsymbol{x} \in \mathcal{S}_\beta^*)$, and $(\mathbb{Q}_\downarrow^{\boldsymbol{x}}, \boldsymbol{x} \in \mathcal{S}_\beta^\downarrow)$, and transition semigroups $(Q_t)_{t \geq 0}$, $(Q_t^*)_{t \geq 0}$, and $(Q_t^\downarrow)_{t \geq 0}$, such that the following hold.*

(i) Almost surely, the sample paths of $\boldsymbol{X}$ (resp. $\boldsymbol{X}^*$, $\boldsymbol{X}^\downarrow$) are basic $\mathcal{S}_\beta$-coalescent evolutions (resp. binary leftward $\mathcal{S}_\beta$-coalescent evolutions, binary $\mathcal{S}_\beta$-coalescent evolutions).

(ii) If $\boldsymbol{x} \in \mathcal{S}_\beta^1 \cap \mathcal{S}^K$ (resp. $\boldsymbol{x} \in \mathcal{S}_\beta^* \cap \mathcal{S}^K$, $\boldsymbol{x} \in \mathcal{S}_\beta^\downarrow \cap \mathcal{S}^K$), then $\boldsymbol{X}$ (resp. $\boldsymbol{X}^*$, $\boldsymbol{X}^\downarrow$) under $\mathbb{Q}^{\boldsymbol{x}}$ (resp. $\mathbb{Q}_*^{\boldsymbol{x}}$, $\mathbb{Q}_\downarrow^{\boldsymbol{x}}$) is a basic (resp. shunted, ranked) $(\mathcal{S}^K, \kappa)$-coalescent process.

(iii) The kernel μ_κ (resp. μ_κ^*, $\mu_\kappa^\downarrow$) is a jump kernel for $\boldsymbol{X}$ (resp. $\boldsymbol{X}^*$, $\boldsymbol{X}^\downarrow$).

(iv) The maps $\boldsymbol{x} \mapsto Q_t(\boldsymbol{x}, \cdot)$, $t \geq 0$, from $(\mathcal{S}_\beta^1, \Delta_\beta)$ into the space of probability measures on $(\mathcal{S}_\beta^1, \Delta_\beta)$ equipped with the topology of weak convergence, and the map $\boldsymbol{x} \mapsto \mathbb{Q}^{\boldsymbol{x}}$ from $(\mathcal{S}_\beta^1, \Delta_\beta)$ to the space of probability measures on $D(\mathbb{R}_+, \mathcal{S}_\beta^1, \Delta_\beta)$ equipped with the topology of weak convergence, are continuous. Analogous continuity results hold with $(\mathcal{S}_\beta^1, \Delta_\beta, Q_t, \mathbb{Q}^{\boldsymbol{x}})$ replaced by $(\mathcal{S}_\beta^*, \delta_\beta, Q_t^*, \mathbb{Q}_*^{\boldsymbol{x}})$ or $(\mathcal{S}_\beta^\downarrow, \delta_\beta, Q_t^\downarrow, \mathbb{Q}_\downarrow^{\boldsymbol{x}})$.

To spell out the meaning of (iii), the kernel called the jump kernel, together with the time parameter as a deterministic additive functional, form a Lévy system for the process in question [9]. For instance, for the basic coalescent, this means that the identity

$$\begin{aligned}\mathbb{Q}^{\boldsymbol{x}}&\left[\sum_{0\le s\le t} f(\boldsymbol{X}(s-),\boldsymbol{X}(s))1(\boldsymbol{X}(s-)\neq\boldsymbol{X}(s))\right]\\&=\mathbb{Q}^{\boldsymbol{x}}\left[\int_0^t\int f(\boldsymbol{X}(s-),\boldsymbol{y})\,\mu_\kappa(\boldsymbol{X}(s-),d\boldsymbol{x})\,ds\right]\end{aligned}\tag{17}$$

holds for all $\boldsymbol{x}\in\mathcal{S}^1_\beta$ and all non-negative measurable f.

The two properties (ii) and (iv) of the theorem uniquely specify each collection of laws $(\mathbb{Q}^{\boldsymbol{x}},\boldsymbol{x}\in\mathcal{S}^1)$, $(\mathbb{Q}^{\boldsymbol{x}}_*,\ \boldsymbol{x}\in\mathcal{S}^*)$, and $(\mathbb{Q}^{\boldsymbol{x}}_\downarrow,\ \boldsymbol{x}\in\mathcal{S}^\downarrow)$. Further path regularity properties of $\boldsymbol{X}$, $\boldsymbol{X}^*$, and $\boldsymbol{X}^\downarrow$ can be read from Proposition 1. It would be interesting to have a more direct characterization of the laws of these processes via a generator or a martingale problem, but we do not pursue that here.

3 Partition-valued coalescents

Let $(\boldsymbol{x}(t),\ t\in\mathbb{I})$ be an $\mathcal{S}$-coalescent evolution with associated tracking functions $(\Psi_{s,t})$. Note from the conservation of mass property (7) that if $s\in\mathbb{I}$, then $\boldsymbol{x}(t)$ for each $t\ge s$ can be recovered from $\boldsymbol{x}(s)$ and the sub-collection of pre-images $\Psi^{-1}_{s,t}(\{j\})$, $j\in\mathbb{N}$, that consists of the non-empty sets. Observe that the non-empty pre-images form a partition of the set $\mathbb{N}$. (Recall that *partition of a set* K is a collection $\{K_\alpha\}$ of non-empty subsets of K such that $K_\alpha\cap K_\beta=\emptyset$ for $\alpha\neq\beta$ and $\bigcup_\alpha K_\alpha=K$; the subsets K_α are called the *components* or *blocks* of the partition.) To construct various $\mathcal{S}$-coalescents, we first construct their associated processes of partitions of $\mathbb{N}$.

Every partition v of of a set K gives rise to an equivalence relation $\sim_v$ on K by declaring that $a\sim_v b$ if a and b belong to the same component. All equivalence relations on K arise this way. Given two partitions v and w of a set K, say that v is a *refinement* of w, or that w is a *coarsening of* v, and write $v\le w$, if every component of w is the union of one or more components of v; that is, $a\sim_v b$ implies $a\sim_w b$. If v is a partition of a set K and J is a subset of K, the *restriction of* v *to* J is the partition of J associated with the restriction to J of the equivalence relation associated with v.

For $f:K\to K$, the collection of subsets $\{f^{-1}(\{k\}):k\in K,\ f^{-1}(\{k\})\neq\emptyset\}$ is a partition of K that we call the *partition induced by* f. Note that for $g:K\to K$ the partition induced by $g \circ f$ is a coarsening of the partition induced by f. For example, if $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ is a $\mathcal{S}$-coalescent evolution with associated tracking functions $(\Psi_{s,t})$, then the composition rule implies that for $s, t, u \in \mathbb{I}$ with $s < t < u$, the partition of $\mathbb{N}$ induced by $\Psi_{s,u}$ is a coarsening of the partition induced by $\Psi_{s,t}$.

3.1 $\mathcal{P}_n$-coalescents

Let $\mathcal{P}_n$ denote the set of partitions of $\mathbb{N}_n := \{1, \dots, n\}$. Say that a $\mathcal{P}_n$-valued function $(w(t),\, t \in \mathbb{I})$ is *càdlàg* if it is right continuous with left limits in the discrete topology on $\mathcal{P}_n$. Say that $(w(t),\, t \in \mathbb{I})$ is a *$\mathcal{P}_n$-coalescent evolution* if it is càdlàg and $w(s) \leq w(t)$ for $s, t \in \mathbb{I}$, $s < t$. Finally, say that a $\mathcal{P}_n$-coalescent evolution $(w(t),\, t \in \mathbb{I})$ is *binary* if whenever $w(t-) \neq w(t)$ the partition $w(t)$ is obtained by coalescing two of the components of of $w(t-)$.

Call a $\mathcal{P}_n$-valued stochastic process $(W_t, t \in \mathbb{I})$ a *$\mathcal{P}_n$-coalescent* (resp. a *binary $\mathcal{P}_n$-coalescent*) if the sample paths are almost surely *$\mathcal{P}_n$-coalescent evolutions* (resp. *binary $\mathcal{P}_n$-coalescent evolutions*).

Let $\boldsymbol{p} := (p_1, \dots, p_n)$ be a sequence of strictly positive numbers. For $I \subseteq \mathbb{N}_n$ set $p_I = \sum_{i \in I} p_i$. Define a binary $\mathcal{P}_n$-coalescent Markov process $(W_t,\, t \geq 0)$ by specifying that two components I, J coalesce into a single component at rate $\kappa(p_I, p_J)$ for some symmetric non-negative collision kernel κ. Regard p_i as the mass of an i^{th} *proto-galaxy*, and interpret the components of W_t as the galaxies present at time t. Thus $(W_t,\, t \geq 0)$ describes a process in which pairs of galaxies coalesce at a rate depending on their masses. Call the $\mathcal{P}_n$-valued Markov process $(W_t,\, t \geq 0)$ the *$(\mathcal{P}_n, \kappa)$-coalescent with proto-galaxy mass distribution $\boldsymbol{p}$*, or the *$(\mathcal{P}_n, \kappa, \boldsymbol{p})$-coalescent* for short. Note that κ and $\boldsymbol{p}$ determine only the transition rates of the $(\mathcal{P}_n, \kappa, \boldsymbol{p})$-coalescent. The initial state W_0 can have any probability distribution over $\mathcal{P}_n$.

Suppose now that $(W_t,\, t \geq 0)$ is a $(\mathcal{P}_n, \kappa, \boldsymbol{p})$-coalescent for a $\boldsymbol{p}$ with $p_1 + \cdots + p_n = 1$. Define an $\mathcal{S}^K$-valued process $\boldsymbol{X} := (\boldsymbol{X}(t),\, t \geq 0)$ by setting $X_i(t) = p_J$ if W_t contains a component J whose least element is i, and $X_i(t) = 0$ otherwise. In particular, $X_i(t) = 0$ for $i > n$. It is easily checked that $\boldsymbol{X}$ is a basic $(\mathcal{S}^K, \kappa)$-coalescent, as defined in Section 2.3. Similarly, define another $\mathcal{S}^K$-valued process $\boldsymbol{X}^* := (\boldsymbol{X}^*(t),\, t \geq 0)$ by setting $X_1^*(t)$

equal to the mass of the galaxy at time t containing the proto-galaxy labeled 1, and $X_2^*(t)$ equal to the mass of the galaxy at time t containing the least numbered proto-galaxy not in the galaxy containing the proto-galaxy labeled 1, and so on. Then $\boldsymbol{X}^* = \texttt{SHUNT} \circ \boldsymbol{X}$ is a shunted $(\mathcal{S}^K, \kappa)$-coalescent. Of course $\boldsymbol{X}^\downarrow := \texttt{RANK} \circ \boldsymbol{X}$ is a ranked $(\mathcal{S}^K, \kappa)$-coalescent.

3.2 The finite additive coalescent.

Call a $(\mathcal{S}', \kappa)$-coalescent with $\mathcal{S}' \subseteq \mathcal{S}^K$ and $\kappa(x, y) = x + y$ an $(\mathcal{S}', +)$-coalescent. Similarly, call a $(\mathcal{P}_n, \kappa, \boldsymbol{p})$-coalescent with $\kappa(x, y) = x + y$ a $(\mathcal{P}_n, +, \boldsymbol{p})$-coalescent. Refer to both such processes as *additive coalescents*. The additive $\mathcal{S}_n^\downarrow$-coalescent has been studied by a number of authors [20, 27, 42, 43] as a particular case of the Marcus-Lushnikov model for which it is possible to make explicit calculations.

As observed by Hendriks et al. [20], the $(\mathcal{S}_n^\downarrow, \kappa)$-coalescent process with collision kernel $\kappa(x, y) = a + b(x + y)$ for constants a and b has the following property. For each $k \in \mathbb{N}$ and all states $\boldsymbol{x}$ with $\#\boldsymbol{x} = k$, the total rate of transitions out of state $\boldsymbol{x}$ and into the set of states $\{\boldsymbol{y} : \#\boldsymbol{y} = k - 1\}$ has the same value $\mu_{a,b}(k) := \binom{k}{2} a + (k-1)\, b$. It is easily seen that this property is shared by any $(\mathcal{S}^K, \kappa)$-coalescent with κ of this form. Consequently, if $(\boldsymbol{X}(t), t \geq 0)$ is such a coalescent, given $\#\boldsymbol{X}(0) = n$ for some fixed n, the process $(\#\boldsymbol{X}(t), t \geq 0)$ is a Markovian death process with death rates $\mu_{a,b}(k)$. Moreover, this death process is independent of the discrete-time jumping chain defined by the sequence of n distinct states through which $\boldsymbol{X}(t)$ passes as $\#\boldsymbol{X}(t)$ decreases by steps of 1 from n to 1.

In particular, if $(\boldsymbol{X}(t), t \geq 0)$ is an $(\mathcal{S}^K, +)$-coalescent, the death rate when there are k galaxies is $\mu_{0,1}(k) = k - 1$. It follows [44, §6.2.1] that given $\#\boldsymbol{X}(0) = \boldsymbol{x}$ for any $\boldsymbol{x}$ with $\#\boldsymbol{x} = m$ there is the identity in distribution

$$(\#\boldsymbol{X}(t) - 1,\, t \geq 0) \stackrel{d}{=} \left(\sum_{i=1}^{m-1} 1(\varepsilon_i > t),\, t \geq 0 \right) \tag{18}$$

where the ε_i are independent exponential variables with rate 1. Consequently [27, 20], the conditional distribution of $\#\boldsymbol{X}(t) - 1$ given $\boldsymbol{X}(0) = \boldsymbol{x}$ with $\#\boldsymbol{x} = m$ is binomial with parameters $m - 1$ and e^{-t}.

Consider now a $(\mathcal{P}_n, +, \boldsymbol{p})$-coalescent $W = (W(t), t \geq 0)$ constructed from the proto-galaxy masses $p_1, \cdots, p_n$ as in Section 3.1, assuming $\sum_i p_i = 1$. For W there is the following straightfoward extension of the results of the previous paragraph. Let $\#w$ denote the number of components of a partition w. Given $\#W(0) = m$ the process $(\#W(t) - 1,\, t \geq 0)$ has the same distribution as the process displayed in (18), and this process is independent of the sequence of distinct partitions embedded in the process $(W(t), t \geq 0)$. The sequence of distinct partitions is a discrete time $\mathcal{P}_n$-valued Markov chain whose one step transition probabilities are as follows: given that the current state is the partition $w = \{B_1, \ldots, B_k\}$ of $\mathbb{N}_n$, the next distinct partition is derived from w by merging the components B_i and B_j with probability $(p_{B_i} + p_{B_j})/(k-1)$. Such a Markov chain, call it a *discrete-time* $(\mathcal{P}_n, +, \boldsymbol{p})$*-coalescent*, will now be constructed following the method of [33].

For a finite or countable set S, call a subset $\mathbf{g}$ of $S \times S$ a *directed graph labeled by* S. Say (s,t) is an *edge of* $\mathbf{g}$ *directed from* s *to* t if $(s,t) \in \mathbf{g}$. Call $\mathbf{g}$ a *rooted forest* if each connected component of $\mathbf{g}$ is a rooted tree, with the convention that edges of a tree are directed away from its root.

Construction 3 Let $X_1, \ldots, X_{n-1}$ be independent random variables with distribution $\boldsymbol{p}$ on $\mathbb{N}_n$. Define a sequence of random rooted forests $(\mathcal{F}_k, 1 \leq k \leq n)$ in reverse order as follows. Let $\mathcal{F}_n$ be the rooted forest with vertex set $\mathbb{N}_n$ with no edges, and n trivial tree components. For $1 \leq k \leq n-1$, given that $\mathcal{F}_n, \ldots, \mathcal{F}_{k+1}$ have been defined so that $\mathcal{F}_{k+1}$ is a rooted forest of $k+1$ trees labeled by S, define $\mathcal{F}_k$ by addition to $\mathcal{F}_{k+1}$ of a single directed edge from X_k to R_k, where given $\mathcal{F}_{k+1}$ and X_k the vertex R_k is picked uniformly at random from the set of k roots of the k trees in $\mathcal{F}_{k+1}$ other than the tree containing X_k. For $1 \leq k \leq n$ let Π_k be the partition of $\mathbb{N}_n$ whose components are the connected components of $\mathcal{F}_k$.

It is obvious by construction that the sequence $(\Pi_n, \Pi_{n-1}, \ldots, \Pi_1)$ is a discrete-time $(\mathcal{P}_n, +, \boldsymbol{p})$-coalescent started at the partition of $\mathbb{N}_n$ into singletons. As shown by an induction in [33], for each $1 \le k \le n$ the forest $\mathcal{F}_k$ has the distribution

$$P\{\mathcal{F}_k = \mathbf{f}\} = \binom{n-1}{k-1}^{-1} \prod_{s=1}^{n} p_s^{C_s \mathbf{f}} \qquad (\mathbf{f} \in \mathbf{F}_{kn}) \tag{19}$$

where $C_s\mathbf{f}$ is the *number of children* or *out-degree* of vertex s in the rooted forest $\mathbf{f}$ and $\mathbf{F}_{kn}$ is the set of all forests of k rooted trees labeled by $\mathbb{N}_n$. Moreover, for each $1 \le k \le n-1$, conditionally given $(\mathcal{F}_j, 1 \le j \le k)$ the forest $\mathcal{F}_{k+1}$ is derived from $\mathcal{F}_k$ by deletion of (X_k, R_k), which is an edge picked uniformly at random from the set of $n-k$ edges of $\mathcal{F}_k$.

For $k = 1$ the fact that probabilities in the distribution (19) sum to 1 amounts to *Cayley's multinomial expansion over trees* [13, 38, 41, 33, 35]:

$$\sum_{\mathbf{t} \in \mathbf{T}_n} \prod_{s=1}^{n} p_s^{C_s \mathbf{t}} = \left(\sum_{s=1}^{n} p_s \right)^{n-1} \tag{20}$$

where $\mathbf{T}_n := \mathbf{F}_{1n}$ is the set of n^{n-1} rooted trees labeled by $\mathbb{N}_n$, and the formula holds as an identity of polynomials in n commuting variables $p_s, 1 \le s \le n$.

The probability of the event $\{\Pi_k = \{S_1, \ldots, S_k\}\}$ is obtained by summing the expression (19) over all forests $\mathbf{f}$ whose tree components are $S_1, \ldots, S_k$. Write the product over $\mathbb{N}_n$ in (19) as the product over $1 \le i \le k$ of products over S_i. The sum of products is then a product of sums, where the ith sum is a sum over all trees labeled by S_i. Each of these sums can be evaluated by Cayley's expansion (20) to obtain

$$P\{\Pi_k = \{S_1, \ldots, S_k\}\} = \binom{n-1}{k-1}^{-1} \prod_{i=1}^{k} p_{S_i}^{|S_i|-1} \tag{21}$$

where $p_A := \sum_{s \in A} p_s$ is the $\boldsymbol{p}$-mass of A and $|A|$ is the number of elements of A. We note as a consequence of this formula the following identity of polynomials in commuting variables $x_s, 1 \le s \le n$: for each $1 \le k \le n$

$$\sum_{\{S_1, \ldots, S_k\}} \prod_{\ell=1}^{k} x_{S_\ell}^{|S_\ell|-1} = \binom{n-1}{k-1} (x_1 + \cdots + x_n)^{n-k} \tag{22}$$

where the sum on the left side is over all unordered partitions of $\mathbb{N}_n$ into k components $\{S_1, \ldots, S_k\}$, and $x_S = \sum_{s \in S} x_s$. See [35] for a review of related identities and their probabilistic and combinatorial interpretations.

The following proposition combines the above results to give an explicit description of the semigroup of the $(\mathcal{P}_n, +, \boldsymbol{p})$-coalescent:

Proposition 4 *Let $(W(t), t \geq 0)$ be a $(\mathcal{P}_n, +, \boldsymbol{p})$-coalescent for a probability distribution $\boldsymbol{p}$. Then for each partition $w = \{R_1, \ldots, R_j\}$ of $\mathbb{N}_n$, and each partition $\{S_1, \ldots, S_k\}$ of $\mathbb{N}_n$ that is a coarsening of w,*

$$P\{W(t) = \{S_1, \ldots, S_k\} \,|\, W(0) = w\} = e^{-(k-1)t}(1 - e^{-t})^{j-k} \prod_{i=1}^{k} p_{S_i}^{\#\{h: R_h \subseteq S_i\} - 1}$$

Proof. Consider first the special case when w is the partition of $\mathbb{N}_n$ into n singletons. Write $\tilde{\Pi}_k$ for the state of $W(t)$ when $\#W(t) = k$. The result follows from the observations that $\#W(t) - 1$ has a binomial distribution with parameters $n-1$ and e^{-t}, the distribution of $\tilde{\Pi}_k$ is identical to that of Π_k displayed in (21), and $\#W(t)$ is independent of $(\tilde{\Pi}_k, 1 \leq k \leq n)$. For a general initial partition w, the only possible states of $W(t)$ are coarsenings of w. Every such coarsening is identified in an obvious way by a partition of $\mathbb{N}_j$. With this identification the $(\mathbb{N}_n, +, \boldsymbol{p})$ coalescent with initial partition w is identified with the $(\mathbb{N}_j, +, \boldsymbol{p}')$ coalescent with initial state the partitition of $\mathbb{N}_j$ into singletons, and $p'_i = p_{R_i}, i \in \mathbb{N}_j$. The general case then follows from the special case.

□

The following variation of Construction 3 will be used in Section 4.

Construction 5 Let $\boldsymbol{p} = (p_\sigma, \sigma \in \Sigma)$ be a finite measure on a finite or countable set Σ, with $p_\sigma > 0$ for all $\sigma \in \Sigma$. Let $(Y_j)_{j=0}^{\infty}$ and $(\varepsilon_\sigma)_{\sigma \in \Sigma}$ be independent random variables, with each Y_j distributed according to $\boldsymbol{p}$, and each ε_σ an exponential variable with rate 1. Let $\mathcal{T}((Y_j)_{j=0}^{\infty})$ be the random rooted tree with vertex set Σ and set of directed edges

$$\{(Y_{j-1}, Y_j) \ : \ Y_j \notin \{Y_0, \ldots, Y_{j-1}\}, \, j \geq 1\}.$$

For $t \geq 0$ let $\mathcal{F}(t; (Y_j)_{j=0}^{\infty}, (\varepsilon_\sigma)_{\sigma \in \Sigma})$, be the random forest with vertex set Σ and directed edge set

$$\{(Y_{j-1}, Y_j) \ : \ Y_j \notin \{Y_0, \ldots, Y_{j-1}\} \text{ and } \varepsilon_{Y_j} \leq t, \, j \geq 1\}.$$

Let $\Pi(t; (Y_j)_{j=0}^{\infty}, (\varepsilon_\sigma)_{\sigma \in \Sigma})$ be the random partition of Σ whose components are the tree components of $\mathcal{F}(t; (Y_j)_{j=0}^{\infty}, (\varepsilon_\sigma)_{\sigma \in \Sigma})$.

Proposition 6 *Suppose* $\Sigma = \mathbb{N}_n$. *Then* $(\Pi(t; (Y_j)_{j=0}^{\infty}, (\varepsilon_\sigma)_{\sigma\in\Sigma}),\ t \geq 0)$ *is a* $(\mathcal{P}_n, +, \boldsymbol{p})$*–coalescent starting from the partition that consists of all singletons.*

Proof. By application of the Markov chain tree theorem [12] [28, §6.1] to the Markov chain formed by a sequence of independent random variables with identical distribution $\boldsymbol{p}$, for each rooted tree $\mathbf{t}$ labeled by $\mathbb{N}_n$

$$P\{\mathcal{T}((Y_j)_{j=0}^{\infty}) = \mathbf{t}\} = c \prod_{s=1}^{n} p_s^{C_s \mathbf{t}}$$

for some constant c. Cayley's multinomial expansion (20) implies that $c = 1$, so the distribution of $\mathcal{T}((Y_j)_{j=0}^{\infty})$ is identical to the distribution of $\mathcal{F}_1$ displayed in (19). The conclusion now follows from the description of the process $(\mathcal{F}_k, 1 \leq k \leq n)$ given below (19), and the independence of the jump times and the discrete skeleton of a $(\mathcal{P}_n, +, \boldsymbol{p})$–coalescent.

□

3.3 $\mathcal{P}_\infty$-coalescents

Recall that $\mathcal{P}_n$ is the set of partitions of $\mathbb{N}_n = \{1, \ldots, n\}$. Let $\mathcal{P}_\infty$ denote the set of partitions of $\mathbb{N}$. For $m \in \mathbb{N}$ and $v \in \mathcal{P}_\infty$, or $v \in \mathcal{P}_n$ with $m \leq n$, write $\pi^m v$ for the restriction of v to $\mathbb{N}_m$. Topologize each $\mathcal{P}_n$ with the discrete topology and equip $\mathcal{P}_\infty$ with the topology generated by the maps $\{\pi^n\}_{n\in\mathbb{N}}$ (that is, the weakest topology with respect to which all of the maps π^n are continuous). Thus a $\mathcal{P}_\infty$-valued function $(w(t),\ t \in \mathbb{I})$ is càdlàg if and only if for all $n \in \mathbb{N}$ the $\mathcal{P}_n$-valued function $(\pi_n \circ w(t),\ t \in \mathbb{I})$ is càdlàg. Say that $(w(t),\ t \in \mathbb{I})$ is a $\mathcal{P}_\infty$*-coalescent evolution* (resp. a *binary* $\mathcal{P}_\infty$*-coalescent evolution*) if for all $n \in \mathbb{N}$ the $\mathcal{P}_n$-valued function $(\pi_n \circ w(t),\ t \in \mathbb{I})$ is a $\mathcal{P}_n$-coalescent evolution (resp. a binary $\mathcal{P}_n$-coalescent evolution). Of course, $(w(t),\ t \in \mathbb{I})$ is a binary $\mathcal{P}_\infty$-coalescent evolution if and only if whenever $w(t-) \neq w(t)$ the partition $w(t)$ is obtained by coalescing two of the components of $w(t-)$. We call a $\mathcal{P}_\infty$-valued stochastic process with such sample paths a $\mathcal{P}_\infty$*-coalescent* or a *binary* $\mathcal{P}_\infty$*-coalescent*, as the case may be.

Example 7 *Kingman's* $\mathcal{P}_\infty$*-coalescent.* Kingman [24, 25] proved the existence of an essentially unique $\mathcal{P}_\infty$-coalescent $(W_t, t \geq 0)$ starting at the partition of $\mathbb{N}$ into singletons such that for each n the process $(\pi^n \circ W_t, t \geq 0)$ is a Markovian $(\mathcal{P}_n, \kappa, \boldsymbol{p})$-coalescent with collision kernel $\kappa(x,y) \equiv 1$ and arbitrary weights $\boldsymbol{p}$. At each time $t > 0$, the partition W_t of $\mathbb{N}$ has an almost surely finite number of components D_t, where $(D_t, t > 0)$ is a pure death process with state space $\mathbb{N}$ coming down from $D_{0+} = \infty$ with an exponential hold at each $k \geq 2$ with mean $\binom{k}{2}^{-1}$ before jumping down to $k-1$. Kingman [24, Theorem 3 and §5], showed that for arbitrary $t > 0$ and $k \in \mathbb{N}$, conditionally given $D_t = k$, each of the k components of W_t has an asymptotic frequency; moreover if these frequencies are listed in ranked order as say $(X_1^{\downarrow}(t), X_2^{\downarrow}(t), \dots X_k^{\downarrow}(t))$ then the conditional distribution of this random vector given $(D_s, s \geq 0)$ with $D_t = k$ is identical to the distribution of the ranked lengths of k subintervals obtained by cutting the unit interval $]0,1[$ at $k-1$ points picked independently and uniformly at random from $]0,1[$. Let $(\mathbf{X}^{\downarrow}(t), t > 0)$ denote the $\mathcal{S}$-valued random process defined by

$$\mathbf{X}^{\downarrow}(t) = (X_1^{\downarrow}(t), X_2^{\downarrow}(t), \dots X_k^{\downarrow}(t), 0, 0, \dots) \text{ if } D_t = k \tag{23}$$

It follows easily from these results of Kingman that $(\mathbf{X}^{\downarrow}(t), t > 0)$ is a ranked $(\mathcal{S}^K, \kappa)$-coalescent process with $\kappa(x,y) = 1$. See Section 7 for a generalization of this construction. Another modification of this construction allows the definition of basic, shunted and ranked $(\mathcal{S}^K, \kappa)$-coalescents for this κ with time parameter $t > 0$, with entrance laws derived from an arbitrary distribution of mass at time $t = 0+$. See [36] for details in the ranked case.

In principle, the distribution of a $\mathcal{P}_\infty$-coalescent $(W_t,\, t \in \mathbb{I})$ is determined by the collection of finite-dimensional distributions of each of the finite state space $\mathcal{P}_n$-coalescents $(\pi^n \circ W_t,\, t \in \mathbb{I})$. The analysis of $(W_t,\, t \in \mathbb{I})$ is greatly simplified if each of these $\mathcal{P}_n$-coalescents is Markovian, as is the case for Kingman's coalescent and the more general class of coalescents considered in the next example. But this method does not extend easily to the kinds of $\mathcal{P}_\infty$-valued coalescents which are subject of this paper. So in the following sections we use other methods to construct these $\mathcal{P}_\infty$-coalescents and their associated $\mathcal{S}$-coalescents.

Example 8 *The Λ-coalescent.* It is shown in [36] that a large class of $\mathcal{P}_\infty$-valued Feller processes $(W_t,\, t \geq 0)$ is obtained by supposing that for each n the process $(\pi^n \circ W_t,\, t \geq 0)$

is Markovian with transition rates of the following form: when the partition of $\mathbb{N}_n$ has b components, each k-tuple of components is merging to form a single component at rate $\lambda_{b,k}$ for some array of rates $(\lambda_{b,k})$. This prescription turns out to be consistent if and only if $\lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx)$ for all $2 \le k \le b$ for some finite measure Λ on $[0,1]$. Kingman's coalescent is obtained for $\Lambda = \delta_0$, a unit mass at 0, while the coalescent recently derived by Bolthausen and Sznitman [11] from Ruelle's probability cascades, in the context of the Sherrington-Kirkpatrick spin glass model in mathematical physics, is obtained for Λ uniform on $[0,1]$. The Λ-coalescent has binary collisions only in Kingman's case $\Lambda = \delta_0$.

The following lemma collects together some simple facts about $\mathcal{P}_\infty$.

Lemma 9 *(i) The metric d on $\mathcal{P}_\infty$ defined by*

$$d(v,w) := \sup_{n\in\mathbb{N}} 2^{-n} 1(\pi^n v \neq \pi^n w)$$

induces the topology on $\mathcal{P}_\infty$.

(ii) The space $\mathcal{P}_\infty$ is compact and totally disconnected.

(iii) A sequence $\{v_k\}_{k\in\mathbb{N}}$ in $\mathcal{P}_\infty$ converges to v if and only if for each $n \in \mathbb{N}$, $\pi^n v_k = \pi^n v$ for all k sufficiently large.

(iv) The algebra of functions $\{f \circ \pi^m : f \in C(\mathcal{P}_m),\ m \in \mathbb{N}\}$ is dense in $C(\mathcal{P}_\infty)$.

4 Construction of infinite partition-valued coalescents

Our aim in this section is to construct for suitable probability distributions $\boldsymbol{p} \in \mathcal{S}^+ := \{\boldsymbol{x} \in \mathcal{S} : x_i > 0, \forall i \in \mathbb{N}\}$ and a suitable kernel κ a binary $\mathcal{P}_\infty$-coalescent Markov process $W_\kappa^{\boldsymbol{p}}$ such that two components I, J coalesce into the single component $I \cup J$ at rate $\kappa(p_I, p_J)$, where $p_I := \sum_{i\in I} p_i$. We usually regard κ as fixed, so the dependence of various objects on κ will be largely supressed in the notation. So we will write for instance $W^{\boldsymbol{p}}$ instead of $W^{\boldsymbol{p}}_{\kappa}$. As in section 3.1 we think of the components of the partition $W^{\boldsymbol{p}}(t)$ as the galaxies present at time t. The elements of a component are the labels of the proto-galaxies that have been merged together to form the galaxy. Define a kernel $\nu^{\boldsymbol{p}}_{\kappa}(w, dv)$ on $\mathcal{P}_\infty$ by declaring that $\nu^{\boldsymbol{p}}_{\kappa}(w, \cdot)$ is the measure that, for each unordered pair I, J of components of w, places mass $\kappa(p_I, p_J)$ on the coarsening of w obtained by coalescing I and J. Our aim then is to construct, for given κ, a Markovian coalescent process with jump kernel $\nu^{\boldsymbol{p}}_{\kappa}$ for as many $\boldsymbol{p} \in \mathcal{S}^+$ as possible.

Theorem 10 *Suppose either that $\kappa(a, b) = K(a + b)$ for some constant $K \geq 0$ and $\beta = 0$, or that κ is subject to the Lipschitz condition* (15) *and $\beta = 1$. For each $\boldsymbol{p} \in \mathcal{S}^+_\beta := \{\boldsymbol{p} \in \mathcal{S}^+ : \sum_k k^\beta p_k < \infty\}$ there is a unique $\mathcal{P}_\infty$-valued Feller process $W^{\boldsymbol{p}}$ with laws $(\mathbb{P}^w_{\boldsymbol{p}},\ w \in \mathcal{P}_\infty)$ such that the following hold.*

- **(i)** *If $w \in \mathcal{P}_\infty$ is such that $n \sim_w n+1 \sim_w n+2 \sim_w \dots$ for some $n \in \mathbb{N}$, then $\pi^n W^{\boldsymbol{p}}$ under $\mathbb{P}^w_{\boldsymbol{p}}$ is a $(\mathcal{P}_n, \kappa, \boldsymbol{p}^{[n]})$-coalescent started at $\pi^n w$, where $p^{[n]}_1 := p_1, \dots, p^{[n]}_{n-1} := p_{n-1}, p^{[n]}_n := \sum_{k=n}^\infty p_k$.*
- **(ii)** *Almost surely, the sample paths of $W^{\boldsymbol{p}}$ are binary $\mathcal{P}_\infty$-coalescent evolutions.*
- **(iii)** *The kernel $\nu^{\boldsymbol{p}}_{\kappa}$ is a jump kernel for $W^{\boldsymbol{p}}$.*

Definition 11 Call the process described in Theorem 10 the $(\mathcal{P}_\infty, \kappa, \boldsymbol{p})$-coalescent.

Note that for given κ and $\boldsymbol{p}$ the laws $(\mathbb{P}^w_{\boldsymbol{p}},\ w \in \mathcal{P}_\infty)$ are uniquely specified by κ and $\boldsymbol{p}$ through part (i) of the theorem and the Feller property. As with Theorem 2, it would be interesting to have a more direct generator or martingale problem characterization of $W^{\boldsymbol{p}}$.

Theorem 10 is a consequence of coupling arguments carried out in Lemmas 14, 16 and 17. Central to these arguments is the following set–up.

Definition 12 A *coupled family of coalescents* is the following collection of ingredients:

- a collision kernel $\kappa : [0, 1]^2 \to \mathbb{R}_+$;

- a subset $\mathcal{S}'$ of $\mathcal{S}^+$;
- for each $\boldsymbol{p} \in \mathcal{S}'$ and $n \in \mathbb{N}$ an associated sub-probability measure $\boldsymbol{p}^{(n)}$ on $\mathbb{N}_n$;
- some fixed, complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$;
- for each $n \in \mathbb{N}$, $w \in \mathcal{P}_\infty$, and $\boldsymbol{p} \in \mathcal{S}'$, a $(\mathcal{P}_n, \kappa, \boldsymbol{p}^{(n)})$-coalescent $W^{n,\boldsymbol{p},w} := (W^{n,\boldsymbol{p},w}(t), t \geq 0)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with $W^{n,\boldsymbol{p},w}(0) = \pi^n w$.

For $m \in \mathbb{N}$, $w \in \mathcal{P}_\infty$, $\boldsymbol{p} \in \mathcal{S}'$, and $t \geq 0$ set

$$N(m, w, \boldsymbol{p}, t) := \inf\{N \geq m : \pi^m W^{n,\boldsymbol{p},w}(s) = \pi^m W^{N,\boldsymbol{p},w}(s), \forall s \in [0,t], \forall n \geq N\}.$$

Definition 13 For $v \in \mathcal{P}_n$ let $\lambda^n v \in \mathcal{P}_\infty$ denote the unique partition of $\mathbb{N}$ that has $\{n+1, n+2, \dots\}$ as a component and satisfies $\pi^n \lambda^n v = v$. In other words,

$$i \sim_{\lambda^n v} j \Leftrightarrow \text{ either } (i, j \in \mathbb{N}_n \text{ and } i \sim_v j) \text{ or } (i, j \in \mathbb{N} \backslash \mathbb{N}_n).$$

Lemma 14 *Consider a coupled family of coalescents. Suppose that the following conditions hold.*

(a) *The collision kernel κ is symmetric and continuous on $[0,1]^2$.*

(b) *For $n, m \in \mathbb{N}$, $\boldsymbol{p} \in \mathcal{S}'$, and $t \geq 0$ there is a constant $\Gamma(n, m, \boldsymbol{p}, t)$ such that $\mathbb{P}\{N(m, w, \boldsymbol{p}, t) > n\} \leq \Gamma(n, m, \boldsymbol{p}, t)$ for all $w \in \mathcal{P}_\infty$ and $\lim_{n\to\infty} \Gamma(n, m, \boldsymbol{p}, t) = 0$ for all fixed $m, \boldsymbol{p}, t$.*

(c) *For all $\boldsymbol{p} \in \mathcal{S}'$, $\lim_{n\to\infty} \sum_{k=1}^n |p_k - p_k^{(n)}| = 0$.*

Then the following conclusions hold.

(i) *For each $w \in \mathcal{P}_\infty$ and $\boldsymbol{p} \in \mathcal{S}'$ there is a càdlàg $\mathcal{P}_\infty$–valued process $(W^{\boldsymbol{p},w}(t),\ t \geq 0)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$\lim_{n\to\infty} \sup_{0 \leq s \leq t} d(\lambda^n W^{n,\boldsymbol{p},w}(s), W^{\boldsymbol{p},w}(s)) = 0 \text{ a.s.} \tag{24}$$

(ii) *Let $\mathbb{P}_{\boldsymbol{p}}^w$ be the law of $W^{\boldsymbol{p},w}$. For each $\boldsymbol{p} \in \mathcal{S}'$ the collection of laws $(\mathbb{P}_{\boldsymbol{p}}^w,\, w \in \mathcal{P}_\infty)$ is that of a $\mathcal{P}_\infty$-valued Feller process $W^{\boldsymbol{p}}$ such that conclusions* (i)-(iii) *of Theorem 10 hold.*

Proof. By hypothesis (b), for each $w \in \mathcal{P}_\infty$ and $\boldsymbol{p} \in \mathcal{S}'$ there exists a set $\Omega^*(w, \boldsymbol{p}) \subseteq \Omega$ with $\mathbb{P}(\Omega^*(w, \boldsymbol{p})) = 1$ such that if $\omega \in \Omega^*(w, \boldsymbol{p})$ then $N(\omega, m, w, \boldsymbol{p}, t) < \infty$ for all $m \in \mathbb{N}$ and all $t \geq 0$. Define a càdlàg $\mathcal{P}_\infty$–valued process $(W^{\boldsymbol{p},w}(t),\, t \geq 0)$ by declaring that

$$\pi^m W^{\boldsymbol{p},w}(\omega, t) = \pi^m W^{N(\omega,m,w,\boldsymbol{p},s),\boldsymbol{p},w}(\omega, t),\ m \in \mathbb{N},\ \text{for } \omega \in \Omega^*(w, \boldsymbol{p})$$

and $W^{\boldsymbol{p},w}(\omega, t) = w$ for $\omega \notin \Omega^*(w, \boldsymbol{p})$. It is immediate that part (i) of the lemma holds, and for $N \geq m$ we have the bound

$$\begin{aligned}
&\mathbb{P}\{\sup_{n \geq N} \sup_{0 \leq s \leq t} d(\lambda^n W^{n,\boldsymbol{p},w}(s), W^{\boldsymbol{p},w}(s)) > 2^{-(m+1)}\} \\
&\quad = \mathbb{P}\{\exists n \geq N,\ \exists 0 \leq s \leq t \ :\ \pi^m \lambda^n W^{n,\boldsymbol{p},w}(s) \neq \pi^m W^{\boldsymbol{p},w}(s)\} \\
&\quad = \mathbb{P}\{\exists n \geq N,\ \exists 0 \leq s \leq t \ :\ \pi^m W^{n,\boldsymbol{p},w}(s) \neq \pi^m W^{\boldsymbol{p},w}(s)\} \\
&\quad \leq \mathbb{P}\{N(m, w, \boldsymbol{p}, t) > N\} \leq \Gamma(N, m, \boldsymbol{p}, t).
\end{aligned} \tag{25}$$

let $(P_t^{n,\boldsymbol{p}})_{t \geq 0}$ denote the semigroup of the $(\mathcal{P}_n, \kappa, \boldsymbol{p}^{(n)})$-coalescent. For $f \in C(\mathcal{P}_\infty)$ set $P_t^{\boldsymbol{p}} f(w) := \mathbb{P}[f(W^{\boldsymbol{p},w}(t))]$. We claim that $P_t^{\boldsymbol{p}} f \in C(\mathcal{P}_\infty)$ and the function $(P_t^{n,\boldsymbol{p}}(f \circ \lambda^n)) \circ \pi^n$ converges in $C(\mathcal{P}_\infty)$ to $P_t^{\boldsymbol{p}} f$ as $n \to \infty$. From Lemma 9 it suffices to take f of the form $g \circ \pi^m$ for some $m \in \mathbb{N}$ and $g \in C(\mathcal{P}_m)$. Observe that in this case $f \circ \lambda^n = g \circ \pi^m$ for $n \geq m$. We know from the above that the uniformly bounded sequence of functions $(P_t^{n,\boldsymbol{p}}(g \circ \pi^m) \circ \pi^n)_{n \geq m}$ converges pointwise as $n \to \infty$, and so it will suffice, by the Arzela-Ascoli theorem, to show that this sequence is equicontinuous. Suppose that $w, v \in \mathcal{P}_\infty$ are such that $d(w, v) \leq 2^{-(N+1)}$ for some $N \geq m$ so that $\pi^N w = \pi^N v$. Of course,

$$P_t^{N,\boldsymbol{p}}(g \circ \pi^m)(\pi^N w) = P_t^{N,\boldsymbol{p}}(g \circ \pi^m)(\pi^N v).$$

We have for $n \geq N$

$$\begin{aligned}
|P_t^{n,\boldsymbol{p}}(g \circ \pi^m)(\pi^n w) - P_t^{N,\boldsymbol{p}}(g \circ \pi^m)(\pi^N w)| &\leq 2\|g\|_{C(\mathcal{P}_m)} \mathbb{P}\{N(m, w, \boldsymbol{p}, t) > N\} \\
&\leq 2\|g\|_{C(\mathcal{P}_m)} \Gamma(N, m, \boldsymbol{p}, t),
\end{aligned}$$

and the same bound holds with w replaced by v. Therefore,

$$|P_t^{n,\boldsymbol{p}}(g\circ\pi^m)(\pi^n w) - P_t^{n,\boldsymbol{p}}(g\circ\pi^m)(\pi^n v)| \le 4\|g\|_{C(\mathcal{P}_m)}\Gamma(N,m,\boldsymbol{p},t),$$

which establishes the required equicontinuity.

For $f_1, f_2 \in C(\mathcal{P}_\infty)$ and $0 \le t_1 \le t_2$, we have from the uniform convergence established above that

$$\begin{aligned}
&\mathbb{P}[f_1(W^{\boldsymbol{p},w}(t_1)) \times f_2(W^{\boldsymbol{p},w}(t_2))] \\
&\quad = \lim_{n\to\infty} \mathbb{P}[f_1(\lambda^n W^{n,\boldsymbol{p},w}(t_1)) \times f_2(\lambda^n W^{n,\boldsymbol{p},w}(t_2))] \\
&\quad = \lim_{n\to\infty} P_{t_1}^{n,\boldsymbol{p}}((f_1\circ\lambda^n) \times P_{t_2-t_1}^{n,\boldsymbol{p}}(f_2\circ\lambda^n))(\pi^n w) \\
&\quad = \lim_{n\to\infty} P_{t_1}^{n,\boldsymbol{p}}((f_1 \times (P_{t_2-t_1}^{n,\boldsymbol{p}}(f_2\circ\lambda^n)\circ\pi^n))\circ\lambda^n)(\pi^n w) \\
&\quad = P_{t_1}^{\boldsymbol{p}}(f_1 \times P_{t_2-t_1}^{\boldsymbol{p}} f_2)(w)
\end{aligned}$$

Similar expressions hold for the higher order finite dimensional distributions of $(W^{\boldsymbol{p},w}(t),\, t \ge 0)$, and hence this process is time-homogeneous Markov with semigroup $(P_t^{\boldsymbol{p}})_{t\ge 0}$. In order to complete the proof that this process is Feller, it suffices to show that $\lim_{t\downarrow 0} P_t^{\boldsymbol{p}} f = f$ pointwise for all $f \in C(\mathcal{P}_\infty)$ (see [10, Remark after Theorem I.9.4]). This, however, is immediate from the right–continuity of the sample paths of $(W^{\boldsymbol{p},w}(t),\, t \ge 0)$.

We now move on to the proof that parts (i)-(iii) of Theorem 10 hold. Let w be as in the statement of part (i) of Theorem 10. For $n' \ge n$ it is easily verified that $\pi^n W^{n',\boldsymbol{p},w}$ is a $(\mathcal{P}_n, \kappa, \boldsymbol{p}^{[n,n']})$-coalescent started at $\pi^n w$, where $p_1^{[n,n']} := p_1^{(n')}, \dots, p_{n-1}^{[n,n']} := p_{n-1}^{(n')}$, $p_n^{[n,n']} := \sum_{k=n}^{n'} p_k^{(n')}$. We know from part (i) of the lemma that $\pi^n W^{\boldsymbol{p},w} = \lim_{n'\to\infty} \pi^n \lambda^{n'} W^{n',\boldsymbol{p},w} = \lim_{n'\to\infty} \pi^n W^{n',\boldsymbol{p},w}$ in $D(\mathbb{R}_+, \mathcal{P}_n)$. On the other hand, for a given starting state, the law of a Markov chain on a finite state space is weakly continuous with respect to its transition rates, and so part (i) of Theorem 10 follows from hypotheses (a) and (c).

To establish part (iii) of Theorem 10, it must be shown that for each $\boldsymbol{p} \in \mathcal{S}'$ and $w \in \mathcal{P}_\infty$ the process $W(s) := W^{\boldsymbol{p},w}(s)$, $s \ge 0$, is such that for every non-negative Borel function f,

$$\mathbb{P}\left[\sum_{0\le s\le t} f(W(s-), W(s))\, 1(W(s-) \ne W(s))\right] = \mathbb{P}\left[\int_0^t \int f(W(s-), v)\, \nu_\kappa^{\boldsymbol{p}}(W(s-), dv)\, ds\right].$$

Note for $M \in \mathbb{N}$ that

$$\sum_{0 \le s \le t} 1(d(W(s-), W(s)) > 2^{-M}) \le M$$

and

$$\nu_\kappa^{\boldsymbol{p}}(v, \{v' \in \mathcal{P}_\infty : d(v, v') > 2^{-M}\}) \le \binom{M}{2} \sup\{\kappa(a,b) : 0 \le a, b \le 1\}$$

for all $v \in \mathcal{P}_\infty$.

Similar bounds hold for $W^{n,\boldsymbol{p},w}$ instead of W. By a passage to the limit as $n \to \infty$ from the corresponding identity for $W^{n,\boldsymbol{p},w}$, the identity for W holds for f of the form $f(v, v') = g(\pi^m v, \pi^m v')\, 1(d(v, v') > 2^{-M})$ with $g \in C(\mathcal{P}_m \times \mathcal{P}_m)$ for some $m \in \mathbb{N}$. The identity for $f(v, v') = h(v, v') 1(d(v, v') > 2^{-M})$ for bounded non-negative Borel h is then obtained by a monotone class argument. The identity for general non-negative Borel f follows by monotone convergence as $M \to \infty$. Finally, it is immediate from the construction that we have built a process whose paths are $\mathcal{P}_\infty$–coalescent evolutions. It follows easily from property (iii) of Theorem 10 that all coalescences are binary, and so property (ii) of Theorem 10 also holds.

□

The next lemma is a refinement of the previous one. While not required for the proof of Theorem 10, this lemma will be used in the proof of Theorem 2.

Lemma 15 *Suppose the conditions of Lemma 14 hold for $\mathcal{S}' = \mathcal{S}_\beta^+$ for some $\beta \ge 0$, and suppose further that the following conditions hold.*

(a) For n, m, t fixed, the function $\boldsymbol{p} \mapsto \Gamma(n, m, \boldsymbol{p}, t)$ from $\mathcal{S}_\beta^+$ into $\mathbb{R}_+$ is continuous in the δ_β metric.

(b) For each fixed $n \in \mathbb{N}$, $\boldsymbol{p} \in \mathcal{S}_\beta^+$, and $w \in \mathcal{P}_\infty$

$$\lim_{\delta_\beta(\boldsymbol{p},\boldsymbol{q}) \downarrow 0} \mathbb{P}\{\exists 0 \le s \le t \,:\, W^{n,\boldsymbol{p},w}(s) \ne W^{n,\boldsymbol{p},w}(s)\} = 0,$$

with the limit taken over $\boldsymbol{q} \in \mathcal{S}_\beta^+$.

Then for all $\epsilon > 0$, $\boldsymbol{p} \in \mathcal{S}_\beta^+$, and $w \in \mathcal{P}_\infty$

$$\lim_{\delta_\beta(\boldsymbol{p},\boldsymbol{q})\downarrow 0,\, d(w,v)\downarrow 0} \mathbb{P}\{\sup_{0\le s\le t} d(W^{\boldsymbol{p},w}(s), W^{\boldsymbol{q},v}(s)) > \epsilon\} = 0,$$

with the limit taken over $\boldsymbol{q} \in \mathcal{S}_\beta^+$ and $v \in \mathcal{P}_\infty$.

Proof. Fix $m \in \mathbb{N}$. We have from (25) that if $d(w,v) \le 2^{-n}$, so that $\pi^n w = \pi^n v$, then

$$\begin{aligned}
&\mathbb{P}\{\sup_{0\le s\le t} d(W^{\boldsymbol{p},w}(s), W^{\boldsymbol{q},v}(s)) > 2^{-(m+1)}\} \\
&\quad = \mathbb{P}\{\exists 0 \le s \le t : \pi^m W^{\boldsymbol{p},w}(s) \ne \pi^m W^{\boldsymbol{q},v}(s)\} \\
&\quad \le \mathbb{P}\{\exists 0 \le s \le t : \pi^m W^{n,\boldsymbol{p},w}(s) \ne \pi^m W^{\boldsymbol{p},w}(s)\} \\
&\qquad\quad + \mathbb{P}\{\exists 0 \le s \le t : \pi^m W^{n,\boldsymbol{q},v}(s) \ne \pi^m W^{\boldsymbol{q},v}(s)\} \\
&\qquad\quad + \mathbb{P}\{\exists 0 \le s \le t : \pi^m W^{n,\boldsymbol{p},w}(s) \ne \pi^m W^{n,\boldsymbol{q},v}(s)\} \\
&\quad \le 2\Gamma(n,m,\boldsymbol{p},t) + |\Gamma(n,m,\boldsymbol{q},t) - \Gamma(n,m,\boldsymbol{p},t)| \\
&\qquad\quad + \mathbb{P}\{\exists 0 \le s \le t : \pi^m W^{n,\boldsymbol{p},w}(s) \ne \pi^m W^{n,\boldsymbol{q},v}(s)\}
\end{aligned}$$

Choosing n sufficiently large makes $\Gamma(n,m,\boldsymbol{p},t)$ arbitrarily small. Once n is fixed, taking $\delta_\beta(\boldsymbol{p},\boldsymbol{q})$ sufficiently small makes both $|\Gamma(n,m,\boldsymbol{q},t) - \Gamma(n,m,\boldsymbol{p},t)|$ and the last term arbitrarily small.

□

The proof of Theorem 10 is completed by the next two lemmas.

Lemma 16 *Suppose that $\kappa(a,b) = K(a+b)$ for some constant $K > 0$. Then it is possible to construct a coupled family of coalescents satisfying the conditions (a)–(c) of Lemma 14 for $\mathcal{S}' = \mathcal{S}^+$ and conditions (a) and (b) of Lemma 15 for $\beta = 0$.*

Proof. We need only consider the case $K = 1$, as the general case reduces to this case by rescaling time.

The collection of processes $W^{n,\boldsymbol{p},w}$, $n \in \mathbb{N}$, $w \in \mathcal{P}_\infty$ will be constructed simultaneously for all $\boldsymbol{p} \in \mathcal{S}^+$ by an adaptation of Construction 5. On some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $\{H_i\}_{i=1}^\infty$ be a a sequence of independent Poisson random measures on $\mathbb{R}_+ \times [0,1]$, each with intensity $dt \otimes du$, and let ε_k, $k = 1, 2, \ldots,$ be a sequence of standard exponential variables independent of the H_i. For $\boldsymbol{p} \in \mathcal{S}^+$ define Poisson random measures $\{B_i^{\boldsymbol{p}}\}_{i=1}^{\infty}$ on $\mathbb{R}_+$ by $B_i^{\boldsymbol{p}}(E) = H_i(E \times [0, p_i])$ for $E \subset \mathbb{R}_+$, so that $B_i^{\boldsymbol{p}}$ has intensity $p_i\, dt$. Let $S_0^{n,\boldsymbol{p}} < S_1^{n,\boldsymbol{p}} < \dots$ to be the successive atoms of $\sum_{i=1}^n B_i^{\boldsymbol{p}}$. Let $\alpha_k^{n,w}$, $1 \le k \le \#(n, w)$, denote the components of $\pi^n w$ listed in increasing order of their least elements. Define a sequence of $\mathbb{N}_{\#(n,w)}$–valued random variables $(Y_j^{n,\boldsymbol{p},w})_{j=0}^{\infty}$ by setting $Y_j^{n,\boldsymbol{p},w} = k$ if the atom of $\sum_{i=1}^n B_i^{\boldsymbol{p}}$ at time $S_j^{n,\boldsymbol{p}}$ is an atom of $B_i^{\boldsymbol{p}}$ for some $i \in \alpha_k^{n,w}$. Thus the sequence $(Y_j^{n,\boldsymbol{p},w})_{j=0}^{\infty}$ is i.i.d. with

$$\mathbb{P}\{Y_j^{n,\boldsymbol{p},w} = k\} = p(\alpha_k^{n,w})/p(\mathbb{N}_n)$$

where $p(A) := \sum_{i \in A} p_i$. Let $\tilde{\boldsymbol{p}}^{(n)}$ be the probability measure on $\mathbb{N}_{\#(n,w)}$ with mass $\tilde{p}_k^{(n)} := p(\alpha_k^{n,w})/p(\mathbb{N}_n)$ at $k \in \mathbb{N}_{\#(n,w)}$. According to Proposition 6, the process $(\Pi(t; (Y_j^{n,\boldsymbol{p},w})_{j=0}^{\infty}, (\varepsilon_k)_{k=1}^{\#(n,w)}),\, t \ge 0)$ is a $(\mathcal{P}_{\#(n,w)}, +, \tilde{\boldsymbol{p}}^{(n)})$–coalescent with starting point the partition of $\mathbb{N}_{\#(n,w)}$ into singletons. Define a random partition $W^{n,\boldsymbol{p},w}(t)$ of $\mathbb{N}_n$ by declaring that i and j belong to the same component of $W^{n,\boldsymbol{p},w}(t))$ if i and j belong to respective components $\alpha_k^{n,w}$ and $\alpha_\ell^{n,w}$ of w such that k and ℓ belong to the same component of $\Pi(t; (Y_j^{n,\boldsymbol{p},w})_{j=0}^{\infty}, (\varepsilon_k)_{k=1}^{\#(n,w)})$. It is clear that $(W^{n,\boldsymbol{p},w}(t),\, t \ge 0)$ is a $(\mathcal{P}_n, +, \boldsymbol{p}^{(n)})$–coalescent with starting state $\pi^n w$, where $\boldsymbol{p}^{(n)}$ is $\boldsymbol{p}$ conditioned on $\mathbb{N}_n$.

Let $V_m^{n,\boldsymbol{p},w}$ be the subset of $\mathbb{N}$ consisting of the vertices of the subtree of $\mathcal{T}((Y_j^{n,\boldsymbol{p},w}))_{j=0}^{\infty})$ that spans $\mathbb{N}_{\#(m,w)}$. Thus $v \in V_m^{n,\boldsymbol{p},w}$ if and only if v lies on the unique path from a to b in the tree $\mathcal{T}((Y_j^{n,\boldsymbol{p},w}))_{j=0}^{\infty})$ for some $a, b \in \mathbb{N}_m$, where edge directions in the tree are ignored in constructing the paths. If $V_m^{n,\boldsymbol{p},w} = V_m^{N,\boldsymbol{p},w}$, then by construction the restrictions to $\mathbb{N}_m$ of $(\Pi(t; (Y_j^{n,\boldsymbol{p},w})_{j=0}^{\infty}, (\varepsilon_k)_{k=1}^{\#(n,w)})$ and $(\Pi(t; (Y_j^{N,\boldsymbol{p},w})_{j=0}^{\infty}, (\varepsilon_k)_{k=1}^{\#(N,w)})$ are identical for all $t \ge 0$, hence so are the restrictions to $\mathbb{N}_m$ of $W^{n,\boldsymbol{p},w}(t)$ and $W^{N,\boldsymbol{p},w}(t)$. For $1 \le k \le \#(n, w)$ let $R_k^{n,\boldsymbol{p}}$ be the time of the first atom of $B_i^{\boldsymbol{p}}$ for some $i \in \alpha_k^{n,w}$. For $1 \le m \le n$ let

$$T_m^{n,w,\boldsymbol{p}} := \bigvee_{k=1}^{\#(m,w)} R_k^{n,\boldsymbol{p}} \text{ and } T_m^{\boldsymbol{p}} := \bigvee_{i=1}^{m} \inf\{t \ge 0 \;:\; B_i^{\boldsymbol{p}}([0,t] > 0\}.$$

By construction, $T_m^{n,w,\boldsymbol{p}}$ is decreasing as n increases, and $T_m^{n,w,\boldsymbol{p}} \le T_m^{\boldsymbol{p}}$ which is a.s. finite for every m and $\boldsymbol{p} \in \mathcal{S}^+$. Now for $n \ge N \ge m$

$$V_m^{n,\boldsymbol{p},w} \subseteq \{Y_j^{n,\boldsymbol{p},w} : S_j^{n,\boldsymbol{p}} \le T_m^{n,\boldsymbol{p},w}\} \subseteq \{Y_j^{n,\boldsymbol{p},w} : S_j^{n,\boldsymbol{p}} \le T_m^{N,\boldsymbol{p},w}\}.$$

Therefore,

$$\begin{aligned}
\mathbb{P}\{N(m,w,\boldsymbol{p},t) > N\} &\le \mathbb{P}\{\exists n > N \ : \ V_m^{n,\boldsymbol{p},w} \neq V_m^{N,\boldsymbol{p},w}\} \\
&\le \mathbb{P}[\mathbb{P}\{\exists i > N \ : \ H_i([0, T_m^{N,\boldsymbol{p},w}] \times [0,p_i]) > 0 \mid T_m^{N,\boldsymbol{p},w}\}] \\
&= 1 - \mathbb{P}\left[\exp\left(-T_m^{N,\boldsymbol{p},w} \bar{p}_{N+1}\right)\right] \text{ where } \bar{p}_{N+1} := \sum_{i=N+1}^{\infty} p_i \\
&\le 1 - \mathbb{P}\left[\exp(-T_m^{\boldsymbol{p}}\, \bar{p}_{N+1})\right] =: \Gamma(N,m,t,\boldsymbol{p})
\end{aligned}$$

It is clear that Γ has property (b) of Lemma 14 and property (a) of Lemma 15 for $\beta = 0$. It is also clear that κ satisfies condition (a) of Lemma 14 and that hypothesis (c) of Lemma 14 holds for this choice of $\boldsymbol{p}^{(n)}$. Finally, observe for $w \in \mathcal{P}_\infty$ and $\boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}^+$ that

$$\begin{aligned}
&\mathbb{P}\{\exists 0 \le s \le t \ : \ W^{n,\boldsymbol{p},w}(s) \neq W^{n,\boldsymbol{q},w}(s)\} \\
&\quad \le \mathbb{P}\{\exists 1 \le i \le n \ : \ B_i^{\boldsymbol{p}}(\cdot \cap [0, T_n^{\boldsymbol{p}} \vee T_n^{\boldsymbol{q}}]) \neq B_i^{\boldsymbol{q}}(\cdot \cap [0, T_n^{\boldsymbol{p}} \vee T_n^{\boldsymbol{q}}])\}
\end{aligned}$$

and it follows that property (b) of Lemma 15 holds for $\beta = 0$.

□

Lemma 17 *Suppose that κ satisfies the Lipschitz condition (15). Then it is possible to construct a coupled family of coalescents satisfying conditions (a)–(c) of Lemma 14 for $\mathcal{S}' = \mathcal{S}_1^+$ and conditions (a) and (b) of Lemma 15 with $\beta = 1$.*

Proof. Once again we will treat the case $K = 1$ and observe that the general case can be reduced to this by rescaling time. On some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ construct a collection $\{D_{ij}\}_{i,j \in \mathbb{N}}$ of independent Poisson random measures on $\mathbb{R}_+ \times [0,1] \times [0,1]$, each having intensity Lebesgue measure on $\mathbb{R}_+ \times [0,1] \times [0,1]$. For $\boldsymbol{p} \in \mathcal{S}^+$ define a Poisson random measure $A_{ij}^{\boldsymbol{p}}$ on $\mathbb{R}_+ \times [0,1]$ by setting $A_{ij}^{\boldsymbol{p}}(E) = D_{ij}(E \times [0,p_i])$ for $E \subset \mathbb{R}_+ \times [0,1]$, so that $A_{ij}^{\boldsymbol{p}}$ has intensity $p_i\, dt \otimes du$.

For $I, J \subseteq \mathbb{N}_n$ set $C_{IJ}^{\boldsymbol{p}} = \sum_{i \in I} A_{i,\min J}^{\boldsymbol{p}}$. Note that if $I_1, \ldots, I_h \subseteq \mathbb{N}$ are pairwise disjoint, then $\{C_{I_k I_\ell}^{\boldsymbol{p}}\}_{1 \le k \neq \ell \le h}$ is a collection of independent Poisson random measures on $\mathbb{R}_+ \times [0,1]$, with $C_{I_k I_\ell}^{\boldsymbol{p}}$ having intensity $p_{I_k}\, dt \otimes du$.

For $n \in \mathbb{N}$, $w \in \mathcal{P}_\infty$ and $\boldsymbol{p} \in \mathcal{S}^+$, define a càdlàg $\mathcal{P}_n$-valued process $(W^{n,\boldsymbol{p},w}(t),\ t \ge 0)$ and stopping times $0 = T_0^{n,\boldsymbol{p},w} < T_1^{n,\boldsymbol{p},w} < T_2^{n,\boldsymbol{p},w} < \ldots$ as follows, where we abbreviate

$W^n(t) := W^{n,\boldsymbol{p},w}(t)$ and $T_k^n := T_k^{n,\boldsymbol{p},w}$. Let $W^n(0) = \pi^n w$,

$$T_{k+1}^n = \inf\Big\{t > T_k^n \ : \exists I_{k+1}^n, J_{k+1}^n \in W^n(T_k),\ I_{k+1}^n \neq J_{k+1}^n,$$
$$C^{\boldsymbol{p}}_{I_{k+1}^n J_{k+1}^n}\Big(]T_k^n, t] \times \Big[0, \frac{\kappa(p_{I_{k+1}^n}, p_{J_{k+1}^n})}{p_{I_{k+1}^n} + p_{J_{k+1}^n}}\Big]\Big) > 0\Big\},$$

$$W^n(t) = W^n(T_k), \quad \text{for } t \in]T_k^n, T_{k+1}^n[,$$

$$W^n(T_{k+1}^n) = \text{the coarsening of } W^n(T_k) \text{ obtained by aggregating } I_{k+1}^n \text{ and } J_{k+1}^n.$$

As a transition of W^n involving the aggregation of I and J may occur either due to a point of C_{IJ} or due to a point of C_{JI}, it is easily verified that that for each $\boldsymbol{p} \in \mathcal{S}^+$ and $w \in \mathcal{P}_\infty$ the process $W^n = W^{n,\boldsymbol{p},w}$ is a $(\mathcal{P}_n, \kappa, \boldsymbol{p}^{(n)})$-coalescent started at $\pi^n w$ for $\boldsymbol{p}^{(n)}$ the restriction of $\boldsymbol{p}$ to $\mathbb{N}_n$. It is also easily seen that for each $n \in \mathbb{N}$, the process $(W^{n,\boldsymbol{p},w}, W^{n+1,\boldsymbol{p},w})$ with state space $\mathcal{P}_n \times \mathcal{P}_{n+1}$ is a time–homogeneous Markov chain whose transition rates depend only on $\boldsymbol{p}$ and not on w.

We will now derive an upper bound on

$$\mathbb{P}\{\exists 0 \le s \le t \ : \ \pi^n W^{n+1,\boldsymbol{p},w}(s) \neq W^{n,\boldsymbol{p},w}(s)\}.$$

Consider first the case that $\{n+1\}$ is not a component of $\pi^{n+1} w$. Write I for the component of $\pi^n w$ that is the intersection with $\mathbb{N}_n$ of the component of w that contains $\{n+1\}$. Observe that if J is another component of w, then

$$C^{\boldsymbol{p}}_{J, I\cup\{n+1\}} = C^{\boldsymbol{p}}_{JI}$$
$$C^{\boldsymbol{p}}_{I\cup\{n+1\}, J} = C^{\boldsymbol{p}}_{IJ} + A^{\boldsymbol{p}}_{n+1, \min J}$$

Now, using the notation $A \triangle B$ for the symmetric difference of two sets A and B,

$$\begin{aligned}
&\lim_{t\downarrow 0} t^{-1}\mathbb{P}\{\exists 0 \le s \le t : \pi^n W^{n+1}(s) \ne W^n(s)\} \\
&\quad = \lim_{t\downarrow 0} t^{-1}\mathbb{P}\Bigg\{\exists J \in w,\ J \ne I : \\
&\qquad (C^{\boldsymbol{p}}_{IJ} + C^{\boldsymbol{p}}_{JI})\Big(\Big([0,t]\times\Big[0, \frac{\kappa(p_I + p_{n+1}, p_J)}{p_I + p_{n+1} + p_J}\Big]\Big)\triangle\Big([0,t]\times\Big[0, \frac{\kappa(p_I, p_J)}{p_I + p_J}\Big]\Big)\Big) > 0 \\
&\qquad\qquad \text{or } A^{\boldsymbol{p}}_{n+1,\min J}([0,t]\times[0,1]) > 0\Bigg\} \\
&\quad = \lim_{t\downarrow 0} t^{-1}\Bigg(1 - \prod_{J\ne I}\Bigg(1 - \exp\Bigg(-t\Bigg|\frac{\kappa(p_I + p_{n+1}, p_J)}{p_I + p_{n+1} + p_J} - \frac{\kappa(p_I, p_J)}{p_I + p_J}\Bigg|(p_I + p_J)\Bigg)\Bigg) \\
&\qquad\qquad \times \prod_{J\ne I}\Big(1 - \exp(-t p_{n+1})\Big)\Bigg) \\
&\quad \le 3 n p_{n+1},
\end{aligned}$$

where we have used the fact that the number of components of w is at most n and the inequality

$$\begin{aligned}
\left|\frac{\kappa(a+c,b)}{a+c+b} - \frac{\kappa(a,b)}{a+b}\right| &\le \frac{\big|\kappa(a+c,b) - \kappa(a,b)\big|(a+b) + c\kappa(a,b)}{(a+c+b)(a+b)} \\
&\le \frac{2c(a+b)}{(a+c+b)(a+b)} \\
&= \frac{2c}{a+c+b} \\
&\le \frac{2c}{a+b}
\end{aligned}$$

for $a, b, c > 0$. It follows from the Markov property of $(W^n(\cdot), W^{n+1}(\cdot))$ that for all $\boldsymbol{p} \in \mathcal{S}^+$ and $w \in \mathcal{P}_\infty$

$$\begin{aligned}
\mathbb{P}\{\exists 0 \le s \le t \ :\ \pi^n W^{n+1}(s) \ne W^n(s)\} &\le 1 - \exp(-3 n p_{n+1} t) \\
&\le 3(n+1) p_{n+1} t,
\end{aligned} \tag{26}$$

Moreover, by the strong Markov property of $(W^n(\cdot), W^{n+1}(\cdot))$, this inequality holds *a fortiori* if $\{n+1\}$ is a component of $\pi^{n+1} w$. For $N \ge m$ we have from (26) that for

$w \in \mathcal{P}_\infty$ and $\boldsymbol{p} \in \mathcal{S}^+$

$$\begin{aligned}\mathbb{P}\{N(m,w,\boldsymbol{p},t) > N\} &\leq \sum_{n=N}^{\infty} \mathbb{P}\{\exists 0 \leq s \leq t : \pi^n W^{n+1,\boldsymbol{p},w}(s) \neq W^{n,\boldsymbol{p},w}(s)\} \\ &\leq 3t \sum_{i=N+1}^{\infty} i p_i =: \Gamma(N,m,\boldsymbol{p},t).\end{aligned}$$

It is clear that Γ has property (b) of Lemma 14 and property (a) of Lemma 15 for $\beta = 1$. It is also clear that κ satisfies condition (a) of Lemma 14 and that hypothesis (c) of Lemma 14 holds for this choice of $\boldsymbol{p}^{(n)}$. Finally, observe for $w \in \mathcal{P}_\infty$ and $\boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}^+$ that

$$\begin{aligned}&\mathbb{P}\{\exists 0 \leq s \leq t \ : \ W^{n,\boldsymbol{p},w}(s) \neq W^{n,\boldsymbol{q},w}(s)\} \\ &\quad \leq \mathbb{P}\{\exists 1 \leq i,j \leq n \ : \ A_{ij}^{\boldsymbol{p}}(\cdot \cap [0,t] \times [0,1]) \neq A_{ij}^{\boldsymbol{q}}(\cdot \cap [0,t] \times [0,1])\}\end{aligned}$$

and it follows that hypothesis (b) of Lemma 15 holds.

$\square$

5 Construction of infinite measure-valued coalescents

In this section we prove Theorem 2 by a development of the results of the previous section. Define a map $\texttt{CLUMP} : \mathbb{N} \times \mathcal{P}_\infty \to \mathbb{N}$ by setting

$$\texttt{CLUMP}(k,w) = \inf\{\ell \in \mathbb{N} : \ell \sim_w k\}.$$

Recall that a map $f : \mathbb{N} \to \mathbb{N}$ *leftward* if $f(k) \leq k$, $k \in \mathbb{N}$. Note that the map $k \mapsto \texttt{CLUMP}(k,w)$ is leftward for every $w \in \mathcal{P}_\infty$. Recall from around (16) the definition of the subspace $\mathcal{S}_\beta$ of $\mathcal{S}$ for $\beta \geq 0$ and the definition of the metrics δ_β and Δ_β on $\mathcal{S}_\beta$. Recall that $\mathcal{S}^+ := \{\boldsymbol{x} \in \mathcal{S} : x_k > 0,\ k \in \mathbb{N}\}$, $\mathcal{S}^1 := \{\boldsymbol{x} \in \mathcal{S} : x_1 > 0\}$, $\mathcal{S}_\beta^+ = \mathcal{S}_\beta \cap \mathcal{S}^+$, and $\mathcal{S}_\beta^1 = \mathcal{S}_\beta \cap \mathcal{S}^1$. Note that $\mathcal{S}_\beta^1$ is a closed subset of $(\mathcal{S}_\beta, \Delta_\beta)$. Define $\texttt{WEIGH} : \mathcal{P}_\infty \times \mathcal{S}^+ \to \mathcal{S}^1$ by letting $\texttt{WEIGH}(w,\boldsymbol{p})$ be the push-forward of $\boldsymbol{p}$ by $\texttt{CLUMP}(\cdot,w)$. Thus, $\texttt{WEIGH}(w,\boldsymbol{p})$ assigns the $\boldsymbol{p}$-mass of each component of w to the smallest element of the component.

Lemma 18 *For each $\beta \geq 0$ the map* $\texttt{WEIGH}$ *from $(\mathcal{P}_\infty, d) \times (\mathcal{S}_\beta^+, \delta_\beta)$ into $(\mathcal{S}_\beta^1, \Delta_\beta)$ is continuous.*

Proof. For $\boldsymbol{p}', \boldsymbol{p}'' \in \mathcal{S}^+$, $w', w'' \in \mathcal{P}_\infty$ we have

$$\begin{aligned}
&\Delta_\beta(\mathtt{WEIGH}(w', \boldsymbol{p}'), \mathtt{WEIGH}(w'', \boldsymbol{p}'')) \\
\leq \quad & \Delta_\beta(\mathtt{WEIGH}(w', \boldsymbol{p}'), \mathtt{WEIGH}(w'', \boldsymbol{p}')) + \Delta_\beta(\mathtt{WEIGH}(w'', \boldsymbol{p}'), \mathtt{WEIGH}(w'', \boldsymbol{p}'')).
\end{aligned}$$

Observe that if $d(w', w'') \leq 2^{-(n+1)}$ so that $\pi^n w' = \pi^n w''$, then

$$\Delta_\beta(\mathtt{WEIGH}(w', \boldsymbol{p}'), \mathtt{WEIGH}(w'', \boldsymbol{p}')) \leq 2^{-(n+1)} + 2 \sum_{k=n+1}^{\infty} k^\beta p'_k.$$

Moreover,

$$\Delta_\beta(\mathtt{WEIGH}(w'', \boldsymbol{p}'), \mathtt{WEIGH}(w'', \boldsymbol{p}'')) \leq \delta_\beta(\boldsymbol{p}', \boldsymbol{p}'').$$

$\square$

Suppose that $\boldsymbol{p} \in \mathcal{S}^+$ and $(w(t),\, t \in \mathbb{I})$ is a $\mathcal{P}_\infty$-coalescent evolution. Define an $\mathcal{S}^1$-valued function $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ by setting

$$\boldsymbol{x}(t) = \mathtt{WEIGH}(w(t), \boldsymbol{p}). \tag{27}$$

It is clear that $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ is a leftward $\mathcal{S}^1$-coalescent evolution. We can take $\Psi_{s,t} := \mathtt{CLUMP}(\,\cdot\,, w(t))$, $s, t \in \mathbb{I}$, $s < t$, as the associated tracking maps. We call $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ the *$\mathcal{S}^1$-coalescent evolution derived from the proto-galaxy mass distribution $\boldsymbol{p}$ via the $\mathcal{P}_\infty$-coalescent evolution* $(w(t),\, t \in \mathbb{I})$. Note that if $(w(t),\, t \in \mathbb{I})$ is binary, then $(\boldsymbol{x}(t),\, t \in \mathbb{I})$ is basic. This method of construction of an $\mathcal{S}^1$-coalescent evolution from a $\mathcal{P}_\infty$-coalescent evolution generalizes the construction in Section 3.1 of a basic $(\mathcal{S}^K, \kappa)$-coalescent from a $(\mathcal{P}_n, \kappa, \boldsymbol{p})$-coalescent for a finite vector $\boldsymbol{p}$ of proto-galaxy masses.

A continuous one-sided inverse for the map `WEIGH` can be defined as follows. Given $\boldsymbol{x} \in \mathcal{S}^1$, write $\{\tau_k(\boldsymbol{x})\}_{k=1}^\infty$ for the ordered list of the elements of $\mathbb{N}$ that are assigned positive mass by $\boldsymbol{x}$. That is,

$$\tau_1(\boldsymbol{x}) = 1; \;\; \tau_{k+1}(\boldsymbol{x}) = \inf\{i > \tau_k(\boldsymbol{x}) : x_i > 0\}. \tag{28}$$

Define $\psi(\boldsymbol{x}) \in \mathcal{S}^+$ as follows:

$$\psi_i(\boldsymbol{x}) = \begin{cases} \frac{2^{-(i-\tau_k(\boldsymbol{x})+1)}}{1-2^{-(\tau_{k+1}(\boldsymbol{x})-\tau_k(\boldsymbol{x}))}} x_{\tau_k(\boldsymbol{x})}, & \text{if } \tau_k(\boldsymbol{x}) \le i < \tau_{k+1}(\boldsymbol{x}) < \infty \text{ for some } k \in \mathbb{N}, \\ 2^{-(i-\tau_k(\boldsymbol{x})+1)} x_{\tau_k(\boldsymbol{x})}, & \text{if } \tau_k(\boldsymbol{x}) \le i < \tau_{k+1}(\boldsymbol{x}) = \infty \text{ for some } k \in \mathbb{N}, \end{cases}$$

where the second case occurs only if $\boldsymbol{x}$ has finite support. Define $\theta(\boldsymbol{x}) \in \mathcal{P}_\infty$ by declaring that $i \sim_{\theta(\boldsymbol{x})} j$ if and only if $\sup\{k : \tau_k(\boldsymbol{x}) \le i\} = \sup\{k : \tau_k(\boldsymbol{x}) \le j\}$. Note that $\mathtt{WEIGH}(\theta(\boldsymbol{x}), \psi(\boldsymbol{x})) = \boldsymbol{x}$. The following result is elementary.

Lemma 19 *For each $\beta \ge 0$ the map $\boldsymbol{x} \mapsto (\psi(\boldsymbol{x}), \theta(\boldsymbol{x}))$ from $(\mathcal{S}^1_\beta, \Delta_\beta)$ into $(\mathcal{S}^+_\beta, \delta_\beta) \times (\mathcal{P}_\infty, d)$ is continuous.*

We turn now to the proof of the basic case of Theorem 2. Let $(W^{\boldsymbol{p},w}(t),\ t \ge 0)$, $w \in \mathcal{P}_\infty$, $\boldsymbol{p} \in \mathcal{S}^+_\beta$, be the collection of processes whose existence is guaranteed by either Lemma 16 or Lemma 17. We now write $W(t; \boldsymbol{p}, w)$ instead of $W^{\boldsymbol{p},w}(t)$ for typographical convenience. Given $\boldsymbol{x} \in \mathcal{S}^1_\beta$, define a càdlàg $(\mathcal{S}^1_\beta, \Delta_\beta)$–valued process $(X(t; \boldsymbol{x}),\ t \ge 0)$ by setting

$$X(t; \boldsymbol{x}) := \mathtt{WEIGH}(W(t; \psi(\boldsymbol{x}), \theta(\boldsymbol{x})), \psi(\boldsymbol{x})).$$

Note that $X(t; \boldsymbol{x})$ takes values in the set $\{\mathtt{WEIGH}(w, \psi(\boldsymbol{x})) : w \le \theta(\boldsymbol{x})\}$. Write $\mathbb{Q}^x$ for the law of $(X(t; \boldsymbol{x}),\ t \ge 0)$.

For $n \in \mathbb{N}$ let $\mathcal{P}^n_\infty$ be the finite set of partitions w with the property that $n \sim_w n+1 \sim_w n+2 \sim_w \cdots$. Note that if $\boldsymbol{x} \in \mathcal{S}^1_\beta \cap \mathcal{S}^K$, then $\theta(\boldsymbol{x}) \in \mathcal{P}^n_\infty$ for some n. It is clear that for such an $\boldsymbol{x}$ the process $(X(t; \boldsymbol{x}),\ t \ge 0)$ is a basic $(\mathcal{S}^K, \kappa)$–coalescent. It follows by Lemmas 18, 19, 14, 15 and the compactness of $\mathcal{P}_\infty$ that for $\boldsymbol{x} \in \mathcal{S}^1_\beta$, $t \ge 0$ and $\epsilon > 0$,

$$\lim_{\Delta_\beta(\boldsymbol{x},\boldsymbol{y}) \downarrow 0} \mathbb{P}\{\sup_{0 \le s \le t} \Delta_\beta(X(s; \boldsymbol{x}), X(s; \boldsymbol{y})) > \epsilon\} = 0. \tag{29}$$

Note for each $\boldsymbol{p} \in \mathcal{S}^+_\beta$ that $\{\mathtt{WEIGH}(w, \boldsymbol{p}) : w \in \mathcal{P}_\infty\}$ is a compact subset of $(\mathcal{S}^1_\beta, \Delta_\beta)$. As the set $\bigcup_{n \in \mathbb{N}} \{\mathtt{WEIGH}(w, \boldsymbol{p}) : w \in \mathcal{P}^n_\infty\}$ is dense in $\{\mathtt{WEIGH}(w, \boldsymbol{p}) : w \in \mathcal{P}_\infty\}$, it follows from (29) that $(\mathbb{Q}^{\boldsymbol{x}},\ \boldsymbol{x} \in \{\mathtt{WEIGH}(w, \boldsymbol{p}) : w \in \mathcal{P}_\infty\})$ is the family of laws of a Feller process on $(\{\mathtt{WEIGH}(w, \boldsymbol{p}) : w \in \mathcal{P}_\infty\}, \Delta_\beta)$. Finally, as $\mathcal{S}^1_\beta = \bigcup_{\boldsymbol{p} \in \mathcal{S}^+_\beta} \{\mathtt{WEIGH}(w, \boldsymbol{p}) : w \in \mathcal{P}_\infty\}$, we have that $(\mathbb{Q}^{\boldsymbol{x}},\ \boldsymbol{x} \in \mathcal{S}^1_\beta)$ is the family of laws of a Hunt process on $(\mathcal{S}^1_\beta, \Delta_\beta)$. We have already established (ii). Claim (iv) is immediate from (29). Claims (i) and (iii) follow from parts (iii) and (iv) of Theorem 10.

The proof of Theorem 2 in the shunted and ranked cases is similar. Just change the definition of $X(t;\boldsymbol{x})$ to

$$\texttt{SHUNT}(\texttt{WEIGH}(W(t;\psi(\boldsymbol{x}),\theta(\boldsymbol{x})),\psi(\boldsymbol{x})))$$

and

$$\texttt{RANK}(\texttt{WEIGH}(W(t;\psi(\boldsymbol{x}),\theta(\boldsymbol{x})),\psi(\boldsymbol{x}))),$$

respectively, and use the fact that SHUNT is continuous from $(\mathcal{S}^1_\beta,\Delta_\beta)$ to $(\mathcal{S}^*_\beta,\delta_\beta)$, and RANK is continuous from $(\mathcal{S}^1_\beta,\Delta_\beta)$ to $(\mathcal{S}^\downarrow_\beta,\delta_\beta)$.

□

6 The Infinite Additive Coalescent.

We present in this section a number of results for infinite additive coalescent processes. We record first the following explicit construction of these processes, which follows easily from Proposition 6, Theorem 2, and Theorem 10.

Corollary 20 *In the notation of Construction 5, suppose that* $\boldsymbol{p}$ *is a probability distribution on* $\mathbb{N}$ *with* $p_i > 0$ *for all* $i \in \mathbb{N}$. *Let* $\Pi(t) := \Pi(t;(Y_j)_{j=0}^{\infty},(\varepsilon_\sigma)_{\sigma\in\mathbb{N}})$. *Then* $(\Pi(t),\, t\geq 0)$ *is a* $(\mathcal{P}_\infty,+,\boldsymbol{p})$*–coalescent starting from the partition that consists of all singletons. The process* $(\texttt{WEIGH}(\Pi(t),\boldsymbol{p}),\, t\geq 0)$ *is a basic additive coalescent with initial state* $\boldsymbol{p}$*, and the processes* $(\texttt{SHUNT}(\texttt{WEIGH}(\Pi(t),\boldsymbol{p})),\, t\geq 0)$ *and* $(\texttt{RANK}(\texttt{WEIGH}(\Pi(t),\boldsymbol{p})),\, t\geq 0)$ *are shunted and ranked additive coalescents with starting states* $\boldsymbol{p}$ *and* RANK$(\boldsymbol{p})$ *respectively.*

6.1 Asymptotics for uniform initial condition

Consider now a shunted $(\mathcal{S}^K,+)$-coalescent $(\boldsymbol{X}^*(t), t\geq 0)$ started with the *uniform* or *monodisperse* initial condition $\boldsymbol{u}_n$ defined by n equal masses of size $1/n$ labeled by $\mathbb{N}_n$. Given $\#\boldsymbol{X}^*(t)=k$ let $\tilde{X}_1(t),\ldots,\tilde{X}_k(t)$ denote the sizes of the k non-zero components of $\boldsymbol{X}^*(t)$ presented in an exchangeable random order. It is known [33], and follows from Proposition 4, that there is the equality of joint distributions

$$(\tilde{X}_i(t), 1 \le i \le k \,|\, \boldsymbol{X}^*(0) = \boldsymbol{u}_n, \#\boldsymbol{X}^*(t) = k) \stackrel{d}{=} \left(\frac{Y_i}{n}, 1 \le i \le k \,\middle|\, \sum_{i=1}^{k} Y_i = n \right) \tag{30}$$

where the random variables Y_i are independent and identically distributed with the *Borel*(1) *distribution*

$$P\{Y_i = m\} = e^{-m} m^{m-1}/m! \quad (m = 1, 2, \dots) \tag{31}$$

We may suppose that the process $(\boldsymbol{X}^*(t),\, t \ge 0)$ has been constructed in the manner of Section 3.1 from an additive $\mathcal{P}_n$-coalescent process $(W_t,\, t \ge 0)$ with proto-galaxy masses $p_1 = \dots = p_n = 1/n$. Due to exchangeability of W_t, the components of $\boldsymbol{X}^*(t)$ are in *size-biased random order* [15, 34]. It follows from the representation (30), and [2, Lemmas 11 and 12] that for each $t > 0$ and $s \ge 0$, as $n \to \infty$ and k varies with n in such a way that $n/k^2 \to s$

$$(X_i^*(t), i \ge 1 \,|\, \boldsymbol{X}^*(0) = \boldsymbol{u}_n, \#\boldsymbol{X}^*(t) = k) \stackrel{d}{\to} \left(\frac{H_i}{\Sigma}, i \ge 1 \,\middle|\, \Sigma = s \right) \tag{32}$$

where $\stackrel{d}{\to}$ denotes convergence in distribution of $(\mathcal{S}^*, \delta_0)$-valued random elements and on the right-hand side $(H_i, i \ge 1)$ is a size-biased random permutation of the points of a Poisson point process on $]0, \infty[$ with intensity measure $x^{-3/2}dx/\sqrt{2\pi}$ and $\Sigma := \sum_i H_i$. Note that Σ has a stable(1/2) density. Let $\Pi(s)$ denote the probability distribution on $\mathcal{S}$ appearing as the limit distribution in (32). A formula for the joint density of the first n components of a random sequence $(V_1(s), V_2(s), \dots)$ with distribution $\Pi(s)$ can be read from [31, Theorem 2.1]. Let $\overline{V}_m(s) := \sum_{i=m}^{\infty} V_i(s)$, so $\overline{V}_1(s) = 1$ and $V_m(s) = \overline{V}_m(s) - \overline{V}_{m+1}(s)$ for $m \ge 1$. As shown in [8], a sequence $(V_m(s), m \ge 1)$ with distribution $\Pi(s)$ is generated by the formula

$$\overline{V}_m(s) = \left(1 + s \sum_{i=1}^{m-1} Z_i^2 \right)^{-1} \tag{33}$$

where $Z_1, Z_2, \dots$ is a sequence of independent standard normal variables. Here Z_i^2 has the same distribution as $1/\Sigma$. Let $\Pi^{\downarrow}(s)$ denote the push-forward of $\Pi(s)$ by the ranking map from $\mathcal{S}^*$ to $\mathcal{S}^\downarrow$. Because the ranking map is continuous [15], the convergence in distribution (32) implies a corresponding result for a ranked additive $\mathcal{S}^K$-coalescent $\mathbf{X}^\downarrow(t)$ instead of $\boldsymbol{X}^*(t)$, with limit $\Pi^\downarrow(s)$. The finite dimensional distributions of $\Pi^\downarrow(s)$ can be described explicitly [30, 37], but they are much more complicated than those of $\Pi(s)$.

Proposition 21 *Let $(\boldsymbol{X}^*(t), t \geq 0)$ be a shunted additive $\mathcal{S}^K$-coalescent with initial state $\boldsymbol{X}^*(0) = \boldsymbol{u}_n$. Let $h_n := \frac{1}{2}\log n$. For each $r \in \mathbb{R}$, as $n \to \infty$ the distribution of $\boldsymbol{X}^*(h_n+r)$ on $(\mathcal{S}, \delta_0)$ converges to $\Pi(e^{2r})$, and the distribution of $\mathbf{X}^\downarrow(h_n+r)$ converges to $\Pi^\downarrow(e^{2r})$.*

Proof. From the binomial $(n-1, e^{-t})$ distribution of $\#\boldsymbol{X}^*(t) - 1$ given $\boldsymbol{X}^*(0) = \boldsymbol{u}_n$, we know that for each real number r,

$$E(\#\boldsymbol{X}^*(h_n+r) \,|\, \boldsymbol{X}^*(0) = \boldsymbol{u}_n) = 1 + (n-1)\exp(-\tfrac{1}{2}\log(n) - r) \sim \sqrt{n}e^{-r} \tag{34}$$

and the variance of $\#\boldsymbol{X}^*(h_n+r)$ given $\boldsymbol{X}^*(0) = \boldsymbol{u}_n$ is of the same order of magnitude. It follows that for each fixed r, as $n \to \infty$ the random variable $\#\boldsymbol{X}^*(h_n+r)/\sqrt{n}$ given $\boldsymbol{X}^*(0) = \boldsymbol{u}_n$ converges in probability to the constant e^{-r}, and hence $n/(\#\boldsymbol{X}^*(h_n+r))^2$ given $\boldsymbol{X}^*(0) = \boldsymbol{u}_n$ converges in probability to the constant e^{2r}. The proposition now follows from (32) and continuity of the ranking map.

□

Suppose now that

$$\boldsymbol{X}^n = (\boldsymbol{X}^n(r),\ -h_n \leq r < \infty) \tag{35}$$

is a shunted additive $\mathcal{S}^K$-coalescent started at time $-h_n$ with $\boldsymbol{X}^n(-h_n) = \boldsymbol{u}_n$. Proposition 21 combined with Theorem 2 shows that as $n \to \infty$, the finite dimensional distributions of $\boldsymbol{X}^n$ converge to those of a limiting $\mathcal{S}^*$-coalescent process

$$\boldsymbol{X}^\infty = (\boldsymbol{X}^\infty(r), -\infty < r < \infty) \tag{36}$$

such that for each real r the distribution of $\boldsymbol{X}^\infty(r)$ on $\mathcal{S}^*$ is $\Pi(e^{2r})$. Thus $\boldsymbol{X}^\infty$ actually takes values in $\mathcal{S}^+$. Moreover, there is weak convergence on the appropriate Skorohod path space, and the limiting process is a strong-Markov process with the shunted additive coalescent semigroup $(Q_t^*, t \geq 0)$ of transition operators. See [8] for another construction of $\boldsymbol{X}^\infty$ based on the combinatorial representation of the additive $\mathcal{P}_n$-coalescent in terms of random trees [33], and Aldous's continuum random tree [5]. That approach yields various distributional properties of the limit process, but not the regularity properties of $\boldsymbol{X}^\infty$ such as the strong-Markov property obtained here. The family of probability measures $\Pi(e^{2r})$ define an *entrance law* for the semigroup (Q_t^*), that is

$$\Pi(e^{2r})Q_t^* = \Pi(e^{2(r+t)}) \qquad (r \in \mathbb{R},\, t \geq 0) \tag{37}$$

There are corresponding results for the ranked rather than shunted additive coalescent. See also [19] for some recent developments.

6.2 Tail thinning

Let $(\boldsymbol{X}^*(t),\, t \geq 0)$ be a shunted additive $\mathcal{S}$-coalescent. Let $\bar{X}_n^*(t) = \sum_{i=n}^\infty X_i^*(t)$. Suppose that $\boldsymbol{X}^*(0)$ has distribution $\Pi(c^{-1})$ on $\mathcal{S}$ for some $c > 0$, for $\Pi(s)$ as in Proposition 21. From the representation (33) of a random element with distribution $\Pi(s)$, the consequence of the law of large numbers that $\sum_{i=1}^{n-1} Z_i^2 \sim n$ almost surely, and the consequence of (37) that $\boldsymbol{X}^*(t)$ has distribution $\Pi(c^{-1}e^{2t})$ on $\mathcal{S}^*$ for each $t > 0$, we have for each $t \geq 0$ that

$$\bar{X}_n^*(t) \sim c\, e^{-2t}\, n^{-1} \text{ almost surely as } n \to \infty \tag{38}$$

where $a_n \sim b_n$ means that $a_n/b_n \to 1$. Let $\bar{X}_n^\downarrow(t) = \sum_{i=n}^\infty X_i^\downarrow(t)$ where $(\mathbf{X}^\downarrow(t),\, t \geq 0)$ is a ranked additive $\mathcal{S}$-coalescent. For $\mathbf{X}^\downarrow(0)$ with distribution $\Pi^\downarrow(c^{-1})$ on $\mathcal{S}^\downarrow$, the distribution of $\mathbf{X}^\downarrow(t)$ is $\Pi^\downarrow(c^{-1}e^{2t})$. It follows [22, (68)] that

$$\bar{X}_n^\downarrow(t) \sim (2/\pi)\, c\, e^{-2t}\, n^{-1} \text{ almost surely as } n \to \infty \tag{39}$$

We conjecture that if the initial state of a shunted additve $\mathcal{S}$-coalescent $(\boldsymbol{X}^*(t),\, t \geq 0)$ (resp. a ranked additive $\mathcal{S}$-coalescent $(\mathbf{X}^\downarrow(t),\, t \geq 0)$) is such that (38) (resp. (39)) holds for $t = 0$, then (38) (resp. (39)) holds for each $t > 0$. As a step towards understanding how tails of a mass distribution are affected by an additive coalescent process, this section presents some results related to (38) and (39) for the basic and shunted additive $\mathcal{S}$-coalescents.

The following simple lemma is no doubt present in the differential equations literature, but we have been unable to find a reference.

Lemma 22 *Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of Borel functions mapping $\mathbb{R}_+$ into $\mathbb{R}_+$ that is uniformly bounded on compacts and satisfies*

$$f_n(t) = b + c\int_0^t f_n(s)\,ds + d_n(t),$$

where $d_n(t)$ converges to 0 uniformly on compacts as $n \to 0$. Then $f_n(t)$ converges to be^{ct} uniformly on compacts as $n \to \infty$.

Proof. Observe for $m, n \in \mathbb{N}$ that

$$|f_m(t) - f_n(t)| \le |c|\int_0^t |f_m(s) - f_n(s)|\,ds + |d_m(t) - d_n(t)|,$$

By Gronwall's lemma,

$$|f_m(t) - f_n(t)| \le \sup_{0\le s\le t} |d_m(s) - d_n(s)|e^{|c|t},$$

and so there exists a function f such that f_n converges to f uniformly on compacts as $n \to \infty$. Clearly,

$$f(t) = b + c\int_0^t f(s)\,ds,$$

and hence $f(t) = be^{ct}$.

□

Recall that $\mathbb{Q}^{\boldsymbol{x}}$ governs the basic additive coalescent with initial state $\boldsymbol{x}$, as constructed in Theorem 2.

Lemma 23 *For $\boldsymbol{x} \in S^1$, $k \in \mathbb{N}$ and $t \ge 0$,*

$$\mathbb{Q}^{\boldsymbol{x}}\left[\sum_{j=k}^{\infty} X_j^2(t)\right] \le e^{4t}\sum_{j=k}^{\infty} x_j^2.$$

Proof. It suffices by part (iv) of Theorem 2 to consider the case when $\boldsymbol{x} \in \mathcal{S}^1 \cap \mathcal{S}^K$.

Put $H = \{(i,j) \in \mathbb{N} \times \mathbb{N} : i < j\}$ and let $\{N_{ij}\}_{(i,j)\in H}$ be a collection of independent Poisson random measures on $[0,1] \times \mathbb{R}_+$ with intensity Lebesgue measure on $[0,1] \times \mathbb{R}_+$.

Define $G_k : \mathcal{S} \times H \times [0,1] \to \mathbb{R}$, $k \in \mathbb{N}$, by

$$G_k(\boldsymbol{x},(i,j),u) = \begin{cases} x_j, & \text{if } k = i,\ 0 \le u \le (x_i + x_j)1(x_i \neq 0, x_j \neq 0), \\ -x_j, & \text{if } k = j,\ 0 \le u \le (x_i + x_j)1(x_i \neq 0, x_j \neq 0), \\ 0, & \text{otherwise.} \end{cases}$$

It is elementary to construct a càdlàg $\mathcal{S}^1 \cap \mathcal{S}^K$-valued solution $\breve{\boldsymbol{X}}$ to the family of SDEs

$$\breve{X}_k(t) = x_k + \sum_{i<j} \int_{[0,1]\times[0,t]} G_k(\breve{X}(s-),(i,j),u)\, N_{ij}(du,ds), \tag{40}$$

and $\breve{\boldsymbol{X}}$ will have law $\mathbb{Q}^{\boldsymbol{x}}$. We may therefore suppose that $\boldsymbol{X}$ has been constructed as a solution to (40).

Applying the "Itô formula" for stochastic integrals against an (uncompensated) Poisson random measure (see, for example, [21, Theorem II.5.1]) and taking expectations, we have that

$$\begin{aligned} \mathbb{Q}^{\boldsymbol{x}}\left[\sum_{j=k}^{\infty} X_j^2(t)\right] &= \sum_{j=k}^{\infty} x_j^2 - \mathbb{Q}^{\boldsymbol{x}}\left[\int_0^t \sum_{i=1}^{k-1}\sum_{j=k}^{\infty} X_j^2(s)\{X_i(s)+X_j(s)\}1(X_i(s)\neq 0, X_j(s)\neq 0)\,ds\right] \\ &+ 2\mathbb{Q}^{\boldsymbol{x}}\left[\int_0^t \sum_{i=k}^{\infty}\sum_{j=i+1}^{\infty} X_i(s)X_j(s)\{X_i(s)+X_j(s)\}1(X_i(s)\neq 0, X_j(s)\neq 0)\,ds\right]. \end{aligned} \tag{41}$$

Observe that

$$2\sum_{i=k}^{\infty}\sum_{j=i+1}^{\infty} X_i(s)X_j(s)\{X_i(s)+X_j(s)\}1(X_i(s)\neq 0, X_j(s)\neq 0) \le 4\sum_{j=k}^{\infty} X_j^2(s),$$

and the result follows from Gronwall's lemma.

□

For $\boldsymbol{x} \in \mathcal{S}$ and $k \in \mathbb{N}$ let $\bar{x}_k := \sum_{i=k}^{\infty} x_i$.

Theorem 24 *Suppose that $\boldsymbol{x} \in \mathcal{S}^1$ satisfies each of the following three conditions:*

(a) $\bar{x}_k > 0$ *for all* $k \in \mathbb{N}$*,*

(b) $\lim_{k\to\infty} \sum_{j=k}^{\infty} x_j^2/(\bar{x}_k)^2 = 0$*,*

(c) $\lim_{k\to\infty} k \sum_{j=k}^{\infty} x_j^2/\bar{x}_k = 0$*.*

Then, for all $T \geq 0$ *and* $\epsilon > 0$*,*

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}} \left\{ \sup_{0\leq t\leq T} |\bar{X}_k(t)/\bar{x}_k - e^{-t}| > \epsilon \right\} = 0.$$

Remark 25 The above conditions on $\boldsymbol{x} = (x_k)$ are satisfied if (x_k) is regularly varying of order $-\alpha$ for some $\alpha > 1$.

Proof. We claim that under $\mathbb{Q}^{\boldsymbol{y}}$ for any $\boldsymbol{y} \in \mathcal{S}^1$, the process $\bar{X}_k$ has the semimartingale decomposition

$$\bar{X}_k(t) = \bar{X}_k(0) + M_k(t) - \int_0^t \sum_{i=1}^{k-1} \sum_{j=k}^{\infty} X_j(s)\{X_i(s) + X_j(s)\} 1(X_i(s) \neq 0,\, X_j(s) \neq 0)\, ds, \tag{42}$$

where M_k is a martingale. As in the proof of Lemma 23, when $\boldsymbol{y}$ has finite support it is elementary to construct a solution to (40) with starting point $\boldsymbol{y}$ and the solution has law $\mathbb{Q}^{\boldsymbol{y}}$. It is then immediate that (42) holds when $\boldsymbol{y}$ is finitely supported. The general case follows from part (iv) of Theorem 2.

Observe that under $\mathbb{Q}^{\boldsymbol{x}}$, for $s \geq 0$ and $\ell \geq k$,

$$\begin{aligned} &\sum_{i=1}^{k-1} \sum_{j=k}^{\infty} X_j(s)\{X_i(s) + X_j(s)\} 1(X_i(s) \neq 0,\, X_j(s) \neq 0) \\ &\quad \leq \sum_{i=1}^{k-1} \sum_{j=k}^{\infty} X_j(s)\{X_i(s) + X_j(s)\} \\ &\quad = \bar{X}_k(s) + (k-1) \sum_{j=k}^{\infty} X_j^2(s). \end{aligned} \tag{43}$$

On the other hand, as

$$X_\ell(s) \leq \bar{X}_\ell(s) \leq \bar{X}_\ell(0) \leq \bar{X}_k(0) = \bar{x}_k,$$

it follows that

$$\begin{aligned}
&\sum_{i=1}^{k-1}\sum_{j=k}^{\infty} X_j(s)\{X_i(s)+X_j(s)\}1(X_i(s)\neq 0,\, X_j(s)\neq 0) \\
&\quad\geq \sum_{i=1}^{k-1}\sum_{j=k}^{\infty} X_j(s)X_i(s)1(X_i(s)\neq 0)1(X_j(s)\neq 0) \\
&\quad= \bar{X}_k(s)\{1-\bar{X}_k(s)\} \\
&\quad\geq \bar{X}_k(s)\{1-\bar{x}_k\}.
\end{aligned} \tag{44}$$

In order to prove the theorem it suffices, by Lemma 22, Lemma 23, (43), (44), and assumption (c), to show that for all $T \geq 0$ and $\epsilon > 0$

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}}\{\sup_{0\leq t\leq T} |M_k(t)/\bar{x}_k| > \epsilon\} = 0. \tag{45}$$

By standard facts about the "angle-brackets" of stochastic integrals against compensated Poisson random measures (see, for example, [21, §II.3] we have for finitely supported $\boldsymbol{y}$ that

$$\mathbb{Q}^{\boldsymbol{y}}[M_k^2(t)] = \mathbb{Q}^{\boldsymbol{y}}\left[\sum_{i=1}^{k-1}\sum_{j=k}^{\infty}\int_0^t X_j^2\{X_i(s)+X_j(s)\}1(X_i(s)\neq 0,\, X_j(s)\neq 0)\, ds\right] \tag{46}$$

It follows from part (iv) of Theorem 2 that (46) holds for all $\boldsymbol{y}$. Thus,

$$\begin{aligned}
&\sum_{i=1}^{k-1}\sum_{j=k}^{\infty} X_j^2\{X_i(s)+X_j(s)\}1(X_i(s)\neq 0,\, X_j(s)\neq 0) \\
&\quad\leq \sum_{j=k}^{\infty} X_j^2(s) + (k-1)\sum_{j=k}^{\infty} X_j^3(s) \\
&\quad\leq \sum_{j=k}^{\infty} X_j^2(s)\{1+(k-1)\bar{x}_k\}.
\end{aligned}$$

Hence, by Lemma 23 and assumptions (b) and (c),

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}}[\{M_k(t)/\bar{x}_k\}^2] = 0,$$

and (45) holds by the L^2 maximal inequality.

□

Set

$$L_k(t) = \#\{1 \le i \le k : X_i(t) \ne 0\},$$

$$\Lambda_k(t) = \tau_k(\boldsymbol{X}(t)) = \inf\{\ell \in \mathbb{N} : L_\ell(t) = k\}.$$

Proposition 26 *Suppose that $\boldsymbol{x} \in \mathcal{S}^1$ is such that*

$$\lim_{k\to\infty} \#\{1 \le i \le k : x_i \ne 0\}/k = \beta.$$

Then, for all $T \ge 0$ and $\epsilon > 0$,

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}}\{\sup_{0\le t\le T} |L_k(t)/k - \beta e^{-t}| > \epsilon\} = 0$$

and

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}}\{\sup_{0\le t\le T} |\Lambda_k(t)/k - \beta^{-1}e^{t}| > \epsilon\} = 0.$$

Proof. The claim for L clearly implies the claim for Λ, so it suffices to prove the former.

Arguing as in the proof of Theorem 24, we have the semimartingale decomposition

$$L_k(t) = L_k(0) + M_k(t) - \int_0^t \sum_{1\le i<j\le k} (X_i(s) + X_j(s))1\{X_i(s) \ne 0, X_j(s) \ne 0\}\, ds,$$

where M_k is a martingale. Observe that

$$\begin{aligned}
&\sum_{1\le i<j\le k} (X_i(s) + X_j(s))1\{X_i(s) \ne 0, X_j(s) \ne 0\} \\
&\quad = \frac{1}{2}\left[\sum_{1\le i,j\le k} (X_i(s) + X_j(s))1\{X_i(s) \ne 0, X_j(s) \ne 0\} - 2\sum_{1\le \ell\le k} X_\ell(s)1\{X_\ell(s) \ne 0\}\right] \\
&\quad = [L_k(s) - 1]\sum_{1\le \ell\le k} X_\ell(s).
\end{aligned}$$

The result will therefore follow from Lemma 22 if we can show that for all $T \ge 0$ and $\epsilon > 0$,

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}}\left\{\sup_{0\le t\le T} |M_k(t)/k| > \epsilon\right\} = 0.$$

However, an "angle brackets" calculation similar to the one in the proof of Theorem 24 gives that

$$\mathbb{Q}^{\boldsymbol{x}}[M_k(t)^2] = \mathbb{Q}^{\boldsymbol{x}}\left[\int_0^t \sum_{1\le i<j\le k} (X_i(s) + X_j(s))1\{X_i(s) \neq 0, X_j(s) \neq 0\}\, ds\right] \le (k-1)t;$$

and an application of the L^2 maximal inequality completes the proof.

$\square$

Corollary 27 *Suppose that $\boldsymbol{x} \in \mathcal{S}^1$ satisfies the conditions of Theorem 24, the sequence $\{\bar{x}_k\}_{k\in\mathbb{N}}$ is regularly varying of order $-\gamma$, $\gamma > 0$, and*

$$\lim_{k\to\infty} \#\{1 \le i \le k : x_i \neq 0\}/k = \beta.$$

Then, for all $T \ge 0$ and $\epsilon > 0$,

$$\lim_{k\to\infty} \mathbb{Q}^{\boldsymbol{x}}\left\{\sup_{0\le t\le T} |(\overline{\mathtt{SHUNT} \circ \boldsymbol{X}})_k(t)/\bar{x}_k - \beta^\gamma e^{-(1+\gamma)t}| > \epsilon\right\} = 0.$$

Thus, if $\boldsymbol{x}^* = \mathtt{SHUNT}(\boldsymbol{x})$,

$$\lim_{k\to\infty} \mathbb{Q}_*^{\boldsymbol{x}^*}\left\{\sup_{0\le t\le T} |\overline{X}_k^*(t)/\bar{x}_k^* - e^{-(1+\gamma)t}| > \epsilon\right\} = 0,$$

Proof. Note that

$$(\mathtt{SHUNT} \circ \boldsymbol{X})_k(t) = X_{\Lambda_k(t)}(t),\ k \in \mathbb{N},$$

by definition, and so

$$(\overline{\mathtt{SHUNT} \circ \boldsymbol{X}})_k(t) = \bar{X}_{\Lambda_k(t)}(t),\ k \in \mathbb{N}.$$

The result now follows immediately from Theorem 24, Proposition 26, and the properties of regularly varying sequences.

$\square$

Remark 28 In view of Remark 25, if the sequence (x_k) is regularly varying of order $-\alpha$, $\alpha > 1$, then for all $T \ge 0$ and $\epsilon > 0$,

$$\lim_{k\to\infty} \mathbb{Q}_*^{\boldsymbol{x}}\left\{\sup_{0\le t\le T} |\overline{X}_k^*(t)/\bar{x}_k - e^{-\alpha t}| > \epsilon\right\} = 0.$$

7 Exchangeable $\mathcal{P}_\infty$-coalescents

Call a $\mathcal{P}_\infty$-valued process $(W_t,\, t \in \mathbb{I})$ *exchangeable*, if the distribution of each of the processes $(\pi^n \circ W_t,\, t \in \mathbb{I})$ is invariant under the action of permutations of $\mathbb{N}_n$ on the space $\mathcal{P}_n$ of partitions of $\mathbb{N}_n$. An example is provided by Kingman's $\mathcal{P}_\infty$-coalescent $(W_t,\, t \geq 0)$ described in Example 7. If $(W_t,\, t \in \mathbb{I})$ is an exchangeable $\mathcal{P}_\infty$-valued process then each of the random partitions W_t is an exchangeable random partition of $\mathbb{N}$, as studied in [23, 24, 4, 32]. Let $2^{\mathbb{N}}$ denote the set of all subsets of $\mathbb{N}$. Define a map $\mathtt{GAL} : \mathbb{N} \times \mathcal{P}_\infty \to 2^{\mathbb{N}}$ as follows. If i is the least element of some component of v, let $\mathtt{GAL}(i, v)$ be that component; otherwise let $\mathtt{GAL}(i, v)$ be the empty set. Thinking of the components of v as representing galaxies, call $\mathtt{GAL}(i, v)$ the *galaxy labeled i in the partition v*. Say that a partition $v \in \mathcal{P}_\infty$ *has frequencies* if the asymptotic frequency

$$\mathtt{FREQ}(i, v) = \lim_{n\to\infty} |\mathtt{GAL}(i, v) \cap \mathbb{N}_n|/n \tag{47}$$

exists for all i. And say that v has *proper frequencies* if also

$$\sum_{i=1}^{\infty} \mathtt{FREQ}(i, v) = 1 \tag{48}$$

The following lemma is elementary:

Lemma 29 *Suppose that v has proper frequencies and that w is a coarsening of v. Then w has proper frequencies given by the formula*

$$\mathtt{FREQ}(j, w) = \sum_i \mathtt{FREQ}(i, v) 1\{\mathtt{GAL}(i, v) \subseteq \mathtt{GAL}(j, w)\} \tag{49}$$

Kingman [23, 24] showed that if W is an exchangeable random partition of $\mathbb{N}$, then W has frequencies. Combined with the above lemma this implies the following proposition:

Proposition 30 *Suppose that $(W_t,\, t \in \mathbb{I})$ is an exchangeable $\mathcal{P}_\infty$-coalescent process such that W_t has proper frequencies almost surely for each $t \in \mathbb{I}$. Then W_t has proper frequencies for all $t \in \mathbb{I}$ almost surely. Let*

$$\boldsymbol{X}(t) = (\mathtt{FREQ}(i, W_t), i \geq 1) \tag{50}$$

Then the process $(\boldsymbol{X}(t),\, t \in \mathbb{I})$ *is a leftward* $\mathcal{S}$*-coalescent process with tracking functions* $\Psi_{s,t}$ *defined by* $\Psi_{s,t}(i) = j$ *if* $\texttt{FREQ}(i, W_s) > 0$ *and* $\texttt{GAL}(i, W_s) \subseteq \texttt{GAL}(j, W_t)$*.*

The transformation (50) is an analog of the transformation (27) applied in Section 5, with masses defined by asymptotic frequencies instead of an arbitrary distribution $\boldsymbol{p}$. Proposition 30 shows how to transform an exchangeable $\mathcal{P}_\infty$-coalescent with proper frequencies into an $\mathcal{S}$-coalescent. The following proposition shows that modulo labeling every $\mathcal{S}$-coalescent has the same distribution as one that has been constructed by this transformation.

Proposition 31 *Given an* $\mathcal{S}$*-coalescent process* $(\boldsymbol{X}(t),\, t \in \mathbb{I})$*, there exists an exchangeable* $\mathcal{P}_\infty$*-coalescent process* $(W_t,\, t \in \mathbb{I})$ *such that*

$$(\texttt{RANK}(\boldsymbol{X}(t)),\, t \in \mathbb{I}) \stackrel{d}{=} (\texttt{RANK}(\texttt{FREQ}(i, W_t), i \geq 1),\, t \in \mathbb{I}) \tag{51}$$

where $\stackrel{d}{=}$ *denotes equality of finite-dimensional distributions.*

Proof. Since the distribution of a $\mathcal{P}_\infty$-coalescent is determined its sequence of restrictions to $\mathcal{P}_n$, by application of the Kolmogorov extension theorem it suffices to prove the existence for each $s \in \mathbb{I}$ of an exchangeable $\mathcal{P}_\infty$-coalescent $(W_t,\, t \in \mathbb{I}, t \geq s)$ such that (51) holds with t restricted to $t \in \mathbb{I}$ with $t \geq s$. Such a $\mathcal{P}_\infty$-coalescent can be constructed as follows. First enlarge the probability space on which $(\boldsymbol{X}(t),\, t \in \mathbb{I})$ is defined to construct a sequence $I_0, I_1, \ldots$ of $\mathbb{N}$-valued random variables which conditionally given $(\boldsymbol{X}(t),\, t \in \mathbb{I})$ are independent with identical distribution $\boldsymbol{X}(s)$. For $t \in \mathbb{I}$ with $t \geq s$ let W_t be the partition of $\mathbb{N}$ generated by the random equivalence relation $m \sim n$ iff $\Psi_{s,t}(I_m) = \Psi_{s,t}(I_n)$ where $\Psi_{s,t}$ is the tracking function associated with $\boldsymbol{X}$. Then, by construction of W_t and the law of large numbers for the sequence $I_0, I_1, \ldots$, the process $(W_t,\, t \in \mathbb{I}, t \geq s)$ is an exchangeable $\mathcal{P}_\infty$-coalescent such that $\texttt{RANK}(\boldsymbol{X}(t)) = \texttt{RANK}(\texttt{FREQ}(i, W_t), i \geq 1)$ almost surely for each $t \in \mathbb{I}$ with $t \geq s$.

□

Let ISBP denote the set of distributions of $\mathcal{S}$-valued random elements which are invariant under the operation of size-biased random permutation. A number of characterizations of ISBP are known [34], one of which is that $\Pi \in$ ISBP if and only if Π is the distribution of $\texttt{SHUNT}(\texttt{FREQ}(i,W), i \geq 1)$ for some exchangeable random partition W of $\mathbb{N}$ with proper frequencies. As an application of the above ideas, there is the following proposition, which generalizes some of the observations regarding the shunted additive coalescent that were made in Section 6.1.

Proposition 32 *Let $(Q_t^*)_{t\geq 0}$ denote the semigroup of a shunted additive $(\mathcal{S}_\beta^*, \kappa)$-coalescent as in Theorem 2. Let Π be a probability distribution on $\mathcal{S}_\beta^*$. If $\Pi \in$ ISBP then $\Pi Q_t^* \in$ ISBP for each $t > 0$.*

See also [36] where the ideas of this section are applied to characterize the entrance boundary of the $\mathcal{S}^\downarrow$-coalescent derived from the $\mathcal{P}_\infty$-valued Λ-coalescent mentioned in Example 8.

8 Open Problems

For each of the shunted and ranked $\mathcal{S}$-coalescent semigroups discussed in this paper, there is the problem of characterizing the possible entrance laws for a Markov process with the given semigroup and time parameter set $\mathbb{I} =]-\infty, \infty[$. See Aldous and Limic [7] and Aldous [6] for a treatment of this problem for the multiplicative coalescent, and related questions. From Section 6.1 we have existence of a non-trivial entrance law for a shunted or ranked additive coalescent with $\mathbb{I} =]-\infty, \infty[$. It can be shown [8, 3] that there are other non-trivial extreme entrance laws in this case which are not just shifts of this entrance law, and an explicit description of all extreme entrance laws can be given. For an adequate discussion of the entrance boundary problem for more general collision kernel κ that satisfy the Lipschitz condition (15) it would seem necessary to first enlarge the statespaces of the shunted and ranked processes to all of $\mathcal{S}^*$ and $\mathcal{S}^\downarrow$, rather than the statespaces $\mathcal{S}_1^*$ and $\mathcal{S}_1^\downarrow$ required in Theorem 2. Much remains to be understood about how the evolution and asymptotic behavior of a coalescent process are affected by initial conditions.

Acknowledgments. We thank D.J. Aldous and J.R. Norris for pointing out errors in an earlier version of this paper.

References

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

[2] D. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures and Algorithms*, 5:487–512, 1994.

[3] D. Aldous and J. Pitman. The entrance boundary of the additive coalescent. Paper in preparation, 1997.

[4] 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.

[5] D.J. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993.

[6] D.J. Aldous. Deterministic and stochastic models for coalescence: a review of the mean-field theory for probabilists. To appear in *Bernoulli*. Available via homepage http://www.stat.berkeley.edu/users/aldous, 1997.

[7] D.J. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. Unpublished. Available via homepage http://www.stat.berkeley.edu/users/aldous, 1997.

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

[9] A. Benveniste and J. Jacod. Systèmes de Lévy des processus de Markov. *Invent. Math.*, 21:183–198, 1973.

[10] R.M. Blumenthal and R.K. Getoor. *Markov Processes and Potential Theory*. Academic Press, New York, 1968.

[11] E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. Preliminary Draft, Sept. 1997.

[12] A. Broder. Generating random spanning trees. In *Proc. 30'th IEEE Symp. Found. Comp. Sci.*, pages 442–447, 1989.

[13] A. Cayley. A theorem on trees. *Quarterly Journal of Pure and Applied Mathematics*, 23:376–378, 1889. (Also in *The Collected Mathematical Papers of Arthur Cayley. Vol XIII*, 26-28, Cambridge University Press, 1897).

[14] D.A. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour, 1991, Springer Lecture Notes in Mathematics, Vol. 1117*, 1994.

[15] P. Donnelly and P. Joyce. Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Processes and their Applications*, 31:89 – 103, 1989.

[16] P. Donnelly and S. Simons. On the stochastic approach to cluster size distribution during particle coagulation I: Asymptotic expansion in the deterministic limit. *J. Phys. A: Math. Gen.*, 26:2755–2767, 1993.

[17] S.N. Ethier and T.G. Kurtz. Fleming-Viot processes in population genetics. *SIAM Journal on Control and Optimization*, 31:345–386, 1993.

[18] S.N. Evans. Infinitely-many-species Lotka-Volterra equations arising from systems of coalescing masses. Technical Report 478, Dept. Statistics, U.C. Berkeley, 1997. To appear in *J. London Math. Soc.*. Available via http://www.stat.berkeley.edu/tech-reports.

[19] S.N. Evans and J. Pitman. Stationary Markov processes related to stable Ornstein-Uhlenbeck processes and the additive coalescent. Technical Report 488, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[20] E.M. Hendriks, J.L. Spouge, M. Eibl, and M. Shreckenberg. Exact solutions for random coagulation processes. *Z. Phys. B - Condensed Matter*, 58:219–227, 1985.

[21] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North Holland/Kodansha, 1989. 2nd edition.

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

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

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

[25] 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.

[26] C. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. *Mon. Not. R. Astron. Soc.*, 262:627–649, 1993.

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

[28] R. Lyons and Y. Peres. Probability on trees and networks. Book in preparation, available at http://www.ma.huji.ac.il/ lyons/prbtree.html, 1996.

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

[30] M. Perman. Order statistics for jumps of normalized subordinators. *Stoch. Proc. Appl.*, 46:267–281, 1993.

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

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

[33] J. Pitman. Coalescent random forests. Technical Report 457, Dept. Statistics, U.C. Berkeley, 1996. Available via http://www.stat.berkeley.edu/users/pitman.

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

[35] J. Pitman. Abel-Cayley-Hurwitz multinomial expansions associated with random mappings, forests and subsets. Technical Report 498, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[36] J. Pitman. Coalescents with multiple collisions. Technical Report 495, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

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

[38] A. Rényi. On the enumeration of trees. In R. Guy, H. Hanani, N. Sauer, and J. Schonheim, editors, *Combinatorial Structures and their Applications*, pages 355–360. Gordon and Breach, New York, 1970.

[39] R.K. Sheth. Merging and hierarchical clustering from an initially Poisson distribution. *Mon. Not. R. Astron. Soc.*, 276:796–824, 1995.

[40] R.K. Sheth and J. Pitman. Coagulation and branching process models of gravitational clustering. *Mon. Not. R. Astron. Soc.*, 289:66–80, 1997.

[41] R. Stanley. Enumerative combinatorics, vol. 2. Book in preparation, to be published by Cambridge University Press, 1996.

[42] H. Tanaka and K. Nakazawa. Stochastic coagulation equation and validity of the statistical coagulation equation. *J. Geomag. Geoelectr.*, 45:361–381, 1993.

[43] H. Tanaka and K. Nakazawa. Validity of the stochastic coagulation equation and runaway growth of protoplanets. *Icarus*, 107:404–412, 1994.

[44] H.M. Taylor and S. Karlin. *An Introduction to Stochastic Modeling (revised ed.)*. Academic Press, Boston, 1994.

[45] P.G.J. van Dongen. Fluctuations in coagulating systems II. *J. Statist. Phys.*, 49:927–975, 1987.

[46] P.G.J. van Dongen and M.H. Ernst. Fluctuations in coagulating systems. *J. Statist. Phys.*, 49:879–926, 1987.

[47] M. von Smoluchowski. Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Physik. Z.*, 17:557–585, 1916.