DOI: 10.1214/EJP.v5-68

Electronic Journal of Probability

Vol. 5 (2000) Paper no. 12, pages 1–50.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol5/paper12.abs.html

COALESCENTS WITH SIMULTANEOUS MULTIPLE COLLISIONS

Jason Schweinsberg

Department of Statistics, University of California, Berkeley
367 Evans Hall #3860, Berkeley, CA 94720-3860, USA
jason@stat.berkeley.edu

Abstract We study a family of coalescent processes that undergo "simultaneous multiple collisions," meaning that many clusters of particles can merge into a single cluster at one time, and many such mergers can occur simultaneously. This family of processes, which we obtain from simple assumptions about the rates of different types of mergers, essentially coincides with a family of processes that Möhle and Sagitov obtain as a limit of scaled ancestral processes in a population model with exchangeable family sizes. We characterize the possible merger rates in terms of a single measure, show how these coalescents can be constructed from a Poisson process, and discuss some basic properties of these processes. This work generalizes some work of Pitman, who provides similar analysis for a family of coalescent processes in which many clusters can coalesce into a single cluster, but almost surely no two such mergers occur simultaneously.

Keywords coalescence, ancestral processes, Poisson point processes, Markov processes, exchangeable random partitions.

AMS subject classification 60J25, 60G09, 60G55, 05A18, 60J75.

Submitted to EJP on May 24, 2000. Final version accepted on July 10, 2000.

1 Introduction

In this paper, we study a family of coalescent processes that undergo "simultaneous multiple collisions," meaning that many clusters of particles can merge into a single cluster at one time, and many such mergers can occur simultaneously. These processes were previously introduced in [13] by Möhle and Sagitov, who obtained them by taking limits of scaled ancestral processes in a population model with exchangeable family sizes. Here we take a different approach to characterizing these processes. The approach is similar to that used by Pitman in [16] for "coalescents with multiple collisions," also called Λ-coalescents, in which many clusters of particles can merge at one time into a single cluster but almost surely no two such mergers occur simultaneously. The family of processes studied here includes the Λ-coalescents as a special case, and we generalize several facts about Λ-coalescents.

Let $\mathcal{P}_n$ denote the set of partitions of $\{1, \dots, n\}$, and let $\mathcal{P}_\infty$ denote the set of partitions of $\mathbb{N} = \{1, 2, \dots\}$. Given $m < n \leq \infty$ and $\pi \in \mathcal{P}_n$, let $R_m \pi$ be the partition in $\mathcal{P}_m$ obtained by restricting π to $\{1, \dots, m\}$. That is, if $1 \leq i < j \leq m$, then i and j are in the same block of the partition $R_m \pi$ if and only if they are in the same block of π. Following [16], we identify each $\pi \in \mathcal{P}_\infty$ with the sequence $(R_1\pi, R_2\pi, \dots) \in \mathcal{P}_1 \times \mathcal{P}_2 \times \dots$. Each $\mathcal{P}_n$ is given the discrete topology and $\mathcal{P}_\infty$ is given the topology that it inherits from the product $\mathcal{P}_1 \times \mathcal{P}_2 \times \dots$, so $\mathcal{P}_\infty$ is compact and metrizable. We equip $\mathcal{P}_\infty$ with the Borel σ-field associated with this topology. We call a $\mathcal{P}_n$-valued process $(\Pi_n(t))_{t \geq 0}$ a *coalescent* if it has right-continuous step function paths and if $\Pi_n(s)$ is a refinement of $\Pi_n(t)$ for all $s < t$. We call a $\mathcal{P}_\infty$-valued process $(\Pi_\infty(t))_{t \geq 0}$ a *coalescent* if it has càdlàg paths and if $\Pi_\infty(s)$ is a refinement of $\Pi_\infty(t)$ for all $s < t$. Equivalently, $(\Pi_\infty(t))_{t \geq 0}$ is a coalescent if and only if for each n, the process $(R_n \Pi_\infty(t))_{t \geq 0}$ is a coalescent.

In [16], Pitman studies "coalescents with multiple collisions," which are $\mathcal{P}_\infty$-valued coalescents $(\Pi_\infty(t))_{t \geq 0}$ with the property that for each $n \in \mathbb{N}$, the process $(R_n \Pi_\infty(t))_{t \geq 0}$ is a $\mathcal{P}_n$-valued Markov chain such that when $R_n \Pi_\infty(t)$ has b blocks, each possible merger of k blocks into a single block is occurring at some fixed rate $\lambda_{b,k}$ that does not depend on n, and no other transitions are possible. It is shown in [16] that given a collection of rates $\{\lambda_{b,k} : 2 \leq k \leq b < \infty\}$, such a process exists if and only if the consistency condition

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1} \tag{1}$$

holds for all $2 \leq k \leq b$. Theorem 1 of [16] shows that (1) holds if and only if whenever $2 \leq k \leq b$, we have

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

for some finite measure Λ on $[0, 1]$. The process is then called the Λ-coalescent. When Λ is a unit mass at zero, we obtain Kingman's coalescent, a process introduced in [10] in which only two blocks can merge at a time and each pair of blocks is merging at rate 1. The case in which Λ is the uniform distribution on $[0, 1]$ was studied by Bolthausen and Sznitman in [3].

In [18], Sagitov obtains all of the Λ-coalescents, up to a time-scaling constant, as limits of ancestral processes in a haploid population model with an exchangeable distribution of family sizes in which there are N individuals in each generation. The ancestral processes are $\mathcal{P}_n$-valued processes obtained by sampling n out of N individuals from the current generation and tracing their ancestors backwards in time. A simpler formulation of this convergence result is presented in [12], and similar results for a diploid population model are given in [14].

An important property of the Λ-coalescent is that the rate at which blocks are merging does not depend on the size of the blocks or on which integers are in the blocks. It is therefore natural to pursue a generalization to a larger class of processes that still have this property but that may undergo "simultaneous multiple collisions." The possibility of such a generalization is mentioned in section 3.3 of [16]. We define a $(b; k_1, \ldots, k_r; s)$*-collision* to be a merger of b blocks into $r + s$ blocks in which s blocks remain unchanged and the other r blocks contain $k_1, \ldots, k_r \geq 2$ of the original blocks. Thus, $b = \sum_{j=1}^r k_j + s$. The order of $k_1, \ldots, k_r$ does not matter; for example, any $(5; 3, 2; 0)$-collision is also a $(5; 2, 3; 0)$-collision. It is easily checked (see equation (11) of [15]) that if $r \geq 1$, $k_1, \ldots, k_r \geq 2$, $s \geq 0$, $b = \sum_{j=1}^r k_j + s$, and l_j is the number of $k_1, \ldots, k_r$ that equal j, then the number of possible $(b; k_1, \ldots, k_r; s)$-collisions is

$$\frac{b!}{s! \prod_{j=2}^b (j!)^{l_j} l_j!} = \binom{b}{k_1 \ldots k_r \ s} \frac{1}{\prod_{j=2}^b l_j!}. \tag{3}$$

We define a *coalescent with simultaneous multiple collisions* to be a $\mathcal{P}_\infty$-valued coalescent process $(\Pi_\infty(t))_{t \geq 0}$ with the property that for each $n \in \mathbb{N}$, the process $(R_n \Pi_\infty(t))_{t \geq 0}$ is a $\mathcal{P}_n$-valued Markov chain such that when $R_n \Pi_\infty(t)$ has b blocks, each possible $(b; k_1, \ldots, k_r; s)$-collision is occurring at some fixed rate $\lambda_{b; k_1, \ldots, k_r; s}$.

In [13], Möhle and Sagitov generalize the proofs in [12] and obtain coalescents with simultaneous multiple collisions as limits of ancestral processes in a haploid population model. We now describe their model and their results. Assume there are N individuals in each generation. For all $a \geq 0$, let $\nu_{1,N}^{(a)}, \ldots, \nu_{N,N}^{(a)}$ denote the family sizes in the ath generation backwards in time, where $\nu_{i,N}^{(a)}$ is the number of offspring of the ith individual in the $(a+1)$st generation backwards in time. Note that $\nu_{1,N}^{(a)} + \ldots + \nu_{N,N}^{(a)} = N$ because the population size is fixed. The random variables $\nu_{1,N}^{(a)}, \ldots, \nu_{N,N}^{(a)}$ are assumed to be independent for different generations. The distribution of $(\nu_{1,N}^{(a)}, \ldots, \nu_{N,N}^{(a)})$, which we denote by μ_N, is assumed to be exchangeable and to be the same for all a. Therefore, we will suppress the superscript in the notation when we are concerned only about the distributions of the family sizes. Möhle and Sagitov consider a random sample of $n \leq N$ distinct individuals from the 0th generation. They define the Markov chain $(\Psi_{n,N}(a))_{a=0}^\infty$, where $\Psi_{n,N}(a)$ is the random partition of $\{1, \ldots, n\}$ such that i and j are in the same block if and only if the ith and jth individuals in the sample have a common ancestor in the ath generation backwards in time. Let $(m)_k = m(m-1)\ldots(m-k+1)$, and let $(m)_0 = 1$. Let c_N be the probability that two individuals chosen randomly from some generation have the same ancestor in the previous generation. ¿From equation (5) of [13], we have

$$c_N = \frac{E[(\nu_{1,N})_2]}{N-1}. \tag{4}$$

The following result is part of Theorem 2.1 in [13].

Proposition 1 *In the population model described above, suppose for all $r \in \mathbb{N}$ and $k_1, \ldots, k_r \geq 2$, the limits*

$$\lim_{N \to \infty} \frac{E[(\nu_{1,N})_{k_1} \ldots (\nu_{r,N})_{k_r}]}{N^{k_1 + \ldots + k_r - r} c_N} \tag{5}$$

exist. Also suppose $\lim_{N \to \infty} c_N = 0$. Then as $N \to \infty$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge in the Skorohod topology to a coalescent process $(\Psi_{n,\infty}(t))_{t \geq 0}$. Here, $\Psi_{n,\infty}(0)$ is the partition *of* $\{1,\ldots,n\}$ *into singletons, and when* $\Psi_{n,\infty}(t)$ *has* b *blocks, each* $(b;k_1,\ldots,k_r;s)$*-collision is occurring at some fixed rate* $\lambda_{b;k_1,\ldots,k_r;s}$*. Moreover, there exists a unique sequence of measures* $(F_r)_{r=1}^{\infty}$ *satisfying the following three conditions:*

A1: each F_r *is concentrated on* $\Delta_r = \{(x_1,\ldots x_r) : x_i \geq 0$ *for all* i *and* $\sum_{i=1}^{r} x_i \leq 1\}$,

A2: each F_r *is symmetric with respect to the* r *coordinates of* Δ_r,

A3: $1 = F_1(\Delta_1) \geq F_2(\Delta_2) \geq \ldots$,

such that all of the collision rates satisfy

$$\lambda_{b;k_1,\ldots,k_r;s} = \sum_{m=r}^{\lfloor r+s/2 \rfloor} \int_{\Delta_m} x_1^{k_1-2} \ldots x_r^{k_r-2} T_{r,s}^{(m)}(x_1,\ldots,x_m) F_m(dx_1,\ldots dx_m), \tag{6}$$

where

$$T_{m,s}^{(m)}(x_1,\ldots,x_m) = (1 - \sum_{i=1}^{m} x_i)^s \tag{7}$$

and for $1 \leq j \leq m$,

$$T_{m-j,s}^{(m)}(x_1,\ldots,x_m) = (-1)^{j+1} \sum_{i_j=2j-1}^{i_{j+1}-2} \cdots \sum_{i_1=1}^{i_2-2} \prod_{k=0}^{j} i_k (1 - \sum_{i=1}^{m-k} x_i)^{i_{k+1}-i_k-2}, \tag{8}$$

when we set $i_0 = -1$ *and* $i_{j+1} = s+1$.

Möhle and Sagitov point out that when $s = 0$, equation (6) gives

$$\lambda_{b;k_1,\ldots,k_r;0} = \int_{\Delta_r} x_1^{k_1-2} \ldots x_r^{k_r-2}\, F_r(dx_1,\ldots,dx_r), \tag{9}$$

so the moments of the measures F_r are collision rates. Also, equation (9) above and equations (16) and (19) of [13] imply that

$$\lambda_{b;k_1,\ldots,k_r;0} = \lim_{N\to\infty} \frac{E[(\nu_{1,N})_{k_1} \ldots (\nu_{r,N})_{k_r}]}{N^{k_1+\ldots+k_r-r} c_N} \tag{10}$$

for all $r \geq 1$, $k_1,\ldots,k_r \geq 2$, and $b = \sum_{j=1}^{r} k_j$.

Note that some condition like the existence of the limits in (5) is needed to relate the distributions μ_N for different values of N. The condition $\lim_{N\to\infty} c_N = 0$ ensures that the limit obtained is a continuous-time process. If instead $\lim_{N\to\infty} c_N = c > 0$, then Theorem 2.1 of [13] states that the limit is a discrete-time Markov chain.

Note that the rates $\lambda_{b;k_1,\ldots,k_r;s}$ calculated in (6) do not depend on n. By Lemma 3.4 of [13] and equation (27) of [13], the rates satisfy the consistency condition

$$\lambda_{b;k_1,\ldots,k_r;s} = \sum_{m=1}^{r} \lambda_{b+1;k_1,\ldots,k_{m-1},k_m+1,k_{m+1},\ldots,k_r;s} + s\lambda_{b+1;k_1,\ldots,k_r,2;s-1} + \lambda_{b+1;k_1,\ldots,k_r;s+1}.$$

In section 4, we will prove Lemma 18, which shows that this condition implies the existence of a $\mathcal{P}_\infty$-valued coalescent $(\Pi_\infty(t))_{t\geq 0}$ with the property that $(R_n\Pi_\infty(t))_{t\geq 0}$ and $(\Psi_{n,\infty}(t))_{t\geq 0}$ have the same distribution for all $n \in \mathbb{N}$. Thus, coalescents with simultaneous multiple collisions can be derived from the ancestral processes studied in [13].

However, Möhle and Sagitov leave open the question of whether every possible coalescent with simultaneous multiple collisions can be obtained as a limit of ancestral processes in their population model. They also do not discuss the questions of which sequences of measures $(F_r)_{r=1}^{\infty}$ satisfying conditions A1, A2, and A3 of Proposition 1 are associated with coalescent processes in the manner described above, and whether there is a natural probabilistic interpretation of the measures F_r, aside from the interpretation of their moments as collision rates.

The primary goals of this paper are to answer these questions, and to establish an alternative characterization of coalescents with simultaneous multiple collisions based on a single measure Ξ on the infinite simplex

$$\Delta = \{(x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots \geq 0, \sum_{i=1}^{\infty} x_i \leq 1\}.$$

We will show that, up to a scaling constant, all coalescents with simultaneous multiple collisions can be obtained as limits of ancestral process as described above, and therefore these coalescents can be characterized either by a single measure Ξ or by a sequence of measures $(F_r)_{r=1}^{\infty}$. One advantage to the characterization based on Ξ is that every finite measure on the infinite simplex is associated with a coalescent process. However, for a sequence of measures $(F_r)_{r=1}^{\infty}$ satisfying conditions A1, A2, and A3 of Proposition 1 to be associated with a coalescent process, we will show that it must satisfy an additional consistency condition that does not appear to be easy to check.

The rest of this paper is organized as follows. In section 2, we summarize the results that establish the two characterizations of coalescents with simultaneous multiple collisions. We also state results that give interpretations of the characterizing measures. In section 3, we give a Poisson process construction of these coalescents. The Poisson process construction is an important tool for studying the coalescents and is used in most of the proofs in the paper. In section 4 we prove the results stated in section 2. In section 5, we build on work done for the Λ-coalescent to establish some further properties of coalescents with simultaneous multiple collisions. We establish there some regularity properties of the coalescents and derive a condition for a coalescent with simultaneous multiple collisions to be a jump-hold Markov process with bounded transition rates. We also present some results related to the question of whether the coalescents "come down from infinity," meaning that only finitely many blocks remain at any time $t > 0$ even if the coalescent is started with infinitely many blocks at time zero. Finally in section 6, we discuss the discrete-time analogs of these processes, which can also arise as limits of ancestral processes in the population models studied in [13].

2 Summary of results characterizing the coalescents

In this section, we summarize the results needed to establish the two characterizations of the coalescents with simultaneous multiple collisions, one involving a single measure Ξ and the other involving a sequence of measures $(F_r)_{r=1}^{\infty}$. The proofs of all propositions and theorems in this section are given in section 4.

We first state the main theorem of this paper, which characterizes coalescents with simultaneous multiple collisions in terms of a measure Ξ on the infinite simplex Δ. In the statement of this result, and throughout the rest of the paper, we refer to the point $(0,0,\ldots) \in \Delta$ as "zero," and we denote a generic point in Δ by $x = (x_1, x_2, \ldots)$.

Theorem 2 *Let $\{\lambda_{b;k_1,\ldots,k_r;s} : r \geq 1,\ k_1,\ldots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$ be a collection of nonnegative real numbers. Then there exists a $\mathcal{P}_\infty$-valued coalescent $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ satisfying:*

B1: $\Pi_\infty(0)$ is the partition of $\mathbb{N}$ into singletons,

B2: for each n, $\Pi_n = R_n\Pi_\infty$ is a Markov chain such that when $\Pi_n(t)$ has b blocks, each $(b;k_1,\ldots,k_r;s)$-collision is occurring at the rate $\lambda_{b;k_1,\ldots,k_r;s}$,

if and only if there is a finite measure Ξ on the infinite simplex Δ of the form $\Xi = \Xi_0 + a\delta_0$, where Ξ_0 has no atom at zero and δ_0 is a unit mass at zero, such that $\lambda_{b;k_1,\ldots,k_r;s}$ equals

$$\int_\Delta \left(\sum_{l=0}^{s} \sum_{i_1\neq\ldots\neq i_{r+l}} \binom{s}{l} x_{i_1}^{k_1}\ldots x_{i_r}^{k_r} x_{i_{r+1}}\ldots x_{i_{r+l}} \Big(1-\sum_{j=1}^{\infty} x_j\Big)^{s-l} \right) \Big/ \sum_{j=1}^{\infty} x_j^2 \,\Xi_0(d\mathbf{x}) + a1_{\{r=1,k_1=2\}} \tag{11}$$

for all $r \geq 1$, $k_1,\ldots,k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$.

Definition 3 *We call a coalescent process satisfying B2 whose collision rates are given by (11) for a particular finite measure Ξ on Δ a* Ξ-coalescent*. We call a Ξ-coalescent satisfying B1 the* standard Ξ-coalescent*.*

Suppose $\pi \in \mathcal{P}_\infty$ and $B_1, B_2, \ldots$ are the blocks of π. If $(\Pi_\infty(t))_{t\geq 0}$ is a standard Ξ-coalescent, then we can define a Ξ-coalescent $(\Pi_\infty^\pi(t))_{t\geq 0}$ satisfying $\Pi_\infty^\pi(0) = \pi$ by defining $i \in B_k$ and $j \in B_l$ to be in the same block of $\Pi_\infty^\pi(t)$ if and only if k and l are in the same block of $\Pi_\infty(t)$. Since any Ξ-coalescent can thus be derived easily from the standard Ξ-coalescent, we will restrict our attention to the standard Ξ-coalescent whenever it is simpler to do so.

For any finite measure Ξ on Δ, a collection of nonnegative collision rates can be defined by (11), so Theorem 2 implies that a standard Ξ-coalescent exists. The following proposition states that the collision rates of a coalescent with simultaneous multiple collisions uniquely determine the associated measure Ξ. Thus, there is a one-to-one correspondence between finite measures Ξ on Δ and coalescent processes satisfying conditions B1 and B2 of Theorem 2.

Proposition 4 *Let Ξ and Ξ' be finite measures on the infinite simplex Δ. Let $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ be a standard Ξ-coalescent. If Π_∞ is also a standard Ξ'-coalescent, then $\Xi = \Xi'$.*

Observe that we can easily recover the Λ-coalescent as a special case of the Ξ-coalescent. Suppose Ξ is concentrated on the subset of Δ consisting of the sequences $(x_1, x_2, \ldots)$ such that $x_i = 0$ for all $i \geq 2$. Then, $\lambda_{b;k_1,\ldots,k_r;s} = 0$ unless $r = 1$. When $r = 1$, the expression inside the double summation in the numerator in the integrand of (11) is zero unless $l = 0$, so

$$\lambda_{b;k_1;b-k_1} = \int_\Delta x_1^{k_1}(1-x_1)^{b-k_1}/x_1^2 \,\Xi_0(dx) + a1_{\{k_1=2\}}.$$

This result agrees with the formula for λ_{b,k_1} given in (2) when Λ is the projection of Ξ onto the first coordinate. Note that when Ξ is a unit mass at zero, the Ξ-coalescent is therefore Kingman's coalescent.

We next work towards giving an interpretation of the measure Ξ. Equation (11) implies that

$$\lambda_{2;2;0} = \int_\Delta \left(\sum_{i=1}^\infty x_i^2 \Big/ \sum_{j=1}^\infty x_j^2 \right) \Xi_0(dx) + a = \Xi_0(\Delta) + a = \Xi(\Delta). \tag{12}$$

If $\Xi(\Delta) = 0$, then all the collision rates are zero. Otherwise (12) implies that $\Xi = \lambda_{2;2;0} G$, where G is a probability measure defined by $G(S) = \Xi(S)/\Xi(\Delta)$ for all measurable subsets S of Δ. From (11), we see that if Ξ is multiplied by a constant, then all of the collision rates are multiplied by the same constant. Therefore, unless $\Xi = 0$, any Ξ-coalescent can be obtained from a G-coalescent, where G is a probability measure, by rescaling time by a constant factor. To interpret G, we first give the following definition.

Definition 5 *Let G be a probability measure on the infinite simplex Δ, and let $S \subset \mathbb{N}$. Let $\tilde{\Theta}$ be an exchangeable random partition of $\mathbb{N}$ such that the ranked sequence of limiting relative frequencies of the blocks of $\tilde{\Theta}$ has distribution G; such partitions are defined in appendix A. Let Θ be the random partition of S such that if $i, j \in S$, then i and j are in the same block of Θ if and only if i and j are in the same block of $\tilde{\Theta}$. A G-*partition of S *is defined to be a random partition of S with the same distribution as Θ.*

The following proposition, which is the natural analog of Theorem 4 of [16], gives an interpretation of G. We write $\#\pi$ for the number of blocks in a partition π.

Proposition 6 *Let G be a probability measure on the infinite simplex Δ. Let $\Pi_\infty = (\Pi_\infty(t))_{t \geq 0}$ be a standard G-coalescent. Let $T = \inf\{t : \#R_2\Pi_\infty(t) = 1\}$ be the collision time of 1 and 2. Let $B_1, B_2, \ldots$ be the blocks of $\Pi_\infty(T-)$. Let Θ be a partition of $\{3, 4, \ldots\}$ on $\{\#\Pi_\infty(T-) = \infty\}$ and a partition of $\{3, 4, \ldots, \#\Pi_\infty(T-)\}$ on $\{\#\Pi_\infty(T-) < \infty\}$ such that i and j are in the same block of Θ if and only if B_i and B_j are in the same block of $\Pi_\infty(T)$. If $P(\#\Pi_\infty(T-) = \infty) > 0$, then conditional on $\{\#\Pi_\infty(T-) = \infty\}$, Θ is a G-partition of $\{3, 4, \ldots\}$. If $P(\#\Pi_\infty(T-) = n) > 0$, then conditional on $\{\#\Pi_\infty(T-) = n\}$, Θ is a G-partition of $\{3, 4, \ldots, n\}$.*

Note that essentially the same result would hold if we defined T to be the time at which two arbitrary fixed integers merged, but we state the result in terms of the collision time of 1 and 2 to simplify notation.

We now turn to the question of whether all coalescents with simultaneous multiple collisions can arise as limits of ancestral processes in a population model of the type discussed in [13]. Since $\Psi_{n,\infty}(0)$, as defined in Proposition 1, equals the partition of $\{1, \ldots, n\}$ into singletons, only standard Ξ-coalescents can arise in this way. Also, it follows from (9) and condition A3 of Proposition 1 that for coalescents obtained from ancestral processes as described in the introduction, we have

$$\lambda_{2;2;0} = F_1(\Delta_1) = 1.$$

It then follows from (12) that Ξ is a probability measure. However, since $\lambda_{2;2;0}$ is just a time-scaling factor and the case $\Xi = 0$ is trivial, Proposition 7 below shows that the family of continuous-time processes that can be obtained from the ancestral processes studied in [13] should be regarded as essentially the same as the family of standard Ξ-coalescents.

Proposition 7 *Let Ξ be a probability measure on Δ. Then there exists a sequence $(\mu_N)_{N=1}^{\infty}$ such that each μ_N is a probability distribution on $\{0,1,2,\dots\}^N$ that is exchangeable with respect to the N coordinates with the property that if for all N, μ_N is the distribution of family sizes in the population model described in the introduction, then for all n, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ in the Skorohod topology to $(R_n \Pi_\infty(t))_{t \geq 0}$, where Π_∞ is a standard Ξ-coalescent.*

We can use Proposition 1 and Proposition 7 to characterize the coalescents with simultaneous multiple collisions by a sequence of measures $(F_r)_{r=1}^{\infty}$. We state this result precisely below. Note that the result can be stated without referring to the population model that originally motivated Möhle and Sagitov to study the measures F_r.

Proposition 8 *Let $\{\lambda_{b;k_1,\dots,k_r;s} : r \geq 1,\ k_1,\dots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^{r} k_j + s\}$ be a collection of nonnegative real numbers. Then there exists a $\mathcal{P}_\infty$-valued coalescent satisfying conditions B1 and B2 of Theorem 2 if and only if there is a sequence of measures $(F_r)_{r=1}^{\infty}$ satisfying*

A1: each F_r is concentrated on $\Delta_r = \{(x_1,\dots x_r) : x_i \geq 0$ for all i and $\sum_{i=1}^{r} x_i \leq 1\}$,

A2: each F_r is symmetric with respect to the r coordinates of Δ_r,

A3′: $F_1(\Delta_1) \geq F_2(\Delta_2) \geq \dots$,

such that (6) holds for all $\lambda_{b;k_1,\dots,k_r;s}$ in the collection. If such a sequence $(F_r)_{r=1}^{\infty}$ exists, then it is unique. Moreover, suppose $(F_r)_{r=1}^{\infty}$ is a sequence of measures satisfying A1, A2, and A3′. Define a collection of real numbers $\{\lambda_{b;k_1,\dots,k_r;s} : r \geq 1,\ k_1,\dots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^{r} k_j + s\}$ by (6). Then, a coalescent process satisfying B1 and B2 exists if and only if $(F_r)_{r=1}^{\infty}$ also satisfies

A4: the right-hand side of (6) is nonnegative for all $r \geq 1$, $k_1,\dots,k_r \geq 2$, and $s \geq 0$.

Conditions A1 and A2 come directly from Proposition 1. The reason for replacing A3 with A3′ is to obtain coalescents for which $\lambda_{2;2;0} \neq 1$. Condition A4, which is clearly necessary for $(F_r)_{r=1}^{\infty}$ to be associated with a coalescent process, can be viewed as a consistency condition on the measures F_r. The following example shows that A4 does not always hold.

Example 9 Consider the sequence of measures $(F_r)_{r=1}^{\infty}$ such that F_1 is a unit mass at 1, F_2 is a unit mass at $(1/2,1/2)$, and $F_r = 0$ for all $r \geq 3$. Then, conditions A1, A2, and A3′ are satisfied. If we define $\lambda_{4;2;2}$ by (6), we get

$$\begin{aligned} \lambda_{4;2;2} &= \int_{\Delta_1} T_{1,2}^{(1)}(x)\, F_1(dx) + \int_{\Delta_2} T_{1,2}^{(2)}(x_1,x_2)\, F_2(dx_1,dx_2) \\ &= T_{1,2}^{(1)}(1) + T_{1,2}^{(2)}(.5,.5). \end{aligned}$$

By (7), $T_{1,2}^{(1)}(x) = (1-x)^2$, so $T_{1,2}^{(1)}(1) = 0$. Using (8) with $j = 1$ and $s = 2$, and recalling that $i_0 = -1$ and $i_2 = i_{j+1} = s+1 = 3$, we obtain

$$\begin{aligned} T_{1,2}^{(2)}(x_1,x_2) &= (-1)^2 \sum_{i_1=1}^{1} \prod_{k=0}^{1} i_k (1 - \sum_{i=1}^{2-k} x_i)^{i_{k+1}-i_k-2} \\ &= (-1)(1-x_1-x_2)^0(1)(1-x_1)^0 = -1. \end{aligned}$$

Hence, $\lambda_{4;2;2} = -1$, so $(F_r)_{r=1}^{\infty}$ can not be associated with a coalescent process.

We next prove a result which interprets the measures F_r as distributions of limiting relative frequencies of blocks of random partitions. This result parallels Proposition 6, which provides a similar interpretation of Ξ. We could prove essentially the same result replacing the integers $1, \ldots, 2r$ with any distinct integers $i_1, \ldots, i_{2r}$.

Proposition 10 *Let $(F_r)_{r=1}^{\infty}$ be a sequence of measures satisfying conditions A1, A2, A3′, and A4 of Proposition 8, and assume $F_1(\Delta_1) > 0$. Let $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ be a coalescent satisfying conditions B1 and B2 of Theorem 2 such that the collision rates $\lambda_{b;k_1,\ldots,k_r;s}$ are defined from $(F_r)_{r=1}^{\infty}$ by (6). Let $T_r = \inf\{t : \#R_{2r}\Pi_\infty(t) < 2r\}$, which is the first time that two integers in $\{1, \ldots, 2r\}$ merge. Let $B_1, B_2, \ldots$ be the blocks of $\Pi_\infty(T_r-)$. Let Θ_r be a partition of $\mathbb{N}$ on $\{\#\Pi_\infty(T_r-) = \infty\}$ and a partition of $\{1, 2, \ldots, \#\Pi_\infty(T_r-)\}$ on $\{\#\Pi_\infty(T_r-) < \infty\}$ such that i and j are in the same block of Θ_r if and only if B_i and B_j are in the same block of $\Pi_\infty(T_r)$. Let E_r be the event that $R_{2r}\Pi_\infty(T_r)$ consists of the r blocks $\{1, 2\}, \{3, 4\}, \ldots, \{2r-1, 2r\}$. Then, the statements $P(E_r) = 0$, $\lambda_{2r;2,\ldots,2;0} = 0$, and $F_r(\Delta_r) = 0$ are equivalent. For all r such that $P(E_r) > 0$, there exists a probability measure Q_r on Δ_r satisfying the following three conditions:*

(a) $F_r = \lambda_{2r;2,\ldots,2;0} Q_r$.

(b) If $P(\#\Pi_\infty(T_r-) = \infty) > 0$, then conditional on the event $E_r \cap \{\#\Pi_\infty(T_r-) = \infty\}$, the restriction of Θ_r to $\{2r+1, 2r+2, \ldots\}$ is exchangeable, and the distribution of $(f_{1,r}, \ldots, f_{r,r})$, where $f_{j,r}$ is the limiting relative frequency of the block of Θ_r containing $2j-1$ and $2j$, equals Q_r.

(c) If $P(\#\Pi_\infty(T_r-) = n) > 0$, then there exists a random partition Θ'_r of $\mathbb{N}$ whose restriction to $\{2r+1, 2r+2, \ldots\}$ is exchangeable such that the distribution of Θ'_r restricted to $\{1, 2, \ldots, n\}$ is the same as the conditional distribution of Θ_r given $E_r \cap \{\#\Pi_\infty(T_r-) = n\}$. Moreover, Θ'_r can be chosen such that if $f'_{j,r}$ is the limiting relative frequency of the block of Θ'_r containing $2j-1$ and $2j$, then the distribution of $(f'_{1,r}, \ldots, f'_{r,r})$ equals Q_r.

Consider again the sequence $(F_r)_{r=1}^{\infty}$ of Example 9 in which F_1 is a unit mass at 1, F_2 is a unit mass at $(.5, .5)$, and $F_r = 0$ for $r \geq 3$. Suppose there were a coalescent process $(\Pi_\infty(t))_{t\geq 0}$ corresponding to $(F_r)_{r=1}^{\infty}$. Define T_r, Θ_r, and E_r as in Proposition 10. If $P(\#\Pi_\infty(T_1-) = \infty) > 0$, then parts (a) and (b) of Proposition 10 imply that conditional on $E_1 \cap \{\#\Pi_\infty(T_1-) = \infty\}$, the restriction of Θ_1 to $\{3, 4, \ldots\}$ is exchangeable, and the block of Θ_1 containing 1 and 2 has a limiting relative frequency of 1 a.s. It follows from Lemma 40 in appendix A that Θ_1 consists of a single block almost surely on $E_1 \cap \{\#\Pi_\infty(T_1-) = \infty\}$. If $P(\#\Pi_\infty(T_1-) = n) > 0$, then parts (a) and (c) of Proposition 10 imply that conditional on $E_1 \cap \{\#\Pi_\infty(T_1-) = n\}$, Θ_1 has the same distribution as the restriction to $\{1, 2, \ldots, n\}$ of a partition Θ'_1, where the restriction of Θ'_1 to $\{3, 4, \ldots\}$ is exchangeable and the block of Θ'_1 containing 1 and 2 has a limiting relative frequency of 1 a.s. Thus, Θ_1 consists of a single block almost surely on $E_1 \cap \{\#\Pi_\infty(T_1-) = n\}$ for all $n \in \mathbb{N}$. Therefore, $\#\Pi_\infty(T_1) = 1$ almost surely on E_1. Since $E_2 \subset E_1$, we have $\#\Pi_\infty(T_1) = 1$ almost surely on E_2. However, on E_2, $R_4\Pi_\infty(T_2)$ consists of the blocks $\{1,2\}$ and $\{3,4\}$. Therefore, $P(E_2)=0$. This contradicts Proposition 10 because $F_2(\Delta_2)>0$. Thus, the interpretation of the sequence $(F_r)_{r=1}^\infty$ given in Proposition 10 provides another way of seeing that condition A4 does not always hold.

We have established in Theorem 2 and Proposition 4 a one-to-one correspondence between finite measures Ξ on the infinite simplex Δ and coalescent processes satisfying B1 and B2. Proposition 8 shows that these coalescent processes are also in one-to-one correspondence with the sequences of measures $(F_r)_{r=1}^\infty$ satisfying A1, A2, A3$'$, and A4. These results, of course, yield a natural one-to-one correspondence between finite measures Ξ on Δ and sequences of measures $(F_r)_{r=1}^\infty$ satisfying A1, A2, A3$'$, and A4. The last result of this section shows how to calculate $(F_r)_{r=1}^\infty$ given the measure Ξ. We are unable to give a simple formula for calculating Ξ directly from $(F_r)_{r=1}^\infty$.

Proposition 11 *Let $\Xi=\Xi_0+a\delta_0$ be a finite measure on the infinite simplex Δ, where Ξ_0 has no atom at zero and δ_0 is a unit mass at zero. Let $(F_r)_{r=1}^\infty$ be the unique sequence of measures satisfying conditions A1, A2, and A3$'$ of Proposition 8 such that the collision rates of a standard Ξ-coalescent are given by (6). Let S be a measurable subset of Δ_r. Then,*

$$F_r(S)=\int_\Delta \sum_{i_1\neq\ldots\neq i_r} x_{i_1}^2\ldots x_{i_r}^2 1_{\{(x_{i_1},\ldots,x_{i_r})\in S\}} \Big/ \sum_{j=1}^\infty x_j^2\,\Xi_0(dx)+a1_{\{r=1,(0,0,\ldots)\in S\}}. \tag{13}$$

3 The Poisson process construction

In [16], Pitman gives a Poisson process construction of the Λ-coalescent when Λ has no atom at zero. Here we generalize this idea to obtain a Poisson process construction of the Ξ-coalescent started from any $\pi\in\mathcal{P}_\infty$ for all finite measures Ξ on the infinite simplex Δ. Our construction does permit Ξ to have an atom at zero. This construction is a useful tool for studying the Ξ-coalescent, in part because making computations using (11) can be tedious.

We will first define a σ-finite measure L on $\mathbb{Z}^\infty$, which is a Polish space when equipped with the product topology. We will then use a Poisson point process $(e(t))_{t\geq 0}$ with characteristic measure L, as defined in appendix B, to construct the Ξ-coalescent. For each $x=(x_1,x_2,\ldots)\in\Delta$, define a probability measure P_x on $\mathbb{Z}^\infty$ to be the distribution of a sequence $\xi=(\xi_i)_{i=1}^\infty$ of independent $\mathbb{Z}$-valued random variables such that for all i we have $P(\xi_i=j)=x_j$ for all $j\in\mathbb{N}$ and $P(\xi_i=-i)=1-\sum_{j=1}^\infty x_j$. Let z_{ij} be the sequence $(z_1,z_2,\ldots)$ in $\mathbb{Z}^\infty$ such that $z_i=z_j=1$ and $z_k=-k$ for $k\notin\{i,j\}$. Then define a measure L on $\mathbb{Z}^\infty$ by

$$L(A)=\int_\Delta\left(P_{\mathbf{x}}(A)\Big/\sum_{j=1}^\infty x_j^2\right)\Xi_0(d\mathbf{x})+a\sum_{i=1}^\infty\sum_{j=i+1}^\infty 1_{\{\mathbf{z}_{ij}\in A\}} \tag{14}$$

for all product measurable $A\subset\mathbb{Z}^\infty$. To show that L is σ-finite, and to establish some facts that will be useful later, we define

$$A_b=\{\xi\in\mathbb{Z}^\infty:\xi_1,\ldots,\xi_b \text{ are not all distinct}\} \tag{15}$$

for all $b \geq 2$ and

$$A_{k,l} = \{\xi \in \mathbb{Z}^\infty : \xi_k = \xi_l\} \tag{16}$$

for all $k \neq l$. Note that if $k \neq l$ then $P_x(\{\xi : \xi_k = \xi_l = j\}) = x_j^2$ for all $j \in \mathbb{N}$, which means $P_x(A_{k,l}) = \sum_{j=1}^\infty x_j^2$. Therefore,

$$L(A_{k,l}) = \int_\Delta 1 \, \Xi_0(dx) + a \sum_{i=1}^\infty \sum_{j=i+1}^\infty 1_{\{z_{ij} \in A_{k,l}\}} = \Xi_0(\Delta) + a = \Xi(\Delta). \tag{17}$$

Also, $A_b = \bigcup_{k=1}^b \bigcup_{l=k+1}^b A_{k,l}$, so

$$L(A_b) \leq \sum_{k=1}^b \sum_{l=k+1}^b L(A_{k,l}) = \binom{b}{2} \Xi(\Delta) < \infty. \tag{18}$$

Define

$$A_\infty = \{\xi \in \mathbb{Z}^\infty : \xi_i = \xi_j \text{ for some } i \neq j\}. \tag{19}$$

Note that $L(A_\infty^c) = 0$. Thus, the union of the sets in the countable collection consisting of A_∞^c and A_b for all $b \geq 2$ equals $\mathbb{Z}^\infty$, and L assigns finite measure to each set in this collection. Hence, L is σ-finite, so we can define a Poisson point process $(e(t))_{t \geq 0}$ with characteristic measure L.

We now use this Poisson point process to construct a Ξ-coalescent starting from an arbitrary $\pi \in \mathcal{P}_\infty$. First, we define for each n a $\mathcal{P}_n$-valued coalescent $\Pi_n^\pi = (\Pi_n^\pi(t))_{t \geq 0}$ as follows. Let $T_{0,n} = 0$ and for $k \geq 1$, define $T_{k,n} = \inf\{t > T_{k-1,n} : e(t) \in A_n\}$. Since $L(A_n) < \infty$ by (18), it follows from part (b) of Lemma 41 in appendix B that $\lim_{k \to \infty} T_{k,n} = \infty$ a.s. Therefore, by the argument used to prove part (c) of Lemma 41 in appendix B, we have $e(T_{k,n}) \in A_n$ for all k almost surely. We will define Π_n^π to have right-continuous step function paths with jumps only possible at the times $T_{k,n}$ for $k \geq 1$. Therefore, it suffices to specify $\Pi_n^\pi(T_{k,n})$ for all $k \geq 0$. Define $\Pi_n^\pi(0) = R_n \pi$. For $k \geq 1$, if $\Pi_n^\pi(T_{k-1,n})$ consists of the blocks $B_1, \ldots, B_b$, where the blocks are ordered by their smallest elements, then $\Pi_n^\pi(T_{k,n})$ is defined to be the partition of $\{1, \ldots, n\}$, each of whose blocks is a union of some of the blocks $B_1, \ldots, B_b$, such that B_i and B_j are in the same block of $\Pi_n^\pi(T_{k,n})$ if and only if $e(T_{k,n})_i = e(T_{k,n})_j$, where $e(T_{k,n})_i$ and $e(T_{k,n})_j$ denote the ith and jth coordinates respectively of $e(T_{k,n})$.

Suppose $m < n$. We claim that $(R_m \Pi_n^\pi(t))_{t \geq 0} = (\Pi_m^\pi(t))_{t \geq 0}$. If $\xi \in A_m$ then $\xi \in A_n$, so the processes $(R_m \Pi_n^\pi(t))_{t \geq 0}$ and $(\Pi_m^\pi(t))_{t \geq 0}$ can only jump at times $T_{k,n}$ for some $k \geq 1$. Thus, to prove the claim, it suffices to show that $R_m \Pi_n^\pi(T_{k,n}) = \Pi_m^\pi(T_{k,n})$ for all $k \geq 0$. We use induction on k. Note first that $R_m \Pi_n^\pi(T_{0,n}) = R_m \Pi_n^\pi(0) = R_m \pi = \Pi_m^\pi(0) = \Pi_m^\pi(T_{0,n})$. Suppose $k \geq 1$ and $R_m \Pi_n^\pi(T_{k-1,n}) = \Pi_m^\pi(T_{k-1,n})$. Let $B_1, \ldots, B_b$ be the blocks of $\Pi_m^\pi(T_{k-1,n})$ and let $B_1', \ldots, B_d'$ be the blocks of $\Pi_n^\pi(T_{k-1,n})$, where blocks are ordered by their smallest elements. Fix $i, j \in \{1, \ldots, m\}$ and define $h(i)$ and $h(j)$ such that $i \in B_{h(i)}$ and $j \in B_{h(j)}$. Since $B_i = \{1, \ldots, m\} \cap B_i'$ for $i = 1, \ldots, b$, we have $i \in B_{h(i)}'$ and $j \in B_{h(j)}'$. Therefore i and j are in the same block of $\Pi_m^\pi(T_{k,n})$ if and only if $e(T_{k,n})_{h(i)} = e(T_{k,n})_{h(j)}$. Likewise, i and j are in the same block of $\Pi_n^\pi(T_{k,n})$, and thus the same block of $R_m \Pi_n^\pi(T_{k,n})$, if and only if $e(T_{k,n})_{h(i)} = e(T_{k,n})_{h(j)}$. We conclude that $\Pi_m^\pi(T_{k,n}) = R_m \Pi_n^\pi(T_{k,n})$, so by induction, $(R_m \Pi_n^\pi(t))_{t \geq 0} = (\Pi_m^\pi(t))_{t \geq 0}$.

Now define a $\mathcal{P}_\infty$-valued process $\Pi_\infty^\pi = (\Pi_\infty^\pi(t))_{t \geq 0}$ such that i and j are in the same block of $\Pi_\infty^\pi(t)$ if and only if i and j are in the same block of $\Pi_n^\pi(t)$ for $n \geq \max\{i, j\}$. Then, $\Pi_n^\pi = R_n \Pi_\infty^\pi$. The proposition below establishes that Π_∞^π is a Ξ-coalescent started from π.

Proposition 12 *Let $\Xi = \Xi_0 + a\delta_0$ be a finite measure on the infinite simplex Δ, where Ξ_0 has no atom at zero and δ_0 is a unit mass at zero. Let $(e(t))_{t\geq 0}$ be a Poisson point process with characteristic measure L, where L is defined by (14). For all $\pi \in \mathcal{P}_\infty$, define a process $\Pi_\infty^\pi = (\Pi_\infty^\pi(t))_{t\geq 0}$ from $(e(t))_{t\geq 0}$ as described above. Then Π_∞^π is a Ξ-coalescent satisfying $\Pi_\infty^\pi(0) = \pi$.*

Proof. That $\Pi_\infty^\pi(0) = \pi$ is clear from the definition of Π_∞^π. Thus, it suffices to show that Π_∞^π satisfies condition B2 of Theorem 2 for the rates defined from Ξ as in (11). Fix $n \geq 2$, and define $\Pi_n^\pi = R_n \Pi_\infty^\pi$. To see that Π_n^π is Markov, choose $\alpha_1, \ldots, \alpha_m \in \mathcal{P}_n$ and choose times $t_1, \ldots, t_{m+1}$ such that $0 < t_1 < \ldots < t_{m+1}$. It follows from the definition of Poisson point processes that the process $(e(t_m + t))_{t>0}$ has the same law as $(e(t))_{t>0}$ and is independent of $(e(t))_{0\leq t\leq t_m}$. Therefore, we see from the construction of Π_n^π that

$$P(\Pi_n^\pi(t_{m+1}) = \alpha_{m+1} | \Pi_n^\pi(t_1) = \alpha_1, \ldots, \Pi_n^\pi(t_m) = \alpha_m) = P(\Pi_n^{\alpha_m}(t_{m+1} - t_m) = \alpha_{m+1}),$$

from which it follows that Π_n^π is Markov.

It remains to show that the collision rates of Π_n^π agree with the rates given in Theorem 2 for the restriction to $\{1, \ldots, n\}$ of the Ξ-coalescent. Let $B_1, \ldots, B_b$ be the blocks of $\Pi_n^\pi(t)$. Let θ be a partition of $\{1, \ldots, b\}$ into s singletons and larger blocks $B_1', \ldots, B_r'$ of sizes $k_1, \ldots, k_r$. Let A_θ consist of all sequences $\xi \in \mathbb{Z}^\infty$ such that if $1 \leq i, j \leq b$, then $\xi_i = \xi_j$ if and only if i and j are in the same block of θ. Note that θ is associated with a $(b; k_1, \ldots, k_r; s)$-collision in which $B_1, \ldots, B_b$ merge in such a way that B_i and B_j end up in the same block if and only if i and j are in the same block of θ. By the above construction and part (a) of Lemma 41 in appendix B, this collision is occurring at rate $L(A_\theta)$. Thus, we must show that $L(A_\theta)$ equals the expression for $\lambda_{b;k_1,\ldots,k_r;s}$ given in (11).

A point $\xi \in \mathbb{Z}^\infty$ is in A_θ if and only if there exist $l \in \{0, 1, \ldots s\}$ and distinct positive integers $i_1, \ldots, i_{r+l}$ such that the following hold:

(a) $\xi_m = i_j$ for all $j \in \{1, \ldots, r\}$ and $m \in B_j'$.

(b) There exist $m_1 < \ldots < m_l \leq b$ such that $\xi_{m_j} = i_{r+j}$ for $j \in \{1, \ldots, l\}$.

(c) $\xi_m < 0$ for the $s - l$ values of m such that $m \leq b$, $m \notin B_j'$ for all $j \in \{1, \ldots, r\}$, and $m \notin \{m_1, \ldots, m_l\}$.

By summing over the possible values for l and the possible distinct integers $i_1, \ldots, i_{r+l}$, and counting the possible values of $m_1, \ldots, m_l$, we get

$$P_{\mathbf{x}}(A_\theta) = \sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r} x_{i_{r+1}} \ldots x_{i_{r+l}} (1 - \sum_{j=1}^{\infty} x_j)^{s-l} \tag{20}$$

for all $x \neq (0, 0, \ldots)$. Note also that

$$\sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} 1_{\{z_{ij} \in A_\theta\}} = 1_{\{r=1, k_1=2\}}. \tag{21}$$

It follows from (14), (20), and (21) that $L(A_\theta)$ equals

$$\int_\Delta \left(\sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r} x_{i_{r+1}} \ldots x_{i_{r+l}} (1 - \sum_{j=1}^{\infty} x_j)^{s-l} \right) \Big/ \sum_{j=1}^{\infty} x_j^2 \, \Xi_0(d\mathbf{x}) + a 1_{\{r=1, k_1=2\}},$$

which is the expression for $\lambda_{b;k_1,\dots,k_r;s}$ given in (11). Hence, $(\Pi_\infty^\pi(t))_{t\geq 0}$ is a Ξ-coalescent satisfying $\Pi_\infty^\pi(0) = \pi$. $\square$

Definition 13 *Let Ξ be a finite measure on the infinite simplex Δ. Suppose $(e(t))_{t\geq 0}$ is a Poisson point process with characteristic measure L, where L is defined in terms of Ξ by (14). If Π_∞ is defined from $(e(t))_{t\geq 0}$ as described above, then we say Π_∞ is a* Ξ-coalescent derived from $(e(t))_{t\geq 0}$.

Remark 14 Note that by taking $\Xi = \delta_0$, we can obtain a Poisson process construction of Kingman's coalescent. In this case, (14) reduces to

$$L(A) = \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} 1_{\{z_{ij}\in A\}}. \tag{22}$$

Equivalently, we have $L(\{z_{ij}\}) = 1$ for all $i, j \in \mathbb{N}$ with $i \neq j$ and $L(\{\xi\}) = 0$ if we do not have $\xi = z_{ij}$ for some $i \neq j$. The coalescent derived from a Poisson point process $(e(t))_{t\geq 0}$ with the characteristic measure L defined in (22) has the property that if the blocks are ordered by their smallest elements, then the ith and jth blocks merge at the times t for which $e(t) = z_{ij}$.

Lemma 15 *Let Ξ be a finite measure on the infinite simplex Δ. Let $(\Pi_\infty(t))_{t\geq 0}$ be a Ξ-coalescent derived from a Poisson point process $(e(t))_{t\geq 0}$ with characteristic measure L, where L is defined by (14). Let A be a subset of $\mathbb{Z}^\infty$ such that $0 < L(A) < \infty$, and let $T_A = \inf\{t : e(t) \in A\}$. Then, $\Pi_\infty(T_A-)$ and $e(T_A)$ are independent.*

Proof. Let $(e'(t))_{t\geq 0}$ be defined such that $e'(t) = \delta$ if $e(t) \in A$ and $e'(t) = e(t)$ otherwise. By parts (d) and (e) of Lemma 41 in appendix B, we have that $e(T_A)$, T_A, and $(e'(t))_{t\geq 0}$ are mutually independent. Since $\Pi_\infty(T_A-)$ is a function of $(e'(t))_{t\geq 0}$ and T_A, it follows that $\Pi_\infty(T_A-)$ is independent of $e(T_A)$. $\square$

4 Proofs of results characterizing the coalescents

4.1 Preliminary Lemmas

In this subsection, we give some preliminary lemmas that will be useful for some of the proofs of results in section 2. We begin with the following result, which can be proved by a straightforward application of the Daniell-Kolmogorov Theorem.

Lemma 16 *Suppose, for each n, that Θ_n is a random partition of $\{1, \dots, n\}$. Suppose $R_m\Theta_n$ and Θ_m have the same distribution for all $m < n$. Then, there exists on some probability space a random partition Θ_∞ of $\mathbb{N}$ such that $R_n\Theta_\infty$ has the same distribution as Θ_n for all n.*

The lemma below will enable us to construct $\mathcal{P}_\infty$-valued coalescents from consistently-defined $\mathcal{P}_n$-valued coalescents. It is proved by an application of the Daniell-Kolmogorov Theorem as on p.40 of [11].

Lemma 17 *Suppose, for each n, that $(\Pi_n(t))_{t\geq 0}$ is a $\mathcal{P}_n$-valued coalescent. Suppose, for all $m < n$, that the processes $(R_m\Pi_n(t))_{t\geq 0}$ and $(\Pi_m(t))_{t\geq 0}$ have the same law. Then, there exists on some probability space a $\mathcal{P}_\infty$-valued coalescent $(\Pi_\infty(t))_{t\geq 0}$ such that $(R_n\Pi_\infty(t))_{t\geq 0}$ has the same law as $(\Pi_n(t))_{t\geq 0}$ for all n.*

The next lemma gives the consistency condition that an array of nonnegative real numbers $\{\lambda_{b;k_1,\dots,k_r;s} : r \geq 1,\ k_1,\dots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$ must satisfy to be the array of collision rates for a coalescent with simultaneous multiple collisions. This condition is the analog of condition (1) for the Λ-coalescent.

Lemma 18 *Let $\{\lambda_{b;k_1,\dots,k_r;s} : r \geq 1,\ k_1,\dots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$ be a collection of nonnegative real numbers. Then there exists a $\mathcal{P}_\infty$-valued coalescent $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ satisfying conditions B1 and B2 of Theorem 2 if and only if*

$$\lambda_{b;k_1,\dots,k_r;s} = \sum_{m=1}^{r} \lambda_{b+1;k_1,\dots,k_{m-1},k_m+1,k_{m+1},\dots,k_r;s} + s\lambda_{b+1;k_1,\dots,k_r,2;s-1} + \lambda_{b+1;k_1,\dots,k_r;s+1} \qquad (23)$$

for all $r \geq 1$, $k_1,\dots,k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$.

When $s = 0$, we say that $s\lambda_{b+1;k_1,\dots,k_r,2;s-1} = 0$ even though $\lambda_{b+1;k_1,\dots,k_r,2;s-1}$ is undefined, so that the right-hand side of (23) makes sense.

Proof. Continuous-time Markov chains on a finite state space can be constructed with arbitrary nonnegative transition rates. Thus, we can define for each n a Markov chain $\Pi_n = (\Pi_n(t))_{t\geq 0}$ with state space $\mathcal{P}_n$ and right-continuous paths such that $\Pi_n(0)$ is the partition of $\{1,\dots,n\}$ into singletons and, when $\Pi_n(t)$ has b blocks, each $(b;k_1,\dots,k_r;s)$-collision is occurring at rate $\lambda_{b;k_1,\dots,k_r;s}$.

Define for each n a process $\Theta_n = (\Theta_n(t))_{t\geq 0}$ by $\Theta_n(t) = R_n\Pi_{n+1}(t)$. Suppose Θ_n and Π_n have the same law. Then $R_m\Pi_n$ and Π_m have the same law for all $m < n$. By Lemma 17, there exists a coalescent process $(\Pi_\infty(t))_{t\geq 0}$ such that $(R_n\Pi_\infty(t))_{t\geq 0}$ has the same law as $(\Pi_n(t))_{t\geq 0}$ for all n. The process $(\Pi_\infty(t))_{t\geq 0}$ satisfies conditions B1 and B2 of Theorem 2. Conversely, suppose $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ satisfies B1 and B2 of Theorem 2. Then, Π_n has the same law as $R_n\Pi_\infty$ and Θ_n has the same law as $R_n(R_{n+1}\Pi_\infty) = R_n\Pi_\infty$, so Θ_n and Π_n have the same law. Thus, we must show that Θ_n and Π_n have the same law for all n if and only if (23) holds.

Let $U = \inf\{t : \{n+1\}$ is not a block of $\Pi_{n+1}(t)\}$. On the event $\{t > U\}$, we have that $\#\Pi_{n+1}(t) = \#\Theta_n(t)$ and Θ_n undergoes a $(b;k_1,\dots,k_r;s)$-collision at time t if and only if Π_{n+1} undergoes a $(b;k_1,\dots,k_r;s)$-collision at time t. Therefore, after time U, if $\Theta_n(t)$ has b blocks, then each $(b;k_1,\dots,k_r;s)$-collision is occurring at rate $\lambda_{b;k_1,\dots,k_r;s}$.

Next, suppose Θ_n undergoes a $(b;k_1,\dots,k_r;s)$-collision at time $t \leq U$. Then, Π_{n+1} could undergo any of $r+s+1$ possible collisions at time t, as can be seen by considering the following three cases:

Case 1: The block $\{n+1\}$ could remain a singleton at time t, in which case Π_{n+1} undergoes a $(b+1;k_1,\dots,k_r;s+1)$-collision at time t.

Case 2: The block $\{n+1\}$ could join one of the s blocks of $\Theta_n(t)$ that consists of a single block of $\Theta_n(t-)$, in which case Π_{n+1} undergoes a $(b;k_1,\ldots,k_r,2;s-1)$-collision at time t.

Case 3: The block $\{n+1\}$ could join one of the r blocks of $\Theta_n(t)$ consisting of two or more blocks of $\Theta_n(t-)$. Then, Π_{n+1} undergoes a $(b+1;k_1,\ldots,k_{m-1},k_m+1,k_{m+1},\ldots,k_r;s)$-collision for some $m\in\{1,\ldots,r\}$.

Thus, before time U, the rate of any $(b;k_1,\ldots,k_r;s)$-collision for the process Θ_n is the same as the sum of the rates at which Π_{n+1} is undergoing one of the $r+s+1$ collisions described above. The definition of Π_{n+1} implies that this rate equals the right-hand side of (23).

It follows from the results proved in the last two paragraphs that if (23) holds for all $r\geq 1$, $k_1,\ldots,k_r\geq 2$, $s\geq 0$, and $b=\sum_{j=1}^r k_j+s$, then Θ_n and Π_n have the same law for all n. Conversely, suppose Θ_n and Π_n have the same law. Then the initial rate at which Π_n is undergoing an $(n;k_1,\ldots,k_r;s)$-collision is $\lambda_{n;k_1,\ldots,k_r;s}$ by definition, and arguments in the previous paragraph imply that the initial rate at which Θ_n is undergoing an $(n;k_1,\ldots,k_r;s)$-collision is given by the right-hand side of (23) with $b=n$. Hence, (23) holds when $b=n$. Thus, if Θ_n and Π_n have the same law for all n, then (23) holds for all $r\geq 1$, $k_1,\ldots,k_r\geq 2$, $s\geq 0$, and $b=\sum_{j=1}^r k_j+s$. □

4.2 Proof of Theorem 2

In this subsection, we prove Theorem 2. Recall that in section 3, we constructed from a Poisson point process a standard Ξ-coalescent for an arbitrary finite measure Ξ on the infinite simplex Δ. This construction proves the "if" part of Theorem 2. However, Lemma 18 provides a way of proving the "if" part of Theorem 2 more directly by checking the consistency of the transition rates defined by (11). We provide this alternative proof below.

Proof of "if" part of Theorem 2. Suppose $\Xi=\Xi_0+a\delta_0$ is a finite measure on the infinite simplex Δ, where Ξ_0 has no atom at zero and δ_0 is a unit mass at zero. Define $\lambda_{b;k_1,\ldots,k_r;s}$ by (11) for all $r\geq 1$, $k_1,\ldots,k_r\geq 2$, $s\geq 0$, and $b=\sum_{j=1}^r k_j+s$. By Lemma 18, to prove the existence of a $\mathcal{P}_\infty$-valued coalescent process satisfying conditions B1 and B2 of Theorem 2, it suffices to verify (23) for all $r\geq 1$, $k_1,\ldots,k_r\geq 2$, $s\geq 0$, and $b=\sum_{j=1}^r k_j+s$.

It suffices to verify (23) separately when $\Xi_0(\Delta)=0$ and when $a=0$. When $\Xi_0(\Delta)=0$, the right-hand side of (23) is $0+0+\lambda_{b+1;k_1,\ldots,k_r;s+1}$, which equals $\lambda_{b;k_1,\ldots,k_r;s}$. When $a=0$, the right-hand side of (23) can be written as

$$\int_\Delta W(\mathbf{x})/\sum_{j=1}^\infty x_j^2\,\Xi(d\mathbf{x}),$$

where

$$\begin{aligned} W(\mathbf{x}) &= \sum_{m=1}^r\sum_{l=0}^s\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s}{l}(x_{i_1}^{k_1}\ldots x_{i_{m-1}}^{k_{m-1}}x_{i_m}^{k_m+1}x_{i_{m+1}}^{k_{m+1}}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^\infty x_j)^{s-l} \\ &+ s\sum_{l=0}^{s-1}\sum_{i_1\neq\ldots\neq i_{r+1+l}}\binom{s-1}{l}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}^2x_{i_{r+2}}\ldots x_{i_{r+1+l}})(1-\sum_{j=1}^\infty x_j)^{s-1-l} \end{aligned}$$

$$+\sum_{l=0}^{s+1}\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s+1}{l}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s+1-l}. \tag{24}$$

Let A, B, and C be the three terms on the right-hand side of (24). We have

$$A=\sum_{l=0}^{s}\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s}{l}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s-l}(\sum_{m=1}^{r}x_{i_m}).$$

Also, we have

$$\begin{aligned}
B &= \sum_{l=1}^{s}\sum_{i_1\neq\ldots\neq i_{r+l}} s\binom{s-1}{l-1}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}^2x_{i_{r+2}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s-l}\\
&= \sum_{l=1}^{s}\sum_{i_1\neq\ldots\neq i_{r+l}} s\binom{s-1}{l-1}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s-l}\frac{(x_{i_{r+1}}+\ldots+x_{i_{r+l}})}{l}\\
&= \sum_{l=1}^{s}\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s}{l}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s-l}(x_{i_{r+1}}+\ldots+x_{i_{r+l}})
\end{aligned}$$

and

$$\begin{aligned}
C &= \sum_{l=0}^{s}\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s}{l}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s+1-l}\\
&+ \sum_{l=1}^{s+1}\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s}{l-1}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s+1-l}\\
&= \sum_{l=0}^{s}\sum_{i_1\neq\ldots\neq i_{r+l}}\binom{s}{l}(x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}x_{i_{r+1}}\ldots x_{i_{r+l}})(1-\sum_{j=1}^{\infty}x_j)^{s-l}(1-\sum_{j=1}^{\infty}x_j+\sum_{i_{r+l+1}\notin S}x_{i_{r+l+1}}),
\end{aligned}$$

where $S=\{i_1,\ldots,i_{r+l}\}$. By adding the above expressions for A, B, and C, we see that $W(x)$ equals the numerator of the integrand on the right-hand side of (11), which implies (23). This completes the proof of the theorem. □

The proof of the "only if" part of Theorem 2 relies heavily on exchangeability arguments. Some well-known results that we will apply are reviewed in appendix A. It will be convenient to make the following additional definition.

Definition 19 *We call a $\mathcal{P}_\infty$-valued process $(\Pi_\infty(t))_{t\geq 0}$* exchangeable *if $(\Pi_\infty(t))_{t\geq 0}$ has the same distribution as $(\hat{\sigma}\Pi_\infty(t))_{t\geq 0}$ for all finite permutations σ of $\mathbb{N}$, where $\sigma(i)$ and $\sigma(j)$ are in the same block of $\hat{\sigma}\Pi_\infty(t)$ if and only if i and j are in the same block of $\Pi_\infty(t)$.*

Note that any coalescent process satisfying conditions B1 and B2 of Theorem 2 for some collection of rates $\{\lambda_{b;k_1,\ldots,k_r;s} : r\geq 1,\ k_1,\ldots,k_r\geq 2,\ s\geq 0,\ b=\sum_{j=1}^{r}k_j+s\}$ is exchangeable.

Lemma 20 *Let $(a_i)_{i=1}^{\infty}$ be a sequence of nonnegative real numbers such that $\sum_{i=1}^{\infty} a_i \leq 1$. Let $(b_i)_{i=1}^{\infty}$ be a bounded sequence of real numbers. Suppose $b_i = b_j$ whenever $a_i = a_j$, and suppose $\sum_{i=1}^{\infty} a_i^k b_i = 0$ for all $k \in \mathbb{N}$. Then $b_i = 0$ for all i such that $a_i > 0$.*

Proof. We may assume, without loss of generality, that $a_1 \geq a_2 \geq \ldots$. Suppose there exists an i such that $b_i \neq 0$ and $a_i > 0$. Let $m = \min\{i : b_i \neq 0\}$, and let $s = \max\{r \geq 0 : a_{m+r} = a_m\}$. We have

$$0 = \sum_{i=1}^{\infty} a_i^k b_i = a_m^k \sum_{i=m}^{m+s} b_i + \sum_{i=m+s+1}^{\infty} a_i^k b_i \tag{25}$$

for all $k \in \mathbb{N}$. Note that

$$\left| a_m^k \sum_{i=m}^{m+s} b_i \right| \geq a_m^k |b_m|$$

because $b_m = b_{m+1} = \ldots = b_{m+s}$ by assumption. Choose $B < \infty$ such that $|b_i| \leq B$ for all i. Then

$$\left| \sum_{i=m+s+1}^{\infty} a_i^k b_i \right| \leq B \sum_{i=m+s+1}^{\infty} a_i^k \leq B a_{m+s+1}^{k-1} \sum_{i=m+s+1}^{\infty} a_i \leq B a_{m+s+1}^{k-1}.$$

We have $a_m^k |b_m| > B a_{m+s+1}^{k-1}$ for sufficiently large k because $a_{m+s+1} < a_m$ and $|b_m| > 0$, which contradicts the fact that the expression on the right-hand side of (25) is zero for all $k \in \mathbb{N}$. $\square$

Proof of the "only if" part of Theorem 2. Suppose we have a collection of nonnegative real numbers $\{\lambda_{b;k_1,\ldots,k_r;s} : r \geq 1,\ k_1, \ldots, k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^{r} k_j + s\}$ such that there exists a coalescent process $\Pi_\infty = (\Pi_\infty(t))_{t \geq 0}$ satisfying conditions B1 and B2 of Theorem 2 with collision rates $\lambda_{b;k_1,\ldots,k_r;s}$. We wish to show that there is a finite measure Ξ on the infinite simplex Δ such that all of the collision rates are given in terms of Ξ by (11).

Let $T = \inf\{t : 1 \text{ and } 2 \text{ are in the same block of } \Pi_\infty(t)\}$. We may assume that $T < \infty$ a.s. because if the rate at which the blocks containing 1 and 2 are merging is zero, then all collision rates are zero and (11) holds with $\Xi = 0$. For $n \geq 2$, let E_n be the event that $1, 2, \ldots, n$ are in distinct blocks of $\Pi_\infty(T-)$. Since Π_∞ is an exchangeable process, the probability that no pair of integers in the set $\{1, 2, \ldots, n\}$ merges before 1 and 2 merge is at least $2/n(n-1)$. Therefore, $P(E_n) > 0$. Let Γ_n be a random partition of $\{1, \ldots, n\}$ whose distribution is the same as the conditional distribution of $\Pi_n(T)$ given E_n. We claim that there exists a random partition Θ_∞ of $\mathbb{N}$ such that $\Theta_n = R_n \Theta_\infty$ has the same distribution as Γ_n for all n. By Lemma 16, it suffices to show that $R_m \Gamma_n$ has the same distribution as Γ_m for all $m < n$.

Fix $m < n$. Let θ be a partition of $\{1, 2, \ldots, m\}$ in which 1 and 2 are in the same block. Let s be the number of singletons in θ, and let $k_1, \ldots, k_r$ be the sizes of the larger blocks of θ. It follows from condition B2 that if $\Pi_m(t)$ consists of m singletons, then the merger of the m singletons into the blocks of the partition θ is occurring at rate $\lambda_{m;k_1,\ldots,k_r;s}$. Also, the total rate of all mergers involving the blocks $\{1\}$ and $\{2\}$ is $\lambda_{2;2;0}$. Thus, $P(\Gamma_m = \theta) = \lambda_{m;k_1,\ldots,k_r;s}/\lambda_{2;2;0}$. If $\Pi_n(t)$ consists of n singletons, then condition B2 implies that the total rate of all mergers of $\{1, 2, \ldots, n\}$ whose restriction to $\{1, 2, \ldots, m\}$ is the merger of m singletons into the blocks of θ is also $\lambda_{m;k_1,\ldots,k_r;s}$. Therefore, we have $P(R_m \Gamma_n = \theta) = \lambda_{m;k_1,\ldots,k_r;s}/\lambda_{2;2;0}$. Thus, $R_m \Gamma_n$ has the same distribution as Γ_m.

Let Θ'_∞ be the restriction of Θ_∞ to $\{3, 4, \ldots\}$. Since Π_∞ is an exchangeable process, Θ'_∞ is an exchangeable random partition of $\{3, 4, \ldots\}$. It follows from Lemma 40 in appendix A that each block of Θ'_∞ has a limiting relative frequency. Let $P_1 \geq P_2 \geq \ldots$ be the ranked sequence of limiting relative frequencies of the distinct blocks of Θ'_∞, where $P_n = 0$ if Θ'_∞ has fewer than n blocks with nonzero limiting relative frequencies. Note that the blocks of Θ_∞ also have limiting relative frequencies, and $(P_j)_{j=1}^\infty$ is the ranked sequence of limiting relative frequencies of distinct blocks of Θ_∞.

We now label the blocks of Θ_∞ having nonzero limiting relative frequencies by $B_1, B_2, \ldots$, where B_j has limiting relative frequency P_j on $\{P_j > 0\}$ and blocks with the same limiting relative frequency are ordered at random, independently of Θ_∞. On $\{P_j = 0\}$, the block B_j is undefined. Define a sequence of random variables $(Z_m)_{m=1}^\infty$ such that $Z_m = i$ on $\{m \in B_i\}$ and $Z_m = 0$ when the block of Θ_∞ containing m has a limiting relative frequency of zero. Let $\mathcal{F}$ denote the σ-field generated by $(P_j)_{j=1}^\infty$. Let $P_0 = 1 - \sum_{j=1}^\infty P_j$. We make the following claim regarding the distribution of the sequence $(Z_m)_{m=1}^\infty$:

Claim. We have $P(Z_m = i|\mathcal{F}) = P_i$ a.s. for all $m \geq 3$ and $i \geq 0$. On $\{P_1 > 0\}$, we have $P(Z_1 = i|\mathcal{F}) = P_i^2 / \sum_{j=1}^\infty P_j^2$ a.s. for all $i \geq 1$. Moreover, the random variables $Z_1, Z_3, Z_4, \ldots$ are conditionally independent given $\mathcal{F}$.

Before proving the claim, we show how we can use the claim to complete the proof of the "only if" part of Theorem 2.

Let θ be a partition of $\{1, 2, \ldots, b\}$ into s singletons and larger blocks $D_1, \ldots, D_r$ of sizes $k_1, \ldots, k_r$ respectively. Assume that 1 and 2 are in the block D_1. Then, as we showed in the third paragraph of this proof,

$$\lambda_{b;k_1,\ldots,k_r;s} = \lambda_{2;2;0} P(\Theta_b = \theta). \tag{26}$$

By Lemma 40, almost surely every block of the exchangeable random partition Θ'_∞ having a limiting relative frequency of zero is a singleton. Therefore, if $i, j \geq 3$, then almost surely i and j are in the same block of Θ_∞ if and only if $Z_i = Z_j \neq 0$. On $\{P_1 > 0\}$, the claim implies that $Z_1 = Z_2 > 0$ a.s., so for all positive integers i and j, we have that almost surely i and j are in the same block of Θ_∞ if and only if $Z_i = Z_j \neq 0$. Therefore, on $\{P_1 > 0\}$, the event that $\Theta_b = \theta$ is the same, up to a null set, as the event that there exist $l \in \{0, \ldots, s\}$ and distinct positive integers $i_1, \ldots, i_{r+l}$ such that the following hold:

(a) For all $j \in \{1, \ldots, r\}$ and $m \in D_j$, we have $Z_m = i_j$.

(b) There exist $m_1 < \ldots < m_l \leq b$ such that $Z_{m_j} = i_{r+j}$ for $j \in \{1, \ldots, l\}$.

(c) $Z_m = 0$ for each of the $s - l$ values of m such that $m \leq b$, $m \notin D_j$ for $j \in \{1, \ldots, r\}$, and $m \notin \{m_1, \ldots, m_l\}$.

Note that l is the number of singletons in Θ_b that are in blocks of Θ_∞ having nonzero limiting relative frequencies. For now, fix $l \in \{0, \ldots, s\}$ and fix distinct positive integers $i_1, \ldots, i_{r+l}$. The claim implies that for $j \in \{2, \ldots, r\}$, we have

$$P(Z_m = i_j \text{ for all } m \in D_j|\mathcal{F}) = P_{i_j}^{k_j}, \text{ a.s. on } \{P_1 > 0\}. \tag{27}$$

Since 1 and 2 are in D_1, it follows from the claim that

$$P(Z_m = i_1 \text{ for all } m \in D_1|\mathcal{F}) = P_{i_1}^{k_1-2} P(Z_1 = i_1|\mathcal{F}) = P_{i_1}^{k_1} / \sum_{j=1}^{\infty} P_j^2, \text{ a.s. on } \{P_1 > 0\}. \quad (28)$$

Also, we have

$$P(Z_{m_j} = i_{r+j} \text{ for } j \in \{1, \ldots, l\}|\mathcal{F}) = P_{i_{r+1}} \ldots P_{i_{r+l}}, \text{ a.s. on } \{P_1 > 0\} \quad (29)$$

and

$$P(Z_m = 0|\mathcal{F}) = P_0, \text{ a.s. on } \{P_1 > 0\} \quad (30)$$

for all $m \geq 3$. By summing over the possible values of l and $i_1, \ldots, i_{r+l}$ and counting the possible values of $m_1, \ldots, m_l$, we can conclude from equations (27) through (30), conditions (a), (b), and (c) above, and the conditional independence of $Z_1, Z_3, Z_4, \ldots$ given $\mathcal{F}$ that

$$P(\Theta_b = \theta|\mathcal{F}) = \left(\sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} P_{i_1}^{k_1} P_{i_2}^{k_2} \ldots P_{i_r}^{k_r} P_{i_{r+1}} \ldots P_{i_{r+l}} P_0^{s-l} \right) \Big/ \sum_{j=1}^{\infty} P_j^2, \text{ a.s. on } \{P_1 > 0\}.$$

On $\{P_1 = 0\}$, the claim implies that $Z_m = 0$ almost surely for all $m \geq 3$, so Θ'_∞ almost surely consists of all singletons. The exchangeability of Π_∞ implies that the probability, conditional on the event that Θ'_∞ consists of all singletons, that 1 and 2 are in the same block as q is the same for all $q \geq 3$, and therefore must be zero. Therefore, on $\{P_1 = 0\}$, almost surely Θ_b consists of the blocks $\{1, 2\}, \{3\}, \{4\}, \ldots, \{b\}$. It follows that $P(\Theta_b = \theta|\mathcal{F}) = 1$ a.s. on $\{P_1 = 0\}$ if $r = 1$ and $k_1 = 2$, and $P(\Theta_b = \theta|\mathcal{F}) = 0$ a.s. on $\{P_1 = 0\}$ otherwise. Therefore, $P(\Theta_b = \theta|\mathcal{F})$ equals

$$\left(\sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} P_{i_1}^{k_1} P_{i_2}^{k_2} \ldots P_{i_r}^{k_r} P_{i_{r+1}} \ldots P_{i_{r+l}} P_0^{s-l} \Big/ \sum_{j=1}^{\infty} P_j^2 \right) 1_{\{P_1 > 0\}} + 1_{\{r=1, k_1=2\}} 1_{\{P_1 = 0\}} \quad (31)$$

almost surely.

By Lemma 40 in appendix A, the point $(P_1, P_2, \ldots)$ is in Δ a.s. Let G be the distribution of $(P_1, P_2, \ldots)$. Write G as $G_0 + a\delta_0$, where G_0 has no atom at zero and δ_0 is a unit mass at zero. By taking the expectation of the expression in (31), we see that $P(\Theta_b = \theta)$ equals

$$\int_\Delta \left(\sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} P_{i_1}^{k_1} P_{i_2}^{k_2} \ldots P_{i_r}^{k_r} P_{i_{r+1}} \ldots P_{i_{r+l}} P_0^{s-l} \right) \Big/ \sum_{j=1}^{\infty} P_j^2 \, dG_0 + \alpha 1_{\{r=1, k_1=2\}}. \quad (32)$$

Let $\Xi = \Xi_0 + a\delta_0$, where $\Xi_0 = \lambda_{2;2;0} G_0$ and $a = \lambda_{2;2;0}\alpha$. Note that Ξ is a finite measure on Δ. It follows from (26) and (32) that the rates $\lambda_{b;k_1,\ldots,k_r;s}$ are given by (11). Hence, Theorem 2 follows from the claim. $\square$

Proof of Claim. Since Θ'_∞ is an exchangeable random partition, the sequence $(Z_m)_{m=3}^\infty$ is exchangeable. It is clear from the definition of this sequence that its limiting empirical distribution is the probability measure with an atom of size P_j at j for all $j \geq 0$. The σ-field generated by this random distribution is $\mathcal{F}$. Lemma 38 in appendix A implies that the random variables $Z_3, Z_4, \ldots$ are conditionally independent given $\mathcal{F}$ and that

$$P(Z_m = i|\mathcal{F}) = P_i \text{ a.s.} \tag{33}$$

for all $m \geq 3$ and $i \geq 0$. By the exchangeability of Π_∞, the sequences $(Z_1, Z_3, Z_4, \ldots)$ and $(Z_1, Z_{\sigma(3)}, Z_{\sigma(4)}, \ldots)$ have the same distribution for all finite permutations σ of $\{3, 4, \ldots\}$. By Lemma 39 in appendix A, Z_1 and $(Z_m)_{m=3}^\infty$ are conditionally independent given $\mathcal{F}$. It follows that the random variables $Z_1, Z_3, Z_4, \ldots$ are conditionally independent given $\mathcal{F}$.

Define $Q_i = P(Z_1 = i|\mathcal{F})$ for all $i \geq 1$. It remains only to show that $Q_i = P_i^2 / \sum_{j=1}^\infty P_j^2$ a.s. on $\{P_1 > 0\}$ for all $i \geq 1$. Fix $k \geq 4$ and $n > k$. Then make the following definitions:

$$\begin{aligned}
T &= \inf\{t : 1 \text{ and } 2 \text{ are in the same block of } \Pi_\infty(t)\}. \\
T_k &= \inf\{t : k-1 \text{ and } k \text{ are in the same block of } \Pi_\infty(t)\}. \\
E_n &= \{1, \ldots, n \text{ are in distinct blocks of } \Pi_\infty(T-)\}. \\
E_{n,k} &= \{1, \ldots, n \text{ are in distinct blocks of } \Pi_\infty(T_k-)\}.
\end{aligned}$$

Also, let $\theta_{k,1}$ be the partition of $\{1, \ldots, k\}$ into the blocks $\{1, 2\}$ and $\{3, 4, \ldots, k\}$, and let $\theta_{k,2}$ be the partition of $\{1, \ldots, k\}$ into the blocks $\{1, \ldots, k-2\}$ and $\{k-1, k\}$. Let $\Pi_{k+1,n}$ denote the restriction of Π_∞ to $\{k+1, \ldots, n\}$ and let $\pi_{k+1,n}$ be any partition of $\{k+1, \ldots, n\}$. The exchangeability of Π_∞ implies that

$$P(\{\Pi_k(T) = \theta_{k,1}\} \cap E_n \cap \{\Pi_{k+1,n}(T) = \pi_{k+1,n}\}) = P(\{\Pi_k(T_k) = \theta_{k,2}\} \cap E_{n,k} \cap \{\Pi_{k+1}(T_k) = \pi_{k+1,n}\}).$$

Since $T = T_k$ on the sets $\{\Pi_k(T) = \theta_{k,2}\} \cap E_n$ and $\{\Pi_k(T_k) = \theta_{k,2}\} \cap E_{n,k}$, it follows that

$$P(\{\Pi_k(T) = \theta_{k,1}\} \cap E_n \cap \{\Pi_{k+1,n}(T) = \pi_{k+1,n}\}) = P(\{\Pi_k(T) = \theta_{k,2}\} \cap E_n \cap \{\Pi_{k+1,n}(T) = \pi_{k+1,n}\})$$

and thus

$$P(\{\Pi_k(T) = \theta_{k,1}\} \cap \{\Pi_{k+1,n}(T) = \pi_{k+1,n}\}|E_n) = P(\{\Pi_k(T) = \theta_{k,2}\} \cap \{\Pi_{k+1,n}(T) = \pi_{k+1,n}\}|E_n).$$

Let $\Theta_{k+1,n}$ denote the restriction of Θ_∞ to $\{k+1, \ldots, n\}$. Since the distribution of Θ_n is the conditional distribution of $\Pi_n(T)$ given E_n, we have

$$P(\{\Theta_k = \theta_{k,1}\} \cap \{\Theta_{k+1,n} = \pi_{k+1,n}\}) = P(\{\Theta_k = \theta_{k,2}\} \cap \{\Theta_{k+1,n} = \pi_{k+1,n}\}). \tag{34}$$

Let $\mathcal{F}_{k+1,n}$ be the σ-field generated by $\Theta_{k+1,n}$. Since equation (34) holds for all partitions $\pi_{k+1,n}$ of $\{k+1, \ldots, n\}$, we have

$$P(\Theta_k = \theta_{k,1}|\mathcal{F}_{k+1,n}) = P(\Theta_k = \theta_{k,2}|\mathcal{F}_{k+1,n}) \text{ a.s.}$$

Let $\mathcal{F}_{k+1}$ be the σ-field generated by the restriction of Θ_∞ to $\{k+1, k+2, \ldots\}$. Since $\mathcal{F}_{k+1,n} \uparrow \mathcal{F}_{k+1}$ as $n \to \infty$, standard martingale convergence arguments (see, for example, Theorem 5.7 in chapter 4 of [5]) give

$$P(\Theta_k = \theta_{k,1}|\mathcal{F}_{k+1}) = P(\Theta_k = \theta_{k,2}|\mathcal{F}_{k+1}) \text{ a.s.}$$

The limiting relative frequencies of the blocks of Θ_∞ can be recovered from the restriction of Θ_∞ to $\{k+1, k+2, \ldots\}$, so the sequence $(P_j)_{j=1}^\infty$ is $\mathcal{F}_{k+1}$-measurable. Thus, $\mathcal{F} \subset \mathcal{F}_{k+1}$, so

$$P(\Theta_k = \theta_{k,1}|\mathcal{F}) = P(\Theta_k = \theta_{k,2}|\mathcal{F}) \text{ a.s.} \tag{35}$$

for all $k \geq 4$.

Using (33) combined with the fact that the random variables $Z_1, Z_3, Z_4, \ldots$ are conditionally independent given $\mathcal{F}$, we obtain

$$
\begin{aligned}
P(\Theta_k = \theta_{k,1}|\mathcal{F}) &= \sum_{i=1}^{\infty} P(Z_3 = \ldots = Z_k = i, Z_1 = Z_2 \neq i|\mathcal{F}) \\
&= \sum_{i=1}^{\infty} P(Z_3 = \ldots = Z_k = i|\mathcal{F}) P(Z_1 \neq i|\mathcal{F}) \\
&= \sum_{i=1}^{\infty} P_i^{k-2}(1 - Q_i) \qquad (36)
\end{aligned}
$$

almost surely for all $k \geq 4$. Likewise, almost surely we have, for all $k \geq 4$,

$$
\begin{aligned}
P(\Theta_k = \theta_{k,2}|\mathcal{F}) &= \sum_{i=1}^{\infty} \sum_{1 \leq j \neq i} P(Z_3 = \ldots = Z_{k-2} = i, Z_{k-1} = Z_k = j, Z_1 = Z_2 = i|\mathcal{F}) \\
&= \sum_{i=1}^{\infty} \sum_{1 \leq j \neq i} P(Z_3 = \ldots = Z_{k-2} = i|\mathcal{F}) P(Z_{k-1} = Z_k = j|\mathcal{F}) P(Z_1 = i|\mathcal{F}) \\
&= \sum_{i=1}^{\infty} \sum_{1 \leq j \neq i} P_i^{k-4} P_j^2 Q_i. \qquad (37)
\end{aligned}
$$

¿From (35), (36) and (37), we obtain, for all $k \geq 4$, the equation

$$
0 = \sum_{i=1}^{\infty} P_i^{k-4}(P_i^2(1 - Q_i) - Q_i \sum_{1 \leq j \neq i} P_j^2) = \sum_{i=1}^{\infty} P_i^{k-4}(P_i^2 - Q_i \sum_{j=1}^{\infty} P_j^2) \text{ a.s.} \qquad (38)
$$

On the set $\{P_i = P_l\}$, we have $Q_i = Q_l$ a.s. because, in the proof of the "only if" part of Theorem 2, the blocks $B_1, B_2, \ldots$ having the same limiting relative frequency were labeled in random order, independently of Θ_∞. Therefore, $P_i^2 - Q_i \sum_{j=1}^{\infty} P_j^2 = P_l^2 - Q_l \sum_{j=1}^{\infty} P_j^2$ a.s. on $\{P_i = P_l\}$. It follows from (38) and Lemma 20 that $P_i^2 - Q_i \sum_{j=1}^{\infty} P_j^2 = 0$ a.s. on $\{P_i > 0\}$. Therefore,

$$
Q_i = P_i^2 / \sum_{j=1}^{\infty} P_j^2 \qquad (39)
$$

almost surely on $\{P_i > 0\}$. It follows that on $\{P_1 > 0\}$, we have $\sum_{\{j: P_j > 0\}} Q_j = 1$ a.s. Therefore, on the event $\{P_1 > 0\} \cap \{P_i = 0\}$, we have $Q_i = 0$ a.s. Thus, (39) holds almost surely on $\{P_1 > 0\}$ for all $i \geq 1$, which completes the proof of the claim. □

4.3 Proofs of Propositions 4 and 6

We can establish Propositions 4 and 6 by a straightforward application of the Poisson process construction in section 3. We prove Proposition 6 first, as it is used in the proof of Proposition 4.

Proof of Proposition 6. Write $G = G_0 + \alpha\delta_0$, where G_0 has no atom at zero. We may assume without loss of generality that Π_∞ is derived from the Poisson point process $(e(t))_{t\geq 0}$ with characteristic measure L given by (14) with G_0 and α in place of Ξ_0 and a. The construction of Π_∞ implies that $T = \inf\{t : e(t) \in A_{1,2}\}$, where $A_{1,2}$ is defined by (16). Let $\tilde{\Theta}$ be the random partition of $\{3, 4, \ldots\}$ such that i and j are in the same block of $\tilde{\Theta}$ if and only if $e(T)_i = e(T)_j$. It follows from the definition of $(e(t))_{t\geq 0}$ that $\tilde{\Theta}$ is an exchangeable random partition. Also, we see from the construction of Π_∞ that $\Theta = \tilde{\Theta}$ on $\{\#\Pi_\infty(T-) = \infty\}$ and Θ is the restriction of $\tilde{\Theta}$ to $\{3, 4, \ldots, \#\Pi_\infty(T-)\}$ on $\{\#\Pi_\infty(T-) < \infty\}$. Since $0 < L(A_{1,2}) < \infty$ by (17), Lemma 15 implies that $\Pi_\infty(T-)$ is independent of $e(T)$ and thus is independent of $\tilde{\Theta}$. Therefore, to show that Θ satisfies the conclusion of Proposition 6, it suffices to show that the ranked sequence of limiting relative frequencies of the blocks of $\tilde{\Theta}$ has distribution G.

For every $\xi = (\xi_i)_{i=1}^\infty \in \mathbb{Z}^\infty$, and $k \in \mathbb{N}$, define

$$N_k(\xi) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n 1_{\{\xi_i = k\}}, \tag{40}$$

provided this limit exists. Define $N(\xi) = (N_1(\xi), N_2(\xi), \ldots)$. If ξ is random with distribution P_x, as defined in section 3, then by the strong law of large numbers, we have $N(\xi) = x$ a.s. Let S be a Borel subset of the infinite simplex Δ, and let $A_{1,2}^S = \{\xi \in A_{1,2} : N(\xi) \in S\}$. By part (d) of Lemma 41 in appendix B, we have

$$P(N(e(T)) \in S) = P(e(T) \in A_{1,2}^S) = L(A_{1,2}^S)/L(A_{1,2}).$$

By (17), $L(A_{1,2}) = G(\Delta) = 1$. Since $P_x(A_{1,2}) = \sum_{j=1}^\infty x_j^2$, we have $P_x(A_{1,2}^S) = (\sum_{j=1}^\infty x_j^2)1_{\{x\in S\}}$. Therefore, by (14),

$$\begin{aligned} L(A_{1,2}^S) &= \int_\Delta \left(\sum_{j=1}^\infty x_j^2\right) 1_{\{x\in S\}} \Big/ \sum_{j=1}^\infty x_j^2 \, G_0(dx) + \alpha \sum_{i=1}^\infty \sum_{j=i+1}^\infty 1_{\{z_{ij} \in A_{1,2}^S\}} \\ &= G_0(S) + \alpha 1_{\{(0,0,\ldots)\in S\}} = G(S). \end{aligned}$$

Thus, $N(e(T))$ has distribution G. Since $N(e(T))$ is the ranked sequence of limiting relative frequencies of blocks of $\tilde{\Theta}$, the proposition follows. □

Proof of Proposition 4. Suppose $(\Pi_\infty(t))_{t\geq 0}$ is both a standard Ξ-coalescent and a standard Ξ'-coalescent. Then, the collision rates associated with the Ξ-coalescent and the Ξ'-coalescent must be the same. By (12), we have $\Xi(\Delta) = \Xi'(\Delta)$. Therefore, it suffices to establish the proposition when Ξ and Ξ', are probability measures.

Define T and Θ as in Proposition 6. First, suppose $P(\#\Pi_\infty(T-) = \infty) > 0$. By Proposition 6, conditional on $\{\#\Pi_\infty(T-) = \infty\}$, the distribution of the ranked sequence of limiting relative frequencies of the blocks of Θ is Ξ. Since Π_∞ is also a Ξ'-coalescent, this distribution must also be Ξ', so $\Xi = \Xi'$.

Next, suppose $P(\#\Pi_\infty(T-) = \infty) = 0$. The second paragraph of the proof of the "only if" part of Theorem 2 implies that $P(\#\Pi_\infty(T-) \geq n) > 0$ for all $n \in \mathbb{N}$. Thus, for all $m \in \mathbb{N}$, there exists an $n \geq m$ such that $P(\#\Pi_\infty(T-) = n) > 0$. Proposition 6 implies that conditional on $\{\#\Pi_\infty(T-) = n\}$, Θ is both a Ξ-partition of $\{3,4,\ldots,n\}$ and a Ξ'-partition of $\{3,4,\ldots,n\}$. Thus, for arbitrarily large n, and therefore for all n, a Ξ-partition of $\{3,4,\ldots,n\}$ and a Ξ'-partition of $\{3,4,\ldots,n\}$ have the same distribution. Thus, a Ξ-partition of $\{3,4,\ldots\}$ has the same distribution as a Ξ'-partition of $\{3,4,\ldots\}$, so $\Xi = \Xi'$, which completes the proof. $\square$

4.4 Proof of Proposition 7

In this subsection, we consider a population model of the type discussed in the introduction, and we adopt the notation used in the introduction. We begin with the following lemma.

Lemma 21 *Let Ξ be a probability measure on the infinite simplex Δ. Denote by $\lambda_{b;k_1,\ldots,k_r;s}$ the collision rates of the Ξ-coalescent, which are defined from Ξ by (11). Suppose, for a population model of the type discussed in the introduction, we have $\lim_{N\to\infty} c_N = 0$ and*

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_{k_1}\ldots(\nu_{r,N})_{k_r}]}{N^{k_1+\ldots+k_r-r}c_N} = \lambda_{b;k_1,\ldots,k_r;0} \tag{41}$$

for all $r \geq 1$, $k_1,\ldots,k_r \geq 2$ and $b = \sum_{j=1}^r k_j$. Then, for all n, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t\geq 0}$ converge as $N \to \infty$ in the Skorohod topology to $(R_n\Pi_\infty(t))_{t\geq 0}$, where Π_∞ is a standard Ξ-coalescent.

Proof. By Proposition 1, for each n the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t\geq 0}$ converge in the Skorohod topology to a coalescent process $(\Psi_{n,\infty}(t))_{t\geq 0}$ as $N \to \infty$. As noted in the introduction, the collision rates defined in (6) for the processes $(\Psi_{n,\infty}(t))_{t\geq 0}$, which we denote by $\lambda'_{b;k_1,\ldots,k_r;s}$, satisfy the consistency condition (23). By Lemma 18, there exists a process $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ satisfying conditions B1 and B2 with collision rates $\lambda'_{b;k_1,\ldots,k_r;s}$, and $(\Psi_{n,\infty}(t))_{t\geq 0}$ has the same distribution as $(R_n\Pi_\infty(t))_{t\geq 0}$ for all n. By Theorem 2, Π_∞ is a Ξ'-coalescent for some finite measure Ξ' on the infinite simplex Δ. It follows from (10) that

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_{k_1}\ldots(\nu_{r,N})_{k_r}]}{N^{k_1+\ldots+k_r-r}c_N} = \lambda'_{b;k_1,\ldots,k_r;0}. \tag{42}$$

Equations (41) and (42) imply that $\lambda_{b;k_1,\ldots,k_r;0} = \lambda'_{b;k_1,\ldots,k_r;0}$ whenever $r \geq 1$, $k_1,\ldots,k_r \geq 2$, and $b = \sum_{j=1}^r k_j$. Following Lemma 3.4 of [13], we write (23) in the form

$$\lambda_{b+1;k_1,\ldots,k_r;s+1} = \lambda_{b;k_1,\ldots,k_r;s} - \sum_{m=1}^r \lambda_{b+1;k_1,\ldots,k_{m-1},k_m+1,k_{m+1},\ldots,k_r;s} - s\lambda_{b+1;k_1,\ldots,k_r,2;s-1}. \tag{43}$$

By induction on s, we can see from (43) and our convention that $s\lambda_{b+1;k_1,\ldots,k_r,2;s-1} = 0$ when $s = 0$ that the collision rates when $s = 0$ determine all of the collision rates uniquely. Thus, $\lambda_{b;k_1,\ldots,k_r;s} = \lambda'_{b;k_1,\ldots,k_r;s}$ for all $r \geq 1$, $k_1,\ldots,k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$. It follows from Proposition 4 that $\Xi = \Xi'$, which proves the lemma. $\square$

We also require the following lemma, stated on p.284 of [8], regarding the factorial moments of the multinomial distribution.

Lemma 22 *Suppose* $(X_1, \dots, X_c)$ *has a multinomial*$(N; p_1, \dots, p_c)$*-distribution. Then for all nonnegative integers* $q_1, \dots, q_c$*, we have* $E[(X_1)_{q_1}(X_2)_{q_2} \dots (X_c)_{q_c}] = (N)_{q_1+\dots+q_c} p_1^{q_1} \dots p_c^{q_c}$.

Proof of Proposition 7. We can write $\Xi = \Xi_0 + a\delta_0$, where Ξ_0 has no atom at zero. Let $\Pi_\infty = (\Pi_\infty(t))_{t\geq 0}$ be a standard Ξ-coalescent, and denote the collision rates of Π_∞ by $\lambda_{b;k_1,\dots,k_r;s}$. For all $N \geq 2$, let $T_N = \inf\{t : \#R_N\Pi_\infty(t) < N\}$, so T_N is the first collision time of $R_N\Pi_\infty$. Let $S_{1,N}, \dots, S_{w,N}$ be the sizes of the blocks of $R_N\Pi_\infty(T_N)$. Let $\tilde{\nu}_{1,N}, \dots, \tilde{\nu}_{N,N}$ be obtained by randomly ordering the elements of the multiset M consisting of $S_{1,N}, \dots, S_{w,N}$ and $N - w$ zeros. Let V_N be a random variable, independent of Π_∞, such that $P(V_N = 1) = 1/N$ and $P(V_N = 0) = 1 - 1/N$. Define $(\nu_{1,N}, \dots, \nu_{N,N}) = (\tilde{\nu}_{1,N}, \dots, \tilde{\nu}_{N,N})$ on $\{V_N = 1\}$ and $(\nu_{1,N}, \dots, \nu_{N,N}) = (1, \dots, 1)$ on $\{V_N = 0\}$. Let μ_N be the distribution of $(\nu_{1,N}, \dots, \nu_{N,N})$. We claim that for all n, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t\geq 0}$ derived from the distributions μ_N as described in the introduction converge in the Skorohod topology to $(R_n\Pi_\infty(t))_{t\geq 0}$ as $N \to \infty$. This claim will establish the proposition. By Lemma 21, it suffices to show that $\lim_{N\to\infty} c_N = 0$ and that for all $r \geq 1$, $k_1, \dots, k_r \geq 2$ and $b = \sum_{j=1}^r k_j$, we have

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_{k_1} \dots (\nu_{r,N})_{k_r}]}{N^{k_1+\dots+k_r-r} c_N} = \lambda_{b;k_1,\dots,k_r;0}. \tag{44}$$

We may assume without loss of generality that Π_∞ is derived from a Poisson point process $(e(t))_{t\geq 0}$ with characteristic measure L, where L is defined from Ξ by (14). Define A_N by (15). Recall from (18) that $L(A_N) < \infty$. It follows from the construction that $T_N = \inf\{t : e(t) \in A_N\}$. Let $\zeta^N = (\zeta_1^N, \dots, \zeta_N^N)$ be the first N coordinates of $e(T_N)$. If $1 \leq i, j \leq N$, then i and j are in the same block of $R_N\Pi_\infty(T_N)$ if and only if $\zeta_i^N = \zeta_j^N$. For $y \in \mathbb{Z}^N$, define

$$S_y = \{\xi \in \mathbb{Z}^\infty : (\xi_1, \dots, \xi_N) = y\}.$$

Whenever $S_y \subset A_N$, we have $P(\zeta^N = y) = L(S_y)/L(A_N)$ by part (d) of Lemma 41 in appendix B. Let $f : \mathbb{Z}^N \to [0, \infty)$ be a function such that $f(y) = 0$ unless $S_y \subset A_N$. Also, let z_{ij}^N denote the first N coordinates of z_{ij}, and let Y_x be the first N coordinates of a random sequence with distribution P_x, where P_x is defined as in section 3. Then,

$$\begin{aligned}
E[f(\zeta^N)] &= \sum_{y\in\mathbb{Z}^N} P(\zeta^N = y) f(y) = \frac{1}{L(A_N)} \sum_{y\in\mathbb{Z}^N} f(y) L(S_y) \\
&= \frac{1}{L(A_N)} \sum_{y\in\mathbb{Z}^N} \left(\int_\Delta f(y) P_x(S_y) \Big/ \sum_{j=1}^\infty x_j^2 \, \Xi_0(dx) + a \sum_{i=1}^\infty \sum_{j=i+1}^\infty f(y) 1_{\{z_{ij}\in S_y\}} \right) \\
&= \frac{1}{L(A_N)} \left(\int_\Delta \sum_{y\in\mathbb{Z}^N} f(y) P_x(S_y) \Big/ \sum_{j=1}^\infty x_j^2 \, \Xi_0(dx) + a \sum_{i=1}^\infty \sum_{j=i+1}^\infty \sum_{y\in\mathbb{Z}^N} f(y) 1_{\{z_{ij}\in S_y\}} \right) \\
&= \frac{1}{L(A_N)} \left(\int_\Delta E[f(Y_x)] \Big/ \sum_{j=1}^\infty x_j^2 \, \Xi_0(dx) + a \sum_{i=1}^N \sum_{j=i+1}^N f(z_{ij}^N) \right).
\end{aligned} \tag{45}$$

Given $y = (y_1, \dots, y_N) \in \mathbb{Z}^N$ and $j \in \mathbb{Z}$, let $h_{j,N}(y)$ be the cardinality of $\{i : y_i = j\}$. Fix $r \geq 1$ and $k_1, \dots, k_r \geq 2$. Define

$$f_N(y) = \frac{1}{(N)_r} \sum_{i_1\neq\dots\neq i_r} (h_{i_1,N}(y))_{k_1} \dots (h_{i_r,N}(y))_{k_r}, \tag{46}$$

where the sum is over distinct positive integers $i_1, \dots, i_r$. We claim that

$$E[(\tilde{\nu}_{1,N})_{k_1} \dots (\tilde{\nu}_{r,N})_{k_r}] = E[f_N(\zeta^N)]. \tag{47}$$

To see (47), recall that we defined $\tilde{\nu}_{1,N}, \dots \tilde{\nu}_{r,N}$ by randomly ordering the elements of the multiset M. By the definitions of M and ζ^N and the properties of the Poisson process construction, M consists of the cardinalities of the w nonempty sets of the form $R_{j,N} = \{i : \zeta_i^N = j\}$ and $N - w$ zeros. Note that $h_{j,N}(\zeta^N)$ is the cardinality of $R_{j,N}$, so M contains $N - w$ zeros as well as $h_{j,N}(\zeta^N)$ for all j such that $h_{j,N}(\zeta^N) \neq 0$. It follows that $(\tilde{\nu}_{1,N})_{k_1} \dots (\tilde{\nu}_{r,N})_{k_r}$ is either zero or equals $(h_{i_1,N}(\zeta^N))_{k_1} \dots (h_{i_r,N}(\zeta^N))_{k_r}$ for some r-tuple $(i_1, \dots, i_r)$ of distinct positive integers. Thus, when $y = \zeta^N$, all possible nonzero values of $(\tilde{\nu}_{1,N})_{k_1} \dots (\tilde{\nu}_{r,N})_{k_r}$ are terms in the sum on the right-hand side of (46). Conditional on the event that $h_{i_1,N}(\zeta^N), \dots, h_{i_r,N}(\zeta^N)$ are nonzero, the probability that these are the first r elements, in order, of the multiset M after a random ordering is $1/(N)_r$. Thus,

$$E[(\tilde{\nu}_{1,N})_{k_1} \dots (\tilde{\nu}_{r,N})_{k_r} | \zeta^N] = f_N(\zeta^N). \tag{48}$$

Equation (47) follows by taking expectations of both side of (48).

Note that for all $x \in \Delta$ and all distinct positive integers $i_1, \dots, i_r$, the distribution of the vector

$$(h_{i_1,N}(Y_x), \dots, h_{i_r,N}(Y_x), N - \sum_{j=1}^{r} h_{i_j,N}(Y_x))$$

is multinomial$(N; x_{i_1}, \dots, x_{i_r}, 1 - \sum_{j=1}^{r} x_{i_j})$. Thus, by applying Lemma 22 with $c = r+1$, $q_i = k_i$ for $i = 1, \dots, r$, and $q_{r+1} = 0$, we obtain

$$E[f_N(Y_x)] = \frac{1}{(N)_r} \sum_{i_1 \neq \dots \neq i_r} (N)_{k_1 + \dots + k_r} x_{i_1}^{k_1} \dots x_{i_r}^{k_r} \sim N^{k_1 + \dots + k_r - r} \sum_{i_1 \neq \dots \neq i_r} x_{i_1}^{k_1} \dots x_{i_r}^{k_r}, \tag{49}$$

where $\sim$ denotes asymptotic equivalence as $N \to \infty$. Also, $f_N(z_{ij}^N) = 0$ unless $r = 1$, $k_1 = 2$, $i \leq N$, $j \leq N$, and $i \neq j$. In this case, $h_{1,N}(z_{ij}^N) = 2$ and $h_{k,N}(z_{ij}^N) \in \{0, 1\}$ for $k \neq 1$, which means

$$f_N(z_{ij}^N) = \frac{1}{N} \sum_{k=1}^{\infty} (h_{k,N}(z_{ij}))_2 = \frac{2}{N}.$$

Therefore, if $r = 1$ and $k_1 = 2$, we have

$$\sum_{i=1}^{N} \sum_{j=i+1}^{N} f_N(z_{ij}^N) = \binom{N}{2} \frac{2}{N} \sim N = N^{k_1 + \dots + k_r - r}. \tag{50}$$

Equations (45), (47), (49), and (50) imply that

$$\begin{aligned} E[(\tilde{\nu}_{1,N})_{k_1} \dots (\tilde{\nu}_{r,N})_{k_r}] &= \frac{1}{L(A_N)} \bigg(\int_\Delta E[f_N(Y_x)] \Big/ \sum_{j=1}^{\infty} x_j^2 \, \Xi_0(dx) + a \sum_{i=1}^{N} \sum_{j=i+1}^{N} f_N(z_{ij}^N) \bigg) \\ &\sim \frac{N^{k_1 + \dots + k_r - r}}{L(A_N)} \bigg(\int_\Delta \sum_{i_1 \neq \dots \neq i_r} x_{i_1}^{k_1} \dots x_{i_r}^{k_r} \Big/ \sum_{j=1}^{\infty} x_j^2 \, \Xi_0(dx) + a 1_{\{r=1, k_1=2\}} \bigg). \end{aligned}$$

By (11), it follows that

$$E[(\tilde{\nu}_{1,N})_{k_1}\dots(\tilde{\nu}_{r,N})_{k_r}]\sim\frac{N^{k_1+\dots+k_r-r}}{L(A_N)}\lambda_{b;k_1,\dots,k_r;0} \tag{51}$$

for all $r\geq 1$, $k_1,\dots,k_r\geq 2$, and $b=\sum_{j=1}^r k_j$. Since $(\nu_{1,N},\dots,\nu_{N,N})=(1,\dots,1)$ with probability $1-1/N$ and $(\nu_{1,N},\dots,\nu_{N,N})=(\tilde{\nu}_{1,N},\dots,\tilde{\nu}_{N,N})$ with probability $1/N$, it follows from (51) that

$$E[(\nu_{1,N})_{k_1}\dots(\nu_{r,N})_{k_r}]\sim\frac{N^{k_1+\dots+k_r-r}}{NL(A_N)}\lambda_{b;k_1,\dots,k_r;0}. \tag{52}$$

When $r=1$ and $k_1=2$, equations (4) and (52) give

$$c_N=\frac{E[(\nu_{1,N})_2]}{N-1}\sim\left(\frac{1}{N-1}\right)\frac{N}{NL(A_N)}\lambda_{2;2;0}\sim\frac{1}{NL(A_N)}\lambda_{2;2;0}. \tag{53}$$

Since $L(A_N)\geq L(A_{1,2})=\Xi(\Delta)=1$ by (17), it follows from (53) that $\lim_{N\to\infty}c_N=0$. Moreover, from (52) and (53), we obtain

$$\lim_{N\to\infty}\frac{E[(\nu_{1,N})_{k_1}\dots(\nu_{r,N})_{k_r}]}{N^{k_1+\dots+k_r-r}c_N}=\frac{\lambda_{b;k_1,\dots,k_r;0}}{\lambda_{2;2;0}}.$$

Since Ξ is a probability measure, we have $\lambda_{2;2;0}=1$ by (12), which establishes (44) and completes the proof of the proposition. ☐

4.5 Proofs of propositions related to $(F_r)_{r=1}^\infty$

In this subsection, we prove Propositions 8, 10, and 11, all of which relate to the characterization of coalescents with simultaneous multiple collisions by a sequence of measures $(F_r)_{r=1}^\infty$.

We will use the following lemma. When $(F_r)_{r=1}^\infty$ is associated with a population model as described in Proposition 1, then this result is exactly Lemma 3.4 of [13]. However, the proof in [13] uses the properties of the population model. Since we wish to apply the result without knowing, a priori, whether the sequence $(F_r)_{r=1}^\infty$ is associated with a population model, we will prove the result for all $(F_r)_{r=1}^\infty$ satisfying conditions A1, A2, and A3′ of Proposition 8.

Lemma 23 *Let $(F_r)_{r=1}^\infty$ be a sequence of measures satisfying conditions A1, A2, and A3′ of Proposition 8. Define real numbers $\{\lambda_{b;k_1,\dots,k_r;s}:r\geq 1,\ k_1,\dots,k_r\geq 2,\ s\geq 0,\ b=\sum_{j=1}^r k_j+s\}$ by (6). Then (23) holds.*

Proof. Verifying (23) is equivalent to verifying (43). Equation (6) implies that (43) is equivalent to

$$\begin{aligned}&\sum_{m=r}^{\lfloor r+(s+1)/2\rfloor}\int_{\Delta_m}x_1^{k_1-2}\dots x_r^{k_r-2}T_{r,s+1}^{(m)}(x_1,\dots,x_m)\,F_m(dx_1,\dots,dx_m)\\ =&\sum_{m=r}^{\lfloor r+s/2\rfloor}\int_{\Delta_m}x_1^{k_1-2}\dots x_r^{k_r-2}(1-\sum_{i=1}^r x_i)T_{r,s}^{(m)}(x_1,\dots,x_m)\,F_m(dx_1,\dots,dx_m)\\ -&\ s\sum_{m=r+1}^{\lfloor r+(s+1)/2\rfloor}\int_{\Delta_m}x_1^{k_1-2}\dots x_r^{k_r-2}T_{r+1,s-1}^{(m)}(x_1,\dots,x_m)\,F_m(dx_1,\dots,dx_m).\end{aligned} \tag{54}$$

It follows from (8) that $T_{m-j,s}^{(m)} = 0$ when $s < 2j$, so $T_{r,s}^{(m)} = 0$ when $s < 2(m-r)$ or, equivalently, when $m > r + s/2$. Thus, (54) would be unchanged if we took the three sums over m in (54) up to infinity. The resulting equality follows from (7) and (8) and is stated in the proof of Lemma 3.6 of [13]. $\square$

Proof of Proposition 8. Suppose $\{\lambda_{b;k_1,\ldots,k_r;s} : r \geq 1,\ k_1,\ldots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$ is a collection of nonnegative real numbers such that there exists a $\mathcal{P}_\infty$-valued coalescent Π_∞ satisfying conditions B1 and B2 of Theorem 2. By Theorem 2, the rates can be obtained from (11) for some finite measure Ξ on the infinite simplex Δ. By (12), we can write $\Xi = \lambda_{2;2;0}G$, where G is a probability measure. Denote the collision rates of the G-coalescent by $\lambda'_{b;k_1,\ldots,k_r;s}$. By Proposition 1 and Proposition 7, there exists a unique sequence of measures $(F'_r)_{r=1}^\infty$ satisfying conditions A1, A2, and A3 of Proposition 1 such that the collision rates $\lambda'_{b;k_1,\ldots,k_r;s}$ are given by (6) with F_m replaced by F'_m on the right-hand side. Let $F_r = \lambda_{2;2;0}F'_r$ for all r. Then $(F_r)_{r=1}^\infty$ satisfies A1, A2, and A3$'$, and the collision rates $\lambda_{b;k_1,\ldots,k_r;s}$ are given by (6). The uniqueness of $(F'_r)_{r=1}^\infty$, which is asserted in Proposition 1, implies the uniqueness of $(F_r)_{r=1}^\infty$.

Next, suppose $\{\lambda_{b;k_1,\ldots,k_r;s} : r \geq 1,\ k_1,\ldots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$ is a collection of nonnegative real numbers, and suppose there is a sequence of measures $(F_r)_{r=1}^\infty$ satisfying A1, A2, and A3$'$ such that (6) holds for all $\lambda_{b;k_1,\ldots,k_r;s}$. Lemma 23 implies that (23) holds, so by Lemma 18, there exists a $\mathcal{P}_\infty$-valued coalescent satisfying B1 and B2 with collision rates $\{\lambda_{b;k_1,\ldots,k_r;s} : r \geq 1,\ k_1,\ldots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$.

Finally, suppose $(F_r)_{r=1}^\infty$ is any sequence of measures satisfying A1, A2, and A3$'$, and define a collection of real numbers $\{\lambda_{b;k_1,\ldots,k_r;s} : r \geq 1,\ k_1,\ldots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^r k_j + s\}$ by (6). If A4 holds, then the argument in the previous paragraph implies that there exists a coalescent process with collision rates $\lambda_{b;k_1,\ldots,k_r;s}$ satisfying B1 and B2. However, if the right-hand side of (6) is negative for some $r \geq 1$, $k_1,\ldots,k_r \geq 2$, and $s \geq 0$, then clearly there can be no coalescent process with collision rates $\lambda_{b;k_1,\ldots,k_r;s}$ satisfying B1 and B2. $\square$

Proof of Proposition 10. Recall that $(F_r)_{r=1}^\infty$ satisfies conditions A1, A2, A3$'$, and A4 of Proposition 8, and Π_∞ is a coalescent satisfying conditions B1 and B2 of Theorem 2 with collision rates $\lambda_{b;k_1,\ldots,k_r;s}$ defined from $(F_r)_{r=1}^\infty$ by (6). To prove the first statement in the proposition, note that $R_{2r}\Pi_\infty(0)$ consists of $2r$ singletons, and $\lambda_{2r;2,\ldots,2;0}$ is the rate at which these singletons are colliding to form the blocks $\{1,2\}$, $\{3,4\}$, $\ldots$, $\{2r-1,2r\}$. Thus, $P(E_r) = 0$ if and only if $\lambda_{2r;2,\ldots,2;0} = 0$. Also, by (9), we have $\lambda_{2r;2,\ldots,2;0} = F_r(\Delta_r)$, so $\lambda_{2r;2,\ldots,2;0} = 0$ if and only if $F_r(\Delta_r) = 0$.

By Theorem 2, Π_∞ must be a standard Ξ-coalescent for some finite measure Ξ. Therefore, we may assume that Π_∞ is derived from a Poisson point process $(e(t))_{t\geq 0}$ with intensity measure L, where L is defined in terms of Ξ by (14). We have $T_r = \inf\{t : e(t) \in A_{2r}\}$, where A_{2r} is defined by (15). Define

$$D_r = \{\xi \in \mathbb{Z}^\infty : \xi_{2j-1} = \xi_{2j} \text{ for } j = 1,\ldots,r \text{ and } \xi_{2i} \neq \xi_{2j} \text{ if } 1 \leq i < j \leq r\}. \tag{55}$$

It follows from the construction of Π_∞ that E_r occurs if and only if $e(T_r) \in D_r$.

Let $\tilde{\Theta}_r$ be the partition of $\mathbb{N}$ such that i and j are in the same block of $\tilde{\Theta}_r$ if and only if $e(T_r)_i = e(T_r)_j$. It follows from the construction that Θ_r equals $\tilde{\Theta}_r$ on $\{\#\Pi_\infty(T_r-) = \infty\}$ and

Θ_r is the restriction of $\tilde{\Theta}_r$ to $\{1, 2, \ldots, \#\Pi_\infty(T_r-)\}$ on $\{\#\Pi_\infty(T_r-) < \infty\}$. Suppose $P(E_r) > 0$. Then let Θ'_r be a random partition whose distribution is the same as the conditional distribution of $\tilde{\Theta}_r$ given E_r. Since Π_∞ is an exchangeable process (see Definition 19), the restriction of Θ'_r to $\{2r+1, 2r+2, \ldots\}$ is exchangeable. Let $f'_{j,r}$ be the limiting relative frequency of the block of Θ'_r containing $2j-1$ and $2j$, which exists by Lemma 40 of appendix A. Let Q_r be the distribution of $(f'_{1,r}, \ldots, f'_{r,r})$.

We first show that Q_r satisfies condition (b) of Proposition 10. Suppose that we have $P(\#\Pi_\infty(T_r-) = \infty) > 0$. By Lemma 15, $\Pi_\infty(T_r-)$ and $e(T_r)$ are independent since $L(A_{2r}) < \infty$. Therefore, the conditional distribution of $\tilde{\Theta}_r$ given E_r equals the conditional distribution of $\tilde{\Theta}_r$ given $E_r \cap \{\#\Pi_\infty(T_r-) = \infty\}$. The conditional distribution of the limiting relative frequencies of the first r blocks of $\tilde{\Theta}_r$ given E_r equals Q_r by definition. Since $\Theta_r = \tilde{\Theta}_r$ on $\{\#\Pi_\infty(T_r-) = \infty\}$, the conditional distribution of the limiting relative frequencies of the first r blocks of $\tilde{\Theta}_r$ given $E_r \cap \{\#\Pi_\infty(T_r-) = \infty\}$ equals the conditional distribution of $(f_{1,r}, \ldots, f_{r,r})$ given $E_r \cap \{\#\Pi_\infty(T_r-) = \infty\}$, where $f_{j,r}$ is the limiting relative frequency of the block of Θ_r containing $2j-1$ and $2j$. Hence, conditional on $E_r \cap \{\#\Pi_\infty(T_r-) = \infty\}$, the distribution of $(f_{1,r}, \ldots, f_{r,r})$ equals Q_r, which is condition (b).

Next, we show that Q_r satisfies condition (c). Suppose $P(\#\Pi_\infty(T_r-) = n) > 0$. The conditional distribution of Θ_r given $E_r \cap \{\#\Pi_\infty(T_r-) = n\}$ equals the conditional distribution of the restriction of $\tilde{\Theta}_r$ to $\{1, \ldots, n\}$ given $E_r \cap \{\#\Pi_\infty(T_r-) = n\}$. By the independence of $\Pi_\infty(T_r-)$ and $e(T_r)$, this equals the conditional distribution of the restriction of $\tilde{\Theta}_r$ to $\{1, \ldots, n\}$ given E_r, which equals the distribution of Θ'_r restricted to $\{1, \ldots, n\}$. Since Q_r was defined to be the distribution of $f'_{1,r}, \ldots, f'_{r,r}$, it follows that Q_r satisfies (c).

It remains to show that Q_r satisfies condition (a) for all r such that $P(E_r) > 0$. For all r, define $F'_r = 0$ if $P(E_r) = 0$ and $F'_r = \lambda_{2r;2,\ldots,2;0} Q_r$ if $P(E_r) > 0$. It suffices to show that $F_r = F'_r$ for all r. We first show that the sequence $(F'_r)_{r=1}^\infty$ satisfies properties A1, A2, and A3$'$ of Proposition 8. That $(F'_r)_{r=1}^\infty$ satisfies A1 is clear from the definition. Note that for all r, the measure Q_r is symmetric with respect to the r coordinates of Δ_r because Π_∞ is exchangeable. Therefore each F'_r is symmetric and so A2 holds. Also, we have $F'_r(\Delta_r) = \lambda_{2r;2,\ldots,2;0}$ for all $r \geq 1$. Whenever $B_1, \ldots, B_{2r}$ merge into the r blocks $B_1 \cup B_2, B_3 \cup B_4, \ldots, B_{2r-1} \cup B_{2r}$, the blocks $B_1, \ldots, B_{2(r-1)}$ merge into $B_1 \cup B_2, B_3 \cup B_4, \ldots, B_{2(r-1)-1} \cup B_{2(r-1)}$. Therefore, we have $\lambda_{2(r-1);2\ldots,2;0} \geq \lambda_{2r;2,\ldots,2;0}$ for all $r \geq 2$, which means $F'_1(\Delta_1) \geq F'_2(\Delta_2) \geq \ldots$. Thus, $(F'_r)_{r=1}^\infty$ satisfies A3$'$.

Let $\lambda'_{b;k_1,\ldots,k_r;s}$ be defined by (6) with F'_m in place of F_m. Since both $(F_r)_{r=1}^\infty$ and $(F'_r)_{r=1}^\infty$ satisfy A1, A2, and A3$'$, the uniqueness assertion in Proposition 8 implies that we can prove $F_r = F'_r$ for all r by showing that

$$\lambda_{b;k_1,\ldots,k_r;s} = \lambda'_{b;k_1,\ldots,k_r;s} \tag{56}$$

for all $r \geq 1$, $k_1, \ldots, k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$. Also, Lemma 23 implies that (43) holds for both collections of collision rates. Therefore, by induction on s, it suffices to verify (56) when $s = 0$. If $P(E_r) = 0$, then $F_r(\Delta) = F'_r(\Delta) = 0$, so when $s = 0$, (56) follows immediately from (9). Therefore, we may assume for the rest of the proof that $P(E_r) > 0$.

Let θ be a partition of $\{1, \ldots, b\}$ into blocks of sizes $k_1, \ldots, k_r \geq 2$ such that $R_{2r}\theta$ consists of the blocks $\{1, 2\}, \{3, 4\}, \ldots, \{2r-1, 2r\}$ and k_j is the size of the block containing $2j-1$ and $2j$. Let A_θ consist of all $\xi \in \mathbb{Z}^\infty$ such that if $1 \leq i, j \leq b$ then $\xi_i = \xi_j$ if and only if i and j are in the same block of θ. Define

$$U_r = \inf\{e(t) \in D_r\}, \tag{57}$$

where D_r is defined by (55). Then, E_r is the event that $1, \dots, 2r$ are in distinct blocks of $\Pi_\infty(U_r-)$, and $T_r = U_r$ on E_r. Since $D_r \subset A_{2r}$, where A_{2r} is defined by (15), we have $L(D_r) < \infty$. Therefore, $\Pi_\infty(U_r-)$ and $e(U_r)$ are independent by Lemma 15. Using this independence for the fourth equality and part (d) of Lemma 41 in appendix B for the fifth equality, we have

$$\begin{aligned} P(R_b\Theta'_r = \theta) &= P(R_b\tilde{\Theta}_r = \theta | E_r) = P(e(T_r) \in A_\theta | E_r) \\ &= P(e(U_r) \in A_\theta | E_r) = P(e(U_r) \in A_\theta) = \frac{L(A_\theta)}{L(D_r)} = \frac{\lambda_{b;k_1,\dots,k_r;0}}{\lambda_{2r;2,\dots,2;0}}. \end{aligned} \tag{58}$$

Define a sequence of random variables $(Z_i)_{i=2r+1}^\infty$ such that $Z_i = j$ if i is in the same block of Θ'_r as $2j$ and $Z_i = 0$ if j is not in the same block of Θ'_r as any of $1, \dots, 2r$. Since the restriction of Θ'_r to $\{2r+1, 2r+2, \dots\}$ is an exchangeable random partition, $(Z_i)_{i=2r+1}^\infty$ is an exchangeable random sequence. By Lemma 40 of appendix A, $(Z_i)_{i=2r+1}^\infty$ is has a limiting empirical measure μ given by $\mu(\{j\}) = f'_j$ for $j = 1, \dots, r$ and $\mu(\{0\}) = 1 - \sum_{j=1}^r f'_j$. Let $\mathcal{F}$ be the σ-field generated by μ. By Lemma 38, the random variables $Z_{2r+1}, Z_{2r+2}, \dots$ are conditionally independent given $\mathcal{F}$, and the conditional distribution of each Z_i given $\mathcal{F}$ is μ. Therefore,

$$P(R_b\Theta'_r = \theta | \mathcal{F}) = (f'_1)^{k_1-2} \dots (f'_r)^{k_r-2}. \tag{59}$$

Taking expectations of both sides of (59), we get

$$P(R_b\Theta'_r = \theta) = \int_{\Delta_r} x_1^{k_1-2} \dots x_r^{k_r-2}\, Q_r(dx_1, \dots, dx_r). \tag{60}$$

¿From (9), (58), and (60), we have

$$\begin{aligned} \lambda_{b;k_1,\dots,k_r;0} &= \lambda_{2r;2,\dots,2;0} P(R_b\Theta'_r = \theta) = \lambda_{2r;2,\dots,2;0} \int_{\Delta_r} x_1^{k_1-2} \dots x_r^{k_r-2}\, Q_r(dx_1, \dots, dx_r) \\ &= \int_{\Delta_r} x_1^{k_1-2} \dots x_r^{k_r-2}\, F'_r(dx_1, \dots, dx_r) = \lambda'_{b;k_1,\dots,k_r;0}, \end{aligned}$$

which completes the proof. □

Remark 24 Recall that to prove Proposition 8, we used Proposition 1 to establish the fact that if there is a coalescent process satisfying conditions B1 and B2 of Theorem 2, then the collision rates must be determined by (6) for a unique sequence of measures $(F_r)_{r=1}^\infty$ satisfying conditions A1, A2, and A3′ of Proposition 8. Thus, the proof of Proposition 8 made use of the population model introduced in [13]. In the proof of Proposition 10, we started with a coalescent process Π_∞ satisfying B1 and B2. Without using the assumption that the collision rates were derived from $(F_r)_{r=1}^\infty$ by (6), we then defined a sequence of measures $(F'_r)_{r=1}^\infty$ satisfying A1, A2, and A3′ such that the collision rates of Π_∞ were given by (6), with F'_m in place of F_m on the right-hand side. The uniqueness of this sequence of measures, which is part of Proposition 1, follows from (9) and the fact that measures on Δ_r are uniquely determined by their moments. Consequently, we now have an alternative proof of Proposition 8 that does not refer to the population model.

Proof of Proposition 11. Let Π_∞ be a standard Ξ-coalescent, and denote the collision rates by $\lambda_{b;k_1,\ldots,k_r;s}$. Fix $r \in \mathbb{N}$. Define E_r as in Proposition 10. First, suppose $P(E_r) = 0$. Then $\lambda_{2r;2,\ldots,2;0} = 0$ and $F_r(\Delta_r) = 0$ by Proposition 10. We see from (11) that when $S = \Delta_r$, the right-hand side of (13) equals $\lambda_{2r;2,\ldots,2;0}$, which is zero. Also, the right-hand side of (13) is nonnegative and takes on its maximum value when $S = \Delta_r$, so it is zero for all S. Thus, (13) holds when $P(E_r) = 0$.

Now, suppose $P(E_r) > 0$. Assume that Π_∞ is derived from a Poisson point process $(e(t))_{t\geq 0}$ whose characteristic measure L is defined by (14). Define T_r, $\tilde{\Theta}_r$, Θ'_r, and $f'_{1,r}, \ldots, f'_{r,r}$ as in the proof of Proposition 10, and let Q_r be the distribution of $f'_{1,r}, \ldots, f'_{r,r}$. By the proof of Proposition 10, we have $F_r(S) = \lambda_{2r;2,\ldots,2;0} Q_r(S)$ for all measurable subsets S of Δ_r. For $\xi \in \mathbb{Z}^\infty$ and $k \in \mathbb{Z}$, define $N_k(\xi)$ as in (40), provided this limit exists. Define D_r as in (55) and U_r as in (57). For each measurable subset S of Δ_r, define

$$A_{S,r} = \{\xi \in D_r : (N_{\xi_2}(\xi), N_{\xi_4}(\xi), \ldots, N_{\xi_{2r}}(\xi)) \in S\}.$$

On E_r, if $f_{j,r}$ denotes the limiting relative frequency of the block of $\tilde{\Theta}_r$ containing $2j-1$ and $2j$, then $(f_{1,r}, \ldots, f_{r,r}) \in S$ if and only if $e(T_r) \in A_{S,r}$. Since Θ'_r is defined to have the conditional distribution of $\tilde{\Theta}_r$ given E_r, we have

$$Q_r(S) = P((f'_{1,r}, \ldots, f'_{r,r}) \in S) = P((f_{1,r}, \ldots, f_{r,r}) \in S | E_r) = P(e(T_r) \in A_{S,r} | E_r).$$

Note that $T_r = U_r$ on E_r. By Lemma 15, $\Pi_\infty(U_r-)$ and $e(U_r)$ are independent. Also, recall from the proof of Proposition 10 that E_r is the event that $1, \ldots, 2r$ are in distinct blocks of $\Pi_\infty(U_r-)$. Using these facts and part (d) of Lemma 41 of appendix B, we get

$$Q_r(S) = P(e(U_r) \in A_{S,r} | E_r) = P(e(U_r) \in A_{S,r}) = \frac{L(A_{S,r})}{L(D_r)}.$$

Since $L(D_r) = \lambda_{2r;2,\ldots,2;0}$, which is nonzero when $P(E_r) > 0$, it follows that $F_r(S) = L(A_{S,r})$. By the strong law of large numbers,

$$P_x(A_{S,r}) = \sum_{i_1 \neq \ldots \neq i_r} x_{i_1}^2 \ldots x_{i_r}^2 1_{\{(x_{i_1}, \ldots, x_{i_r}) \in S\}}.$$

Also, $z_{ij} \in A_{S,r}$ if and only if $\{i, j\} = \{1, 2\}$, $r = 1$, and $(0, 0, \ldots) \in S$. Thus, by (14), $L(A_{S,r})$ equals the right-hand side of (13), which completes the proof. $\square$

5 Further properties of the Ξ-coalescent

In this section, we establish some properties of the Ξ-coalescent, most of which are straightforward extensions of properties that have been proved in [16] or [19] for the Λ-coalescent.

5.1 Regularity Properties of the Ξ-coalescent

In this subsection, we prove some regularity properties of the Ξ-coalescent. The analogous results for the Λ-coalescent are given as part of Theorem 1 in [16]. A consequence of the Feller property proved below is that the Ξ-coalescent satisfies the strong Markov property, a fact we will use later.

Recall from the introduction that we can identify $\mathcal{P}_\infty$ with the product space $\mathcal{P}_1 \times \mathcal{P}_2 \times \dots$. Since each $\mathcal{P}_n$ is finite, the associated product topology is compact and metrizable and has a countable basis. Also, one can check that this topology is induced by the metric $d(\sigma, \tau) = 2^{-n}$, where $n = \inf\{m \geq 1 : R_m\sigma \neq R_m\tau\}$.

Proposition 25 *Let $P^{\Xi,\pi}$ denote the law of a Ξ-coalescent Π_∞ started with $\Pi_\infty(0) = \pi$. Then the collection of laws $(P^{\Xi,\pi}, \pi \in \mathcal{P}_\infty)$ defines a Feller process.*

Proof. Let $C(\mathcal{P}_\infty)$ be the set of continuous real-valued functions defined on $\mathcal{P}_\infty$ with the topology induced by the norm $\|f\| = \sup_{\sigma \in \mathcal{P}_\infty} |f(\sigma)|$. Let $\mathcal{A}_n$ consist of all real-valued functions f defined on $\mathcal{P}_\infty$ such that $f(\sigma) = f(\tau)$ whenever $R_n\sigma = R_n\tau$. Since, for $f \in \mathcal{A}_n$, the value of $f(\sigma)$ is determined by $R_n\sigma$, we can associate with each $f \in \mathcal{A}_n$ a function $f^{(n)}$ defined on $\mathcal{P}_n$ such that $f(\sigma) = f^{(n)}(R_n\sigma)$ for all $\sigma \in \mathcal{P}_\infty$. Let $\mathcal{A} = \bigcup_{n=1}^\infty \mathcal{A}_n$. Then $\mathcal{A}$ is a subalgebra of $C(\mathcal{P}_\infty)$ which contains constants and separates points, so by the Stone-Weierstrass Theorem, $\mathcal{A}$ is dense in $C(\mathcal{P}_\infty)$.

We now define a family of operators $(P_t, t \geq 0)$ on $C(\mathcal{P}_\infty)$. For all $n \in \mathbb{N}$, let $(P_t^n, t \geq 0)$ be the transition semigroup associated with the restriction to $\{1, \dots, n\}$ of a Ξ-coalescent. Then for $f \in \mathcal{A}_n$, let $P_t f(\sigma) = P_t^n f^{(n)}(R_n\sigma)$. Note that $P_t f$ is in $\mathcal{A}_n$ and thus is continuous. We must check this definition for consistency; that is, if $f \in \mathcal{A}_m$ and $m < n$, we must check that $P_t^n f^{(n)}(R_n\sigma) = P_t^m f^{(m)}(R_m\sigma)$ for all $\sigma \in \mathcal{P}_\infty$. For $\alpha \in \mathcal{P}_n$, let Π_n^α denote the restriction to $\{1, \dots, n\}$ of a Ξ-coalescent Π_∞^α satisfying $R_n\Pi_\infty^\alpha(0) = \alpha$. For $\alpha, \beta \in \mathcal{P}_n$, let $p_t^n(\alpha, \beta) = P(\Pi_n^\alpha(t) = \beta)$. Property B2 of Theorem 2 implies that for $\alpha \in \mathcal{P}_n$ and $\tau \in \mathcal{P}_m$, we have

$$p_t^m(R_m\alpha, \tau) = \sum_{\{\beta \in \mathcal{P}_n : R_m\beta = \tau\}} p_t^n(\alpha, \beta).$$

Therefore, for all $f \in \mathcal{A}_m$ and $\sigma \in \mathcal{P}_\infty$, we have

$$\begin{aligned}
P_t^n f^{(n)}(R_n\sigma) &= \sum_{\alpha \in \mathcal{P}_n} p_t^n(R_n\sigma, \alpha) f^{(n)}(\alpha) = \sum_{\tau \in \mathcal{P}_m} \sum_{\{\beta \in \mathcal{P}_n : R_m\beta = \tau\}} p_t^n(R_n\sigma, \beta) f^{(n)}(\beta) \\
&= \sum_{\tau \in \mathcal{P}_m} f^{(m)}(\tau) \Bigg(\sum_{\{\beta \in \mathcal{P}_n : R_m\beta = \tau\}} p_t^n(R_n\sigma, \beta) \Bigg) \\
&= \sum_{\tau \in \mathcal{P}_m} p_t^m(R_m\sigma, \tau) f^{(m)}(\tau) = P_t^m f^{(m)}(R_m\sigma),
\end{aligned}$$

so $P_t f$ is consistently defined for all $f \in \mathcal{A}$. Now let $f \in C(\mathcal{P}_\infty)$ be arbitrary, and let $(f_m)_{m=1}^\infty$ be a sequence of functions in $\mathcal{A}$ converging uniformly to f. Note that $\|P_t f_m - P_t f_n\| \leq \|f_m - f_n\| \to 0$ as $m, n \to \infty$, so the sequence $(P_t f_m)_{m=1}^\infty$ converges uniformly to some continuous function that we define to be $P_t f$.

We claim that $(P_t, t \geq 0)$ is a Feller semigroup on $C(\mathcal{P}_\infty)$. Clearly P_0 is the identity operator. Since each $(P_t^n, t \geq 0)$ is the transition semigroup of a Markov process, we have $\|P_t^n\| \leq 1$ for all t and n, which implies $\|P_t\| \leq 1$ for all t. Fix $f \in C(\mathcal{P}_\infty)$ and let $(f_m)_{m=1}^\infty$ be a sequence in $\mathcal{A}$ converging to f. Then $P_{t+s}f = \lim_{m\to\infty} P_{t+s} f_m = \lim_{m\to\infty} P_t P_s f_m = P_t P_s f$, so $(P_t, t \geq 0)$

is a semigroup. Let $\epsilon > 0$, and fix m such that $\|f_m - f\| < \epsilon/3$. Let n be an integer such that $f_m \in \mathcal{A}_n$. There exists a constant $\lambda_n < \infty$ such that the total rate of all collisions for the restriction to $\{1, \dots, n\}$ of any Ξ-coalescent is at most λ_n. Thus, if $\|f_m\| > 0$, we can choose $\delta > 0$ such that $P(\Pi_n^\alpha(t) \neq \Pi_n^\alpha(0)) < \epsilon/(6\|f_m\|)$ for all $t < \delta$ and $\alpha \in \mathcal{P}_n$. Then, for all $t < \delta$ and $\sigma \in \mathcal{P}_\infty$, we have

$$|P_t f_m(\sigma) - f_m(\sigma)| = |P_t^n f_m^{(n)}(R_n\sigma) - f_m^{(n)}(R_n\sigma)| \leq 2\|f_m^{(n)}\| P(\Pi_n^{R_n\sigma}(t) \neq \Pi_n^{R_n\sigma}(0)) < \epsilon/3.$$

Hence,

$$\|P_t f - f\| \leq \|P_t f - P_t f_m\| + \|P_t f_m - f_m\| + \|f_m - f\| < \epsilon.$$

Therefore, $\lim_{t\downarrow 0} \|P_t f - f\| = 0$. It follows (see Definition 2.1 in chapter III of [17]) that $(P_t, t \geq 0)$ is a Feller semigroup on $C(\mathcal{P}_\infty)$. Therefore (see Theorem 1.5 and Proposition 2.2 in chapter III of [17]), there exists a $\mathcal{P}_\infty$-valued Feller process starting from the partition of $\mathbb{N}$ into singletons with $(P_t, t \geq 0)$ as its transition semigroup. Since $P_t f(\sigma) = P_t^n f^{(n)}(R_n\sigma)$ for all $f \in \mathcal{A}_n$ and $\sigma \in \mathcal{P}_\infty$, this Feller process has the property that its restriction to $\{1, \dots, n\}$ has the same law as Π_n. It follows that $(P_t, t \geq 0)$ is the transition semigroup for the Ξ-coalescent, so the collection of laws $(P^{\Xi,\pi}, \pi \in \mathcal{P}_\infty)$ defines a Feller process, as claimed. □

We now work towards proving that the law $P^{\Xi,\pi}$ depends continuously on the measures Ξ and π. To formulate this result precisely, we need to define a topology on Δ. We will use the weakest topology making all of the coordinate functions $x \mapsto x_i$ continuous; this topology is discussed in section 3 of [9]. With this topology, Δ is compact and metrizable, and a sequence $(x^{(n)})_{n=1}^\infty$ in Δ converges to x if and only if $\lim_{n\to\infty} x_i^{(n)} = x_i$ for all $i \in \mathbb{N}$.

Lemma 26 *For all $r \geq 1$ and $k_1, \dots, k_r \geq 2$, define the function $g_{k_1,\dots,k_r} : \Delta \to \mathbb{R}$ by*

$$g_{k_1,\dots,k_r}(x) = \sum_{i_1 \neq \dots \neq i_r} x_{i_1}^{k_1} \dots x_{i_r}^{k_r} \Big/ \sum_{j=1}^\infty x_j^2$$

for $x \neq (0, 0, \dots)$, and $g_{k_1,\dots,k_r}((0, 0, \dots)) = 1_{\{r=1,k_1=2\}}$. Then $g_{k_1,\dots,k_r}$ is bounded and continuous.

Proof. For all $x \in \Delta$, we have

$$\sum_{i_1 \neq \dots \neq i_r} x_{i_1}^{k_1} \dots x_{i_r}^{k_r} \leq \sum_{i_1,\dots,i_r=1}^\infty x_{i_1}^{k_1} \dots x_{i_r}^{k_r} = \prod_{j=1}^r \left(\sum_{i_j=1}^\infty x_{i_j}^{k_j} \right). \tag{61}$$

Since $\sum_{i=1}^\infty x_i \leq 1$ by the definition of Δ, we have $\sum_{i=1}^\infty x_i^k \leq 1$ for all $k \in \mathbb{N}$, so

$$\sum_{i_1 \neq \dots \neq i_r} x_{i_1}^{k_1} \dots x_{i_r}^{k_r} \leq \sum_{i_1=1}^\infty x_{i_1}^{k_1} \leq \sum_{i=1}^\infty x_i^2.$$

Thus, $g_{k_1,\dots,k_r}(x) \leq 1$ for all $x \in \Delta$, which means $g_{k_1,\dots,k_r}$ is bounded.

Next, we show that $g_{k_1,\ldots,k_r}$ is continuous. For $n = 1, 2, \ldots, \infty$, define

$$f^{(n)}_{k_1,\ldots,k_r}(x) = \sum_{i_1 \neq \ldots \neq i_r = 1}^{n} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r}$$

for all $x \in \Delta$. Since $x_i \leq 1/i$ for all $x \in \Delta$, we have

$$\begin{aligned} |f^{(\infty)}_{k_1,\ldots,k_r}(x) - f^{(n)}_{k_1,\ldots,k_r}(x)| &\leq \sum_{j=1}^{r} \sum_{i_j = n+1}^{\infty} \sum_{i_1 \neq \ldots \neq i_{j-1} \neq i_{j+1} \neq \ldots \neq i_r} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r} \leq r \sum_{i_1 = n+1}^{\infty} \sum_{i_2 \neq \ldots \neq i_r} x_{i_1}^2 \ldots x_{i_r}^2 \\ &\leq r \Bigg(\sum_{i_1 = n+1}^{\infty} x_{i_1}^2 \Bigg) \prod_{j=2}^{r} \Bigg(\sum_{i_j = 1}^{\infty} x_{i_j}^2 \Bigg) \leq r \sum_{i_1 = n+1}^{\infty} x_{i_1}^2 \leq r \sum_{i = n+1}^{\infty} \frac{1}{i^2} \leq \frac{r}{n} \end{aligned}$$

for all $x \in \Delta$. Thus, the sequence of functions $(f^{(n)}_{k_1,\ldots,k_r})_{n=1}^{\infty}$ converges uniformly to $f^{(\infty)}_{k_1,\ldots,k_r}$. Since $f^{(n)}_{k_1,\ldots,k_r}$ is continuous for all $r \geq 1$, $k_1, \ldots, k_r \geq 2$, and $n < \infty$, it follows that $f^{(\infty)}_{k_1,\ldots,k_r}$ is continuous for all $r \geq 1$ and $k_1, \ldots, k_r \geq 2$. Thus, $g_{k_1,\ldots,k_r}$ is a ratio of two continuous functions and is therefore continuous wherever the denominator is nonzero.

It remains only to show that $g_{k_1,\ldots,k_r}$ is continuous at zero. If $r = 1$ and $k_1 = 2$, then $g_{k_1,\ldots,k_r}(x) = 1$ for all $x \in \Delta$, so $g_{k_1,\ldots,k_r}$ is continuous at zero. Let $(x^{(n)})_{n=1}^{\infty}$ be a sequence in Δ converging to zero. If $r \geq 2$, then (61) implies

$$g_{k_1,\ldots,k_r}(x^{(n)}) \leq \prod_{j=1}^{r} \Bigg(\sum_{i_j=1}^{\infty} (x_{i_j}^{(n)})^{k_j} \Bigg) \Bigg/ \sum_{j=1}^{\infty} (x_j^{(n)})^2 \leq \prod_{j=1}^{2} \Bigg(\sum_{i_j=1}^{\infty} (x_{i_j}^{(n)})^2 \Bigg) \Bigg/ \sum_{j=1}^{\infty} (x_j^{(n)})^2 = \sum_{i=1}^{\infty} (x_i^{(n)})^2.$$

Given $\epsilon > 0$, choose $N > 2/\epsilon$, and choose M such that for $n > M$, we have $x_i^{(n)} < (\epsilon/2N)^{1/2}$ for all $i = 1, \ldots, N$. For $n > M$, we have

$$\sum_{i=1}^{\infty} (x_i^{(n)})^2 \leq \sum_{i=1}^{N} \frac{\epsilon}{2N} + \sum_{i=N+1}^{\infty} \frac{1}{i^2} \leq \frac{\epsilon}{2} + \frac{1}{N} < \epsilon.$$

Thus, $\lim_{n \to \infty} g_{k_1,\ldots,k_r}(x^{(n)}) = 0$. If $r = 1$ and $k_1 \geq 3$, then

$$g_{k_1,\ldots,k_r}(x^{(n)}) = \sum_{i=1}^{\infty} (x_i^{(n)})^{k_1} \Bigg/ \sum_{j=1}^{\infty} (x_j^{(n)})^2 \leq (x_1^{(n)})^{k_1 - 2} \sum_{i=1}^{n} (x_i^{(n)})^2 \Bigg/ \sum_{j=1}^{n} (x_j^{(n)})^2 = (x_1^{(n)})^{k_1 - 2},$$

which approaches 0 as $n \to \infty$. Hence $g_{k_1,\ldots,k_r}$ is continuous at zero. □

Proposition 27 *Let $P^{\Xi,\pi}$ denote the law of a Ξ-coalescent Π_∞ started with $\Pi_\infty(0) = \pi$. Equip the space of finite measures on Δ with the topology of weak convergence. Let $D_{\mathcal{P}_\infty}[0,\infty)$ denote the set of càdlàg $\mathcal{P}_\infty$-valued paths with the Skorohod topology, and give the space of probability measures on $D_{\mathcal{P}_\infty}[0,\infty)$ the topology of weak convergence. Then, the map $(\Xi, \pi) \mapsto P^{\Xi,\pi}$ is continuous.*

Proof. Let $(\Xi_m)_{m=1}^\infty$ be a sequence of finite measures on Δ converging weakly to Ξ, and let $(\pi_m)_{m=1}^\infty$ be a sequence in $\mathcal{P}_\infty$ converging to π. Since Δ is compact and metrizable, the space of finite measures on Δ with the topology of weak convergence is metrizable, as shown in section 5 of chapter VIII of [4]. Also, $\mathcal{P}_\infty$ is metrizable. Therefore, it suffices to show that $(P^{\Xi_m,\pi_m})_{m=1}^\infty$ converges weakly to $P^{\Xi,\pi}$.

For all $r \geq 1$ and $k_1, \dots, k_r \geq 2$, define $g_{k_1,\dots,k_r}$ as in Lemma 26. By (11),

$$\lambda_{b;k_1,\dots,k_r;0} = \int_\Delta g_{k_1,\dots,k_r}(x)\, \Xi(dx)$$

for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, and $b = \sum_{j=1}^r k_j$. Therefore, by Lemma 26 and the definition of weak convergence, if the rate of a $(b; k_1, \dots, k_r; s)$-collision for a standard Ξ_m-coalescent is denoted by $\lambda^m_{b;k_1,\dots,k_r;s}$, then

$$\lim_{m\to\infty} \lambda^m_{b;k_1,\dots,k_r;0} = \lambda_{b;k_1,\dots,k_r;0} \tag{62}$$

for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, and $b = \sum_{j=1}^r k_j$. It follows from (43) by induction on s that all collision rates $\lambda_{b;k_1,\dots,k_r;s}$ can be obtained by taking finite sums and differences of collision rates with $s = 0$. Hence (62) implies that

$$\lim_{m\to\infty} \lambda^m_{b;k_1,\dots,k_r;s} = \lambda_{b;k_1,\dots,k_r;s} \tag{63}$$

for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$.

Let Π^m_∞ be a Ξ_m-coalescent such that $\Pi^m_\infty(0) = \pi_m$. Let $\Pi^m_n = R_n \Pi^m_\infty$ for all m and n, and let $\Pi_n = R_n \Pi_\infty$. Since $\mathcal{P}_\infty$ is complete, separable, and compact, it follows from Theorem 7.2 in chapter 3 of [6] that to show that $(P^{\Xi_m,\pi_m})_{m=1}^\infty$ is relatively compact, it suffices to show that for all $\epsilon > 0$ and $T > 0$, there exists $\delta > 0$ such that

$$\sup_m P(w'(\Pi^m_\infty, \delta, T) \geq \epsilon) \leq \epsilon, \tag{64}$$

where

$$w'(\Pi^m_\infty, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} d(\Pi^m_\infty(s), \Pi^m_\infty(t))$$

and $\{t_i\}$ ranges over partitions of the form $0 = t_0 < t_1 \ldots < t_{n-1} < T \leq t_n$ with $t_i - t_{i-1} > \delta$ for $i = 1, \dots, n$. Choose n such that $2^{-n} < \epsilon$. Then $w'(\Pi^m_\infty, \delta, T) < \epsilon$ as long as no two jumps of Π^m_n between times 0 and T occur in any time interval of length δ. Condition (63) implies that the total rate of all jumps of Π^m_n is bounded uniformly in m, so for sufficiently small δ, we have $P(w'(\Pi^m_\infty, \delta, T) < \epsilon) > 1 - \epsilon$ for all m, which implies (64). Hence, $(\mathcal{P}^{\Xi_m,\pi_m})_{m=1}^\infty$ is relatively compact.

By Theorem 7.8 in chapter 3 of [6], to show that $(P^{\Xi_m,\pi_m})_{m=1}^\infty$ converges weakly to $P^{\Xi,\pi}$, it now suffices to show that if $0 \leq t_1 < \ldots < t_k < \infty$, then $(\Pi^m_\infty(t_1), \dots, \Pi^m_\infty(t_k))$ converges weakly to $(\Pi_\infty(t_1), \dots, \Pi_\infty(t_k))$ as $m \to \infty$. By the Portmanteau Theorem, which is Theorem 3.1 in chapter 3 of [6], it suffices to show that

$$\lim_{m\to\infty} P((\Pi^m_\infty(t_1), \dots, \Pi^m_\infty(t_k)) \in G) = P((\Pi_\infty(t_1), \dots, \Pi_\infty(t_k)) \in G) \tag{65}$$

for all open subsets G of the k-fold product $\mathcal{P}^k_\infty = \mathcal{P}_\infty \times \ldots \times \mathcal{P}_\infty$. For $\sigma = (\sigma_1, \dots, \sigma_k) \in \mathcal{P}^k_\infty$ and $\tau = (\tau_1, \dots, \tau_k) \in \mathcal{P}^k_\infty$, define $R_n\sigma = (R_n\sigma_1, \dots, R_n\sigma_k)$ and $\rho(\sigma, \tau) = \max_{1\leq i\leq k} d(\sigma_i, \tau_i)$.

Then, ρ is a metric on $\mathcal{P}_\infty^k$ which induces the product topology. Thus, all open subsets of $\mathcal{P}_\infty^k$ are unions of sets of the form

$$B(\sigma, \epsilon) = \{\tau \in \mathcal{P}_\infty^k : \rho(\sigma, \tau) < \epsilon\} = \{\tau \in \mathcal{P}_\infty^k : R_n\tau = R_n\sigma\},$$

where $n = \sup\{j \in \mathbb{N} : 2^{-j} \geq \epsilon\}$. Thus, every open subset of $\mathcal{P}_\infty^k$ can be written in the form

$$\bigcup_{j=1}^{\infty} \{\tau \in \mathcal{P}_\infty^k : R_{n_j}\tau = \sigma_j\}, \tag{66}$$

where the n_j are integers and $\sigma_j \in \mathcal{P}_{n_j}^k$ for all $j \in \mathbb{N}$. If $M < \infty$, then there exists a subset S of $\mathcal{P}_N^k$, where $N = \max_{1 \leq j \leq M} n_j$, such that

$$\bigcup_{j=1}^{M} \{\tau \in \mathcal{P}_\infty^k : R_{n_j}\tau = \sigma_j\} = \{\tau \in \mathcal{P}_\infty^k : R_N\tau \in S\}. \tag{67}$$

The sets in (67) increase to the set in (66) as $M \to \infty$. Therefore, to show (65), it suffices to show that for all $N < \infty$ and all $S \subset \mathcal{P}_N^k$, we have

$$\lim_{m \to \infty} P((\Pi_N^m(t_1), \ldots, \Pi_N^m(t_k)) \in S) = P((\Pi_N(t_1), \ldots, \Pi_N(t_k)) \in S). \tag{68}$$

Equation (68) follows from (63) and the fact that for sufficiently large m, we have the equality $\Pi_N^m(0) = R_N\pi_m = R_N\pi = \Pi_N(0)$. Hence, $(P^{\Xi_m, \pi_m})_{m=1}^\infty$ converges weakly to $P^{\Xi,\pi}$, which completes the proof. $\square$

5.2 Some Formulas

For $b \geq 2$, let λ_b denote the total rate of all collisions when the Ξ-coalescent has b blocks. Let $N(b; k_1, \ldots, k_r; s)$ be the number of possible $(b; k_1, \ldots, k_r; s)$-collisions, which was given in equation (3) of the introduction. We have

$$\lambda_b = \sum_{r=1}^{\lfloor b/2 \rfloor} \sum_{\{k_1, \ldots, k_r\}} N(b; k_1, \ldots, k_r; s)\lambda_{b;k_1,\ldots,k_r;s}, \tag{69}$$

where $s = b - \sum_{j=1}^r k_j$ and the inner sum is over multisets $\{k_1, \ldots, k_r\}$ because we do not have a separate term for each possible ordering of $k_1, \ldots, k_r$. We can obtain another formula for λ_b using the Poisson process construction of section 3. Define A_b as in (15). Then, we have $A_b^c = \{\xi \in \mathbb{Z}^\infty : \xi_1, \ldots, \xi_b \text{ are distinct}\}$. By considering the Poisson process construction and applying part (a) of Lemma 41 of appendix B, we see that $\lambda_b = L(A_b)$. Note that $\xi \in A_b^c$ if and only if there exist $l \in \{0, \ldots, b\}$ and distinct positive integers $i_1, \ldots, i_l$ such that for some $m_1 < \ldots < m_l \leq b$, we have $\xi_{m_j} = i_j$ for $j \in \{1, \ldots, l\}$ and $\xi_m < 0$ for all $m \leq b$ such that $m \notin \{m_1, \ldots, m_l\}$. Summing over the possible values of l and $i_1, \ldots, i_l$, and counting the possible values of $m_1, \ldots, m_l$, we get, for all $x \in \Delta$,

$$P_x(A_b) = 1 - P_x(A_b^c) = 1 - \sum_{l=0}^{b} \sum_{i_1 \neq \ldots \neq i_l} \binom{b}{l} x_{i_1} \ldots x_{i_l} (1 - \sum_{j=1}^{\infty} x_j)^{b-l}.$$

Also, $z_{ij} \in A_b$ if and only if $i \leq b$, $j \leq b$, and $i \neq j$, so from (14), we get

$$\lambda_b = \int_\Delta \left(1 - \sum_{l=0}^{b} \sum_{i_1 \neq \ldots \neq i_l} \binom{b}{l} x_{i_1} \ldots x_{i_l} (1 - \sum_{j=1}^{\infty} x_j)^{b-l}\right) \Big/ \sum_{j=1}^{\infty} x_j^2 \, \Xi_0(dx) + a \binom{b}{2}. \tag{70}$$

Note that for the Λ-coalescent, when Ξ is concentrated on $\{x \in \Delta : x_i = 0 \text{ for all } i \geq 2\}$, only the $l = 0$ and $l = 1$ terms in the integrand of (70) are nonzero. Therefore, (70) reduces to

$$\lambda_b = \int_\Delta (1 - (1 - x_1)^b - b x_1 (1 - x_1)^{b-1})/x_1^2 \, \Xi_0(dx) + a \binom{b}{2},$$

which agrees with equation (6) of [16] because

$$\lim_{x_1 \to 0} \frac{(1 - (1 - x_1)^b - b x_1 (1 - x_1)^{b-1})}{x_1^2} = \binom{b}{2}.$$

Equation (15) of [13] gives another formula for λ_b in terms of the sequence of measures $(F_r)_{r=1}^{\infty}$ associated with Ξ as in Proposition 11.

Let γ_b denote the total rate at which the number of blocks is decreasing when the coalescent has b blocks. Each $(b, k_1, \ldots, k_r, s)$-collision decreases the number of blocks by $b - r - s$, so for $b \geq 2$ we have

$$\gamma_b = \sum_{r=1}^{\lfloor b/2 \rfloor} \sum_{\{k_1, \ldots, k_r\}} (b - r - s) N(b; k_1, \ldots, k_r; s) \lambda_{b; k_1, \ldots, k_r; s}. \tag{71}$$

We record one simple lemma regarding the γ_b, which is proved in [19] for the Λ-coalescent by a direct calculation.

Lemma 28 *The sequence $(\gamma_b)_{b=2}^{\infty}$ is increasing.*

Proof. Let Π_∞ be a standard Ξ-coalescent, and let $\Pi_n = R_n \Pi_\infty$ for all $n \in \mathbb{N}$. Fix $m < n$. Then γ_n is the initial rate at which the number of blocks of Π_n is decreasing, and γ_m is the initial rate at which the number of blocks of Π_m is decreasing. For all $t \in \mathbb{R}$, we have

$$\#\Pi_n(t-) - \#\Pi_n(t) \geq \#\Pi_m(t-) - \#\Pi_m(t).$$

That is, whenever Π_m undergoes a collision, Π_n undergoes a collision at the same time, and the collision reduces the number of blocks of Π_n by at least as much as it reduces the number of blocks of Π_m. It follows that the rate at which the number of blocks of Π_n is decreasing is always greater than or equal to the rate at which the number of blocks of Π_m is decreasing. Hence $\gamma_n \geq \gamma_m$, which proves the lemma. $\square$

5.3 Jump-hold coalescents

Pitman has shown in subsection 2.1 of [16] that a standard Λ-coalescent is a Markov process of jump-hold type with bounded transition rates and step function paths if and only if

$$\int_0^1 x^{-2} \, \Lambda(dx) < \infty. \tag{72}$$

His proof is based on the observation that a standard Λ-coalescent is a Markov process of jump-hold type with bounded transition rates and step function paths if and only if the sequence $(\lambda_b)_{b=2}^{\infty}$ is bounded. He then uses an explicit formula for λ_b to show that $(\lambda_b)_{b=2}^{\infty}$ is bounded if and only if (72) holds. Here, we obtain the following result for the Ξ-coalescent using the Poisson process construction.

Proposition 29 *Let Ξ be a finite measure on the infinite simplex Δ. The standard Ξ-coalescent is a jump-hold Markov process with bounded transition rates and step function paths if and only if Ξ has no atom at zero and*

$$\int_{\Delta} 1/\sum_{j=1}^{\infty} x_j^2 \, \Xi(dx) < \infty. \tag{73}$$

Proof. Write $\Xi = \Xi_0 + a\delta_0$, where Ξ_0 has no atom at zero. As for the standard Λ-coalescent, the standard Ξ-coalescent is a jump-hold Markov process with bounded transition rates and step-function paths if and only if $(\lambda_b)_{b=2}^{\infty}$ is bounded. As observed in subsection 5.2, we have $\lambda_b = L(A_b)$, where L is defined by (14) and A_b is defined by (15). The sets A_b increase to the set A_∞ defined in (19). Therefore, $\lim_{b\to\infty} L(A_b) = L(A_\infty)$, so $(\lambda_b)_{b=2}^{\infty}$ is bounded if and only if $L(A_\infty) < \infty$.

We have $P_x(A_\infty) = 1$ for all $x \in \Delta$ such that $x \neq (0, 0, \ldots)$, and $z_{ij} \in A_\infty$ for all $i, j \in \mathbb{N}$ with $i \neq j$. Therefore,

$$L(A_\infty) = \int_{\Delta} 1/\sum_{j=1}^{\infty} x_j^2 \, \Xi_0(dx) + a \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} 1, \tag{74}$$

which is finite if and only if Ξ has no atom at zero and (73) holds. □

5.4 Proper Frequencies

Here, the results for the standard Λ-coalescent discussed in [16] carry over to the standard Ξ-coalescent without much difficulty. If Π_∞ is a standard Ξ-coalescent, then since Π_∞ is an exchangeable coalescent (see Definition 19), $\Pi_\infty(t)$ is an exchangeable random partition of $\mathbb{N}$ for all $t > 0$. Let $B_1(t), B_2(t), \ldots$ be the blocks of $\Pi_\infty(t)$, ordered by their smallest elements, where $B_j(t) = \emptyset$ if $\Pi_\infty(t)$ has fewer than j blocks. By Lemma 40 in appendix A, almost surely

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{i \in B_j(t)\}}$$

exists for all $j \in \mathbb{N}$, and we denote this limit by $f_j(t)$. We say that $\Pi_\infty(t)$ has *proper frequencies* if $\sum_{j=1}^{\infty} f_j(t) = 1$ a.s. The following result is the analog of Lemma 25 of [16].

Proposition 30 *Let $\Xi = \Xi_0 + a\delta_0$ be a finite measure on the infinite simplex Δ, where Ξ_0 has no atom at zero. Let Π_∞ be a standard Ξ-coalescent, and fix $t > 0$. Then $\Pi_\infty(t)$ has proper frequencies if and only if $a > 0$ or*

$$\int_{\Delta} \left(\sum_{j=1}^{\infty} x_j \Big/ \sum_{j=1}^{\infty} x_j^2 \right) \Xi_0(dx) = \infty. \tag{75}$$

Proof. As noted in the proof of Lemma 25 of [16], Lemma 40 of appendix A implies that $\Pi_\infty(t)$ has proper frequencies if and only if the singleton set $\{1\}$ is almost surely not a block of $\Pi_\infty(t)$. Assume that Π_∞ is derived from the Poisson point process $(e(t))_{t\geq 0}$. Let

$$A^1_\infty = \{\xi \in \mathbb{Z}^\infty : \xi_1 = \xi_i \text{ for some } i > 1\},$$

and let $T_1 = \inf\{t : e(t) \in A^1_\infty\}$. The construction of Π_∞ implies that $\{1\}$ is a block of $\Pi_\infty(t)$ if and only if $T_1 > t$. By Lemma 41 in appendix B, we have $P(T_1 > t) = 0$ if and only if $L(A^1_\infty) = \infty$. Note that $P_x(A^1_\infty) = \sum_{j=1}^\infty x_j$ for all $x \in \Delta$. If $i < j$, then $z_{ij} \in A^1_\infty$ if and only if $i = 1$. Thus,

$$L(A^1_\infty) = \int_\Delta \left(\sum_{j=1}^\infty x_j \bigg/ \sum_{j=1}^\infty x_j^2 \right) \Xi_0(dx) + a \sum_{j=2}^\infty 1,$$

which is infinite if and only if $a > 0$ or (75) holds. □

5.5 Coming down from infinity

Let Π_∞ be a standard Ξ-coalescent. By definition, $\#\Pi_\infty(0) = \infty$. We say the Ξ-coalescent *comes down from infinity* if $\#\Pi_\infty(t) < \infty$ a.s. for all $t > 0$. We say the Ξ-coalescent *stays infinite* if $\#\Pi_\infty(t) = \infty$ a.s. for all $t > 0$. The problem of whether the Λ-coalescent comes down from infinity has been studied thoroughly. Results of Bolthausen and Sznitman in [3] imply that the Λ-coalescent stays infinite when Λ is the uniform distribution on $[0, 1]$, and Sagitov shows in [18] that if $\Lambda(dx) = (1-\alpha)x^{-\alpha}dx$ for $0 < \alpha < 1$, then the Λ-coalescent comes down from infinity. Pitman shows in [16] that the Λ-coalescent comes down from infinity if Λ has an atom at zero and stays infinite if $\int_0^1 x^{-1}\,\Lambda(dx) < \infty$. Pitman also shows in Proposition 23 of [16] that if Λ has no atom at 1, then the Λ-coalescent either comes down from infinity or stays infinite. It is then shown in [19] that the Λ-coalescent comes down from infinity if and only if $\sum_{b=2}^\infty \gamma_b^{-1} < \infty$. This result does not fully generalize to the Ξ-coalescent, but we provide some partial results in this subsection. The problem of finding a necessary and sufficient condition for a Ξ-coalescent to come down from infinity remains open.

Let $\Delta_f = \{x \in \Delta : x_1 + \ldots + x_n = 1 \text{ for some } n\}$. Write Ξ as $\Xi_1 + \Xi_2$, where Ξ_1 is the restriction of Ξ to Δ_f and $\Xi_2 = \Xi - \Xi_1$. Define L_1 and L_2 by (14), replacing Ξ by Ξ_1 and Ξ_2 respectively. Let X_1 and X_2 be independent Poisson random measures on $[0, \infty) \times \mathbb{Z}^\infty$ with intensity measures $\lambda \times L_1$ and $\lambda \times L_2$ respectively, where λ denotes Lebesgue measure. Let $X = X_1 + X_2$, which is a Poisson random measure with intensity $\lambda \times L$, where $L = L_1 + L_2$. Define Poisson point processes $(e^{(1)}(t))_{t\geq 0}$, $(e^{(2)}(t))_{t\geq 0}$, and $(e(t))_{t\geq 0}$ from X_1, X_2, and X respectively, as described in appendix B. Let $\Pi^{(1)}_\infty$ be a standard Ξ_1-coalescent derived from $(e^{(1)}(t))_{t\geq 0}$, and let $\Pi^{(2)}_\infty$ be a standard Ξ_2-coalescent derived from $(e^{(2)}(t))_{t\geq 0}$. We may assume without loss of generality that Π_∞ is derived from $(e(t))_{t\geq 0}$.

Let $A_f = \{\xi \in \mathbb{Z}^\infty : \{\xi_1, \xi_2, \ldots\} \text{ is finite}\}$, and let $T_f = \inf\{t : e(t) \in A_f\}$. For $x \in \Delta_f$, we have $P_x(A_f) = 1$, which means $P_x(A_f^c) = 0$. Since Ξ_1 has no atom at zero, it follows from (14) that

$$L_1(A_f^c) = \int_\Delta P_x(A_f^c) \bigg/ \sum_{j=1}^\infty x_j^2 \, \Xi_1(dx) = 0$$

and

$$L_1(A_f) = \int_\Delta 1/\sum_{j=1}^\infty x_j^2 \, \Xi_1(dx).$$

We consider the following three cases:

Case 1: Suppose $L_1(A_f) = \infty$. Then $L(A_f) = \infty$. By Lemma 41 in appendix B, $T_f = 0$ a.s. Since $\#\Pi_\infty(t) < \infty$ for all $t > T_f$ by properties of the Poisson process construction, the standard Ξ-coalescent comes down from infinity.

Case 2: Suppose $0 < L_1(A_f) < \infty$. Then, we have $0 < T_f < \infty$ a.s. by part (a) of Lemma 41. Since $\#\Pi_\infty(t) < \infty$ for all $t \geq T_f$, the standard Ξ-coalescent does not stay infinite. For all $t < T_f$, we have $e^{(1)}(t) = \delta$ and therefore $e(t) = e^{(2)}(t)$. Thus, $\Pi_\infty(t) = \Pi_\infty^{(2)}(t)$ for all $t < T_f$, so the standard Ξ-coalescent comes down from infinity if and only if the standard Ξ_2-coalescent comes down from infinity.

Case 3: Suppose $L_1(A_f) = 0$. Then, $L_1(\Delta) = 0$, so $e^{(1)}(t) = \delta$ for all t. Therefore, the standard Ξ-coalescent is the same as the standard Ξ_2-coalescent.

We have now reduced the problem to the case when $\Xi_1 = 0$. We begin our analysis of this case with the following straightforward generalization of Proposition 23 of [16].

Lemma 31 *Suppose $\Xi(\Delta_f) = 0$. Then either the standard Ξ-coalescent stays infinite or the standard Ξ-coalescent comes down from infinity.*

Proof. We essentially follow the proof of Proposition 23 of [16] but write out more details. Define $T = \inf\{t : \#\Pi_\infty(t) < \infty\}$. Let $p = P(T = 0)$. Suppose $p > 0$. By the Markov property of Π_∞, we have $p = P(T = t | \#\Pi_\infty(t) = \infty)$ for all $t > 0$. By considering $t = k\epsilon/n$ for $k = 1, \ldots, n$, we see that $P(T \leq \epsilon) \geq 1 - (1-p)^n$ for all $\epsilon > 0$ and $n \in \mathbb{N}$. Therefore, $P(T = 0) = 1$.

Suppose $p = 0$. We wish to show that this implies $P(T = \infty) = 1$. It suffices to show that $P(0 < T < \infty) = 0$. Note that T is a stopping time with respect to the completed natural filtration of Π_∞ (see Theorem 2.17 in chapter III of [17]). Therefore, if

$$P(\{0 < T < \infty\} \cap \{\#\Pi_\infty(T) = \infty\}) > 0,$$

then the strong Markov property of Π_∞, which follows from Proposition 25, implies that

$$P(\inf\{t : \#\Pi_\infty(T+t) < \infty\} = 0 | \{0 < T < \infty\} \cap \{\#\Pi_\infty(T) = \infty\}) = p = 0,$$

which contradicts the definition of T. Thus, $P(\{0 < T < \infty\} \cap \{\#\Pi_\infty(T) = \infty\}) = 0$. On the set $\{0 < T < \infty\} \cap \{\#\Pi_\infty(T) < \infty\} \cap \{\#\Pi_\infty(T-) = \infty\}$, the time T is a collision time. However, Proposition 6 implies that when $\Xi(\Delta_f) = 0$, almost surely no single collision takes Π_∞ down from an infinite number of blocks to a finite number of blocks. Thus,

$$P(\{0 < T < \infty\} \cap \{\#\Pi_\infty(T) < \infty\}) \cap \{\#\Pi_\infty(T-) = \infty\}) = 0.$$

It remains to show that $P(\{0 < T < \infty\} \cap \{\#\Pi_\infty(T-) < \infty\}) = 0$. On $\{\#\Pi_\infty(T-) < \infty\}$, let $n_1, \ldots, n_b$ be the smallest elements in the blocks of $\Pi_\infty(T-)$. Then let $T_0 = 0$, and for $m \geq 1$, let T_m be the time at which the smallest integer that is not in any of the blocks containing one of the integers $n_1, \ldots, n_b$ at time T_{m-1} merges with a block containing one of $n_1, \ldots, n_b$. Note that $T_0 < T_1 < \ldots < T$ a.s. However, for any fixed integers $s_1, \ldots, s_b$, the rate at which any particular block is colliding with one of the blocks containing $s_1, \ldots, s_b$ is at most $\lambda_{b+1} < \infty$. Since only countably many sets $\{s_1, \ldots, s_b\}$ are possible, it follows that $T_m \uparrow \infty$ a.s. as $m \to \infty$. Thus,

$$P(\{0 < T < \infty\} \cap \{\#\Pi_\infty(T-) < \infty\}) = 0,$$

which completes the proof. $\square$

Proposition 32 *Suppose $\Xi(\Delta_f) = 0$ and $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$. Then the Ξ-coalescent comes down from infinity.*

Proof. By Lemma 31, it suffices to show that $E[T_\infty] < \infty$, where $T_\infty = \inf\{t : \#\Pi_\infty(t) = 1\}$. Since $(\gamma_b)_{b=2}^{\infty}$ is increasing by Lemma 28, we can show that $E[T_\infty] < \infty$ by the same argument used to prove Lemma 6 of [19]. $\square$

For the Ξ-coalescent, we are only able to establish a converse to Proposition 32 when an additional condition is satisfied.

Proposition 33 *For all $\epsilon > 0$, let $\Delta^\epsilon = \{x \in \Delta : \sum_{i=1}^{\infty} x_i \leq 1 - \epsilon\}$. Suppose $\Xi(\Delta_f) = 0$ and $\sum_{b=2}^{\infty} \gamma_b^{-1} = \infty$. Also, suppose*

$$\int_{\Delta \setminus \Delta^\epsilon} 1/\sum_{j=1}^{\infty} x_j^2 \, \Xi(dx) < \infty \tag{76}$$

for some $\epsilon > 0$. Then, the Ξ-coalescent stays infinite.

Proof. Let Π_∞ be a standard Ξ-coalescent derived from a Poisson point process $(e(t))_{t \geq 0}$ with characteristic measure L defined by (14). Let $\Pi_n = R_n \Pi_\infty$. Let $T_n = \inf\{t : \#\Pi_n(t) = 1\}$. As observed for the Λ-coalescent in equation (31) of [16], we have

$$0 = T_1 < T_2 \leq T_3 \leq \ldots \uparrow T_\infty \leq \infty.$$

By the argument used to prove Proposition 5 of [19], it suffices to show that $\lim_{n \to \infty} E[T_n] = \infty$. Fix $\epsilon > 0$ such that (76) holds. Let Ξ_1 be the restriction of Ξ to Δ^ϵ, and let Ξ_2 be the restriction of Ξ to $\Delta \setminus \Delta_\epsilon$. By (76) and Proposition 29, the Ξ_2-coalescent is a jump-hold Markov process with bounded transition rates. Therefore, by the argument used to prove Lemma 8 of [19], it suffices to show that the Ξ_1-coalescent stays infinite. We will therefore assume for the remainder of the proof that $\Xi_2 = 0$.

We now follow the idea of the proof of Lemma 7 of [19]. Let $|S|$ denote the cardinality of a set S. Fix an integer $M > 1/\epsilon$. For positive integers b and l such that $b > M^l$, define $R_{b,l} = \{\xi \in \mathbb{Z}^\infty : |\{\xi_1, \ldots, \xi_b\}| \leq M^l\}$ and $S_{b,l} = \{\xi \in \mathbb{Z}^\infty : M^{l-1} + 1 \leq |\{\xi_1, \ldots, \xi_b\}| \leq M^l\}$. Let $g_{b,l}(\xi) = \inf\{i : |\{\xi_1, \ldots, \xi_i\}| + (b - i) = M^l\}$, so $g_{b,l}(\xi) \leq b$ if and only if $\xi \in R_{b,l}$. Let

$x \in \Delta^\epsilon$. Suppose $\xi = (\xi_1, \xi_2, \dots)$ has distribution P_x, where P_x is as defined in section 3. Then $\xi_1, \xi_2, \dots$ are independent and each is negative with probability at least ϵ. Also, no two negative ξ_i are the same. Fix $d \leq b$. On the event $\{g_{b,l}(\xi) = d\}$, we have $\xi \in S_{b,l}$ whenever at least $(M^{l-1}+1) - (M^l - (b-d))$ of the random variables $\xi_{d+1}, \dots, \xi_b$ are negative. Thus, $P(\xi \in S_{b,l} | g_{b,l}(\xi) = d)$ is greater than or equal to the probability that the sum of $b-d$ independent Bernoulli random variables with success probability ϵ is at least $(M^{l-1}+1) - (M^l - (b-d))$. This is greater than or equal to the probability that the sum of M^l independent Bernoulli random variables with success probability ϵ is at least $M^{l-1}+1$. Since $\epsilon > 1/M$, this probability approaches 1 as $l \to \infty$ and is therefore bounded below by some constant $C > 0$ which does not depend on b, d, or l. Therefore, $P(\xi \in S_{b,l} | g_{b,l}(\xi) = d) \geq C$ for all $d \leq b$, which means $P(\xi \in S_{b,l} | \xi \in R_{b,l}) \geq C$. Since $S_{b,l} \subset R_{b,l}$, we have $P(\xi \in S_{b,l}) \geq CP(\xi \in R_{b,l})$, and therefore $P_x(S_{b,l}) \geq CP_x(R_{b,l})$ for all $x \in \Delta^\epsilon$. Also, if $z_{ij} \in R_{b,l}$, then since at least $b-1$ of the first b coordinates of z_{ij} are distinct, we must have $z_{ij} \in S_{b,l}$. Thus, since $C \leq 1$, we have

$$L(S_{b,l}) \geq \int_\Delta CP_x(R_{b,l}) \Big/ \sum_{j=1}^\infty x_j^2 \, \Xi_0(dx) + a \sum_{i=1}^\infty \sum_{j=i+1}^\infty 1_{\{z_{ij} \in R_{b,l}\}} \geq CL(R_{b,l})$$

for all b and l such that $b > M^l$.

Now fix $n \in \mathbb{N}$. For l such that $M^l \leq n$, let D_l be the event that $M^{l-1}+1 \leq \#\Pi_n(t) \leq M^l$ for some t. Assume for now that $M^l < n$. For all $b \geq 1$, let $U_b = \inf\{t : \#\Pi_n(t) \leq b\}$. When $\Pi_n(t)$ has b blocks and $b > M^l$, the total rate of all collisions that take Π_n down to M^l or fewer blocks is $L(R_{b,l})$. The total rate of all collisions that take Π_n down to between $M^{l-1}+1$ and M^l blocks is $L(S_{b,l})$. Therefore, for all $b > M^l$, the strong Markov property and part (d) of Lemma 41 of appendix B imply that if $P(\#\Pi_n(U_b) = b) > 0$, then

$$P(\#\Pi_n(U_{b-1}) \leq M^l | \#\Pi_n(U_b) = b) = L(R_{b,l})/\lambda_b \tag{77}$$

and

$$P(M^{l-1}+1 \leq \#\Pi_n(U_{b-1}) \leq M^l | \#\Pi_n(U_b) = b) = L(S_{b,l})/\lambda_b. \tag{78}$$

Note that the event $\{\#\Pi_n(U_{M^l}-) = b\}$ is the same as $\{\#\Pi_n(U_b) = b\} \cap \{\#\Pi_n(U_{b-1}) \leq M^l\}$. Also, the events $D_l \cap \{\#\Pi_n(U_{M^l}-) = b\}$ and $\{\#\Pi_n(U_b) = b\} \cap \{M^{l-1}+1 \leq \#\Pi_n(U_{b-1}) \leq M^l\}$ are the same. Therefore, by (77) and (78), we have $P(D_l | \#\Pi_n(U_{M^l}-) = b) = L(S_{b,l})/L(R_{b,l}) \geq C$ for all $b > M^l$ such that $P(\#\Pi_n(U_{M^l}-) = b) > 0$. Hence, $P(D_l) \geq C$ whenever $M^l < n$. Also, we have $P(D_l) = 1 \geq C$ when $M^l = n$.

As in the proof of Lemma 6 of [19], we recursively define times $R_0, R_1, \dots, R_{n-1}$ by:

$$\begin{aligned}
&R_0 = 0 \\
&R_i = \inf\{t : \#\Pi_n(t) < \#\Pi_n(R_{i-1})\} && \text{if } i \geq 1 \text{ and } \#\Pi_n(R_{i-1}) > 1. \\
&R_i = R_{i-1} && \text{if } i \geq 1 \text{ and } \#\Pi_n(R_{i-1}) = 1.
\end{aligned}$$

Note that $R_{n-1} = T_n$. For $i = 0, 1, \dots, n-1$, let $N_i = \#\Pi_n(R_i)$. For $i = 1, 2, \dots, n-1$, define $L_i = R_i - R_{i-1}$ and $J_i = N_{i-1} - N_i$. Suppose $n = M^k$ for some $k \in \mathbb{N}$. For $j = 2, 3, \dots, n$, let $L_n(j) = \min\{s \geq j : \#\Pi_n(t) = s \text{ for some } t\}$. If $N_{i-1} \geq j > N_i$, or equivalently if

$N_i + J_i \geq j > N_i$, then $L_n(j) = N_{i-1}$. Therefore, using equation (12) of [19], which is valid for the Ξ-coalescent as well as for the Λ-coalescent, to get the first equality, we have

$$E[T_n] = \sum_{i=1}^{n-1} E[\gamma_{N_{i-1}}^{-1} J_i] = \sum_{j=2}^{n} E[\gamma_{L_n(j)}^{-1}] = \sum_{l=1}^{k} \sum_{j=M^{l-1}+1}^{M^l} E[\gamma_{L_n(j)}^{-1}].$$

Since $(\gamma_b)_{b=2}^{\infty}$ is increasing by Lemma 28 and $L_n(j) \leq M^{l+1}$ on D_{l+1} for all $j \leq M^l$, we have

$$\begin{aligned} E[T_n] &\geq \sum_{l=1}^{k-1} \sum_{j=M^{l-1}+1}^{M^l} E[\gamma_{L_n(j)}^{-1}] \geq \sum_{l=1}^{k-1} \sum_{j=M^{l-1}+1}^{M^l} P(D_{l+1}) \gamma_{M^{l+1}}^{-1} \\ &\geq C \sum_{l=1}^{k-1} (M^l - M^{l-1}) \gamma_{M^{l+1}}^{-1} = \frac{C}{M^2} \sum_{l=1}^{k-1} (M^{l+2} - M^{l+1}) \gamma_{M^{l+1}}^{-1}. \end{aligned}$$

Therefore, using the monotonicity of $(T_n)_{n=1}^{\infty}$ for the first equality, we have

$$\begin{aligned} \lim_{n\to\infty} E[T_n] &= \lim_{k\to\infty} E[T_{M^k}] \geq \lim_{k\to\infty} \frac{C}{M^2} \sum_{l=1}^{k-1} (M^{l+2} - M^{l+1}) \gamma_{M^{l+1}}^{-1} \\ &= \frac{C}{M^2} \sum_{l=1}^{\infty} (M^{l+2} - M^{l+1}) \gamma_{M^{l+1}}^{-1} \geq \frac{C}{M^2} \sum_{b=M^2}^{\infty} \gamma_b^{-1} = \infty, \end{aligned}$$

which completes the proof. $\square$

Suppose there exists $K < \infty$ such that Ξ is concentrated on $\{x : x_i = 0 \text{ for all } i > K\}$. Then almost surely the Ξ-coalescent never undergoes more than K multiple collisions at one time. Note that if $\sum_{i=1}^{K} x_i > 1/2$, then $\sum_{i=1}^{\infty} x_i^2 \geq \max_{1 \leq i \leq K} x_i^2 \geq 1/4K^2$. Therefore, if $\epsilon = 1/2$ then the left-hand side of (76) is at most $4K^2 \Xi(\Delta) < \infty$, so (76) holds. Thus, Theorem 1 of [19] can be deduced from Propositions 32 and 33. However, (76) can fail if the Ξ-coalescent can undergo arbitrarily many collisions simultaneously.

To show that the conclusion of Proposition 33 does not necessarily hold when (76) fails, we give the following example of a Ξ-coalescent for which $\Xi(\Delta_f) = 0$ and $\sum_{b=2}^{\infty} \gamma_b^{-1} = \infty$ but the Ξ-coalescent comes down from infinity.

Example 34 For all $n \in \mathbb{N}$, let $y_n \in \Delta$ be the point whose first $2^n - 1$ coordinates equal 2^{-n} and whose remaining coordinates equal zero. Let Ξ be the probability measure on Δ with an atom of size 2^{-n} at y_n for all $n \in \mathbb{N}$. Define A_b and $A_{k,l}$ as in (15) and (16) respectively. For all $k \neq l$, we have $P_{y_n}(A_{k,l}) \leq 2^{-n}$, so $P_{y_n}(A_b) \leq \min\{1, b^2 2^{-n}\}$ for $b \geq 2$. Since Ξ has no atom at zero, (14) implies that for all $b \geq 2$ we have

$$\begin{aligned} \lambda_b &= L(A_b) = \int_\Delta P_x(A_b) / \sum_{j=1}^{\infty} x_j^2 \, \Xi(dx) = \sum_{n=1}^{\infty} \frac{1}{2^n} \left(\frac{P_{y_n}(A_b)}{(2^n - 1) 2^{-2n}} \right) \\ &= \sum_{n=1}^{\infty} \frac{2^n}{2^n - 1} P_{y_n}(A_b) \leq \sum_{n=1}^{\infty} 2 \min\{1, b^2 2^{-n}\} \leq 2 \left(2 \log_2 b + \sum_{n=\lceil 2 \log_2 b \rceil}^{\infty} b^2 2^{-n} \right) \\ &= 2(2 \log_2 b + b^2 2^{-(\lceil 2 \log_2 b \rceil - 1)}) \leq 4(\log_2 b + b^2 2^{-2 \log_2 b}) = 4(\log_2 b + 1) \leq 8 \log_2 b. \end{aligned}$$

We have $\gamma_b \leq b\lambda_b$ for all $b \geq 2$ by (69) and (71). Therefore,

$$\sum_{b=2}^{\infty} \gamma_b^{-1} \geq \sum_{b=2}^{\infty} (b\lambda_b)^{-1} \geq \sum_{b=2}^{\infty} (8b \log_2 b)^{-1} = \infty.$$

We now show that the Ξ-coalescent comes down from infinity. For all $b \geq 2$, define the set $V_b = \{\xi \in \mathbb{Z}^\infty : |\{\xi_1, \dots, \xi_b\}| \leq b^{3/4}\}$. Note that if $\xi = (\xi_i)_{i=1}^\infty$ has distribution P_{y_n}, then $P(\xi_i < 0) = 2^{-n}$ for all i. Therefore, under P_{y_n}, the expected number of $\xi_1, \dots, \xi_b$ that are negative is $2^{-n}b$. Choose $M \geq 64$ large enough that if $b \geq M$ then $b^{2/3} \leq (b^{3/4} - b^{2/3})/2$. Suppose $b \geq M$ and $b^{1/3} \leq 2^n \leq b^{2/3}$. Then

$$2^{-n}b \leq b^{2/3} \leq \frac{1}{2}(b^{3/4} - b^{2/3}) \leq \frac{1}{2}(b^{3/4} - (2^n - 1)).$$

Therefore, by Markov's inequality, we have

$$P_{y_n}(\{\xi \in \mathbb{Z}^\infty : |\{i \leq b : \xi_i < 0\}| \leq b^{3/4} - (2^n - 1)\}) \geq 1/2.$$

Under P_{y_n} the random variables ξ_i can take on only $2^n - 1$ different positive values. Therefore, $P_{y_n}(V_b) \geq 1/2$ if $b \geq M$ and $b^{1/3} \leq 2^n \leq b^{2/3}$. Thus, if a Ξ-coalescent has $b \geq M$ blocks, then the total rate of all collisions that take the coalescent down to $b^{3/4}$ or fewer blocks is

$$\begin{aligned} L(V_b) &= \int_\Delta P_x(V_b) / \sum_{j=1}^{\infty} x_j^2 \, \Xi(dx) \geq \sum_{n=\lceil (1/3) \log_2 b \rceil}^{\lfloor (2/3) \log_2 b \rfloor} \frac{1}{2^n} \left(\frac{P_{y_n}(V_b)}{(2^n - 1)2^{-2n}} \right) \\ &\geq \frac{1}{2} \left(\frac{1}{3} \log_2 b - 1 \right) \geq \frac{1}{12} \log_2 b, \end{aligned}$$

where for the last inequality we used the fact that $b \geq 64$. For $k \in \mathbb{N}$, let

$$S_k = \{m \in \mathbb{N} : M^{(4/3)^{k-1}} \leq m < M^{(4/3)^k}\}.$$

Then, if a Ξ-coalescent has $b \in S_k$ blocks, the expected time before the number of blocks is no longer in S_k is at most

$$\left(\frac{1}{12} \log_2 M^{(4/3)^{k-1}} \right)^{-1} = 12 \left(\frac{3}{4} \right)^{k-1} (\log_2 M)^{-1}.$$

For a standard Ξ-coalescent Π_∞, let $T_n = \inf\{t : \# R_n \Pi_\infty(t) = 1\}$. Then, for all n,

$$E[T_n] \leq \sum_{k=1}^{\infty} 12 \left(\frac{3}{4} \right)^{k-1} (\log_2 M)^{-1} + E[T_M] = \frac{48}{\log_2 M} + E[T_M] < \infty.$$

Thus, $(E[T_n])_{n=1}^\infty$ is bounded, which means $E[T_\infty] < \infty$. Hence, by Lemma 31, the Ξ-coalescent comes down from infinity.

6 The discrete-time Ξ-coalescent

We have shown that the continuous-time processes $(\Psi_{n,\infty}(t))_{t\geq 0}$ obtained in [13] as limits of ancestral process all have the same distribution as some Ξ-coalescent restricted to $\{1,\dots,n\}$. However, Möhle and Sagitov also obtain some discrete-time Markov chains as limits. They show, as part of Theorem 2.1 of [13], that if the conditions of Proposition 1 are satisfied except that $\lim_{N\to\infty} c_N = c > 0$, then the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t\geq 0}$ converge as $N \to \infty$ to a Markov chain $(\Psi_{n,\infty}(t))_{t\geq 0}$ that jumps only at times $t = cm$ for $m \in \mathbb{N}$. Suppose η and θ are partitions of $\{1,\dots,n\}$ such that θ contains b blocks, s blocks of η consist of a single block of θ, and the remaining r blocks of η are unions of $k_1,\dots,k_r$ blocks of θ. Theorem 2.1 of [13] then states that

$$P(\Psi_{n,\infty}(c(m+1)) = \eta | \Psi_{n,\infty}(cm) = \theta) = c\lambda_{b;k_1,\dots,k_r;s} \tag{79}$$

for all $m \in \mathbb{N}$, where each $\lambda_{b;k_1,\dots,k_r;s}$ is defined by (6) for a unique sequence of measures $(F_r)_{r=1}^{\infty}$ satisfying conditions A1, A2, and A3 of Proposition 1. In section 5 of [13], Möhle and Sagitov give an example in which the discrete-time ancestral process of the Wright-Fisher population model arises as a limit in this way.

Regarding discrete-time processes of this type, we have the following analog of Theorem 2.

Proposition 35 *Let $\{p_{b;k_1,\dots,k_r;s} : r \geq 1,\ k_1,\dots,k_r \geq 2,\ s \geq 0,\ b = \sum_{j=1}^{r} k_j + s\}$ be a collection of nonnegative real numbers. Then there exists a $\mathcal{P}_\infty$-valued Markov chain $(Y_m)_{m=0}^{\infty}$ satisfying*

C1: Y_0 is the partition of $\mathbb{N}$ into singletons,

C2: for each n, $(R_n Y_m)_{m=0}^{\infty}$ is a Markov chain with the property that if η and θ are distinct partitions of $\{1,\dots n\}$ such that θ contains b blocks, s blocks of η consist of a single block of θ, and the remaining r blocks of η are unions of $k_1,\dots,k_r \geq 2$ blocks of θ, then $P(R_n Y_{m+1} = \eta | R_n Y_m = \theta) = p_{b;k_1,\dots,k_r;s}$,

if and only if there is a finite measure Ξ on the infinite simplex Δ with no atom at zero such that

$$\int_\Delta 1/\sum_{j=1}^{\infty} x_j^2\, \Xi(dx) \leq 1 \tag{80}$$

and

$$p_{b;k_1,\dots,k_r;s} = \int_\Delta \left(\sum_{l=0}^{s} \sum_{i_1\neq\dots\neq i_{r+l}} \binom{s}{l} x_{i_1}^{k_1}\dots x_{i_r}^{k_r} x_{i_{r+1}}\dots x_{i_{r+l}} (1-\sum_{j=1}^{\infty} x_j)^{s-l} \right) \Big/ \sum_{j=1}^{\infty} x_j^2\, \Xi(d\mathbf{x}) \tag{81}$$

for all $r \geq 1$, $k_1,\dots,k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^{r} k_j + s$.

Proof. Suppose the nonnegative real numbers $p_{b;k_1,\dots,k_r;s}$ are defined such that there exists a Markov chain $(Y_m)_{m=0}^{\infty}$ satisfying C1 and C2. Let $J_0 = 0$, and let $(J_i)_{i=1}^{\infty}$ be a sequence of independent random variables, each having an exponential distribution with rate 1. For all $t \geq 0$, define $K_t = \max\{i : J_0 + J_1 + \dots + J_i \leq t\}$. Define a $\mathcal{P}_\infty$-valued process $(\Pi_\infty(t))_{t\geq 0}$ by $\Pi_\infty(t) = Y_{K_t}$. Then $\Pi_\infty(0)$ is the partition of $\mathbb{N}$ into singletons. Also, $(R_n\Pi_\infty(t))_{t\geq 0}$ is a jump-hold Markov process such that when $R_n\Pi_\infty(t)$ has b blocks, each $(b;k_1,\ldots,k_r)$-collision is occurring at rate $p_{b;k_1,\ldots,k_r;s}$. By Theorem 2, there exists a finite measure Ξ on Δ such that the rates $p_{b;k_1,\ldots,k_r;s}$ are given by (11). We also have

$$\sum_{\eta\neq\theta} P(R_nY_{m+1}=\eta|R_nY_m=\theta)\leq 1 \tag{82}$$

for all $n\in\mathbb{N}$ and $\theta\in\mathcal{P}_\infty$. If θ has b blocks, then the left-hand side of (82) equals

$$\sum_{r=1}^{\lfloor b/2\rfloor}\sum_{\{k_1,\ldots,k_r\}} N(b;k_1,\ldots,k_r;s)p_{b;k_1,\ldots,k_r;s},$$

which by (69) equals the total rate λ_b of collisions for a (continuous-time) Ξ-coalescent with b blocks. Thus, we have $\lambda_b\leq 1$ for all $b\geq 2$. Since $(\lambda_b)_{b=2}^\infty$ is increasing, we have $\lim_{b\to\infty}\lambda_b\leq 1$. Let L be defined from Ξ by (14). Then, $\lambda_b=L(A_b)$, where A_b is defined by (15). Since the sets A_b increase to the set A_∞ defined in (19), we have $\lim_{b\to\infty}\lambda_b=\lim_{b\to\infty}L(A_b)=L(A_\infty)$. The expression for $L(A_\infty)$ in (74) implies that $\lim_{b\to\infty}\lambda_b\leq 1$ if and only if Ξ has no atom at zero and (80) holds. Since Ξ has no atom at zero when (80) holds, the expression for $p_{b;k_1,\ldots,k_r;s}$ can be simplified to the right-hand side of (81).

Conversely, suppose there is a finite measure Ξ on the infinite simplex Δ with no atom at zero such that (80) and (81) hold. Let Π_∞ be a standard Ξ-coalescent, derived from a Poisson point process $(e(t))_{t\geq 0}$ with characteristic measure L. By (74) and (80), we have $L(A_\infty)\leq 1$. Therefore, if we define $T_0=0$ and

$$T_k=\inf\{t>T_{k-1}:e(t)\in A_\infty\}$$

for $k\geq 1$, then $0=T_0<T_1<\ldots$ a.s. by part (b) of Lemma 41 in appendix B. Define a Markov chain $(Z_m)_{m=0}^\infty$ by $Z_m=\Pi_\infty(T_m)$. Let η and θ be partitions of $\{1,\ldots n\}$ such that θ contains b blocks, s blocks of η consist of a single block of θ, and the remaining r blocks of η are unions of $k_1,\ldots,k_r$ blocks of θ. By the strong Markov property and part (d) of Lemma 41 in appendix B, we have

$$P(R_nZ_{m+1}=\eta|R_nZ_m=\theta)=p_{b;k_1,\ldots,k_r;s}/L(A_\infty).$$

Let $(I_i)_{i=1}^\infty$ be a sequence of independent Bernoulli random variables that take on the value 1 with probability $L(A_\infty)$. Assume the sequence is independent of $(Z_m)_{m=0}^\infty$. Let $V_m=I_1+\ldots+I_m$. Define a Markov chain $(Y_m)_{m=0}^\infty$ by $Y_m=Z_{V_m}$ for all m. Then

$$P(R_nY_{m+1}=\eta|R_nY_m=\theta)=L(A_\infty)P(R_nZ_{V_m+1}=\eta|R_nZ_{V_m}=\theta)=p_{b;k_1,\ldots,k_r;s},$$

which completes the proof. $\square$

Definition 36 We call a discrete-time Markov chain satisfying C2 with transition probabilities given by (81) for a particular finite measure Ξ a *discrete-time* Ξ*-coalescent*. A discrete-time Ξ-coalescent satisfying C1 is called a *standard discrete-time* Ξ*-coalescent*.

Propositions 29 and 35 imply that if Ξ is a finite measure on Δ for which a standard discrete-time Ξ-coalescent exists, then a standard (continuous-time) Ξ-coalescent Π_∞ is a jump-hold Markov process. Therefore, we can define the jump chain $(X_m)_{m=0}^\infty$ associated with Π_∞ by defining X_m to be the value of Π_∞ at the time of its mth jump, unless Π_∞ consists of just a single block after fewer than m jumps, in which case X_m is defined to be $\{\mathbb{N}\}$. Then $X_{m+1} \neq X_m$ a.s. on $\{\#X_m > 2\}$. This chain is different from the standard discrete-time Ξ-coalescent $(Y_m)_{m=0}^\infty$ because $P(Y_{m+1} \neq Y_m | \#Y_m = b) = \lambda_b$ for all $b \geq 2$ and $P(Y_{m+1} \neq Y_m | \#Y_m = \infty) = \lim_{b\to\infty} \lambda_b = L(A_\infty)$.

Finally, we prove the analog of Proposition 7 for discrete-time Ξ-coalescents. Note that although the time-scaling conventions in [13] are such that we can only obtain a continuous-time Ξ-coalescent as a limit of ancestral processes when Ξ is a probability measure, the proposition below shows that any nontrivial discrete-time Ξ-coalescent arises as a limit of ancestral processes in a population model of the type studied in [13].

Proposition 37 *Let Ξ be a finite nonzero measure on Δ with no atom at zero such that (80) holds. Then there exists a sequence $(\mu_N)_{N=1}^\infty$ such that each μ_N is a probability distribution on $\{0, 1, 2, \dots, \}^N$ that is exchangeable with respect to the N coordinates with the property that if for all N, μ_N is the distribution of family sizes in the population model described in the introduction, then the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t\geq 0}$ converge as $N \to \infty$ in the Skorohod topology to a process $(\Psi_{n,\infty}(t))_{t\geq 0}$ satisfying:*

(a) For all n, $(\Psi_{n,\infty}(t))_{t\geq 0}$ jumps only at times cm for $m \in \mathbb{N}$, where $c = \lim_{N\to\infty} c_N$.

(b) $(\Psi_{n,\infty}(cm))_{m=0}^\infty$ has the same distribution as $(R_n Y_m)_{m=0}^\infty$, where $(Y_m)_{m=0}^\infty$ is a standard discrete-time Ξ-coalescent.

Proof. Define the measure L from Ξ as in (14). Define A_N for $N \in \mathbb{N}$ by (15), and define A_∞ by (19). We have $\lim_{N\to\infty} L(A_N) = L(A_\infty)$. Also, $L(A_\infty) \leq 1$ by (74) and (80), and $L(A_\infty) > 0$ because Ξ is nonzero.

We now follow essentially the same argument used to prove Proposition 7. First, let Π_∞ be a continuous-time Ξ-coalescent. Define $\tilde{\nu}_{1,N}, \dots, \tilde{\nu}_{N,N}$ from Π_∞ as in the proof of Proposition 7. However, define V_N such that $P(V_N = 1) = L(A_\infty)$ and $P(V_N = 0) = 1 - L(A_\infty)$. Then define $(\nu_{1,N}, \dots, \nu_{N,N})$ from $(\tilde{\nu}_{1,N}, \dots, \tilde{\nu}_{N,N})$ and V_N as in the proof of Proposition 7.

Denote the transition probabilities for the discrete-time Ξ-coalescent by $p_{b;k_1,\dots,k_r;s}$. These are the same as the transition rates for the the continuous-time Ξ-coalescent. Therefore, we can follow the argument in the proof of Proposition 7 to see that (51) still holds in the setting of Proposition 37 when $\lambda_{b;k_1,\dots,k_r;0}$ is replaced by $p_{b;k_1,\dots,k_r;0}$ on the right-hand side. Equation (52) remains true if we make this change and also replace the factor of $1/N$ on the right-hand side, which comes from the definition of V_N, by $L(A_\infty)$. That is, we have

$$E[(\nu_{1,N})_{k_1} \dots (\nu_{r,N})_{k_r}] \sim \frac{L(A_\infty)N^{k_1+\dots+k_r-r}}{L(A_N)} p_{b;k_1,\dots,k_r;0} \sim N^{k_1+\dots+k_r-r} p_{b;k_1,\dots,k_r;0}. \tag{83}$$

Therefore,

$$c_N = \frac{E[(\nu_{1,N})_2]}{N-1} \sim \frac{N p_{2;2;0}}{N-1} \sim p_{2;2;0}, \tag{84}$$

so if $c = p_{2;2;0}$, then $\lim_{N\to\infty} c_N = c > 0$. By (83) and (84),

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_{k_1} \dots (\nu_{r,N})_{k_r}]}{N^{k_1+\dots+k_r-r} c_N} = \frac{p_{b;k_1,\dots,k_r;s}}{p_{2;2;0}}. \tag{85}$$

By the remarks at the beginning of this section, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t\geq 0}$ converge in the Skorohod topology as $N \to \infty$ to a process $(\Psi_{n,\infty}(t))_{t\geq 0}$ such that condition (a) of this proposition is satisfied and the transition probabilities of $(\Psi_{n,\infty}(cm))_{m=0}^{\infty}$ are given by (79), where each $\lambda_{b;k_1,\dots,k_r;s}$ is defined by (6) for a unique sequence of measures $(F_r)_{r=1}^{\infty}$ satisfying conditions A1, A2, and A3 of Proposition 1. We thus must show that $c\lambda_{b;k_1,\dots,k_r;s} = p_{b;k_1,\dots,k_r;s}$ for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$. It is shown in [13] that equation (10) remains valid when $\lim_{N\to\infty} c_N = c > 0$. Using (10) and (85), we obtain

$$c\lambda_{b;k_1,\dots,k_r;0} = c\left(\frac{p_{b;k_1,\dots,k_r;0}}{p_{2;2;0}}\right) = p_{b;k_1,\dots,k_r;0}$$

for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, and $b = \sum_{j=1}^r k_j$. Lemma 3.4 of [13] implies that (43) holds for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$. By Lemma 18, we also have

$$p_{b+1;k_1,\dots,k_r;s+1} = p_{b;k_1,\dots,k_r;s} - \sum_{m=1}^{r} p_{b+1;k_1,\dots,k_{m-1},k_m+1,k_{m+1},\dots,k_r;s} - s p_{b+1;k_1,\dots,k_r,2;s-1}.$$

Hence, by induction on s, we obtain $c\lambda_{b;k_1,\dots,k_r;s} = p_{b;k_1,\dots,k_r;s}$ for all $r \geq 1$, $k_1, \dots, k_r \geq 2$, $s \geq 0$, and $b = \sum_{j=1}^r k_j + s$, which completes the proof. □

APPENDIX A

Exchangeable Sequences and Exchangeable Random Partitions.

We review here some well-known results about exchangeable sequences and exchangeable random partitions. We first recall a version of de Finetti's Theorem, combining Proposition 12 and Theorem 13 in chapter 27 of [7].

Lemma 38 *Let $(Z_i)_{i=1}^{\infty}$ be an exchangeable sequence of $\mathbb{R}$-valued random variables. Then, there exists a random probability measure μ on the Borel subsets of $\mathbb{R}$, called the* limiting empirical distribution *of $(Z_i)_{i=1}^{\infty}$, such that for each Borel set B,*

$$\mu(B) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{Z_i \in B\}}$$

almost surely. Let $\mathcal{F}$ denote the σ-field generated by the random measure μ. Then the random variables $Z_1, Z_2, \dots$ are conditionally independent given $\mathcal{F}$, and the conditional distribution of each Z_i given $\mathcal{F}$ is μ.

The following lemma is part of Proposition 3.8 of [1].

Lemma 39 *Let $(Z_i)_{i=1}^{\infty}$ be an exchangeable sequence with limiting empirical distribution μ. Suppose V is a random variable such that $(V, Z_1, Z_2, \ldots)$ and $(V, Z_{\sigma(1)}, Z_{\sigma(2)}, \ldots)$ have the same distribution for all finite permutations σ of $\mathbb{N}$. Then $(Z_i)_{i=1}^{\infty}$ and V are conditionally independent given the σ-field generated by μ.*

Given a partition π of a finite or countable set S and a finite permutation σ of S, let $\hat{\sigma}\pi$ be the partition of S such that $\sigma(i)$ and $\sigma(j)$ are in the same block of $\hat{\sigma}\pi$ if and only if i and j are in the same block of π. Following [10], we say that a random partition Π of S is *exchangeable* if $\hat{\sigma}\Pi$ has the same distribution as Π for all finite permutations σ of S.

Given a point $x = (x_1, x_2, \ldots)$ in the infinite simplex Δ, let P^x be the distribution of a random partition Π obtained by first defining an i.i.d. sequence of random variables $(Z_i)_{i=1}^{\infty}$ such that $P(Z_i = j) = x_j$ for $j \geq 1$ and $P(Z_i = 0) = 1 - \sum_{j=1}^{\infty} x_j$, and then declaring i and j to be in the same block of Π if and only if $Z_i = Z_j \geq 1$. In [9] and [10], Kingman establishes that all exchangeable random partitions are mixtures of random partitions that can be constructed in this way. A simpler proof of Kingman's result, using de Finetti's Theorem, is given in section of 11 of [1]. We state below a version of this result, which is essentially Theorem 2 of [10].

Lemma 40 *Let Π be an exchangeable random partition of $\mathbb{N}$. Let B_i be the block of Π containing the integer i. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{\infty} 1_{\{j \in B_i\}}$$

exists almost surely and is called the limiting relative frequency *of the block B_i. Let $X_1 \geq X_2 \geq \ldots$ be the ranked sequence of limiting relative frequencies of distinct blocks of Π, where X_n is defined to be zero if Π has fewer than n blocks with nonzero limiting relative frequencies. Then $X = (X_1, X_2, \ldots)$ is almost surely in the infinite simplex Δ. Moreover, the conditional distribution of Π given X is P^X, and therefore*

$$P(\Pi \in B) = \int_{\Delta} P^x(B)\, G(dx)$$

for all Borel subsets B of $\mathcal{P}_{\infty}$, where G is the distribution of X.

In [10], Kingman defines $X_r = \lim_{n\to\infty} n^{-1}\lambda_r(n)$, where $\lambda_r(n)$ is the size of the rth-largest block of $R_n\Pi$. The observation that this definition is equivalent to the one given in Lemma 40 above is made in the introduction of [15].

APPENDIX B

Poisson Point Processes.

We review here some basic facts about Poisson point processes, most of which are stated in section 0.5 of [2]. Let L be a σ-finite measure on a Polish space E. We can construct a Poisson random measure X on $[0, \infty) \times E$ with intensity measure $\lambda \times L$, where λ denotes Lebesgue measure on $[0, \infty)$. Almost surely $X(\{t\} \times E)$ equals 0 or 1 for all $t \geq 0$. Therefore, we can define a process $(e(t))_{t \geq 0}$ taking values in $E \cup \{\delta\}$, where δ is an isolated point that we add to the state space, by defining $e(t) = \delta$ if $X(\{t\} \times E) = 0$ and $e(t) = x$ if the restriction of X to $\{t\} \times E$ is a unit mass at x. The process $(e(t))_{t \geq 0}$ is called a *Poisson point process with characteristic measure L*. We record below some useful facts about Poisson point processes.

Lemma 41 *Let A be a Borel subset of E, and let $T_A = \inf\{t : e(t) \in A\}$. Then $T_A = 0$ a.s. if $L(A) = \infty$. Suppose $0 < L(A) < \infty$. Then the following hold:*

(a) T_A has an exponential distribution with rate parameter $L(A)$.

(b) For all $t < \infty$, almost surely $e(s) \in A$ for only finitely many $s < t$.

(c) $e(T_A) \in A$ a.s.

(d) $e(T_A)$ is independent of T_A, and if B is a Borel subset of A, then $P(e(T_A) \in B) = L(B)/L(A)$.

(e) The process $(e'(t))_{t \geq 0}$ defined such that $e'(t) \in \delta$ if $e(t) \in A$ and $e'(t) = e(t)$ otherwise is a Poisson point process whose characteristic measure is the restriction of L to A^c. Also, $(e'(t))_{t \geq 0}$ is independent of $(T_A, e(T_A))$.

Proof. If $L(A) = \infty$ and $\epsilon > 0$, then $(\lambda \times L)([0,\epsilon) \times A) = \infty$, so almost surely $X([0,\epsilon) \times A) = \infty$. Thus, $e(t) \in A$ for some $t < \epsilon$ a.s., and so $T_A = 0$ a.s. Next, suppose $0 < L(A) < \infty$. Conditions (a), (d), and (e) are part of Proposition 2 in section 0.5 of [2]. To prove (b), fix $t < \infty$. Since $(\lambda \times L)([0,t) \times A) < \infty$, we have $X([0,t) \times A) < \infty$ a.s., so $e(s) \in A$ for only finitely many $s < t$. Finally, to prove (c), note that for all $\epsilon > 0$, we almost surely do not have $e(s) \in A$ for infinitely many values of s in $(T_A, T_A + \epsilon)$. The definition of T_A thus implies that $e(T_A) \in A$ a.s. $\square$

Acknowledgments

The author would like to thank his advisor Jim Pitman for many helpful discussions regarding this work and detailed comments on earlier drafts of this paper.

References

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

[2] J. Bertoin. *Lévy Processes.* Cambridge University Press, 1996.

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

[4] J. L. Doob. *Measure Theory.* Springer-Verlag, New York, 1994.

[5] R. Durrett. *Probability: Theory and Examples.* 2nd ed. Duxbury Press, Belmont, CA, 1996.

[6] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.

[7] B. Fristedt and L. Gray. *A Modern Approach to Probability Theory.* Birkhauser, Boston, 1997.

[8] N. L. Johnson and S. Kotz. *Discrete Distributions.* Wiley, New York, 1969.

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

[10] J. F. C. Kingman. The coalescent. *Stoch. Proc. Appl.*, 13: 235-248, 1982.

[11] J. F. C. Kingman. On the genealogy of large populations. In J. Gani and E. J. Hannan, editors, *Essays in Statistical Science, Papers in honour of P. A. P. Moran. J. Appl. Prob. Special Volume 19A*, 27-43, 1982.

[12] M. Möhle and S. Sagitov. A characterization of ancestral limit processes arising in haploid population genetics models. Berichte zur Stochastik und verwandten Gebieten, Johannes Gutenberg-Universität Mainz, 1999.

[13] M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. Available via `http://www.math.chalmers.se/~serik/coal.html`, 1999.

[14] M. Möhle and S. Sagitov. Coalescent patterns in exchangeable diploid population models. Berichte zur Stochastik und verwandten Gebieten, Johannes Gutenberg-Universität Mainz, 1999.

[15] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields*, 102:145-158, 1995.

[16] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27:1870-1902, 1999.

[17] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* 3rd ed. Springer-Verlag, Berlin, 1999.

[18] S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36: 1116-1125, 1999.

[19] J. Schweinsberg. A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.*, 5:1-11, 2000.